Université de Lille Nord de France
Université des Sciences et Technologies de Lille I
Ecole Doctorale Biologie-Santé de Lille

Thèse de doctorat

Présentée par

Stéphan HARDIVILLE

En vue de l'obtention du titre de

Docteur de l'Université de Lille I

Spécialité : Sciences de la Vie et de la Santé

La delta-lactoferrine : un facteur de transcription régulé par GlcNAcylation

Soutenue le 26 mars 2010 devant la commission d'examen :

Président :	Dr Jean-Claude MICHALSKI
Rapporteurs :	Dr Anca ROSEANU
	Dr Xiaoyong YANG
Examinateurs :	Dr Nathalie THERET
	Dr Joël MAZURIER
Directeur de thèse :	Pr Annick PIERCE

A la mémoire de mon arrière-grand-mère et de ma grand-mère,

A mes parents, mon frère et Pauline qui ont toujours cru en moi

« Le succès c'est d'aller d'échec en échec sans perdre son enthousiasme »
Sir Winston Churchill

Mes premiers remerciements vont tout d'abord au Pr. Annick Pierce qui a dirigé ces travaux. Merci de m'avoir confié ce sujet aussi passionnant et de m'avoir accompagné pendant ces quatre années. Je te remercie de m'avoir transmis ta passion pour la recherche et de me l'avoir fait partager jour après jour. Merci encore pour ton soutien, pour tes conseils avisés et les différents échanges tant scientifiques que culturels que tu m'as donnée l'opportunité de vivre.

Je tiens à remercier le Dr. Jean-Claude Michalski pour m'avoir accueilli dans son laboratoire et d'avoir accepté de présider mon jury de thèse.

Je remercie, également, le Dr. Joël Mazurier de m'avoir accueilli dans son groupe de recherche et d'avoir accepté d'être examinateur de mes travaux.

Je souhaite remercier le Dr. Anca Roseanu et le Dr. Xiaoyong Yang pour m'avoir fait l'honneur de faire partie de mon jury et d'avoir accepté d'être rapporteurs de mes travaux. Mes remerciements vont également au Dr. Nathalie Theret pour avoir accepté d'être examinatrice.

Je remercie tout particulièrement M^{me} Monique Benaïssa. Je te suis très reconnaissant de m'avoir formé à la biologie moléculaire et aussi pour ta gentillesse, pour les moments passés à faire la queue à « Ma campagne », mais également pour les « ça peut servir » qui servent !

Je tiens également à remercier le Dr. Christophe Mariller. Merci de m'avoir initié à la biochimie de protéines, pour ton aide et tes conseils.

Je souhaite remercier également Agnès, Dominique, Fabrice et Mathieu pour leur sympathie et leur soutien tout au long de ces années.

Je remercie le Pr. Hiroyuki Tsuda de m'avoir si bien accueilli dans son laboratoire et de m'avoir fait partager la culture japonaise. Mes remerciements vont également au Dr. David B. Alexander pour sa collaboration et pour m'avoir initié aux vecteurs adénoviraux.

Ce n’est pas tous les jours faciles de partager un bureau uniquement avec la « gente féminine ». Mais, je tiens à remercier chaleureusement Adeline, Audrey, Aurélie, Esthelle, Fanny, et Marie-Estelle. Merci pour votre amitié, votre bonne humeur et nos franches parties de rigolade. Je souhaite également la bienvenue à Soccoro. Je vous souhaite à toutes bonne chance pour l’avenir et j’espère que nos chemins continueront à se croiser…

Je remercie également le Dr. Yohann Demont pour son amitié et pour les midis détente au Quick. Je te souhaite toute la réussite que tu mérites.

Je tiens à exprimer toute ma gratitude à mes parents, à mon frère et à Pauline qui ont su me soutenir et m’encourager. Vous avez toujours été présents au cours de mes longues années d’études et sans vous je n’en serais pas arrivé là aujourd’hui. Une pensée particulière à Pauline pour sa patience et son stoïcisme envers les heures tardives passées au labo mais également pour la relecture de ce manuscrit qui pour elle pouvait paraître incompréhensible.

Table des matières

========== Table des matières

========== Table des matières

Publication 1 : Mariller, C., Benaïssa, M., Hardivillé, S., Breton, M., Pradelle, G., Mazurier, J., et Pierce, A. « Human delta-lactoferrin is a transcription factor which enhances Skp1 (S-phase kinase associated protein) gene expression », 2007, *FEBS J.* 274(8) : 2038-53

Préface

Préface

Ces travaux ont été réalisés sous la direction du Professeur Annick PIERCE au sein de l'équipe « Régulation de la réaction inflammatoire et du cancer par les glycoconjugués » du Dr. Joël MAZURIER, dans l'Unité de Glycobiologie Structurale et Fonctionnelle (UMR CNRS 8576 dirigée par le Dr. Jean-Claude MICHALSKI).

Cette étude a été financée par le Ministère de l'Education Nationale et de la Recherche sous la forme d'une allocation MENRT

La *O*-N-acétylglucosaminylation ou GlcNAcylation est une glycosylation atypique car elle est cytosolique et met en jeu le branchement dynamique et réversible d'un simple monosaccharide non ramifié. C'est une modification post-traductionnelle contrôlée par un couple unique d'enzymes : l'OGT transfert un résidu de N-acétylglucosamine sur l'hydroxyle d'une sérine ou d'une thréonine alors que l'OGA hydrolyse la liaison Ser/Thr-*O*-GlcNAc. Cette glycosylation, souvent comparée à la phosphorylation avec laquelle elle peut d'ailleurs entrer en compétition, est intimement dépendante des niveaux d'UDP-GlcNAc intracellulaires. Sous l'influence du métabolisme du glucose, *via* la voie de biosynthèse des hexosamines, ces derniers subissent des variations extrêmes en réponse à de nombreux facteurs environnementaux et nutritifs. La GlcNAcylation régit de nombreuses fonctions cellulaires régulant l'activité de facteurs de transcription, la demi-vie de protéines cytosoliques et nucléaires ainsi que le trafic nucléocytoplasmique. Plusieurs travaux rapportent une altération des profils de GlcNAcylation dans le cas des maladies neurodégénératives, du diabète et de certains cancers. Depuis sa découverte en 1984, par C.R. Torres et G.W. Hart, de nombreuses protéines ont été décrites comme porteuses de cette modification post-traductionnelle, notamment des protéines impliquées dans la carcinogénèse (c-Myc, v-erb-A, p53 et Hic1). Ce n'est, néanmoins, que récemment que les variations du taux de GlcNAcylation ont pu être reliées à la progression du cycle cellulaire et qu'une diminution de ce taux dans le cancer a pu être démontré. Cependant, à ce jour, les modalités de régulation de protéines impliquées dans le cancer sont encore mal comprises et de nouvelles cibles moléculaires restent à découvrir.

La delta-lactoferrine, un suppresseur de tumeur potentiel, a servi de modèle à cette étude. Le gène de la lactoferrine produit deux isoformes : une isoforme sécrétée appelée lactoferrine et une isoforme cytosolique appelée delta-lactoferrine. Depuis sa découverte en 1960, la lactoferrine a été décrite pour avoir de nombreuses propriétés telles que des activités anti-prolifératives, anti-inflammatoires et anti-oxydantes. Différents travaux montrent que les activités anti-cancéreuses de cette molécule s'exercent par sa capacité à activer certaines cellules immunitaires clés dans la lutte de l'organisme contre le cancer. Cependant, d'autres études, réalisées par l'équipe du Dr. Son ou au laboratoire, tendent à prouver que certaines voies cellulaires anti-prolifératives sont activées en présence de lactoferrine et ce en condition de monoculture donc sans recours aux cellules du système immunitaire. En outre, lors de son endocytose, il a été démontré que la lactoferrine, associée à la nucléoline (son récepteur soluble), est capable de passer la membrane et de se retrouver dans le cytoplasme cellulaire. Elle est, alors, co-localisée avec la nucléoline dans le noyau où elle pourrait interagir avec des séquences nucléotidiques spécifiques, décrites par He et Furmanski, retrouvées dans le

promoteur de certains gènes. En 1997, Siebert et Huang ont décrit pour la première fois la delta-lactoferrine. De part sa localisation cytoplasmique, cette isoforme pourrait également interagir avec l'ADN et induire un effet anti-prolifératif comme celui observé avec la lactoferrine. Récemment, plusieurs études ont montré que la delta-lactoferrine, tout comme son homologue secrété, est réprimée dans les cas de cancer et que cette répression peut être associée à des aberrations chromosomiques ou à une méthylation des promoteurs du gène de la lactoferrine. Par ailleurs, son expression est corrélée à un facteur de bon pronostic dans les cas de cancer du sein. Cette activité de la delta-lactoferrine a trouvé un début d'explication depuis qu'il a pu être démontré que sa surexpression conduit à un arrêt du cycle cellulaire après 24 heures.

A mon arrivée en Master II Recherche au laboratoire, en 2005, il paraissait clair que la delta-lactoferrine était un facteur de transcription mais plusieurs points restaient encore mal compris. Bien que les travaux en cours montraient, sans équivoque, que cette isoforme de la lactoferrine était capable d'induire la surexpression de la protéine Skp1 *via* la présence d'un élément de réponse spécifique dans son promoteur, la fixation de la delta-lactoferrine sur l'ADN ainsi que son adressage nucléaire demeuraient méconnus. De plus, l'analyse de sa séquence primaire avait permis de mettre en évidence la présence de très nombreux sites de phosphorylation et de quatre sites de GlcNAcylation putatifs. La présence de sites potentiels de modifications post-traductionnelles laisse penser que la delta-lactoferrine devrait être hautement régulée. Le Pr. Annick PIERCE m'a chargé d'étudier les modalités de l'adressage nucléaire de la delta-lactoferrine et de démontrer sa fixation à l'ADN *in vivo*, de confirmer son rôle de facteur de transcription par la description de nouveaux gènes cibles mais surtout de mieux comprendre le mode de régulation de cette protéine notamment par la GlcNAcylation. D'autre part, la difficulté d'isoler en grande quantité la delta-lactoferrine m'a conduit à m'intéresser aux mécanismes de sa dégradation.

Après avoir passé en revue les connaissances actuelles concernant les lactoferrines, la GlcNAcylation des protéines et l'ubiquitinylation, puisque la delta-lactoferrine est également ubiquitinylée, je détaillerai les résultats obtenus au cours de mon année de Master II Recherche et de mes trois années de thèse. Le chapitre « travaux personnels » sera composé de deux parties : une partie traitant des cibles transcriptionnelles de la delta-lactoferrine et une partie traitant de la régulation de son activité et de sa dégradation par la GlcNAcylation.

Dans la première partie, nous verrons que la delta-lactoferrine peut agir sur trois phénotypes cellulaires distincts grâce à son activité de facteur de transcription : la progression du cycle cellulaire, la dégradation des ARNm et l'apoptose. En effet, nous avons pu

démontrer que l'expression transitoire de la delta-lactoferrine dans différentes lignées cellulaires induit une surexpression des protéines Skp1 (cycle cellulaire), DcpS (dégradation des ARNm) et Bax (apoptose). Cette surexpression se fait *via* la fixation de la delta-lactoferrine sur son élément de réponse présent au sein des promoteurs de ces trois gènes.

Dans la seconde partie, nous verrons que la delta-lactoferrine subit différentes modifications post-traductionnelles : GlcNAcylation, phosphorylation et ubiquitinylation. En effet, nous avons pu démontrer que parmi les quatre sites de *O*-GlcNAc prédits seuls trois sont modifiés et qu'en plus des sites de GlcNAcylation, la delta-lactoferrine est ubiquitinylée sur les résidus de lysine 379 et 391. Par ailleurs, la mise en place de ces modifications post-traductionnelles permet une régulation fine de l'activité transcriptionnelle et de la stabilité de la delta-lactoferrine. Bien que distant de sa séquence de dégradation PEST, le site sérine 10 est au cœur de ces régulations. En effet, la glycosylation de ce site confère à la protéine une protection face à la dégradation protéasomale alors que sa phosphorylation antagoniste conduit à la transactivation de ses gènes cibles.

Au cours de ces différents travaux, nous avons également pu mieux définir certains domaines structuraux de la delta-lactoferrine. En effet, nous avons pu clore de manière négative le débat sur la présence du domaine d'interaction à l'ADN en position N-terminale de la delta-lactoferrine. Mais, nous avons surtout pu mettre en évidence la présence sur le lobe C-terminal d'une séquence NLS bipartite fonctionnelle ainsi que d'une séquence de dégradation rapide PEST.

L'ensemble des travaux que j'ai mené au cours de ma thèse a fait l'objet des publications et communications suivantes :

- **Publications soumises à comité de lecture**

1. Mariller, C., Benaïssa, M., **Hardivillé, S**., Breton, M., Pradelle, G., Mazurier, J., et Pierce, A. « Human delta-lactoferrin is a transcription factor which enhances Skp1 (S-phase kinase associated protein) gene expression », 2007, *FEBS J.* 274(8) : 2038-53
2. Mariller, C., **Hardivillé, S.**, Hoedt, E., Benaïssa, M., Mazurier, J., et Pierce, A. « Proteomic approach to the identification of novel delta-lactoferrin target genes : characterization of DcpS, an mRNA scavenger decapping enzyme», 2009, *Biochimie* 91(01) : 109-22
3. Long, N., **Hardivillé, S**., Shirai, T., Pierce, A., Fukamachi, K., Futakuchi, M., Alexander, D. B., et Tsuda, H. « Construction of a Multi-Functional Helper-Dependent Adenovirus

Based System », 2009, *Asian Pac J Cancer Prev* 10(5) : 939-60

4. Hoedt, E., **Hardivillé, S**., Mariller, C., Elass, E., Perraudin, J.-P., et Pierce, A., « Discrimination and evaluation of lactoferrin and delta-lactoferrin gene expression levels in cancer cells and under inflammatory stimuli using TaqMan real time PCR », 2010, *Biometals*. DOI 10.1007
5. **Hardivillé, S.**, Hoedt, E., Mariller, C., Benaïssa, M., et Pierce, A. « A cycling between *O*-GlcNAcylation and phosphorylation at Ser10 controlled both delta-lactoferrin transcriptional activity and stability », en révision pour *J. Biol. Chem*. (2010)
6. **Hardivillé, S**., Hoedt, E., Mariller, C., Benaïssa, M., et Pierce, A., « Delta-lactoferrin induces cell death *via* the mitochondria death signaling pathway by upregulating Bax expression », en préparation pour *Eur. J. Cell. Biol. Chem.* (2010)

- **Publications sans comité de lecture**

1. Mariller, C., Benaïssa, M., **Hardivillé, S.**, Breton, M., Pradelle, G., Perraudin, J.-P., Mazurier, J., et Pierce, A. « Human delta-lactoferrin regulates cell cycle progression », 2007, in « Lactoferrin : New perspectives of lactoferrin research and application » Proceedings of the 2nd Lactoferrin Forum, Tokyo, Japan
2. Mariller, C., **Hardivillé, S**., Hoedt, E., Benaïssa, M., Perraudin, J.-P., Mazurier, J. et Pierce, A., « Identification of DcpS, an mRNA scavenger decapping enzyme, as a novel delta-lactoferrin target gene », 2009, in « Lactoferrin 2009 » Proceedings of the 3rd Lactoferrin Forum, Nihon-Igakukan Edition, Tokyo, Japan
3. Pierce, A., **Hardivillé, S**., Hoedt, S., Mariller, S., Benaïssa, M., Legrand, D., et Mazurier J., « Structure-function of lactoferrin isoforms », 2009, in « Lactoferrin 2009 » Proceedings of the 3rd Lactoferrin Forum, Édition Nihon-Igakukan, Tokyo, Japan

- **Communications orales**

1. Mariller C., **Hardivillé, S.,** Hoedt, E., Benaïssa, M., Mazurier, J., et Pierce, A, « Delta-lactoferrin is a transcription factor regulating cell progression and survival », 8th International Conference on Lactoferrin : Structure, fonction and applications, Nice, France, 21 au 25 octobre 2007
2. **Hardivillé, S.**, Benaïssa, M., Mariller, C., Mazurier, J., et Pierce, A., « Regulation of delta-lactoferrin activity by the *O*-linked-β-N-acetylglucosamine/phosphorylation interplay », 18th Joint Meeting, Lesquin, France, 4 au 6 novembre 2007
3. **Hardivillé, S.**, Benaïssa, M., Mariller, C., Mazurier, J., et Pierce, A., « Delta-lactoferrin a

putative tumor suppressor gene rules by *O*-GlcNAc/Phosphorylation interplay », plenary conference at Nagoya City University, Graduate School of Medical Sciences, Japon, 15 septembre 2008

4. Pierce, A., **Hardivillé, S.**, Mariller, C., Benaïssa, M., et Hoedt, E., « Human delta-lactoferrin is a transcription factor that regulates cell survival », 9th International Conference on Lactoferrin : Structure, fonction and applications, Pekin, Chine, 18 au 22 octobre 2009
5. **Hardivillé, S.,** Hoedt, E., Mariller, C., Benaïssa, M., Mazurier, J., et Pierce, A., « The *O*-GlcNAc/phosphate interplay regulates both delta-lactoferrin stability and transcriptional activity », 9th International Conference on Lactoferrin : Structure, fonction and applications, Pekin, Chine, 18 au 22 octobre 2009 (Obtention d'un financement « Student Awards » et Lauréat du « Geneviève Spik Award » récompensant, parmi les lauréats des « Student Awards », la meilleure présentation orale)

- **Communications par affiche**

1. **Hardivillé, S.**, Mariller, C., Benaïssa, M., Breton, M., Pradelle, G., Mazurier, J., et Pierce, A., « The transcriptional activity of human delta-lactoferrin is regulated by the *O*-*N*-acetylglucosaminylation/phosphorylation balance », 16th Meeting of Methods in Protein Structural Analysis, Villeneuve d'Ascq, 29 août au 2 septembre 2006

2. Mariller C., Benaïssa M., **Hardivillé, S.**, Breton M., Pradelle G., Perraudin J.-P., Mazurier, J., et Pierce, A, « Human delta-lactoferrin regulates cell cycle progression », 2nd Lactoferrin Forum, Tokyo, Japan, 25 novembre 2006

3. **Hardivillé, S.**, Benaïssa, M., Mariller, C., Hoedt, E., Mazurier, J., et Pierce, A., « Regulation of delta-lactoferrin activity by the *O*-linked-β-N-acetylglucosamine/phosphorylation interplay », 8th Lactoferrin Conference : Structure, fonction and applications, Nice, 21 au 25 octobre 2007

4. **Hardivillé, S.,** Benaïssa, M., Mariller, C., Hoedt, E., Mazurier, J., et Pierce, A., « Régulation de la stabilité de la delta-lactoferrine par la *O*-N-acétylglucosaminylation », 8ème Journée André VERBERT Colloque annuel des Doctorants, Lille, Faculté de Médecine, 18 septembre 2008

5. Hoedt, E., **Hardivillé, S.,** Germain, E., Mariller, C., Perraudin, J.-P., et Pierce, A., « The use of *TaqMan* real time PCR to discriminate and evaluate the variation in lactoferrine and delta-lactoferrin expression levels », 9th International Conference on Lactoferrin : Structure, fonction and applications, Pekin, Chine, 18 au 22 octobre 2009

Généralités

La delta-lactoferrine

La delta-lactoferrine est l'isoforme intracellulaire de la lactoferrine. Si la lactoferrine est connue depuis 50 ans, la delta-lactoferrine a été mise en évidence pour la première fois en 1997 par Siebert et Huang. Leur gène est issu d'une duplication du gène de la transferrine qui est apparue avec les mammifères (Baldwin, 1993). Les recherches réalisées depuis la découverte de la lactoferrine ont permis de montrer que cette protéine est multifonctionnelle (Figure 1). En effet, la lactoferrine a des effets immunomodulateurs, anti-cancéreux, anti-bactériens, sur l'homéostasie du fer ou encore sur la croissance osseuse (pour revues voir : Legrand *et al.*, 2008; Pierce *et al.*, 2009). Par contre, la delta-lactoferrine ne semblerait posséder qu'un rôle anti-prolifératif (Breton *et al.*, 2004). Bien que différents mécanismes permettant la régulation de la croissance cellulaire aient été décrits pour la lactoferrine, ceux de la delta-lactoferrine sont encore méconnus.

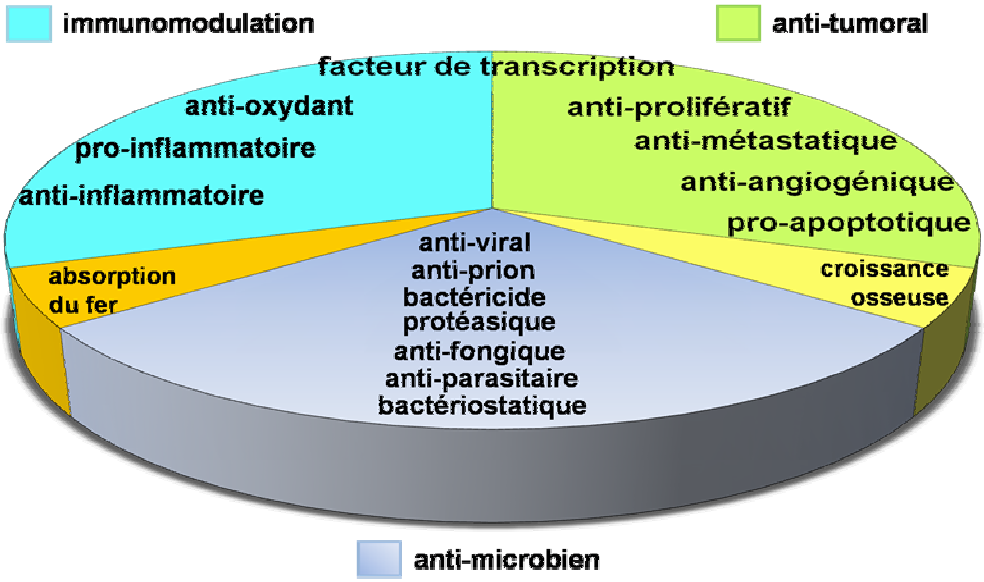


Figure 1 : Récapitulatif des différents rôles de la lactoferrine. Les rôles multiples de la lactoferrine sont regroupés par catégories : rôles anti-microbiens, immunomodulateurs, anti-cancéreux, dans l'absorption du fer et dans la croissance osseuse (d'après Pierce et al., 2009).

I. Localisation et biosynthèse des lactoferrines

A. Gène et régulation génique des lactoferrines

a) Le gène de la lactoferrine

Le gène de la lactoferrine humaine, localisé sur le chromosome 3 (en q21.31) (McCombs *et al.*, 1988), a une longueur totale de 29 kpb (Kim *et al.*, 1998; Siebert et Huang, 1997). Il est fragmenté en 17 exons de 48 à 210 pb et 16 introns de 300 pb à 3,3 kpb. Il est constitué de deux promoteurs : un promoteur P1 conduisant à la lactoferrine de sécrétion et un promoteur alternatif P2 permettant la synthèse de la delta-lactoferrine (Figure 2 A). Deux transcrits sont donc issus du gène de la lactoferrine humaine : l'ARNm de la lactoferrine et l'ARNm de la delta-lactoferrine. Des expériences de 5'RACE ont mis en évidence

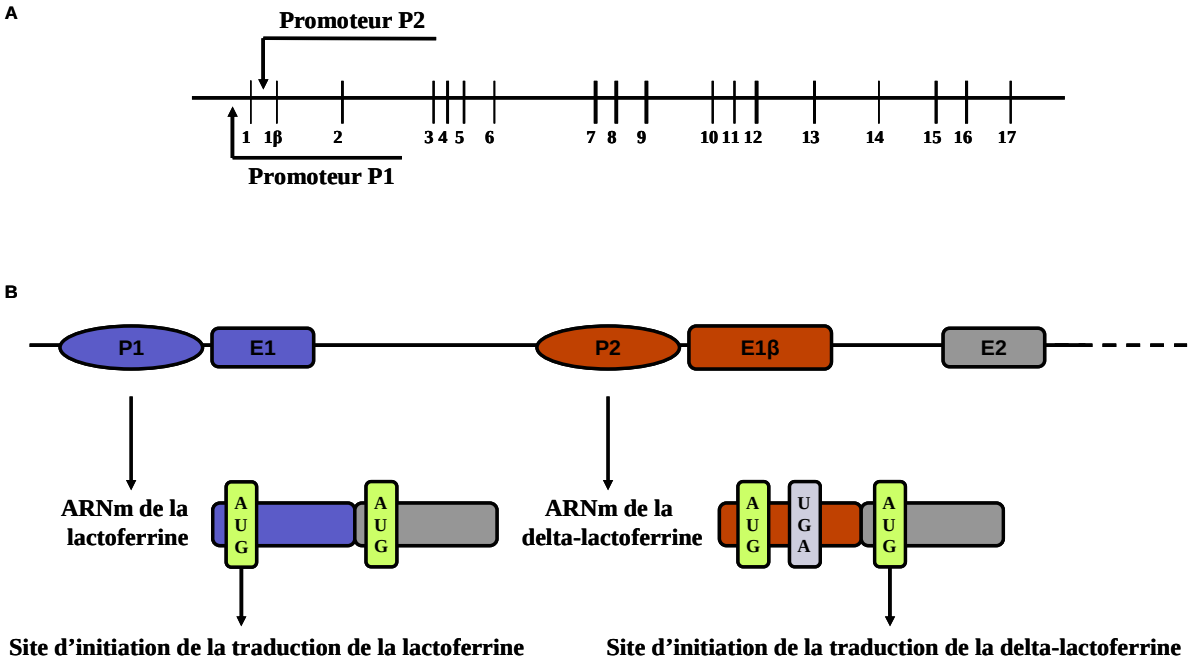


Figure 2 : Représentation schématique du gène des lactoferrines humaines et de son « épissage » alternatif à partir du promoteur P1 et du promoteur P2. (A) Le gène de la lactoferrine est constitué de 17 exons et de 16 introns dont la transcription est sous la dépendance de deux promoteurs distincts. (B) La transcription du gène de la lactoferrine à partir du promoteur P1 permet la synthèse de l'isoforme sécrétée grâce à la traduction du peptide signal encodé par l'exon 1 (Liu et al., 2003). Le promoteur alternatif P2 conduit à la synthèse d'une isoforme cytoplasmique appelée delta-lactoferrine (Siebert et Huang, 1997).

l'existence d'un troisième transcrit, dans les cellules de carcinome de poumon A549, nommé gamma-lactoferrine. Ce transcrit serait exprimé à partir d'un promoteur P3 localisé en amont du promoteur P1 (communication personnelle du Dr. David B. Alexander). Cependant, des travaux complémentaires seront nécessaires pour confirmer la présence de ce transcrit dans d'autres lignées cellulaires et surtout dans les tissus sains.

La lactoferrine et la delta-lactoferrine possèdent en commun les exons 2 à 17 et ne diffèrent qu'au niveau de leur premier exon (Liu *et al.*, 2003; Siebert et Huang, 1997). En effet, la transcription issue de l'utilisation du promoteur P2 remplace l'exon 1 par l'exon 1β plus long d'environ 100 pb et dont le codon d'initiation est immédiatement suivi d'un codon stop. Il existe, cependant, un site d'initiation de la traduction en phase avec le cadre de lecture ouvert de la lactoferrine dans l'exon 2. La traduction débute donc à partir de l'exon 2 conduisant à une protéine identique à la lactoferrine tronquée de ses 26 premiers résidus d'acides aminés (Siebert et Huang, 1997) (Figure 2 B, page 8). Contrairement à l'exon 1β, l'exon 1 permet la traduction d'une séquence signal de sécrétion. La lactoferrine, issue des transcrits comportant l'exon 1, est secrété alors que la delta-lactoferrine est cytoplasmique.

b) La régulation transcriptionnelle

La présence de ces deux promoteurs dans le gène de la lactoferrine humaine suggère qu'il puisse exister des mécanismes de régulation coordonnés et différentiels de leur transcription. Ainsi, Liu *et al.* ont montré en étudiant l'expression différentielle de gènes rapporteurs que ces deux promoteurs étaient régulés différemment selon les lignées cellulaires (Liu *et al.*, 2003). De plus, ils ont montré que le facteur de transcription Ets stimule différemment ces deux promoteurs. Il est à noter que la méthylation de ces deux promoteurs est différente, le promoteur P1 est sous-méthylé à la différence du promoteur P2, conduisant alors à une expression différentielle des deux transcrits (Teng *et al.*, 2004).

1. La régulation du promoteur P1

La région 5' flanquante du gène de la lactoferrine humaine contient plusieurs séquences promotrices situées juste en amont du site d'initiation de la transcription (Figure 3 A, page 11) : une séquence ATAAA (TATA box imparfaite) en position -32 à -27 pb, une séquence CAAT imparfaite (CAAC) en position -55 à -52 pb. A proximité de la TATA box, deux séquences GGAA fonctionnelles sont retrouvées pouvant servir à la fixation des facteurs de transcription Ets (Liu *et al.*, 2003). La région -26 à 1 pb est très riche en GC

(77%). Elle porte également des séquences de régulation de l'initiation de la transcription. Une séquence consensus pour Sp1 a été localisée de -50 à -45 pb ainsi qu'une PU box (Pu1/Sp1). Un module de réponse aux œstrogènes (Teng *et al.*, 1992), constitué d'un élément de réponse composite ERE/COUP, est présent en position -342 à -326 pb. Au niveau de cet élément, la régulation dépend de la compétition entre les facteurs de transcription spécifiques de ces séquences. En outre, le promoteur contient deux modules composites supplémentaires : le premier est un élément de réponse à l'acide rétinoïque (Lindberg *et al.*, 2003) chevauchant un élément de réponse aux oestrogènes (ERE) (Lee *et al.*, 1995), le deuxième est formé d'une séquence EGF-RE de réponse à l'EGF/TGFα et d'un élément CRE de réponse à l'AMPc. La présence d'un élément régulateur négatif dans le promoteur, fixant le facteur CDP/cut ("CCAAT displacement protein") (Khanna-Gupta *et al.*, 1997), permettrait la suppression de l'activité de base du promoteur de ce gène dans les cellules ne synthétisant pas de lactoferrine. Récemment, notre équipe a pu démontrer une surexpression des transcrits de la lactoferrine humaine en condition pro-inflammatoire (traitement des cellules aux LPS) (Hoedt *et al.*, 2010; annexe 1) suggérant que certains éléments de réponse putatifs cRel/NF-κB et STAT3, présents en amont du promoteur P1 et connus pour leurs implications dans la réponse inflammatoire (Baldwin, 2001; Leonard et O'Shea, 1998), seraient fonctionnels.

2. La régulation du promoteur P2

Les séquences régulatrices du promoteur P2 (Figure 3 B, page 11) ont été identifiées par Siebert et Huang (1997) puis confirmées par l'équipe de C. Teng (Liu *et al.*, 2003). Une boite TATA inversée se situe à -29 nucléotides du site d'initiation de la transcription. Des éléments potentiels de régulation spécifiques des mammifères ont été identifiés. Ils incluent une séquence consensus Sp1 et l'élément de réponse aux glucocorticoïdes (GRE), une séquence de liaison d'une protéine activatrice trouvée en amont des gènes NF-k, une boîte GA retrouvée dans le gène c-Myc, une boîte NF-E2, "CCAAT-binding protein", un site de liaison de l'IRF (Interferon Regulatory Factor) et des séquences de fixation pour les facteurs de transcription telles que les XRE (Xenobiotic Response Element), les ½ HRE (Hormone Response Element half site) et AML-1a (Acute Myeloid Leukemia-1). Trois éléments de réponse GGAA de fixation à Ets, dont la fonctionnalité a pu être démontrée, sont retrouvés à proximité de la TATA box (Liu *et al.*, 2003).

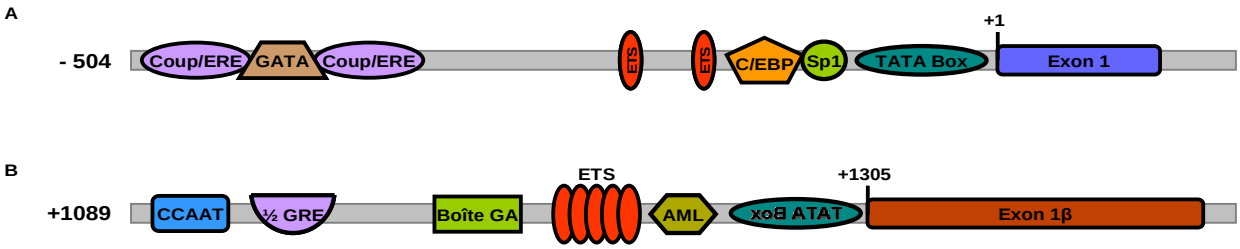


Figure 3 : Représentation schématique des différents éléments de réponse du promoteur P1 (A) (Lee et al., 1995; Teng et al., 1992) et du promoteur P2 (B) (Liu et al., 2003; Siebert et Huang, 1997) du gène de la lactoferrine humaine.

B. Localisation

a) La lactoferrine

La lactoferrine est présente dans la plupart des secrétions corporelles. Elle est retrouvée dans les sécrétions des organes reproducteurs, du système respiratoire, du tractus digestif et gastro-intestinal (Bennett *et al.*, 1973; Masson *et al.*, 1965; Masson *et al.*, 1968). La lactoferrine de sécrétion est issue des épithéliums glandulaires. Ainsi, la lactoferrine du lait est synthétisée par les cellules épithéliales des canaux glandulaires de la glande mammaire (Campbell *et al.*, 1992). Sa concentration dans le lait humain varie au cours de la lactation : le colostrum contient environ 6 à 7 $g.L^{-1}$ de lactoferrine, le lait de transition à une concentration d'environ 3,5 $g.L^{-1}$, alors que le lait mature n'en contient plus qu'environ 2 $g.L^{-1}$ (Hirai *et al.*, 1990). La lactoferrine présente dans les sécrétions vaginales est synthétisée par les cellules épithéliales glandulaires situées dans la couche basale de l'endomètre (Masson *et al.*, 1968). La lactoferrine est présente tout le long du tractus digestif et est produite par les glandes salivaires au niveau des canaux intercalaires, par les canaux striés et des acini séreux, par les cellules muqueuses gastriques de l'antre pylorique et du corps gastrique (Luqmani *et al.*, 1991), par l'épithélium glandulaire duodénal et par la vésicule biliaire (Saito et Nakanuma, 1992).

La lactoferrine est également retrouvée dans une moindre concentration dans le système circulatoire. Sa concentration dans les neutrophiles est comprise entre 70 et 170 µg pour 100 millions de cellules (Barton *et al.*, 1988). La lactoferrine est synthétisée au cours de la maturation des neutrophiles dans la moelle osseuse. Absente des promyélocytes, elle est transcrite dans les myélocytes (Rado *et al.*, 1984; Rado *et al.*, 1987). Une fois traduite, elle est

stockée dans les granules secondaires (Masson *et al.*, 1969) et tertiaires (Saito *et al.*, 1993) des neutrophiles. La lactoferrine est présente dans le plasma à faible concentration (environ 150 $\mu g.ml^{-1}$) (Maacks *et al.*, 1990). Sa concentration est en corrélation avec le nombre de neutrophiles circulants (Baynes *et al.*, 1986; Hansen *et al.*, 1975). Il est à noter que le taux plasmatique de la lactoferrine peut varier de manière significative et augmenter d'un facteur 7 en cas d'infection bactérienne (Maacks *et al.*, 1990). En outre, ce taux est sous l'influence des activités endocrines comme, par exemple, au cours de la grossesse (Sykes *et al.*, 1982).

La lactoferrine est également retrouvée dans le système nerveux central et plus particulièrement dans la substance grise. La protéine est produite par les cellules gliales et est retrouvée dans les neurones dopaminergiques (Fillebeen *et al.*, 2001). Cependant, les neurones dopaminergiques n'expriment pas les ARNm de la lactoferrine mais il apparaît que ces derniers endocyteraient la lactoferrine sécrétée par les cellules gliales. Le taux de lactoferrine cérébrale augmente avec l'âge mais également dans le cas des pathologies neurodégénératives (Kawamata *et al.*, 1993; Leveugle *et al.*, 1994).

La lactoferrine a également été décrite chez différentes espèces de mammifères comme le cochon d'Inde, la vache, la chèvre, le porc, la jument et la souris. Cependant, elle est absente de certaines autres espèces telles que le rat, le lapin et le chien (Masson et Heremans, 1971) (Tableau 1).

Tableau 1 : Concentration en lactoferrine dans le lait de différentes espèces (Teng, 2002)

Concentration dans le lait ($mg.mL^{-1}$)			
1-5	0.1-1	0.01-0.1	<0.05
Homme	Cochon d'Inde	Vache	Rat
	Souris	Chèvre	Lapin
	Jument	Truie	Chien

b) La delta-lactoferrine

De part l'absence de peptide signal, la delta-lactoferrine est localisée dans le cytoplasme. Sa faible expression ainsi que l'absence d'anticorps performant permettant sa détection font que cette isoforme est plus difficile à détecter. En effet, la différence entre la séquence polypeptidique de la delta-lactoferrine et celle de la lactoferrine réside uniquement dans la perte des 26 premiers résidus d'acides aminés pour la delta-lactoferrine ce qui ne permet pas de disposer d'anticorps spécifiques dirigés exclusivement contre la

delta-lactoferrine et donc de différencier les deux isoformes. L'expression simultanée des deux protéines dans certains tissus complique la mise en évidence spécifique d'une des isoformes. Cependant, les techniques de biologie moléculaire permettent de les différencier grâce à leur premier exon qui est spécifique à chaque isoforme. Les transcrits de la delta-lactoferrine sont exprimés par tous les tissus sains analysés impliquant une expression ubiquitaire (Siebert et Huang, 1997). En effet, ces transcrits sont retrouvés en quantité variable dans la glande mammaire, les ovaires, l'utérus, le placenta, les testicules, la prostate, le foie adulte et fœtal, l'intestin grêle, le pancréas, le thymus, les leucocytes, les reins, la rate, le cerveau adulte et fœtal, les poumons et les muscles squelettiques. Une étude récente réalisée au laboratoire a permis de montrer, par la technologie de qPCR TaqMan, que la delta-lactoferrine est très fortement exprimée dans les cellulaires immunitaires tels que les leucocytes et les lymphocytes (Hoedt *et al.*, 2010; annexe 1). En outre, bien que la lactoferrine soit également exprimée par la plupart de ces tissus, il ne paraît pas y avoir de corrélation avec l'expression de la delta-lactoferrine ce qui suggère une expression différentielle de chaque isoforme (Hoedt *et al.*, 2010; Siebert et Huang, 1997).

II. Structure des lactoferrines

La lactoferrine humaine est une glycoprotéine bilobée possédant une masse moléculaire apparente voisine de 80 kDa. Les séquences polypeptidiques de la lactoferrine et de la delta-lactoferrine étant identiques, les deux isoformes devraient avoir la même conformation spaciale même si la delta-lactoferrine ne devrait pas posséder de ponts disulfure du à l'environnement réducteur du cytoplasme (communication personnelle des Dr. Edouard et Heather Baker).

A. Structure primaire

Historiquement, la détermination de la séquence peptidique de lactoferrine humaine fut établie à partir de fragments issus de l'hydrolyse de la lactoferrine purifiée du lait (Metz-Boutigue *et al.*, 1984). Cette séquence a ensuite été confirmée par le clonage de l'ADNc en 1990 pour la lactoferrine (Powell et Ogden, 1990; Rey *et al.*, 1990) et en 1997 pour la delta-lactoferrine (Siebert et Huang, 1997). La lactoferrine de sécrétion est constituée de 692 résidus d'acides aminés et possède une masse de 80,6 kDa. La delta-lactoferrine, quant à elle, est formée de 666 résidus d'acides aminés pour une masse de 73 kDa. L'étude comparative de la séquence peptidique de la lactoferrine humaine et de la transferrine humaine met en

évidence 59 % d'homologie (Mazurier *et al.*, 1983; Metz-Boutigue *et al.*, 1984), ce qui suggère une origine commune aux deux protéines. De plus, il existe 37 % d'homologie entre les lobes N- et C-terminaux suggérant que lactoferrines et la transferrines sont issues de la duplication d'un gène ancestral (Lambert *et al.*, 2005; Metz-Boutigue *et al.*, 1984).

Les ADNc de la lactoferrine de sécrétion de nombreuses autres espèces de mammifères telles que la vache (Goodman et Schanbacher, 1991; Mead et Tweedie, 1990; Pierce *et al.*, 1991), le porc (Alexander *et al.*, 1992; Lydon *et al.*, 1992) ou la souris (Pentecost et Teng, 1987), par exemple, ont également été séquencés. L'alignement des séquences des lactoferrines de ces différentes espèces montre que le gène est conservé phylogénétiquement. En effet, les lactoferrines bovine et porcine présentent respectivement 69 et 70 % d'homologie avec la lactoferrine humaine (Lydon *et al.*, 1992; Pierce *et al.*, 1991).

La delta-lactoferrine n'a pas encore été décrite chez d'autres espèces que l'homme. Cependant, l'étude *in silico* des séquences introniques et promotrices du gène de la lactoferrine chez la vache suggère l'existence d'un promoteur contenu dans l'intron 1 ainsi qu'un exon 1β. Si ces séquences sont fonctionnelles, elles pourraient être à l'origine de la synthèse d'une delta-lactoferrine bovine. En outre, les résultats préliminaires que j'ai obtenus lors d'une étude chez la souris laissent penser qu'il existe une forme d'ARNm de la lactoferrine amputée de l'exon 1 (Figure 4). En effet, mes résultats de 5'RACE, permettant de connaître la séquence en 5'd'un transcrit, ont montré l'existence d'un second transcrit de taille inférieure à celui de la lactoferrine de sécrétion murine (Figure 4 A). Son clonage et son séquençage ont permis de déterminer qu'il s'agissait effectivement d'un transcrit issu du gène murin de la lactoferrine. Ce transcrit plus court de 157 nucléotides par rapport à celui de

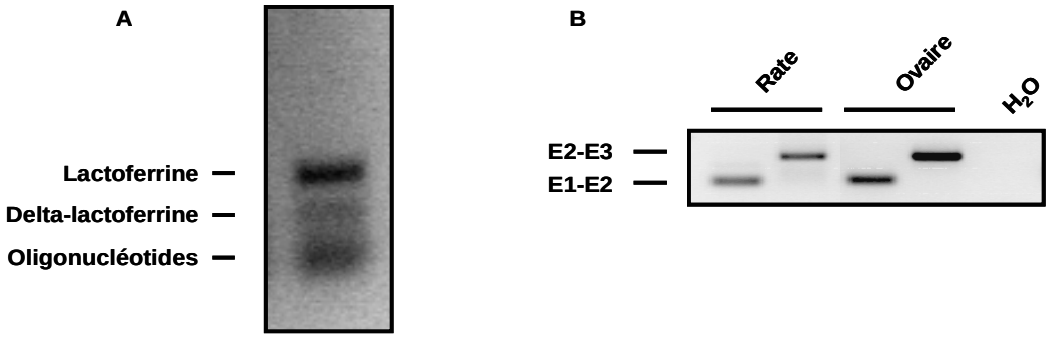


Figure 4 : La delta-lactoferrine existerait chez la souris. (A) L'existence d'un second transcrit issu du gène de la lactoferrine est mise en évidence par 5'RACE. (B) L'amplification des jonctions exon1-exon2 (E1-E2) et exon2-exon3 (E2-E3) montre une différence de quantité suggérant l'existence de transcrits dépourvus d'exon 1.

la lactoferrine de souris ne possède pas l'exon 1, lequel comporte les régions codantes pour le signal peptide. Contrairement au transcrit de la delta-lactoferrine humaine, ce transcrit semble dépourvu d'exon 1β. En outre, l'amplification de régions différentes grâce à des amorces spécifiques des jonctions exon1-exon2 et exon2-exon3 suggère que les exons 2 et 3 sont plus représentés dans la rate et les ovaires que l'exon1 (Figure 4 B, page 14), ce qui confirme l'existence de transcrits ne possédant pas d'exon 1. Ces travaux suggèrent fortement l'existence d'une isoforme de la lactoferrine murine dont la traduction devrait permettre la synthèse d'une protéine cytoplasmique tronquée des 24 premiers acides aminés par rapport à son homologue sécrétée. Ces résultats préliminaires devront, cependant, être confirmés par des expériences de qPCR et de Northern blot et être étendus aux différents tissus de souris.

B. Conformation tridimensionnelle

L'étude cristallographique de la lactoferrine humaine réalisée par l'équipe du Dr. Baker (Anderson *et al.*, 1987; Anderson *et al.*, 1989) a permis d'établir sa structure tridimensionnelle (Figure 5) qui a ensuite été confirmée par l'étude de celles d'autres espèces (Karthikeyan *et al.*, 2000; Khan *et al.*, 2001; Moore *et al.*, 1997).

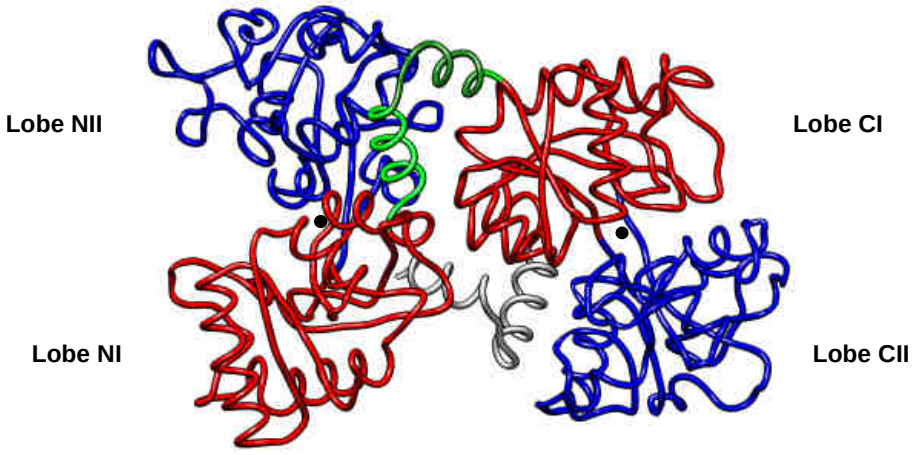


Figure 5 : Représentation schématique de la structure tridimensionnelle de la lactoferrine humaine saturée en fer (Anderson et al., 1989). Les lobes NI et CI et les lobes NII et CII sont respectivement représentés en rouge et en bleu. La région inter-lobe apparaît en vert. Les deux points noirs représentent les atomes de fer chélatés par la lactoferrine. La modélisation moléculaire a été réalisée en utilisant le programme Chimera (http://www.cgl.ucsf.edu/chimera/).

La lactoferrine est une protéine bilobée comprenant 16 ponts disulfure : le lobe N-terminal est constitué des résidus d'acides aminés 1 à 333 (résidus 1 à 307 pour la delta-lactoferrine) et le lobe C-terminal comprend les résidus d'acides aminés 345 à 692 (résidus 319 à 666 pour la delta-lactoferrine). Les deux lobes sont reliés entre eux par une courte hélice α formée par les acides aminés 334 à 344 (de 308 à 318 pour la delta-lactoferrine). Le lobe N-terminal peut se superposer au lobe C-terminal après une rotation de 180° et une translation de 25 Å (Anderson *et al.*, 1989). Après digestion trypsique ménagée, les deux lobes de la lactoferrine (30 kDa et 50 kDa) ainsi séparés sont capables de se réassocier en complexe non covalent en proportion équimolaire. Ce complexe possède alors des propriétés électrophorétiques et spectroscopiques proches de celles de la lactoferrine native (Legrand *et al.*, 1986).

Chaque lobe de la lactoferrine humaine est structuré en deux sous-domaines d'environ 160 résidus d'acides aminés à l'interface desquels se situe un site de fixation du fer. Chaque domaine possède une organisation structurale similaire basée sur l'alternance de feuillets β et d'hélices α (Anderson *et al.*, 1987; Anderson *et al.*, 1989). Le lobe N-terminal est constitué des domaines NI (résidus de 1 à 90 et de 252 à 320 pour la lactoferrine, de 1 à 64 et de 226 à 294 pour la delta-lactoferrine) et NII (résidus de 91 à 251 pour la lactoferrine, de 65 à 225 pour la delta-lacoferrine) tandis que le lobe C-terminal est constitué du domaine CI (résidus de 345 à 433 et 596 à 663 pour la lactoferrine, de 319 à 407 et 570 à 637 pour la delta-lactoferrine) et du domaine CII (résidus de 434 à 595 pour la lactoferrine, de 408 à 569 pour la delta-lactoferrine) (Figure 5, page 15).

La lactoferrine de sécrétion humaine peut adopter deux conformations : une conformation ouverte rencontrée préférentiellement dans la forme apo-protéine et une conformation fermée majoritairement rencontrée lorsque celle-ci est saturée en fer. Ce phénomène est dû à la mobilité du domaine NII qui peut pivoter de 53° par rapport au domaine NI. Cette rotation s'effectue autour des résidus His^{91} et Pro^{251} localisés entre deux feuillets β et deux hélices α de la région inter-domaine (Anderson *et al.*, 1990).

C. Structure des glycannes

Des études réalisées au sein de notre laboratoire sur la structure des glycannes de la lactoferrine de sécrétion ont mis en évidence une grande diversité entre les espèces au niveau de leur nombre et de leur localisation mais également au niveau de leur structure primaire (Spik *et al.*, 1994; Spik *et al.*, 1988; Spik *et al.*, 1982). La lactoferrine isolée du lait humain possède trois sites potentiels de *N*-glycosylation : un au niveau du lobe N-terminal sur le

résidu Asn^{138} (domaine NII) et deux au niveau du lobe C-terminal sur les résidus Asn^{479} (domaine CII) et Asn^{624} (domaine CI). Ces sites font partie d'une séquence consensus de *N*-glycosylation de type Asn-X-Thr/Ser. La *N*-glycosylation s'effectue préférentiellement (85% des molécules) sur les résidus Asn^{138} et Asn^{479}. Dans une moindre proportion, la lactoferrine de sécrétion ne portera qu'un glycanne en Asn^{479} ou en portera trois sur les positions précédemment citées (van Berkel *et al.*, 1996). Les *N*-glycannes de la lactoferrine humaine sont de type *N*-acétyllactosaminique mono- ou disialylés et fucosylés. Les résidus de fucose sont branchés en α-1,6 sur la *N*-acétylglucosamine du point d'attache ou en α-1,3 sur la *N*-acétylglucosamine 5' de l'antenne liée en α-1,6 (Spik *et al.*, 1982). Il est intéressant de noter qu'il existe des variations dans la structure des glycannes des lactoferrines humaines selon leurs lieux d'expression. En effet, la lactoferrine de sécrétion leucocytaire diffère de celle isolée du lait par l'absence de fucose (Derisbourg *et al.*, 1990).

En termes de différences glycanniques, le cas de la delta-lactoferrine est le plus marquant. En effet, de part sa localisation cytoplasmique, la delta-lactoferrine ne sera pas modifiée par *N*-glycosylation avec des glycannes complexes comme la lactoferrine de sécrétion. Cependant, l'étude *in silico* de sa structure primaire par le programme YingOYang 1.2 (http://www.cbs.dtu.dk/services/YinOYang/) montre qu'elle pourrait être modifiée par *O*-glycosylation cytosolique ou GlcNAcylation sur les résidus Ser^{10}, Ser^{227}, Ser^{472} et Thr^{559}. Contrairement à la *N*-glycosylation, la GlcNAcylation consiste en un transfert d'un monosaccharide (*N*-acétyl-β-D-glucosamine) lié sur l'hydroxyle d'un résidu Ser ou Thr (voir le chapitre : « Modifications post-traductionnelles » page 38).

D. Domaines

a) Les caractéristiques structurales de la lactoferrine

La plupart des fonctions proposées et démontrées de la lactoferrine, excepté l'activité de fixation du fer de la lactoferrine de sécrétion, dépendent de sa capacité à se fixer à d'autres macromolécules (protéines, LPS, glycosaminoglycannes, ADN). Cette capacité de fixation est dépendante des caractéristiques de surface de la lactoferrine. La lactoferrine est une protéine très basique puisque son point isoélectrique est de 9. La nature cationique de la lactoferrine, qui la différencie de la transferrine, est impliquée dans sa capacité à se lier à de nombreux types cellulaires et de nombreuses molécules anioniques. La délétion des 26 premiers résidus d'acides aminés fait que la delta-lactoferrine possède un pHi plus faible de l'ordre de 8. Cette différence de pHi entre les isoformes sécrétée et cytoplasmique s'explique par une forte proportion d'acides aminés basiques présents à l'extrémité N-terminale de la forme sécrétée.

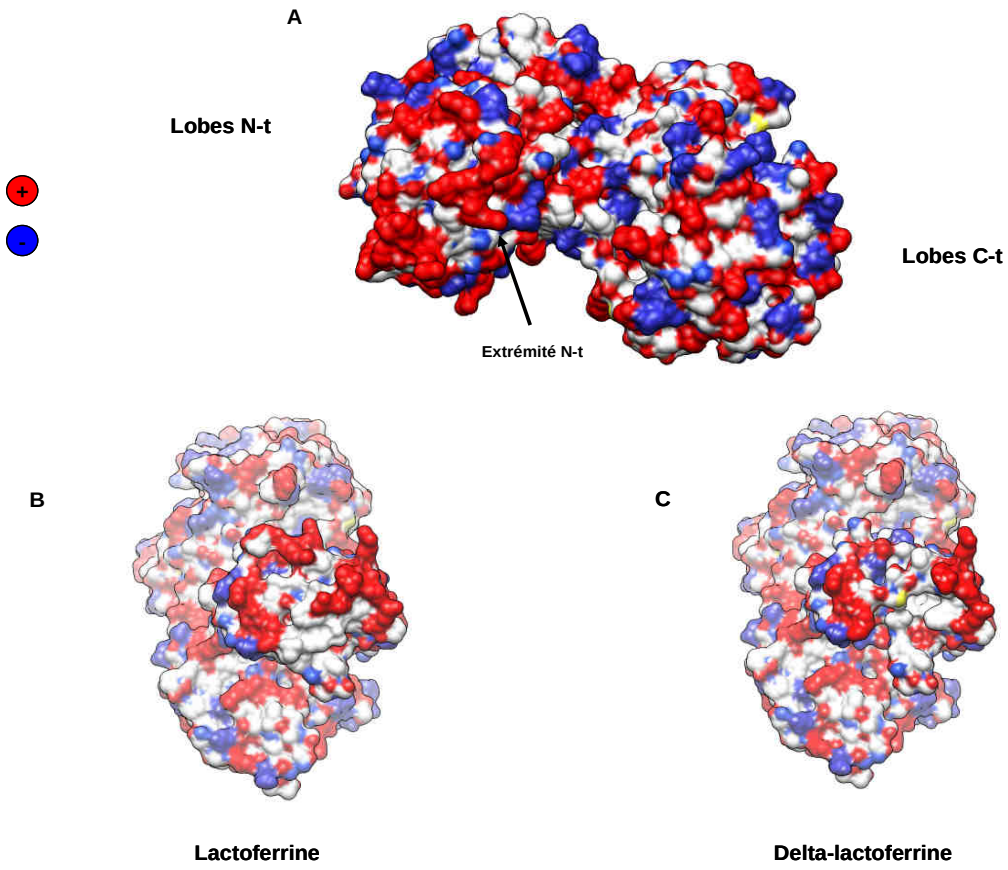


Figure 6 : Modèle tridimensionnel de la répartition des charges de surface de la lactoferrine humaine (A, B) et de la delta-lactoferrine (C). Les charges positives sont représentées en rouge et les charges négatives en bleu. La photographie présentée en A montre la surface de la molécule exposée au solvant avec la distribution des charges. B et C, les molécules sont orientées de manière à mettre en évidence la différence de charges entre les lobes N-terminaux de la lactoferrine (B) et de la delta-lactoferine (C). La modélisation moléculaire, basée sur la structure tridimensionnelle définie par Anderson et al. (1987), a été réalisée en utilisant le programme Chimera (http://www.cgl.ucsf.edu/chimera/).

En effet, la répartition des charges de surface, comme le révèle la structure cristalline de la lactoferrine (Figure 6 A), est très inégale. Trois régions notables de la lactoferrine regroupent une forte concentration de charges positives : les résidus 1 à 7 à l'extrémité N-terminale (absents de la delta-lactoferrine), la partie exposée vers l'extérieur de la première hélice formée des résidus 13 à 30 (partiellement délétée dans la delta-lactoferrine) et la région inter-lobe (Anderson *et al.*, 1989; Baker et Baker, 2005). L'absence de l'extrémité N-terminale chez la delta-lactoferrine conduit à une diminution des charges positives de son lobe NI (Figure 6 C).

b) Les domaines fonctionnels de la lactoferrine

La forte concentration de charges positives des résidus 1 à 5 de l'extrémité N-terminale et de l'extrémité C-terminale de la première hélice (résidus 27 à 30) constitue le site de liaison proposé pour l'ADN, les LPS, l'héparine et les glycosaminoglycannes (Mann *et al.*, 1994; van Berkel *et al.*, 1997). Ces caractéristiques sont partagées par d'autres lactoferrines, notamment par la lactoferrine bovine qui est la molécule la plus utilisée dans les essais *in vitro* et *in vivo*.

En plus de l'extrémité N-terminale, la région inter-lobe, également constituée de résidus basiques, apparaît comme une deuxième région potentielle de fixation à l'ADN. Cette région séparant les deux lobes confère à la lactoferrine l'espacement suffisant pour encercler la molécule d'ADN comme le suggèrent des résultats d'expériences de « docking » (Communication personnelle du Dr. Christophe Mariller et du Dr. Gérard Vergoten).

Il a été démontré que la séquence GRRRR localisée à l'extrémité N-terminale de la lactoferrine humaine, constituée de 4 résidus consécutifs d'arginine serait une séquence de localisation nucléaire (Penco *et al.*, 2001) (Figure 7 A, page 20). Les auteurs ont, en effet, démontré qu'un pentapeptide de même séquence est internalisé indépendamment de la voie de l'endocytose. Une fois internalisé, ce peptide est retrouvé dans le noyau cellulaire. Les auteurs ont suggéré que ce peptide pourrait être une séquence d'adressage au noyau (NLS).

La première hélice de la lactoferrine humaine, constituée des résidus 12 à 31, constitue la majeure partie du domaine bactéricide, nommé lactoferricine (Figure 7 A, page 20), identifié par Bellamy *et al.* (Bellamy *et al.*, 1992b). L'hydrolyse pepsique de la lactoferrine libère ce petit peptide naturel qui devient alors un agent bactérien puissant à spectre large dont l'activité est supérieure à celle de la lactoferrine native (Bellamy *et al.*, 1992a; Bellamy *et al.*, 1992b; Kuwata *et al.*, 1998). La protéolyse de la lactoferrine bovine libère également la lactoferricine (lactoferricine B) dont l'activité est supérieure à la lactoferricine humaine (Bellamy *et al.*, 1992a; Bellamy *et al.*, 1992b; Gifford *et al.*, 2005). Cette région participe au caractère antibactérien de la lactoferrine native probablement de la même manière que la lactoferricine par formation de structures amphiphiles qui perturbent les membranes bactériennes (Gifford *et al.*, 2005).

c) Les domaines putatifs de la delta-lactoferrine

Bien que non caractérisées, la delta-lactoferrine possède différentes régions putatives intéressantes (Figure 7 B, page 20). Comme la lactoferrine de sécrétion, la delta-lactoferrine a un pHi basique. De plus, l'extrémité N-terminale conserve l'extrémité C-terminale de la

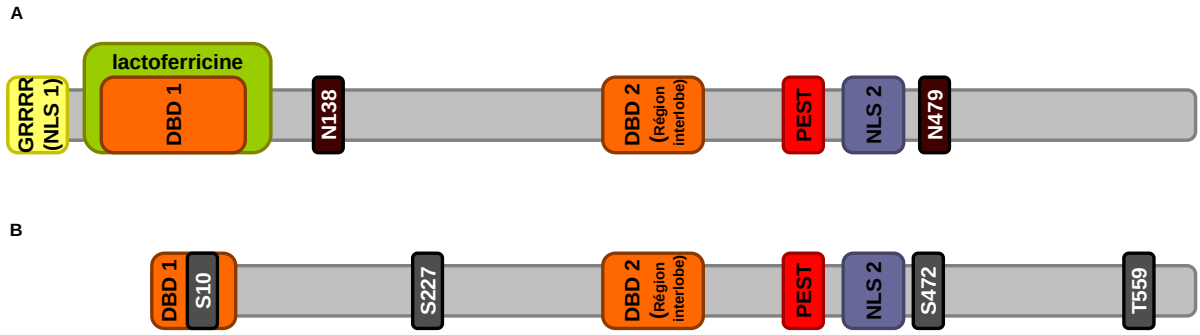


Figure 7 : Localisation des différents domaines de la lactoferrine humaine (A) et de la delta-lactoferrine (B). Représentation schématique de la structure primaire de la lactoferrine (A) et de la delta-lactoferrine (B) présentant les différents domaines ainsi que leurs positions respectives. Les sites de N- et O-glycosylation respectivement pour la lactoferrine et la delta-lactoferrine sont également indiqués.

première hélice (résidus 27 à 30 de la lactoferrine, 2 à 5 pour la delta-lactoferrine) qui constitue le site de liaison proposé pour l'ADN. Il apparaît alors que les résidus de lysine 13 et 260 ainsi que les résidus d'arginine 14 et 28 se trouvent dans l'environnement immédiat du premier domaine d'interaction à l'ADN (DBD) putatif et forment une structure basique saillante susceptible d'interagir avec la double hélice de l'ADN. La séquence GRRRR de l'isoforme sécrétée est absente de la delta-lactoferrine. Cependant, il a été démontré que, malgré une localisation majoritairement cytoplasmique, cette dernière est retrouvée dans le noyau (Breton *et al.*, 2004; Liu *et al.*, 2003). Sa masse moléculaire l'empêchant de subir une diffusion passive vers le noyau, elle devrait y être adressée par un mécanisme actif. En effet, l'analyse *in silico* de sa séquence primaire révèle la présence d'une séquence putative NLS bipartite conservée parmi les lactoferrines de différentes espèces : RRSDTSLTWNSVKGKK localisée sur le lobe C de la protéine en position 417 à 432. Cette séquence putative NLS est proche des séquences NLS fonctionelles retrouvées pour Rb (Efthymiadis *et al.*, 1997), IL-5 (Jans *et al.*, 1997) et la nucléoplasmine (Robbins *et al.*, 1991).

La delta-lactoferrine possède également un domaine de dégradation rapide (PEST) potentiel de séquence PVLAENYKSQQSSDPDPNCVD en position C-terminal (résidus 391 à 405). Cette séquence de dégradation rapide qui pourrait être reconnue par le système ubiquitine/protéasome, est conservée phylogénétiquement parmi les lactoferrines. La difficulté à isoler la delta-lactoferrine même lorsqu'elle est surexprimée suggère que cette séquence pourrait être pleinement fonctionnelle.

Il est intéressant de noter que les sites de GlcNAcylation se situent à proximité d'éléments fonctionnels putatifs de la delta-lactoferrine (Figure 7 B, page 20). En effet, le site sérine 10 se trouve au cœur du DBD1 et pourrait affecter la fixation de cette dernière à l'ADN. De même, le site sérine 472 se situe à proximité des séquences NLS et PEST. Ce site pourrait alors moduler la dégradation, le trafic nucléo-cytoplasmique de la delta-lactoferrine ou les deux.

III. Les lactoferrines : des facteurs de transcription ?

La lactoferrine peut se fixer à l'ADN (Garre *et al.*, 1992). Cette fixation conduit à l'activation de l'expression d'un gène rapporteur (He et Furmanski, 1995) ou son inhibition comme cela a pu être démontré pour le promoteur du gène de la protéine GM-CSF (Granulocyte Macrophage Colony-Stimulating Factor) (Penco *et al.*, 1995). La lactoferrine est capable d'interagir *in vitro* spécifiquement avec trois séquences nucléotidiques (Tableau 2) identifiées pour la première fois par He et Furmanski en 1995. Les auteurs ont isolé ces

Tableau 2 : Séquences d'ADN reconnues spécifiquement par la lactoferrine humaine (He et Furmanski, 1995).

Nom	Séquence	Fonctionnalité
S1	GGCACTT(G/A)C	++
S2	TAGA(A/G)GATCAAA	-
S3	ACTACAGTCTACA	+

séquences après enrichissement de fragments d'ADN sur colonne de lactoferrine et amplification avec des oligonucléotides dégénérés. Parmi les trois séquences ainsi isolées, seules deux permettent l'expression d'un gène rapporteur CAT (Chloramphénicol Acétyl Transférase) en présence de lactoferrine ou d'une construction de lactoferrine sans le peptide signal (proche de la delta-lactoferrine) (He et Furmanski, 1995).

La séquence S1 a été retrouvée fonctionnelle dans le promoteur du gène de l'IL-1β humain en position -3202 à -3193 (Son *et al.*, 2002). En plus de cette séquence, quatre autres séquences S1 imparfaites sont retrouvées dans ce même promoteur en position plus proximale (entre -3137 et -1043). L'étude menée par Son *et al*. en 2002 a permis de mieux définir le DBD de la lactoferrine. Pour cette étude, les auteurs ont réalisé et testé l'activité transcriptionnelle, vis-à-vis d'un gène rapporteur contenant la séquence S1, de différentes constructions tronquées de lactoferrine humaine (Figure 8, page 22).

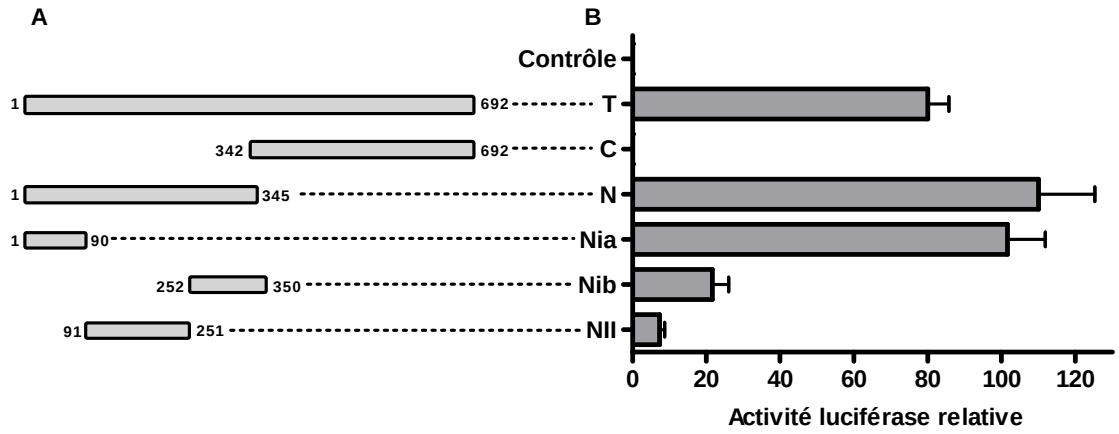


Figure 8 : Activité transcriptionnelle de différentes constructions de lactoferrine tronquées (inspiré de Son et al., 2002). L'activité transcriptionnelle (B) des différentes constructions tronquées de la lactoferrine humaine (A) est estimée grâce à l'expression d'un gène rapporteur dépendant de l'élément de réponse S1 de la lactoferrine.

Hormis la construction exprimant la lactoferrine entière (T), les constructions possédant une activité transcriptionnelle sont les constructions N (des résidus 1 à 345) et Nia (des résidus 1 à 90). Ces constructions contiennent la séquence GRRRR de domiciliation nucléaire, le DBD situé en N-terminal et le DBD putatif de la région charnière pour la construction N. La lactoferrine comporterait donc un domaine de liaison à l'ADN localisé parmi les 90 premiers résidus d'acides aminés (Son *et al.*, 2002) ce qui confirme les données prédites par l'analyse de sa structure (Anderson *et al.*, 1987). De manière remarquable, une construction de 98 acides aminés contenant uniquement la région charnière (Nib) présente une activité transcriptionnelle résiduelle mais non négligeable ce qui suggère la fonctionnalité du DBD2. Une construction délétée des 90 premiers résidus d'acides aminés manque cependant à cette étude, elle aurait permis de confirmer la fonctionnalité du DBD en N-terminal.

La localisation cytoplasmique de la delta-lactoferrine fait de cette isoforme un meilleur candidat pour exercer l'activité transcriptionnelle des lactoferrines humaines. La delta-lactoferrine devrait reconnaître les séquences nucléotidiques décrites par He et Furmanski (Tableau 2, page 21). Elle pourrait donc avoir une activité de facteur ou de cofacteur de transcription et un rôle dans la régulation de l'expression de certains gènes.

IV. Les lactoferrines : des suppresseurs de tumeur ?

La lactoferrine et la delta-lactoferrine sont potentiellement des suppresseurs. Bien que la fonction de la delta-lactoferrine ne soit pas connue, sa surexpression semble indiquer qu'elle a une activité anti-prolifération (Breton *et al.*, 2004). La lactoferrine de sécrétion, quant à elle, possède de très nombreuses fonctions biologiques toutes impliquées dans la défense de l'organisme (Figure 1, page 7) (voir pour revues : Legrand *et al.*, 2008; Pierce *et al.*, 2009). Parmi ces fonctions seule la protection face au cancer est développée dans ce chapitre. En effet, de nombreux travaux réalisés sur des rongeurs ont démontré que cette molécule a un rôle dans la prévention et la régression de différents types de tumeur.

A. Activités anti-cancéreuses

L'action de la lactoferrine de sécrétion sur le développement de tumeurs et de métastases a fait l'objet de diverses études. Cette action est différente selon les types de tumeur et les lignées cellulaires utilisées au cours des expérimentations. De nombreuses études réalisées chez le rat montrent que l'administration orale de lactoferrine bovine réduit très significativement la taille des tumeurs induites chimiquement dans différents organes (Tsuda *et al.*, 2000). L'activité anti-tumorale de la lactoferrine s'exerce par différents mécanismes notamment par son potentiel immunomodulateur. La lactoferrine stimule la prolifération des lymphocytes et favorise la cytoxicité des cellules NK (Natural Killer) (Artym *et al.*, 2003; Damiens *et al.*, 1998). L'augmentation de la réactivité du système immunitaire est la première ligne de défense apportée par la lactoferrine face à la tumorigénèse.

L'effet anti-mitogène *in vitro* de la lactoferrine a essentiellement été mis en évidence sur les cellules mammaires. En effet, Rejman *et al*. (1992) ont dans un premier temps observé que la lactoferrine bovine inhibe la prolifération d'une lignée bovine de cellules alvéolaires mammaires immortalisées (MAC-T) par transfection stable du virus SV40. Cette inhibition est détectée en présence ou en absence de 10% de sérum de veau fœtal. Elle est d'autant plus importante que la concentration de lactoferrine est élevée (la concentration variant de 19,5 à 625 µg/ml). Cette étude a été confirmée par Hurley *et al*. (1994). Ces auteurs ont montré que la lactoferrine induit l'inhibition de la prolifération d'une lignée de cellules épithéliales mammaires bovines transformées mais non tumorigènes (MAC-E) (Hurley *et al.*, 1994). L'effet de la lactoferrine humaine a également été recherché sur les lignées de cellules mammaires humaines cancéreuses hormono-dépendantes (MCF-7) ou

hormono-indépendantes (MDA-MB-231). Il a été observé que la lactoferrine humaine a un effet inhibiteur sur la prolifération des cellules MCF-7 et MDA-MB-231 lorsqu'une forte concentration (500 µg/ml) est utilisées (Hurley *et al.*, 1994).

En outre, Sekine et *al*. (1997) ont montré des effets anti-prolifératifs de la lactoferrine bovine *in vivo* chez le rat. Ainsi, l'administration de lactoferrine bovine à la concentration de 0,2 à 2 % réduit significativement l'incidence et le nombre d'adénocarcinomes du colon chez les rats F344 traités par l'azoxyméthane. La réduction de la formation de foyers de cryptes intestinales anormales pourrait être liée à une activation de la cytotoxicité anti-tumorale des cellules NK chez ces animaux (Sekine *et al.*, 1997b). De même, l'administration orale de lactoferrine bovine inhibe la carcinogénèse chimiquement induite du colon ou d'autres organes chez le rat et la souris. L'inhibition de la tumorigénèse induite par la lactoferrine est due à l'augmentation de la production de cytokines, telles que l'interleukine 18 et l'interféron gamma, et à l'activation des cellules du système immunitaire qui en résulte (Iigo *et al.*, 2009).

Différents auteurs ont étudié l'influence de la lactoferrine de sécrétion dans l'évolution des tumeurs et des métastases chez des modèles animaux. Bien que l'utilisation d'un système hétérologue rend difficile l'extrapolation de l'effet de la lactoferrine en système homologue, une étude de Iigo *et al.* (1999) a permis de démontrer que l'administration orale de lactoferrine bovine diminue significativement l'implantation de métastases pulmonaires après xénogreffe d'une lignée agressive de carcinomes de colon (Co 26Lu) chez des souris BALB/c (Iigo *et al.*, 1999). En outre, Yoo *et al.* (1997) ont constaté que la lactoferrine bovine inhibe *in vivo*, chez la souris, la formation de métastases, issues des lignées cancéreuses murines L5178Y-ML25 et de cellules B16-BL6, aux poumons et au foie (Yoo et al., 1997; Yoo et al., 1998). La lactoferrine humaine empêche également la formation de métastases pulmonaires induites par l'injection chez la souris de cellules de mélanome B16-F10 (Bezault *et al.*, 1994). Cependant, dans ce même modèle expérimental, elle n'inhibe pas la croissance de cellules de mélanome B16-BL6 ni leur capacité à former des métastases pulmonaires (Yoo *et al.*, 1997).

B. Régulation de la prolifération cellulaire

Les lactoferrines régulent la croissance des cellules tumorales en induisant un arrêt du cycle cellulaire. Il a été démontré, notamment au laboratoire, de manière *in vitro* que la delta-lactoferrine (Breton *et al.*, 2004) ainsi que la lactoferrine (Damiens *et al.*, 1999; Xiao *et al.*, 2004) ont un effet anti-prolifératif.

En effet, l'expression de la delta-lactoferrine dans les cellules HEK 293 inhibe la prolifération cellulaire et induit un arrêt du cycle cellulaire en G1/S (Breton *et al.*, 2004). Cependant, contrairement à la lactoferrine, les voies moléculaires impliquées par l'activité de la delta-lactoferrine ne sont pas connues.

Pour la lactoferrine, cet arrêt, est provoqué par un ensemble de processus régulant l'activité des acteurs moléculaires qui contrôlent cette transition. Il semble que la lactoferrine régule des voies différentes selon le modèle cellulaire étudié. Ainsi, dans la lignée cellulaire MDAMB-231, la lactoferrine humaine inhibe la voie des MAP kinases. Cette inhibition est concomitante à une diminution des activités cdk2 et cdk4 et à une augmentation de p21 (Damiens *et al.*, 1999). Cette augmentation de p21 est également retrouvée après traitement des lignées cellulaires humaines H1299, MCF7 et HEK 293 avec la lactoferrine bovine (Son *et al.*, 2006). Xiao *et al.* ont mis en évidence un arrêt du cycle dans quatre lignées cellulaires, issues de résections de tumeurs de la tête et du cou, induit par la lactoferrine humaine recombinante (Xiao *et al.*, 2004). Cet arrêt est associé à une augmentation des taux de la protéine p27 et de la cycline E ainsi qu'à une diminution de la phosphorylation de Rb. Au cours de cette étude, les auteurs ont également rapporté une diminution de la phosphorylation d'Akt (Xiao *et al.*, 2004). En plus d'une hypophosphorylation de Rb rapportée par Xiao *et al.* en 2004, l'équipe de Choi a observé une surexpression de Rb après traitement avec la lactoferrine (Son *et al.*, 2006). La progression du cycle et la traduction des acteurs cellulaires de la transition G1/S nécessitent la dissociation du complexe Rb/E_2F par phosphorylation de Rb. Ainsi, la surexpression de Rb et son hypophosphorylation contribuent à l'arrêt du cycle observé en G1/S après traitement des cellules à la lactoferrine. En outre, la surexpression de la lactoferrine par les cellules HeLa induit une surexpression du suppresseur de tumeur p53, mais également de mdm2 et de p21. La surexpression de ces gènes est due à l'activation de facteurs de transcription NF-κB (Oh *et al.*, 2004). La lactoferrine de sécrétion agit donc sur plusieurs voies contrôlant la transition G1/S (voir Figure 9, page 26).

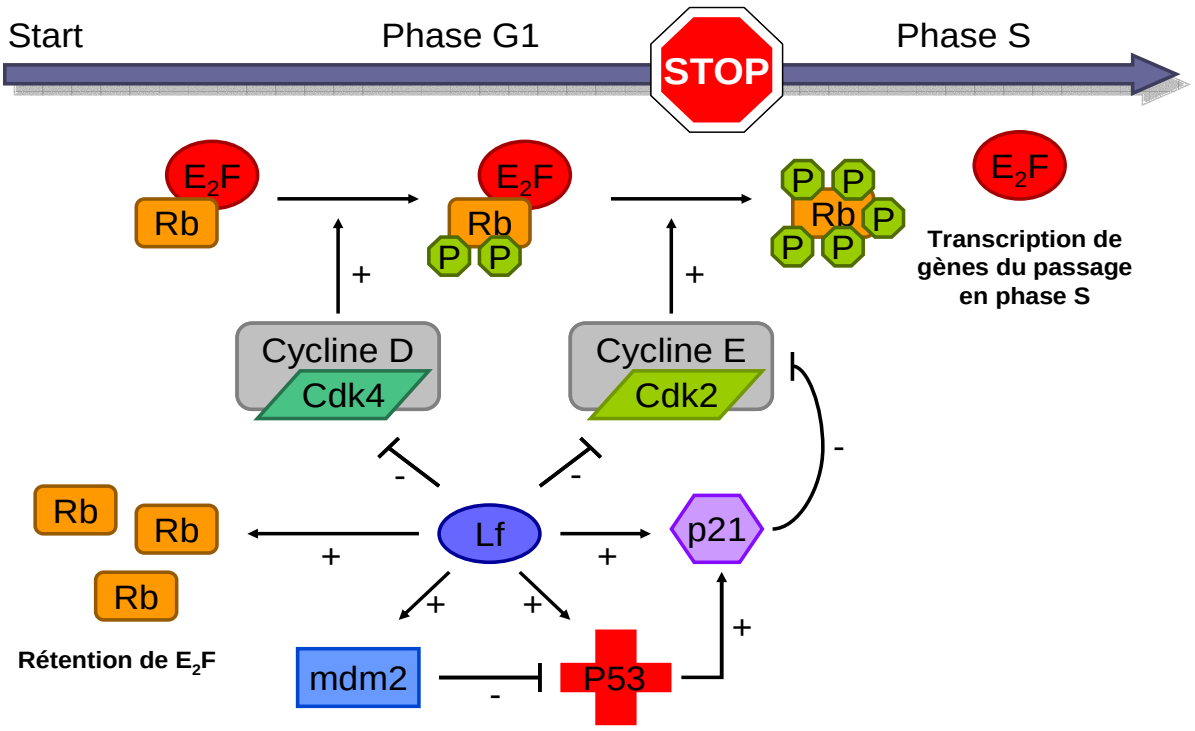


Figure 9 : Effets de la lactoferrine sur la régulation de la transition G1/S. La lactoferrine intervient à différents niveaux de régulation de cette transition. La lactoferrine induit la surexpression de Rb qui s'accumule majoritairement sous une forme hypophosphorylée conduisant à la rétention de E_2F. Elle induit également la surexpression de p53 qui en retour conduit à celle de p21. La lactoferrine comme p21 inhibe alors l'activité des complexes Cycline D/Cdk4 et Cycline E/Cdk2 maintenant ainsi indirectement la protéine Rb sous la forme hypophosphorylée (d'après Pierce et al., 2009).

C. Les lactoferrines : un facteur de bon pronostic

Puisque la lactoferrine de sécrétion est potentiellement anti-tumorale, le niveau d'expression de la lactoferrine humaine a été déterminé dans différents types de cancer. Les études sur la distribution cellulaire et tissulaire de la lactoferrine ont montré que le gène de la lactoferrine est exprimé dans de nombreux tissus et que cette expression est régulée différemment selon le type cellulaire et le phénotype normal ou malin de ces tissus. Ainsi son expression est régulée négativement dans les cancers de la glande mammaire oestrogéno-dépendants (Campbell *et al.*, 1992), les adénocarcinomes du col de l'utérus (Farley *et al.*, 1997) et dans la leucémie myéloïde chronique (Rado *et al.*, 1987). Il semblerait donc qu'il y ait une dérégulation de l'expression du gène de la lactoferrine dans certaines formes de cancers.

La sous-expression de certains gènes dans les cellules cancéreuses, par hyperméthylation de leurs promoteurs, est une caractéristique principale des gènes suppresseurs de tumeur. La méthylation aberrante des îlots CpG des régions promotrices est impliquée dans l'extinction de l'expression des anti-oncongènes (Herman et Baylin, 2003).
Le promoteur des lactoferrines comporte des régions riches en GC. Comme pour d'autres promoteurs, une méthylation de base est retrouvée et constitue un premier niveau de régulation de l'expression. Cependant, des niveaux de méthylation différents sont retrouvés entre les tissus sains et les tissus cancéreux lymphoïde et mammaire (Panella *et al.*, 1991). Les régions promotrices de la lactoferrine de sécrétion ou celles de la delta-lactoferrine sont hyperméthylées dans les tissus tumoraux de patients alors qu'elles ne le sont peu ou pas dans les mêmes tissus sains. Des méthylations différentes sont retrouvées dans toutes les régions avec une augmentation globale sur les îlots CpG. La présence de méthylation en dehors des îlots CpG a été retrouvée dans les échantillons tumoraux de lymphocyte et de glande mammaire (Teng *et al.*, 2004).

La littérature relate que la région chromosomique 3q21 contenant le gène de la lactoferrine subit différents réarrangements dans les cas de cancer. Les modifications chromosomiques de cette région ont principalement été décrites dans les cas de leucémie. En effet, différents réarrangements des régions 3q21-26 ont été répertoriés dans les cas de leucémie myéloïde aiguë (Lahortiga *et al.*, 2004). En outre, la translocation de la région 3q21 est retrouvée dans les cas de leucémie monocytaire (McGrattan *et al.*, 2009). La translocation de la région 3q21 sur le chromosome 7 est associée à un facteur de mauvais pronostic. Une inversion de la partie allant de q21 à q26 est retrouvée dans les cas de leucémie myéloblastique (Testoni *et al.*, 1999). Par ailleurs, des aberrations chromosomiques de la

région 3q21 à q27 ont été décrites dans la lignée cellulaire MCF-10A1h cl13 qui correspond à une lignée de carcinome mammaire invasif (Worsham *et al.*, 2006). En outre, Kemp *et al.* (2006) ont montré que le polymorphisme de cette région est associé à la susceptibilité familiale au cancer colorectal (Kemp *et al.*, 2006).

Les études effectuées par Siebert et Huang (1997) et celles réalisées dans notre laboratoire (Benaïssa *et al.*, 2005; Hoedt *et al.*, 2010) sur plusieurs lignées cellulaires dérivées de cancer ont montré une sous-expression du transcrit de la lactoferrine et de la delta-lactoferrine par rapport aux tissus sains correspondants. Une étude réalisée dans notre laboratoire concernant l'analyse de l'expression des ARNm des lactoferrines dans 99 biopsies de glande mammaire cancéreuses montrent qu'il existe une corrélation positive entre les profils d'expression de ces deux ARNm et la survie des patientes. Cette étude rétrospective a permis d'établir un seuil au-delà duquel aucune récidive n'a été diagnostiquée et ce, après un suivi de 4000 jours. Il existe donc une corrélation positive entre la présence des transcrits de la lactoferrine et de la delta-lactoferrine et la survie des patientes. Cette survie est d'autant plus longue que le transcrit de la delta-lactoferrine est en concentration plus importante, faisant ainsi de la delta-lactoferrine un facteur de bon pronostic (Benaïssa *et al.*, 2005).

Bien que le niveau d'expression des récepteurs aux œstrogènes et à la progestérone accompagne celui de la lactoferrine (Lee *et al.*, 1995), une étude menée chez 78 patientes atteintes de cancer mammaire a montré que la sous-expression du gène de la lactoferrine était corrélée à une forte expression du récepteur aux œstrogènes (Penco *et al.*, 1999). Le dosage de l'expression du gène de la lactoferrine couplé à celui du récepteur aux œstrogènes pourrait constituer un bon indicateur dans la réussite des traitements au tamoxifène (utilisé dans le traitement de cancers du sein) (Giancotti, 2006).

Modifications post-traductionnelles

Le protéome est deux à trois fois plus complexe (supérieur à 1 000 000 protéines) que les régions codantes de l'ADN peuvent le prédire. Il existe deux mécanismes majeurs permettant d'étendre la capacité codante du génome (environ 26 500 gènes chez l'humain) à la diversité du protéome (Venter *et al.*, 2001). Le premier mécanisme générant la diversification des protéines se trouve au niveau transcriptionnel, *via* l'épissage des ARNm, aboutissant à la synthèse de différents transcrits. Le second mécanisme permettant d'accroître la diversité du protéome est la modification covalente des protéines sur un ou plusieurs sites. Pendant la traduction ou une fois la synthèse peptidique terminée, les protéines subissent différentes modifications dites post-traductionnelles (Figure 10). Ces

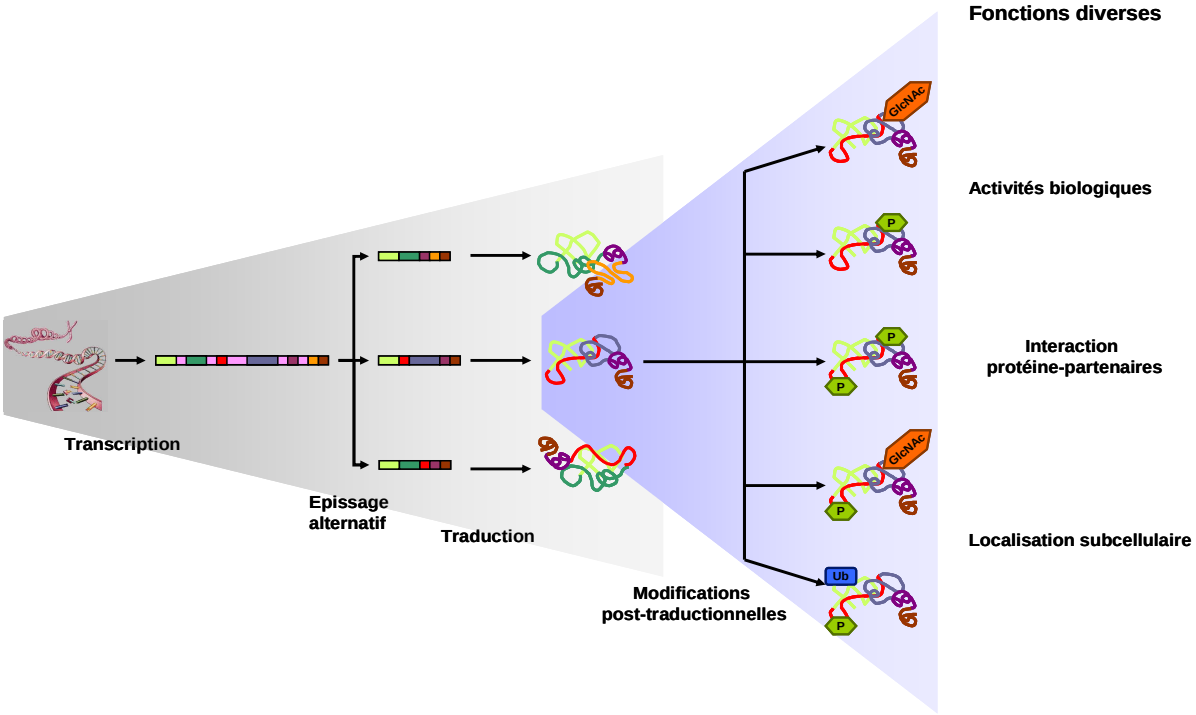


Figure 10 : Du gène aux protéines. Schéma représentant la mise en place de la complexité du protéome s'établissant au cours des étapes successives conduisant à la synthèse des protéines. Une protéine est issue d'un gène unique. Cependant, un gène unique est à l'origine de plusieurs protéines différentes qui auront plusieurs activités biologiques. Le gène est transcrit en pré-ARNm qui grâce au phénomène d'épissage pourra donner plusieurs ARNm matures. La traduction des ces ARNm conduira à la synthèse de protéines diverses qui pourront être modifiées de différentes manières telles que par clivage protéolytique ou par ajout de groupements particuliers (phosphate, GlcNAc, ubiquitine, etc…) pouvant se combiner sur plusieurs sites. In fine, le gène sera à l'origine d'une innombrable diversité fonctionnelle.

modifications, contribuant à augmenter la complexité du protéome, ont pour but de réguler l'activité des protéines, de les intégrer à une cascade de signalisation, de les adresser vers un compartiment cellulaire, de les ancrer à la membrane ou de les « marquer » afin qu'elles soient reconnues par d'autres molécules ou par les systèmes de dégradation cellulaire. Bien que présentes chez les procaryotes, les modifications post-traductionnelles sont plus largement rencontrées chez les eucaryotes aussi bien en terme de type de modifications que de répartition. Environ 5% du génome des eucaryotes supérieurs sont alloués à la synthèse des enzymes réalisant les modifications post-traductionnelles du protéome (pour revue voir : Walsh *et al.*, 2005).

Il existe deux catégories de modifications post-traductionnelles. La première comprend les clivages de la chaîne polypeptidique de la protéine d'origine par l'action de protéases ou par une action autocatalytique. La seconde regroupe les additions covalentes d'un groupement chimique sur un résidu de la chaîne polypeptidique de la protéine. L'environnement peptidique du site modifié est généralement riche en électrons et agit en tant que nucléophile lors du transfert. Cette seconde catégorie comprend les modifications post-traductionnelles dynamiques, *i.e.* ajout puis hydrolyse du résidu ajouté, qui seront développées dans cette partie. Parmi les modifications post-traductionnelles répertoriées à ce jour, une trentaine d'entres elles sont plus communément répandues (Seo et Lee, 2004) (Tableau 3, page 31). A côté de la phosphorylation qui reste à ce jour la plus étudiée des modifications post-traductionnelles et dont les rôles sont quasiment établis, sont retrouvées la méthylation, l'acétylation, l'ubiquitinylation, la SUMOylation, la formation des ponts disulfure et la glycosylation.

Une description rapide de l'ubiquitinylation sera faite et une attention particulière sera portée à la *O*-N-acétylglucosaminylation (GlcNAcylation).

Tableau 3 : Principaux types de modifications post-traductionnelles rencontrées par acide aminé (d'après Walsh et al., 2005). Il n'y a pas de modification connue à ce jour pour les acides aminés Leu, Ile, Val et Phe.

Résidu concerné	Type de modification post-traductionnelle	Résidu concerné	Type de modification post-traductionnelle
Arg	N-méthylation	His	Phosphorylation
	N-ADP-ribosylation		Aminocarboxypropylation
Asn	N-glycosylation		N-méthylation
	N-ADP-ribosylation		
	Epissage protéique	Lys	N-méthylation
Asp	Phosphorylation		N-acylation
	Isomérisation en isoAsp		C-hydroxylation
Cys	S-hydroxylation (S-OH)		Ubiquitinylation
	Pont disulfure		Sumoylation
	Phosphorylation		Néddylation
	S-acylation	Met	Oxydation
	S-prénylation	Pro	C-hydroxylation
		Ser	Phosphorylation
			GlcNAcylation
			Phosphopantéthéinylation
Gln	Transglutamination		Autoclivage
Glu	Méthylation		
	Carboxylation	Thr	Phosphorylation
	Polyglycination		GlcNAcylation
	Polyglutamylation		
		Trp	C-mannosylation
Gly	C-hydroxylation	Tyr	Phosphorylation

I. Ubiquitinylation des protéines

En 2004, le prix Nobel de chimie fut décerné à Avram Hershko, Aaron Ciechanover et Irwin Rose pour avoir décrit la fonction d'une protéine ubiquitaire : l'ubiquitine. Ces lauréats, ainsi que d'autres, ont constaté que l'ubiquitine est attachée de manière covalente à des protéines cibles les conduisant ainsi à la dégradation protéasomale. D'autres fonctions ont été attribuées à l'ubiquitine comme l'internalisation des protéines membranaires, l'adressage aux endosomes et aux lysosomes, l'activation des processus de réparation de l'ADN ou l'activation de certains facteurs de transcription (Schnell et Hicke, 2003; Welchman *et al.*, 2005). L'ubiquitine est la première protéine décrite de la famille des UBLs pour « ubiquitin-like modifiers ». C'est une petite protéine de 76 acides aminés hautement

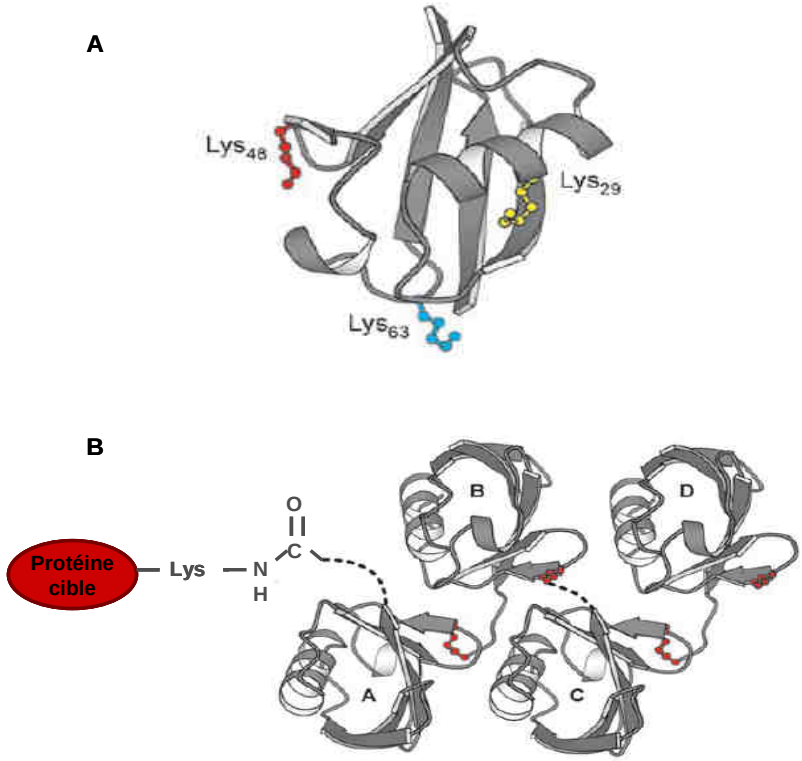


Figure 11 : Représentation tridimensionnelle de l'ubiquitine. (A) Structure tridimensionnelle de la molécule d'ubiquitine présentant les résidus de lysine cible de la polymérisation. Une protéine modifiée par une chaîne d'ubiquitine dont les résidus d'ubiquitine sont reliés entre eux par les résidus de lysine 48 ou 29 sera conduite vers le protéasome pour sa dégradation. La polymérisation des résidus d'ubiquitine au niveau des résidus de lysine 63 est un signal impliqué dans toutes les autres fonctions que la dégradation protéasomale. (B) Représentation schématique d'une protéine modifiée par quatre résidus d'ubiquitine qui est le motif minimum reconnu par le protéasome. (d'après Walsh et al., 2005)

conservée au cours de l'évolution. La structure tridimensionnelle de l'ubiquitine est élucidée et présente une conformation globulaire (Vijay-Kumar *et al.*, 1987). La structure de l'ubiquitine se compose de cinq feuillets β et d'une unique d'hélice α. L'extrémité C-terminale de la molécule est exposée permettant ainsi sa fixation covalente sur la protéine cible (Figure 11 A, page 32).

A. Le cycle de l'ubiquitine

L'ajout d'une molécule d'ubiquitine ne se fait pas de manière directe par l'action d'une seule enzyme (Figure 12). Le branchement d'une ubiquitine sur la protéine est réalisé en au moins trois étapes : la première étape consiste en l'activation du résidu de glycine C-terminal de l'ubiquitine par l'enzyme E1 qui catalyse la formation de l'AMP-ubiquitine. L'ubiquitine est alors transférée sur un résidu de cystéine du site catalytique de cette même enzyme pour former le complexe E1-S-ubiquitine. A ce stade, une seconde enzyme (E2) entre en jeu. L'ubiquitine est transférée sur E2 par réaction de transthiolation. La dernière réaction est catalysée par un complexe multiprotéique (E3) qui reconnaît E2-Ub ainsi que la protéine cible. E3 catalyse le transfert de l'ubiquitine sur le résidu de lysine cible de la protéine à ubiquitinyler.

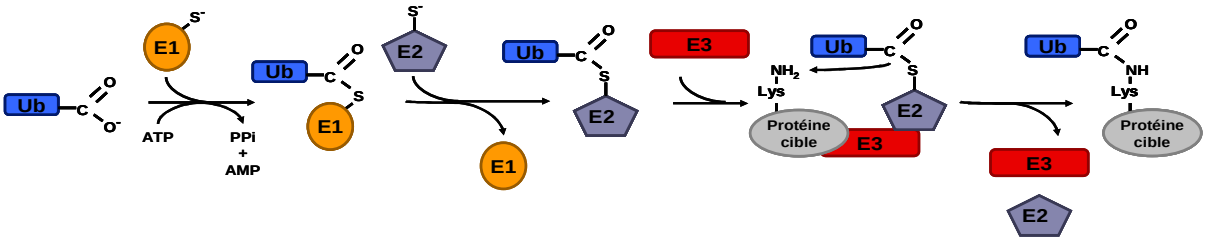


Figure 12 : Mécanisme enzymatique permettant l'ajout d'un résidu d'ubiquitine. L'activation de l'extrémité C-terminale de l'ubiquitine (résidu Gly76), par E1 par la formation d'un ubiquitine-adénylate, est suivie par son transfert sur un résidu de cystéine de E2 au niveau du site actif conduisant la formation d'une Ub-S-E2. E3 peut alors, en tant que chaperon, recruter la protéine cible afin de permettre son ubiquitinylation sur un résidu de lysine. Une fois ubiquitinylée, la protéine cible peut entrer dans un nouveau cycle afin d'être polyubiquitinylée.

B. La polyubiquitinylation

La polyubiquitinylation se réfère à la construction enzymatique d'une chaîne de molécules d'ubiquitine. L'addition en chaîne d'un monomère d'ubiquitine sur un autre se fait sur un résidu de lysine de l'ubiquitine précédente (Figure 11, page 32). Bien que possédant 7 résidus de lysine, la construction d'une chaîne de polyubiquitine se fait préférentiellement par le branchement de l'ubiquitine n+1 sur le résidu de lysine 48 de l'ubiquitine précédente. La structure aux rayons X d'une chaîne de quatre résidus d'ubiquitine (Ub_4) est présentée en Figure 11 B (page 32) (Walsh *et al.*, 2005). Le processus permettant de construire de manière processive une chaîne à n résidus d'ubiquitine sur une protéine n'est pas encore clairement connu (Hochstrasser, 2006). Une chaîne d'un minimum de quatre résidus d'ubiquitine est nécessaire à la reconnaissance par la sous-unité 26S du protéasome. La protéine ainsi polyubiquitinylée est alors dégradée. L'ubiquitine est recyclée par l'action d'une désubiquitinase (DUB) (Schnell et Hicke, 2003) (Figure 13).

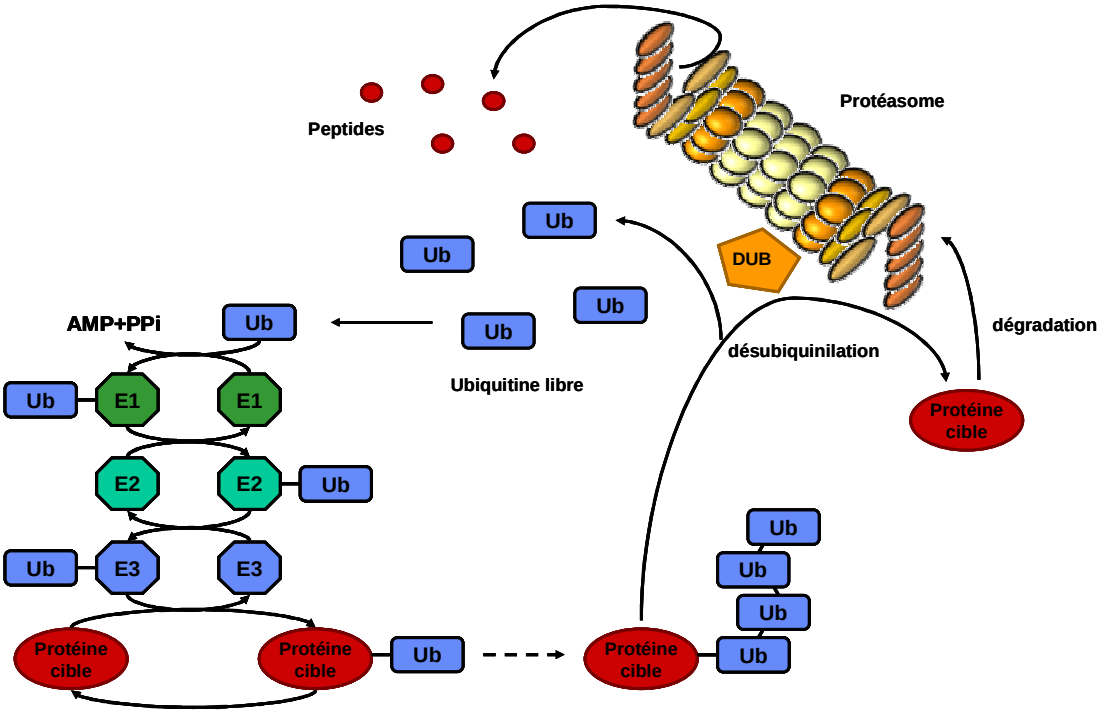


Figure 13 : Cycle de l'ubiquitinylation des protéines. Après l'ajout de plusieurs résidus d'ubiquitine sur une protéine, cette dernière est adressée au protéasome pour y être dégradée. Une désubiquitinase hydrolyse les liaisons protéine-ubiquitine et ubiquitine-ubiquitine tamdis que la protéine est prise en charge par le protéasome pour sa dégradation. L'ubiquitine ainsi libérée est recyclée.

C. Les enzymes de l'ubiquitinylation

Bien qu'agissant en complexe, particulièrement pour les enzymes E2 et E3, chaque étape de l'ubiquitinylation des protéines est réalisée par des enzymes spécifiques permettant en fin de chaîne de sélectionner le substrat. Il existe : deux enzymes d'activation E1 et 18 E2 décrites. La classe des enzymes E3 comporte, quant à elle, beaucoup de représentants dont certains restent encore à identifier (Pickart, 2001).

a) L'enzyme E1

A ce jour, il n'a été décrit chez l'homme que 9 enzymes d'activation E1 dont 2 dédiées spécifiquement à l'activation de l'ubiquitine (Groettrup *et al.*, 2008). La première enzyme E1 de l'ubiquitine nommée UBA1 (Ubiquitin activating enzyme) fut décrite en 1982 par l'équipe de Hershko (Ciechanover *et al.*, 1982). Le gène *UBA1*, porté chez l'homme par le chromosome X en p11.23, produit deux variants d'ARNm dont la traduction permet la synthèse de UBA1a (117 kDa) et UBA1b (110 kDa) (Cook et Chock, 1992). En plus de leur différence de taille, UBA1a est phosphorylée alors que UBA1b ne l'est pas (Cook et Chock, 1995). La phosphorylation d'UBA1a varie au cours du cycle cellulaire et permettrait son adressage nucléaire de la protéine en phase G2 (Stephen *et al.*, 1996).

Ce n'est que 25 ans plus tard qu'une nouvelle E1 humaine a été caractérisée par trois équipes à quelques mois d'intervalle. Cette nouvelle enzyme d'activation de l'ubiquitine humaine fut nommée UBE1L2 par Pelzer *et al*., E1-L2 par Chiu *et al*. ou UBA6 par Jin *et al*. (2007). Selon la nomenclature des enzymes d'activation des UBLs, UBA6 sera utilisée par la suite comme le recommandent Groettrup *et al.* (2008). Cette nouvelle E1 possède 40 % d'homologie avec UBA1 et est capable de complémenter cette dernière *in vitro*. Chez la souris, il apparaît qu'UBA6 est principalement exprimée dans les testicules (Pelzer *et al.*, 2007), suggérant une régulation de l'activation de l'ubiquitine tissu spécifique (Zhu *et al.*, 2004). De plus, chez la souris, le KO d'UBA6 est létal au stade embryonnaire (Chiu *et al.*, 2007) conférant à cette enzyme un rôle primordial au cours de l'embryogénèse. Contrairement à UBA1, UBA6 est capable d'activer, en plus de l'ubiquitine, la protéine apparentée à l'ubiquitine FAT10 (Chiu *et al.*, 2007).

Comme nous l'avons vu précédemment, UBA1 et UBA6 activent l'ubiquitine et la transfèrent sur E2. Cependant, elles peuvent avoir des substrats spécifiques. Ainsi, l'E2 Use1 ne recevrait l'ubiquitine que des UBA6 alors que les E2s Cdc34A et Cdc34B ne la recevraient que de UBA1 (Jin *et al.*, 2007).

b) L'enzyme de conjugaison E2

Toutes les protéines E2 possèdent un domaine conservé de 150 acides aminés. Certaines E2 présentent en plus en N- ou en C-terminal de ce domaine des extensions permettant de faciliter les interactions avec des E3 spécifiques. En effet, chaque E3 reconnaît spécifiquement une ou plusieurs E2 particulières (Hershko et Ciechanover, 1998; Hochstrasser, 1996).

Les enzymes E2 se lient fortement aux molécules E1-S-ubiquitine (Haas *et al.*, 1988) mais ils ne se lient que faiblement à l'ubiquitine ou aux E1 libres (Hershko *et al.*, 1983; Miura *et al.*, 1999). Par contre, les E2 interagissent fortement avec les complexes E3s (Bailly *et al.*, 1994). Il semble probable alors que l'enzyme E1 activée se lie à un complexe E2-E3. Le transfert de l'ubiquitine de E1 sur E2 puis sur la protéine cible se ferait alors très rapidement (Haas *et al.*, 1988).

c) L'ubiquitine ligase E3

La dernière étape de l'ubiquitinylation est réalisée par le complexe E3. Ce complexe est constitué de nombreuses protéines différentes. La réaction catalysée par E3 implique au moins deux étapes distinctes : l'interaction avec le substrat (protéine cible) et le transfert covalent d'une ou des molécules d'ubiquitine sur le substrat. L'étape réalisée par l'E3 est considérée comme celle responsable de la reconnaissance spécifique du substrat. E3 possède des sous-unités variables répertoriées au nombre de 500 à 1000 protéines différentes à ce jour. Ces protéines ont été classées parmi quatre grandes catégories sur la base de motifs structuraux : le groupe « RING-finger » ou à « doigt de zinc », le groupe à domaine « HECT », celui de type U-box et celui à doigt PHD. Le groupe à « doigt de zinc » est considéré comme celui regroupant le plus de protéines et est subdivisé en sous-familles. Parmi elles, la plus représentée est celle des protéines de type culline nommées Cul suivie d'un chiffre (leur nom vient de l'anglais « to cull » qui signifie réformer ou massacrer). Il y a sept complexes E3 contenant une protéine de type culline, notamment l'APC/C (Anaphase-promoting complex/cyclosome) et le SCF (Skp1-Cul1-F-box-protein) (Castro *et al.*, 2005; Harper *et al.*, 2002) qui sont tous deux impliqués dans la protéolyse des acteurs moléculaires régulant la progression du cycle cellulaire.

L'organisation de ces complexes est similaire. Le complexe APC/C se compose d'éléments de base invariables dont APC11 (apparenté à la protéine à « doigt de zinc » Rbx1), APC2 (apparenté à Cul1), d'éléments variables connus comme activateurs et d'au moins 11

autres composants dont le rôle reste encore mal compris.

Parmi ces deux complexes, les fonctions du complexe SCF ubiquitine ligase sont les mieux caractérisées. Au cœur de ce complexe, la sous-unité Cul1 agit en tant que squelette moléculaire, interagissant simultanément à l'extrémité N-terminale avec Skp1 (S-phase-kinase-associated protein-1) et à l'extrémité C-terminale avec RBX1 (protéine à « doigt de zinc », appelée aussi Roc1 ou Roc2) et une enzyme E2 spécifique telle que Ubc3, Ubc4 ou Ubc5. En retour, Skp1 interagit avec l'une des nombreuses protéines à domaine F-box ou FBP. Chaque FBP semble fixer sélectivement un nombre défini de substrats (Figure 14) *via* un domaine d'interaction protéine-protéine. Environ 70 FBP ont été identifiées chez l'homme. Elles ont été nommées selon la structure de leur domaine de fixation au substrat. Ces domaines sont presque toujours en C-terminal du domaine F-box et il semble qu'il n'y ait pas plus d'un domaine F-box par FBP (pour revue voir : Cardozo et Pagano, 2004). Les FBP se repartissent en trois catégories : celles à répétition de séquence WD40 (FbxW), celles riche en leucine (FbxL) ou celles à domaine FbxO (Jin *et al.*, 2004). Le

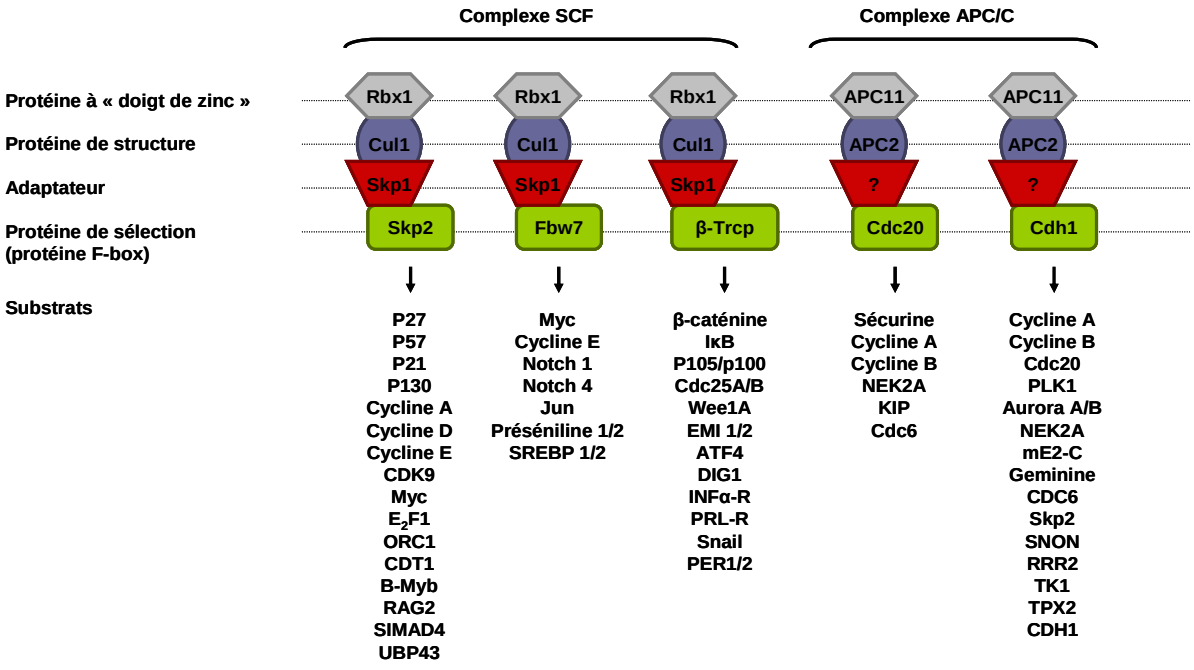


Figure 14 : Les complexes SCF et APC/C. Représentation schématique de ces complexes montrant qu'ils sont tous deux constitués de trois sous-unités communes et d'une sous-unité variable de reconnaissance et de sélection du substrat. (Castro et al., 2005; Harper et al., 2002; Nakayama et Nakayama, 2005).

domaine WD40 des FBP est organisé en hélice β[1]. L'organisation structurale de ces domaines, rencontrés dans de nombreux contextes de fixation protéine-protéine (Smith *et al.*, 1999), semble permettre la reconnaissance spécifique d'un résidu sérine ou thréonine phosphorylé (pS/pT) au sein du consensus DpSGXXX(X)pS de Fbw1 (ou β-Trcp1) et au sein du consensus variable L[I/L/P][pS/pT]P de Fbw7. Le domaine riche en résidus de leucine est constitué de feuillets β qui s'organisent en forme d'arc. Ces structures sont similaires à celle des domaines d'interaction de certains récepteurs membranaires (Kobe et Kajava, 2001). Les FBP ne possédant pas de domaine WD40 ou riche en résidus leucine constituent la dernière catégorie. Ces FBP ont des domaines d'interaction protéine-protéine de type CASH (interaction carbohydrate), homologue au boîte cycline, CH (homologue aux calponines), TDL (Traf domain-like), Sec7, « doigt de zinc » et riche en proline (Cardozo et Pagano, 2004).

Les substrats ciblés pour l'ubiquitinylation sont reconnus par les différents complexes E3 *via* des motifs spécifiques. Le motif PEST (riche en résidus proline, glutamate, sérine et thréonine) est retrouvé dans de nombreux substrats reconnus par le complexe SCF tel que la cycline D1, IκB, Fos, Jun, Myc ou p53 (Rechsteiner et Rogers, 1996). La phosphorylation des résidus de sérine ou de thréonine des séquences PEST, nécessaire à l'interaction du substrat et du complexe SCF, constitue un signal de dégradation de la protéine (Rogers *et al.*, 1986). Par ailleurs, les séquences D-Box (de consensus RXXLXXXXN/D/E) et KEN box (de consensus KENXXXD/N) ont été identifiées dans les substrats du complexe APC/C et sont indispensables à la dégradation des protéines concernées (Glotzer *et al.*, 1991; Pfleger et Kirschner, 2000; Zachariae et Nasmyth, 1999). Bien que certains substrats possèdent les deux motifs, comme cela a pu être démontré pour la protéine Cdc6 ou la cycline A (Jacobs *et al.*, 2001; Petersen *et al.*, 2000), il semble que le motif KEN box soit spécifiquement reconnue par le complexe APC/C^{cdh1} (Zachariae et Nasmyth, 1999).

II. La *O*-N-acétylglucosaminylation

La glycosylation prend une part importante dans la modification des protéines et des lipides. A côté de la phosphorylation, elle est la modification post-traductionnelle plus fréquemment rencontrée. Il existe plusieurs types de glycosylation consistant toutes en un transfert de structures mono- ou oligosaccharidiques sur des protéines données effectué par différentes glycosyltransférases. La glycosylation modifie les protéines par fixation covalente

[1] Les brins β organisés en feuillets peuvent s'enrouler eux-mêmes en hélice. Ces structures sont souvent appelées improprement hélice β. Elles peuvent contenir deux ou trois brins par tour. Elles peuvent former des hélices gauches ou droites.

d'un ou des glycannes au niveau de certains acides aminés (Tableau 4). La glycosylation joue des rôles divers et majeurs au sein de la physiologie cellulaire. En effet, elle permet de stabiliser la structure tridimensionnelle des protéines sous leur forme biologiquement active et de les protéger des attaques protéolytiques. Elle détermine la demi-vie des glycoprotéines plasmatiques. Les glycannes sont par exemple des sites récepteurs d'hormones et de protéines, est interviennent dans la reconnaissance cellule-cellule et font partie des épitopes des antigènes du soi et du non-soi.

Les voies de biosynthèse des glycoprotéines membranaires et secrétées sont aujourd'hui élucidées, avec notamment la mise en place de glycosyltransférases orientées vers la lumière du réticulum endoplasmique ou de l'appareil de Golgi (Hirschberg et Snider, 1987). Bien que les glycosylations les plus connues soient localisées au niveau des membranes ou dans l'environnement extracellulaire, il existe en réalité plusieurs autres types de glycosylation présents au sein des compartiments nucléaires et cytoplasmiques. Il est notamment possible de citer : l'alpha-glucosyl attaché à un groupement hydroxyle d'un résidu

Tableau 4 : Principaux types de glycosylation.

		Monosaccharide engagé dans la liaison	Acide aminé portant la glycosylation
N-Glycosylprotéines		GlcNAc	Asn
O-Glycosylprotéines			
Exemples :	Mucine	GalNAc	Ser/Thr
	Protéoglycanne	Xyl	Ser
	Collagène	Gal	OH-Lys
	Extensine	Ara	OH-Pro
	Glycogène	Glc	Tyr
	Protéine de la levure	Man	Ser/Thr
	Protéine nucléaire et cytoplasmique	GlcNAc	Ser/Thr
C-Glycosylprotéines		Man	Trp
O-Fucosylprotéines		Fuc	Ser/Thr

de tyrosine de la glycogénine (servant d'amorce à la synthèse du glycogène) (Blumenfeld et Krisman, 1986; Rodriguez et Whelan, 1985), les résidus de mannose *O*-liés des protéines cytosoliques (Hiller *et al.*, 1987), des protéoglycannes contenant des résidus de *O*-Man retrouvés exclusivement dans le cytoplasme de cellules du cerveau de rat (Margolis *et al.*, 1976; Margolis *et al.*, 1979) ou encore des résidus de *O*-N-acétylglucosamine ou *O*-GlcNAc liés sur l'hydroxyle de résidus sérine ou thréonine des protéines cytoplasmiques et nucléaires (Holt et Hart, 1986; Holt *et al.*, 1987). La *O*-N-acétylglucosaminylation, ou communément appelée GlcNAcylation, est une modification post-traductionnelle mise en évidence pour la première fois en 1984 (Torres et Hart, 1984). C'est une glycosylation particulière et originale car elle ne fait intervenir qu'un unique monosaccharide, la N-acétyl-D-glucosamine, lié au niveau d'un acide aminé hydroxylé (sérine ou thréonine) d'une protéine par l'intermédiaire d'une liaison β (Figure 15). La modification par *O*-GlcNAc est majoritairement présente dans le cytosol et le noyau (Holt et Hart, 1986; Kearse et Hart, 1991b) où elle est particulièrement abondante au niveau du pore nucléaire (Hanover *et al.*, 1987). Elle a été, récemment, identifiée dans la mitochondrie (Hu *et al.*, 2009). La GlcNAcylation est aujourd'hui la glycosylation nucléocytoplasmique la plus étudiée.

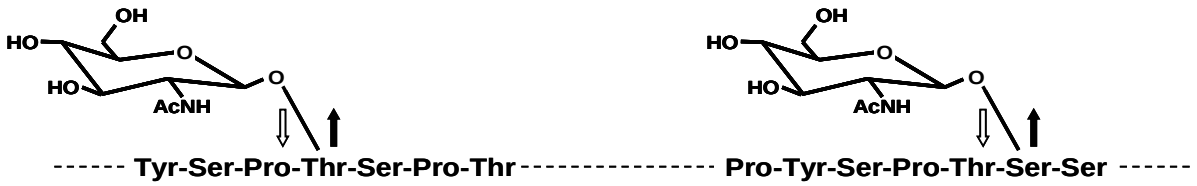


Figure 15 : Représentation schématique d'un polypeptide modifié par GlcNAcylation. La GlcNAcylation consiste en l'ajout d'un unique résidu de N-acétylglucosamine sur des résidus de sérine ou de thréonine.

A. Une modification conservée et ubiquitaire

Depuis sa découverte, plus de 600 protéines ont été identifiées comme modifiées par GlcNAcylation. Ces protéines sont de nature variée, regroupant des facteurs de transcription, des kinases, des protéines du cytosquelette, des protéines de choc thermique, des enzymes métaboliques, des protéines de stucture, etc… Le nombre et la diversité des protéines ainsi modifiées suggèrent que la GlcNAcylation intervient dans différents processus physiologiques à tous les niveaux de régulation de la vie cellulaire. Bien que le rôle de la GlcNAcylation ne soit pas encore complètement compris, de nombreuses études démontrent qu'elle est

impliquée dans la régulation de la phosphorylation et de l'activité enzymatique, dans les interactions protéine-protéine, dans le contrôle de la durée de demi-vie des protéines et qu'elle pourrait être associée au trafic nucléocytoplasmique. Elle permet ainsi la régulation de nombreux processus cellulaires tels que la transcription, la réponse au stress, le cycle cellulaire et le développement.

Son implication dans les mécanismes fondamentaux de la vie cellulaire laisse penser que la GlcNAcylation est apparue très tôt dans l'évolution. En effet, la conservation phylogénétique de la GlcNAcylation ne fait aucun doute puisqu'elle est décrite dans de nombreux organismes. Ainsi, elle est retrouvée dans quasiment tous les *phyla*. Cette glycosylation est présente depuis les procaryotes comme la bactérie *Lysteria monocytogenes* (Schirm *et al.*, 2004) jusqu'aux métazoaires incluant les mammifères, les amphibiens (Fang et Miller, 2001; Slawson *et al.*, 2002). Elle est retrouvée chez les eucaryotes unicellulaires comme le parasite *Plasmodium falciparum* (Dieckmann-Schuppert *et al.*, 1993), chez le ver nématode *Caernorhabditis elegans* (Hanover et al., 2005) ou chez les insectes tels que la drosophile (Kelly et Hart, 1989). Elle a également été décrite chez les végétaux (Swain *et al.*, 2001). Elle pourrait être présente chez les archéobactéries comme *Pyrococcus furiosus* (Mizanur et al., 2005). Elle est même retrouvée à la frontière du vivant chez certains virus tels que SV40 (Medina *et al.*, 1998) et le cytomégalovirus (Greis *et al.*, 1994).

B. Les sites de la GlcNAcylation

La GlcNAcylation ne possède pas de séquence consensus proprement dite. Néanmoins, il existe certaines exigences nécessaires à la reconnaissance du substrat par l'enzyme de transfert du monosaccharide. La présence de séquences de type proline-sérine/thréonine-sérine/thréonine, proline-valine-sérine/thréonine ou riches en acides aminés hydroxylés (sérine et/ou thréonine) est favorable à cette glycosylation.

Certains sites de GlcNAcylation ont été identifiés par spectrométrie de masse comme, par exemple, la sérine 16 du récepteur murin aux œstrogènes (Cheng et Hart, 2001) ou la thréonine 58 de c-Myc (Chou *et al.*, 1995a). L'analyse des cibles de la *O*-GlcNAc transférase (Haltiwanger *et al.*, 1990), l'enzyme permettant le transfert du résidu de GlcNAc sur la protéine cible, ont permis d'établir que les sites de GlcNAcylation sont similaires à ceux reconnus par certaines kinases plus particulièrement les kinases proline-dépendantes (Vulliet *et al.*, 1989), suggérant que la GlcNAcylation et la phosphorylation seraient réciproques.

C. Les enzymes de la GlcNAcylation

a) L'uridine diphospho-*N*-acétylglucosamine : polypeptide β-*N*-acétylglucosaminyltransférase (OGT)

La GlcNAcylation est une glycosylation particulière et, à ce titre, nécessite une glycosyltransférase particulière différente de celle de la voie de sécrétion. En effet, lors de sa découverte, aucune enzyme connue n'était capable d'effectuer le transfert du monosaccharide sur les protéines cibles. Au début des années 90, Haltiwanger *et al.* ont décrit (Haltiwanger *et al.*, 1990) puis isolé et caractérisé (Haltiwanger *et al.*, 1992) l'OGT à partir du foie de rat. L'extinction génique du gène de l'OGT est létale chez la souris au stade embryonnaire et provoque également la mort des cellules souches embryonnaires (Shafi *et al.*, 2000) suggérant que cette enzyme occupe une fonction cellulaire primordiale.

1. Le gène de l'OGT

L'ADNc de l'OGT a été cloné quelques années plus tard (Kreppel *et al.*, 1997; Lubas *et al.*, 1997). Le gène *Ogt* code pour une « nouvelle » glycosyltransférase n'ayant pas de similarité de structure ou de séquence avec les autres glycosyltransférases (Kreppel *et al.*, 1997). Ce gène est fortement conservé au cours de l'évolution puisque l'*Ogt* humain et celui de *C. elegans* présentent 80 % d'homologie (Lubas *et al.*, 1997). Contrairement aux autres animaux, le gène de l'OGT est présent en deux exemplaires chez le poisson zèbre : *Ogta* et *Ogtb* (Webster *et al.*, 2009). Chez les plantes, il en existe également deux, homologues à l'*Ogt* animal, nommés *SPY* pour « spindly » (Robertson *et al.*, 1998) et *SEC* pour « secret agent » (Hartweck *et al.*, 2002).

Chez les mammifères, le gène *Ogt* est localisé sur le chromosome X dans la région q13.1 (Nolte et Muller, 2002; Shafi *et al.*, 2000). Il est composé de 23 exons et de 21 introns dont la transcription est sous la dépendance de deux promoteurs alternatifs (Hanover *et al.*, 2003). Trois transcrits sont issus de la transcription du gène *Ogt* : deux à partir du promoteur P1 dont l'épissage alternatif permet la synthèse de la ncOGT (isoforme nucléaire et cytoplasmique) et de la mOGT (isoforme mitochondriale), et un, issu du promoteur P2, dont la traduction débute à partir d'une méthionine interne et produit la sOGT de plus petite taille (Figure 16 A, page 43). Les exons 1 à 4 codent pour 3 domaines d'interaction TPR (tetratricopeptide repeat). L'intron 4 contient le promoteur alternatif P2 ainsi que l'exon alternatif 5 encodant le domaine d'adressage à la mitochondrie (MTS). Les exons 7 à 13 permettent la traduction de 9 domaines TPR. L'utilisation d'une méthionine en phase interne

pour la traduction des transcrits issus du promoteur P2 conduit à la présence uniquement de deux domaines TPR pour la sOGT.

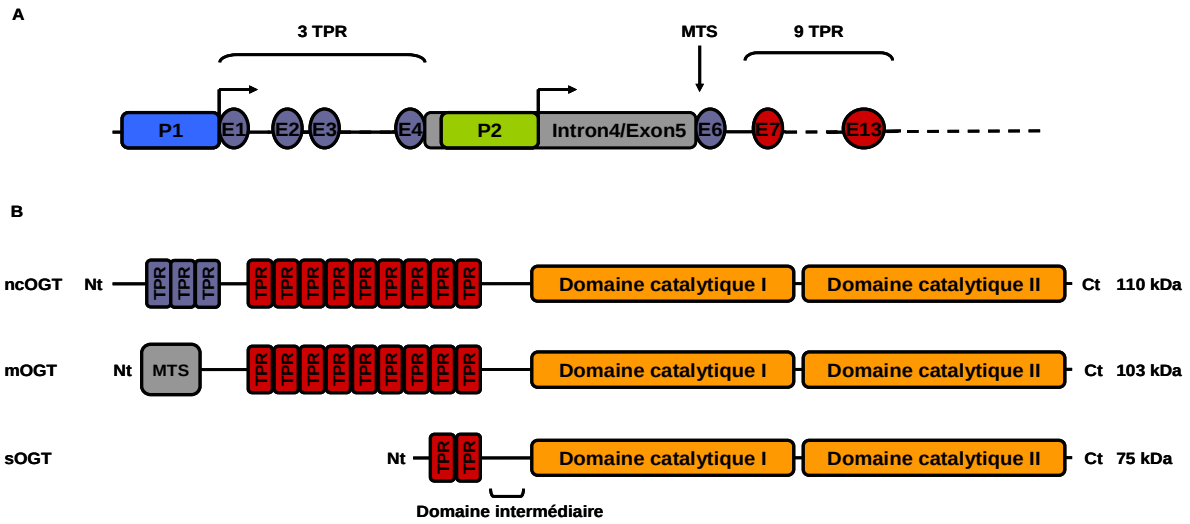


Figure 16 : Les trois isoformes de l'OGT humaine sont codées par un gène unique.
(A) Représentation schématique de la partie 5' du gène de l'OGT dont la transcription et l'épissage alternatif conduiront aux trois isoformes de l'OGT. (B) Représentation schématique des trois isoformes de l'OGT. La ncOGT est l'isoforme de pleine taille issue du promoteur P1. Les transcrits de la mOGT, également issus du promoteur P1, subissent un épissage permettant l'introduction d'une séquence d'adressage à la mitochondrie (MTS). Les transcrits de la sOGT sont issus du promoteur P2. Leur traduction nécessite l'utilisation d'une méthionine en phase interne conduisant à la synthèse d'une protéine tronquée. (Inspiré de Hanover et al., 2003).

2. Structure et caractéristiques de l'OGT

L'OGT est une protéine multimérique de masse apparente de 340 kDa (Haltiwanger *et al.*, 1992). Elle existe principalement sous la forme d'un homotrimère constitué de trois sous-unités de nOGT (Kreppel *et al.*, 1997). Cependant, elle peut être retrouvée dans certains tissus sous la forme d'un hétérotrimère composé de deux sous-unités de 110 kDa, correspondant à l'isoforme ncOGT, et d'une sous-unité de 78 kDa, correspondant à la sOGT (Haltiwanger *et al.*, 1992). En effet, contrairement à la ncOGT dont l'expression est ubiquitaire, l'expression de la sOGT n'est retrouvée que dans le foie, les reins et les muscles. Cette sOGT qui n'a aucune activité catalytique *in vitro* (Lazarus *et al.*, 2006) pourrait n'exercer qu'un rôle de régulateur tissu spécifique. La mOGT, de localisation mitochondriale, possède une activité enzymatique *in vitro* (Love *et al.*, 2003) mais ce n'est que récemment

que son activité *in vivo* a pu être démontrée par la mise en évidence de la GlcNAcylation de protéines codées par le génome mitochondrial (Hu *et al*., 2009).

Quelque soit l'isoforme considérée, l'OGT est constituée de deux domaines distincts reliés par un domaine intermédiaire (Figure 16, page 43). Ce domaine médian est flexible et adopte une conformation en sillon superhélicoïdal (Martinez-Fleites *et al.*, 2008). La partie C-terminale porte l'activité enzymatique. En effet, la délétion de cette extrémité conduit à une perte totale de l'activité enzymatique (Lubas et Hanover, 2000). Ce domaine est divisé en deux sous-domaines : les domaines catalytiques I et II (Roos et Hanover, 2000). Le domaine I porte l'activité catalytique proprement dite alors que le domaine II permettrait la liaison à l'UDP-GlcNAc grâce à une activité de type lectinique (Wrabl et Grishin, 2001). Les données cinétiques suggèrent que l'OGT fonctionne selon un mécanisme enzymatique bi-bi-aléatoire, c'est-à-dire que les deux sites de fixation des substrats (UDP-GlcNAc et protéine) sont accessibles sur l'enzyme libre et indépendants.

La partie N-terminale contient de nombreux motifs TPR. Les TPR sont des séquences de 34 acides aminés contenant le consensus WLGYAFAP (Das *et al.*, 1998; Lamb *et al.*, 1995). Les domaines TPR sont retrouvés dans de nombreuses protéines chez différentes espèces. Ils sont impliqués dans la reconnaissance protéine-protéine et la fixation des substrats (Das *et al.*, 1998) et pourraient être responsables en partie de la sélectivité de ces substrats (Hanover *et al.*, 2003). En outre, leurs délétions ont permis de montrer qu'ils sont impliqués dans la multimérisation de l'enzyme (Kreppel et Hart, 1999). Les TPR influencent donc la multimérisation de l'OGT, la sous-unité p110 pouvant former un homodimère en l'absence de la sous-unité p78.

3. Régulation de l'activité de l'OGT

Les mécanismes régulant l'activité de l'OGT et lui conférant ses spécificités de substrats sont encore mal compris. L'OGT est sa propre cible puisqu'elle est, elle-même, modifiée par GlcNAcylation (Kreppel *et al.*, 1997). Elle est également la cible de kinases et il a été démontré que son activité nécessite une phosphorylation réalisée par la calmoduline kinase IV (Song *et al.*, 2008).

Par ailleurs, l'activité de l'OGT pourrait être modulée par ses interactions avec différentes protéines telles que GRIF-1 et OIP106 (Iyer *et al.*, 2003), la phosphatase PP1 (Wells *et al.*, 2004), le répresseur transcriptionnel mSin3A (Yang *et al.*, 2002) et l'ataxine10 (Andrali *et al.*, 2005). En effet, l'interaction de l'OGT avec une phosphatase permettrait la

déphosphorylation du site cible du substrat préalable à sa GlcNAcylation. Le recrutement de l'OGT par mSin3A permettrait une redistribution de l'OGT au niveau des complexes transcriptionnels ce qui lui permettrait d'avoir une activité extrêmement ciblée en terme de localisation et d'agir sur la transcription des gènes. En outre son interaction avec l'ataxine 10 modulerait son activité.

La multimérisation de l'enzyme n'est pas indispensable à son activité mais influence son affinité pour son substrat. Par ailleurs, il semble que l'activité de l'enzyme est dépendante des taux en UDP-GlcNAc intracellulaires. En effet, il a été démontré que l'OGT recombinante de rat possède *in vitro* trois Km différents (6 µM, 35 µM et 217 µM) pour des concentrations en UDP-GlcNAc croissantes (Kreppel et Hart, 1999).

b) La *N*-acétyl β-D-glucosaminidase (OGA)

L'OGA est l'enzyme qui hydrolyse le résidu de GlcNAc conférant ainsi son caractère de réversibilité à la GlcNAcylation. L'OGA est connue depuis longtemps sous le nom d'hexosaminidase C (Braidman *et al.*, 1974) mais ce n'est qu'en 1994 qu'elle a été isolée et caractérisée à partir de la rate du rat en tant qu'OGA (Dong et Hart, 1994). L'enzyme clive spécifiquement les résidus GlcNAc dont le carbone C1 est engagé dans une liaison d'anomérie β.

1. Le gène de l'OGA

L'ADNc de l'OGA a été cloné en 2001 (Gao *et al.*, 2001). Chez l'homme, il est localisé sur le chromosome 10 dans la région q24.1 à q24.3 (Gao *et al.*, 2001; Heckel *et al.*, 1998). Le gène de l'OGA est ubiquitaire. il est présent et conservé depuis les procaryotes jusqu'à l'homme (Dennis *et al.*, 2006). Les OGA de mammifères conservent plus de 80 % d'homologie entre elles et 55 % avec celle de *C. elegans* (Gao *et al.*, 2001). L'OGA ne possède pas d'homologie de séquence avec d'autres glycosidases. Cependant, de faibles homologies sont respectivement retrouvées entre les extrémités codant pour les régions N- et C-terminales avec des hyaluronidases et des acétyltransférases (Schultz et Pils, 2002). Chez l'homme, l'OGA est exprimée dans tous les tissus (Comtesse *et al.*, 2001; Gao *et al.*, 2001) et après épissage alternatif, il y a production de deux transcrits dont la traduction conduit aux deux isoformes identifiées : l'isoforme de pleine taille (130 kDa) et l'isoforme de 75 kDa (Comtesse *et al.*, 2001; Gao *et al.*, 2001; Heckel *et al.*, 1998) (Figure 17, page 46). L'isoforme de petite taille est due au maintien de l'exon 10, contenant un codon stop en phase. Les isoformes de l'OGA auraient des localisations subcellulaires préférentielles puisque dans une

lignée cellulaire de glioblastome, l'OGA de 75 kDa est principalement localisée dans le noyau alors que son homologue de pleine longueur est plutôt cytoplasmique (Comtesse *et al.*, 2001; Gao *et al.*, 2001; Wells *et al.*, 2002). Chez le rat, il apparaît que l'épissage alternatif des transcrits du gène de l'OGA permet la traduction de deux isoformes supplémentaires de 90 et 84 kDa (Toleman *et al.*, 2004).

2. Structure et caractéristiques de l'OGA

Toutes les isoformes d'OGA connues possèdent une activité *in vitro*. Cependant, l'OGA de pleine taille est beaucoup plus active que l'isoforme de 75 kDa (Kim *et al.*, 2006). Peu de données sont présentes dans la littérature à propos des caractéristiques structurales de l'enzyme. Il apparaît cependant que l'activité catalytique est portée par l'extrémité N-terminale (Toleman *et al.*, 2004) (Figure 17). Pour l'OGA humaine, des expériences de mutagénèse ont montré l'importance des résidus d'aspartate 174 et 175 dans l'activité enzymatique (Cetinbas *et al.*, 2006).

L'OGA serait une enzyme bifonctionnelle puisqu'en plus de son activité *O*-GlcNAcase, elle possède une activité histone acétyltransférase (HAT) (Toleman *et al.*, 2004). Des expériences réalisées avec des constructions d'OGA tronquée montrent que cette activité HAT serait due à la présence d'un domaine structuré en « doigt de zinc » localisé en position C-terminale de la protéine (Toleman *et al.*, 2006) (Figure 17). La mutagénèse des résidus d'aspartate 853 et 884 et de la tyrosine 891 a montré leur importance dans l'activité HAT de l'OGA (Toleman *et al.*, 2004).

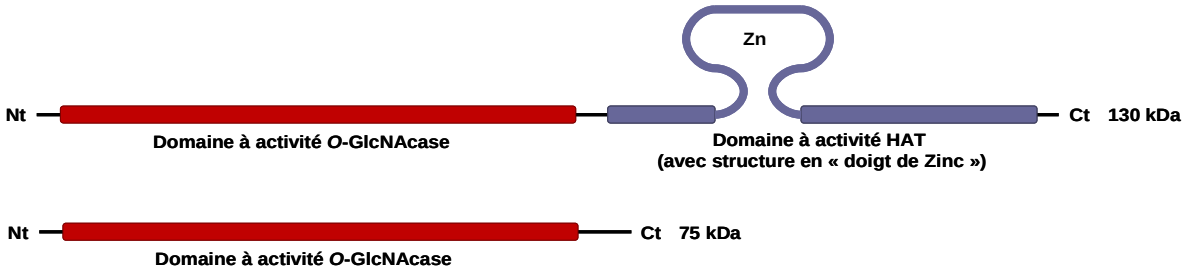


Figure 17 : Représentation schématique des deux isoformes de l'OGA et des domaines O-GlcNAcase et HAT.

3. Régulation de l'OGA

La régulation de l'OGA reste aujourd'hui encore énigmatique. A l'instar de l'OGT, elle subit différentes modifications post-traductionnelles. En effet, l'OGA est GlcNAcylée (Lazarus *et al.*, 2006) suggérant une boucle d'autorégulation entre l'OGT et l'OGA. En outre, l'OGA peut être clivée par la caspase 3 au niveau du résidu d'aspartate 413 (Butkinaree *et al.*, 2008; Wells *et al.*, 2002), libérant ainsi un fragment de 65 kDa qui possède toujours les activités *O*-GlcNAcase et HAT (Toleman *et al.*, 2004). Lors du processus d'apoptose, de nombreuses protéines sont la cible de différentes caspases. Le clivage par ces protéines a pour effet de moduler leur activité de différentes manières. Ainsi, le clivage de l'OGA pourrait modifier son activité lors du déclenchement de l'apoptose. Elle pourrait alors être impliquée dans le phénomène de mort cellulaire en déGlcNAcylant certaines protéines clés.

Par ailleurs, l'isolement de l'OGA a abouti à la purification d'un complexe multiprotéique de 600 kDa au sein duquel elle est associée à différents partenaires tels que des protéines de choc thermique (Hsc 70 et Hsp 110), des transducteurs de signaux intracellulaires tels que la calcineurine ou encore à une protéine impliquée dans l'ubiquitinylation des protéines telle que la culline (Gao *et al.*, 2001).

D. L'UDP-GlcNAc, substrat de l'OGT

Le sucre donneur du résidu *O*-GlcNAc est l'uridine 5'-diphospho-*N*-acétylglucosamine ou UDP-GlcNAc qui est issue de la voie de biosynthèse des hexosamines (Figure 18, page 48). L'activité de l'OGT est dépendante de la synthèse de l'UDP-GlcNAc et donc de l'activité de cette voie de biosynthèse. Une fois transporté dans la cellule, le glucose est rapidement converti en glucose-6-phosphate puis en glucose-1-phosphate pour entrer dans la voie de synthèse du glycogène ou être isomérisé en fructose-6-phosphate. Ce dernier entre préférentiellement dans la voie de la glycolyse. Néanmoins, 2 à 5 % du glucose suivent la voie de biosynthèse des hexosamines. La glutamine : fructose-6-phosphate amido transférase ou GFAT est alors responsable de la conversion du fructose-6-phosphate en *N*-acétylglucosamine-6-phosphate de manière concomitante à la transamination de la glutamine en glutamate. La *N*-acétylglucosamine-6-phosphate est ensuite convertie en *N*-acétylglucosamine-1,6-diphosphate puis en uridine 5'-diphospho-*N*-acétylglucosamine ou UDP-GlcNAc (Marshall *et al.*, 1991). L'UDP-GlcNAc pourra alors servir de donneur pour la glycosylation des lipides et de glycoprotéines sécrétées ou membranaires, pour l'ancrage GPI et pour la GlcNAcylation des protéines nucléocytoplasmiques.

La GFAT est l'enzyme limitante de la voie de biosynthèse des hexosamines. La régulation de l'activité de cette enzyme n'est pas totalement comprise. La GFAT subit une inhibition de rétrocontrôle par l'UDP-GlcNAc (Kornfeld, 1967) et partiellement par la glucosamine-6-phosphate (Broschat *et al.*, 2002). Il est possible d'agir sur cette enzyme pour moduler la GlcNAcylation. En effet, l'utilisation de la glucosamine permet d'activer la synthèse d'UDP-GlcNAc car la porte d'entrée de cette dernière dans la voie de biosynthèse des hexosamines est juste en aval de la GFAT (Champattanachai *et al.*, 2007). A l'opposé, l'azasérine et le DON (6-diazo-5-oxo-norleucine) sont deux inhibiteurs de la GFAT fréquemment utilisés pour diminuer les niveaux d'UDP-GlcNAc (Dauphinee *et al.*, 2005).

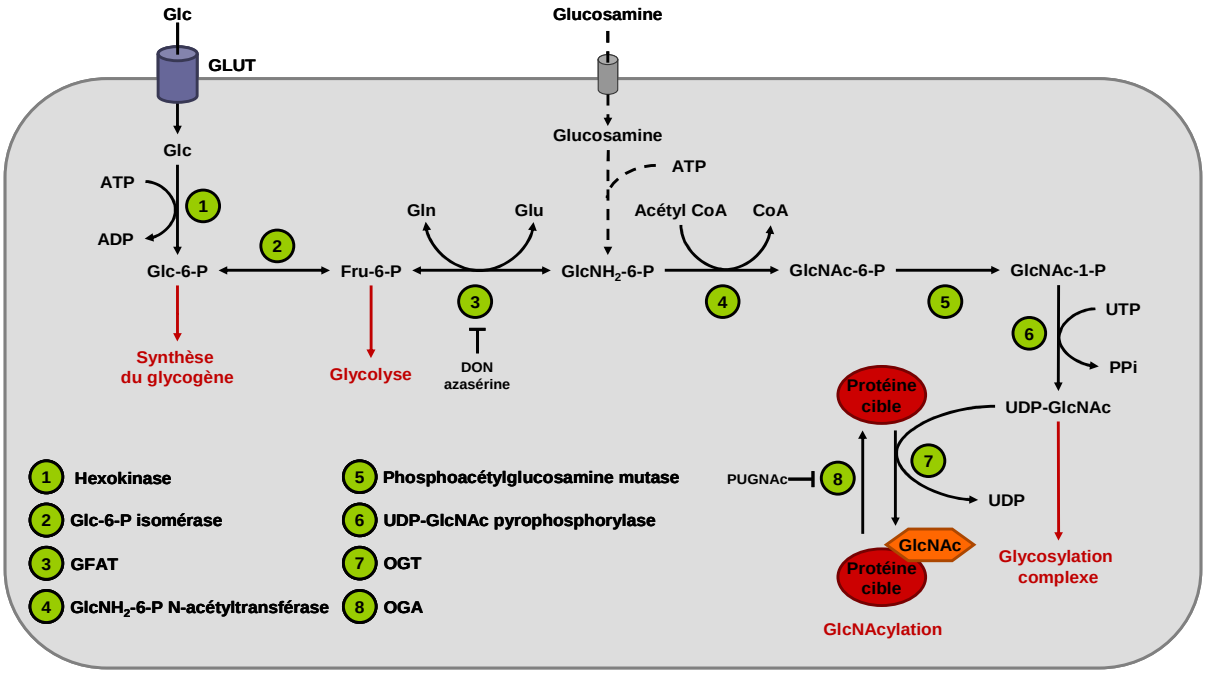


Figure 18 : La voie de biosynthèse des hexosamines.

E. GlcNAcylation et phosphorylation, homologie et antagonisme

La GlcNAcylation, tout comme la phosphorylation, est une modification post-traductionnelle dynamique et réversible (Figure 19, page 49). La demi-vie du motif *O*-GlcNAc étant beaucoup plus courte que la demi-vie de la protéine qu'il modifie (Chou *et al.*, 1992; Kearse et Hart, 1991a; Roquemore *et al.*, 1994), les protéines GlcNAcylées pourraient subir des cycles de GlcNAcylation/déGlcNAcylation.

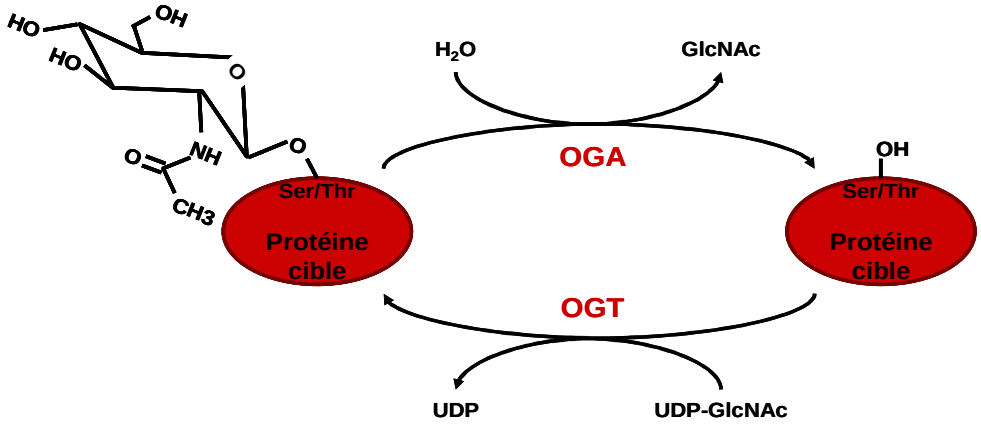


Figure 19 : La GlcNAcylation est une modification post-traductionnelle dynamique et réversible contrôlée par un couple d'enzymes.

Toutes les protéines portant un résidu de *O*-GlcNAc décrites à ce jour sont également des phosphoprotéines. Ainsi, une protéine pourra être soit glycosylée, soit phosphorylée, voire les deux. Pour certaines d'entre elles, il a d'ailleurs pu être démontré qu'un même résidu d'acide aminé hydroxylé ou qu'un résidu hydroxylé adjacent pouvait être soit GlcNAcylé soit phosphorylé. La proximité immédiate de ces sites de modification fait que ces modifications post-traductionnelles s'excluent l'une l'autre par encombrement stérique ou modification des charges de surface. En effet, la modification pharmacologique des taux de phosphorylation affecte de manière opposée ceux de la GlcNAcylation (Griffith et Schmitz, 1999; Lefebvre *et al.*, 1999). En outre, il a été démontré que la déphosphorylation due à l'activité du complexe phosphatase PP1/OGT était suivie d'une GlcNAcylation du substrat (Wells *et al.*, 2004).

Il y a peu d'exemples de cette réciprocité communément appelée réciprocité GlcNAcylation/phosphorylation ou « Ying-*O*-Yang » dans la littérature. Cette réciprocité a pu être démontrée pour le récepteur β murin aux œstrogènes (Cheng et Hart, 2001) pour lequel elle s'exerce sur le résidu de sérine 16 ; pour l'ARN polymérase II (Comer et Hart, 2001) où le « Ying-*O*-Yang » cible sur le domaine C-terminal ; pour la nitrique oxyde synthase endothéliale (eNOS) (Musicki *et al.*, 2005) pour laquelle le site ciblé est le résidu de sérine 1177 ; pour p53 (Yang *et al.*, 2006) qui voit la GlcNAcylation de son résidu sérine 149 inhiber la phosphorylation du résidu de thréonine 155 ; ou pour les résidus de sérine 111 et 112 de l'antigène T du virus SV40 qui peuvent alternativement être modifiés par GlcNAcylation et phosphorylation (Medina *et al.*, 1998).

Cette relation antagoniste entre GlcNAcylation et phosphorylation apporte un haut degré de complexité des modifications d'une protéine. En prenant l'exemple simple d'une protéine modifiée sur deux sites distincts par GlcNAcylation et/ou phosphorylation et ne

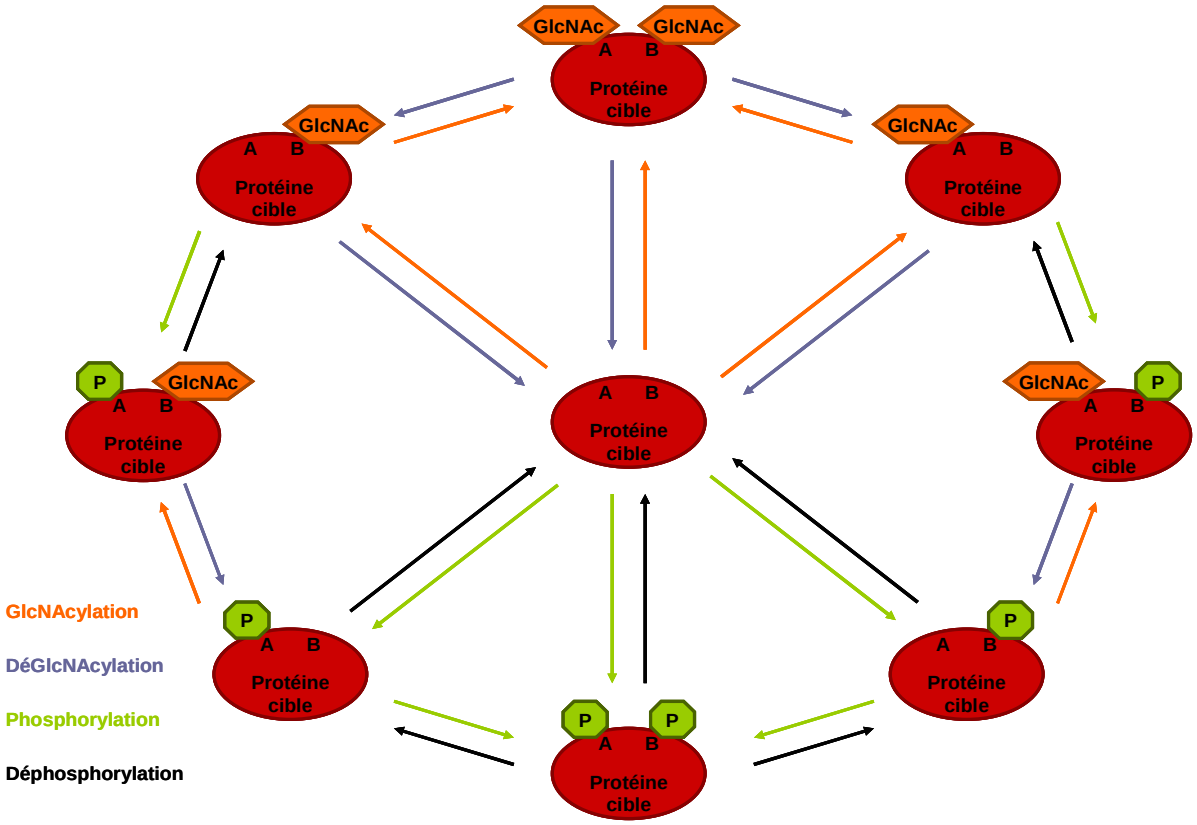


Figure 20 : Modifications par GlcNAcylation et/ou phosphorylation d'une protéine ne possédant que deux sites cibles indépendants (A et B).

présentant aucune autre modification post-traductionnelle, 9 isoformes différentes peuvent être obtenues (Figure 20). Aux vues du grand nombre d'isoformes obtenues, il paraît évident que le « Ying-*O*-Yang » permet d'exercer un haut degré de régulation des protéines où chacune des isoformes protéiques peut avoir des propriétés physiques et biologiques différentes. En plus de pouvoir exercer un rôle propre, la GlcNAcylation peut jouer un rôle d'antagoniste à la phosphorylation. En effet, en masquant l'accès du site aux kinases, la GlcNAcylation empêche la phosphorylation. Par exemple, il est couramment admis aujourd'hui que la phosphorylation des séquences PEST déclenche l'ubiquitinylation puis la dégradation de la protéine. En bloquant cette phosphorylation, la GlcNAcylation va pouvoir protéger cette protéine de la dégradation protéasomale. Ce mécanisme a été mis en évidence grâce aux travaux effectués sur le récepteur β murin aux œstrogènes (Cheng et Hart, 2001). En effet, cette protéine est modifiée par GlcNAcylation ou phosphorylation sur le résidu de sérine 16 qui se trouve au cœur de la séquence PEST. Un autre exemple marquant est celui de

la phosphorylation du domaine C-terminal de l'ARN polymérase II. La GlcNAcylation de ce domaine inhibe sa phosphorylation mais permet l'amarrage de l'ARN ploymérase II au promoteur (phase d'initiation). L'hydrolyse des résidus *O*-GlcNAc permet la phosphorylation de ce domaine nécessaire à l'assemblage du complexe d'élongation, ce qui conduit à l'entrée en phase d'élongation et à la transcription des ARNm (Comer et Hart, 2001).

F. Rôle de la GlcNAcylation

La dynamique de la GlcNAcylation permet de réguler les interactions protéine-protéine, l'activité et la demi-vie de protéines ainsi que leur localisation nucléocytoplasmique (Comer et Hart, 1999), certaines de ces fonctions étant associées à l'antagonisme GlcNAcylation / phosphorylation.

a) Rôle dans le transport nucléaire

L'hypothèse d'un rôle de la GlcNAcylation dans le transport de protéines intracellulaires a été émise depuis sa découverte sur des protéines du pore nucléaire (Hanover *et al.*, 1987; Holt *et al.*, 1987). Un anticorps dirigé spécifiquement contre le motif *O*-GlcNAc est, d'ailleurs, issu fortuitement de recherches sur les nucléoporines (Park *et al.*, 1987). Le rôle de la GlcNAcylation dans le transport nucléocytoplasmique reste, aujourd'hui, sujet à caution. En effet, l'utilisation de WGA (lectine liant le motif *O*-GlcNAc) inhibe le transport nucléaire (Finlay *et al.*, 1987). Cependant, cette inhibition est à attribuer à un encombrement stérique du pore nucléaire plutôt qu'au masquage du résidu modifié par *O*-GlcNAc (Miller et Hanover, 1994). Différentes études sur des facteurs de transcription tels que Pax-6 (Lefebvre *et al.*, 2002), Stat5A (Gewinner *et al.*, 2004; Nanashima *et al.*, 2005) et Sp1 (Dauphinee *et al.*, 2005) ont permis de montrer des localisations différentes entre les isoformes porteuses d'un résidu *O*-GlcNAc et celles non glycosylées, l'isofome GlcNAcylée étant préférentiellement nucléaire. Cependant, l'observation de ce phénomène pourrait être due à une activité accrue de l'OGA dans le compartiment cytoplasmique (Wells *et al.*, 2002). L'exemple de c-Myc est le plus démonstrateur en faveur d'un transport nucléaire. En effet, des expériences de microscopie ont permis de montrer un adressage nucléaire d'un peptide de c-Myc modifié par GlcNAcylation (Kamemura *et al.*, 2002). De même, l'acide okadaïque (inhibiteur de phosphatases) induit une relocalisation nucléaire de la protéine Tau de manière concomitante à une augmentation de sa GlcNAcylation (Lefebvre *et al.*, 2003).

b) Régulation de la stabilité des protéines par la GlcNAcylation

Le protéasome qui est responsable de la dégradation enzymatique des protéines est GlcNAcylé sur plusieurs de ses sous-unités. En effet, des études menées d'abord chez la drosophile ont montré que 5 des 19 protéines formant le 19 S et que 9 des 14 protéines constituant le 20 S sont GlcNAcylées (Sumegi *et al.*, 2003). Ces résultats ont ensuite été confirmés chez les mammifères (Zhang *et al.*, 2003). Cette étude a également permis de mettre en évidence une régulation de l'activité du protéasome par sa GlcNAcylation. En effet, cette modification du protéasome induit une inhibition de l'activité de ce dernier en affectant sa fonction ATPasique. L'ubiquitinylation des protéines, responsable de leur adressage au protéasome, est également régulée par la GlcNAcylation. Une étude réalisée au laboratoire a permis de montrer que l'enzyme responsable de l'activation de l'ubiquitine, UBA1, est GlcNAcylée et qu'une inhibition de la GlcNAcylation cellulaire conduit à une augmentation de l'ubiquitinylation des protéines de manière globale (Guinez *et al.*, 2008). Les résultats de cette étude suggèrent que la voie de l'ubiquitinylation pourrait être régulée en amont par GlcNAcylation.

De manière plus spécifique, il a été démontré que le facteur de transcription Sp1 sous forme hypoGlcNAcylé voit sa durée de demi-vie réduite alors que sa GlcNAcylation a l'effet opposé (Han et Kudlow, 1997; Lee *et al.*, 2008). Cet effet protecteur de la GlcNAcylation des protéines a également été suggéré pour mER-β. La sérine 16 de ce facteur de transcription est la cible d'une GlcNAcylation ainsi que d'une phosphorylation. Située au cœur d'une séquence PEST, la phosphorylation de ce résidu induit la dégradation du récepteur alors que sa GlcNAcylation augmente sa durée de demi-vie (Cheng et Hart, 2001). Par ailleurs, il a également été démontré que la GlcNAcylation de la sérine 149 de p53 induit une diminution de son ubiquitinylation par inhibition de la phosphorylation de la thréonine 155 (Yang *et al.*, 2006).

c) Implication de la GlcNAcylation dans les interactions protéine-protéine

La GlcNAcylation modulerait les interactions entre les partenaires protéiques en induisant des changements de conformation comme cela a été observé pour p53 (Shaw *et al.*, 1996). Ces modifications conformationnelles pourraient être dues à la relation antagoniste entre GlcNAcylation et phosphorylation. En effet, la présence d'un groupement phosphate apporte localement une charge anionique à la protéine alors qu'un groupement *O*-GlcNAc (de charge neutre) apporte un encombrement stérique. La littérature relate quelques exemples

d'inhibition ou d'activation des interactions entre les protéines modifiées par GlcNAcylation et leurs partenaires. C'est notamment le cas pour Sp1 où la glycosylation de son domaine de transactivation inhibe les interactions avec ces partenaires TAF110 et holo-Sp1 (Roos *et al.*, 1997). Les interactions entre CREB et TAFII130 sont inhibées par la GlcNAcylation de CREB (Lamarre-Vincent et Hsieh-Wilson, 2003). De même, la présence d'un résidu de GlcNAc sur YY1 empêche le recrutement de la protéine Rb (Hiromura *et al.*, 2003). Par ailleurs, il a été démontré que l'activateur de transcription CBP interagit avec le facteur de transcription STAT5 lorsque ce dernier est sous sa forme GlcNAcylée (Gewinner *et al.*, 2004). De manière plus générale, la présence d'un résidu *O*-GlcNAc permet la formation de complexes multiprotéiques entre les protéines et les protéines de choc thermique comme cela a pu être démontré au laboratoire. En effet, les protéines de choc thermique Hsc 70 et Hsp 70 sont douées d'une activité lectinique (Guinez *et al.*, 2004; Lefebvre *et al.*, 2001) qui leur permet de s'associer à différentes protéines porteuses d'un résidu *O*-GlcNAc. Cette activité lectinique de Hsp 70 augmente, d'ailleurs, en cas de stress nutritionnel (Guinez *et al.*, 2006). Dans certaines conditions, la formation de complexe entre les protéines GlcNAcylées et les protéines de choc thermique permettrait de les protéger d'une éventuelle dénaturation.

d) Action sur l'activité des facteurs de transcription

La régulation de l'expression des gènes par la GlcNAcylation a été suggérée depuis que le CTD de l'ARN polymérase II a été décrit comme porteur de résidus *O*-GlcNAc (Kelly *et al.*, 1993). En effet, la relation de « Ying-*O*-Yang » s'exerçant sur cette enzyme permet de réguler son activité (Comer et Hart, 2001). L'interaction de l'OGT avec le répresseur transcriptionnel mSin3A relocalise l'OGT au niveau de la chromatine et faciliterait la GlcNAcylation de l'ARN polymérase II, la maintenant ainsi en phase d'initiation (Yang *et al.*, 2002). La formation de ce complexe conduirait également à la désacétylation des histones par le recrutement d'histones désacétylases, *via* l'activité de mSin3A, et à une condensation de la chromatine. *A contrario*, les activités *O*-GlcNAcase et HAT de l'OGA permettraient de contrecarrer l'action du complexe OGT-mSin3A.

La transcription des gènes peut être régulée de manière plus spécifique par la GlcNAcylation des facteurs de transcription. A ce jour, plusieurs facteurs de transcription ont été identifiés comme régulés par GlcNAcylation (Tableau 5, page 54). Différentes fonctions des facteurs de transcription, nécessaires à leur activité, sont régulées par cette modification. En effet, en plus d'intervenir dans la régulation de leur interaction avec différents partenaires protéiques, leur trafic ou leur stabilité, leur GlcNAcylation peut également réguler leur liaison

Tableau 5 : GlcNAcylation et facteurs de transcription. Le tableau présente l'effet activateur (+) ou inhibiteur (-) de la GlcNAcylation sur l'activité, la stabilité et la liaison à l'ADN de différents facteurs de transcription.

Facteurs de transcription	Activité	Stabilité	Liaison à l'ADN	Références
Sp1	+			Goldberg *et al.*, 2006; Han et Kudlow, 1997; Jackson et Tjian, 1988
	-			Donadio *et al.*, 2008; Roos *et al.*, 1997; Yang *et al.*, 2001
		+		Han et Kudlow, 1997
			Sans effet	Jackson et Tjian, 1988
			-	Vicart *et al.*, 2006
FoxO1	+	nd	nd	Housley *et al.*, 2008; Housley *et al.*, 2009
c-Myc	-	nd	nd	Chou *et al.*, 1995b
mER-β	-	+	nd	Cheng et Hart, 2001
p53	+			Fiordaliso *et al.*, 2001
		+		Yang *et al.*, 2006
			+	Shaw *et al.*, 1996
NFκB	+	nd	nd	James *et al.*, 2002
Pax6	nd	nd	+	Lefebvre *et al.*, 2002
PDX-1	nd	nd	+	Gao *et al.*, 2003

(nd signifie non documenté)

à l'ADN. En fonction du facteur de transcription considéré, l'effet de cette modification post-traductionnelle sur son activité sera différent. La GlcNAcylation augmente l'activité transcriptionnelle de certains d'entre eux, notamment FoxO1 (Kuo *et al*., 2008), p53 (Fiordaliso *et al*., 2001; Shaw *et al*., 1996) ou NF-κB (James et al., 2002), mais l'inhibe pour d'autres tels que c-Myc (Kamemura et Hart, 2003) ou mER-β (Cheng et Hart, 2001). De plus, le rôle de la GlcNAcylation semble beaucoup plus compliqué qu'il n'y paraît. En effet, la régulation de Sp1 par GlcNAcylation, la plus étudiée à ce jour, est complexe et est dépendante de co-facteurs particuliers (Han et Kudlow, 1997; Yang *et al.*, 2001). Bien que la

GlcNAcylation du domaine de transactivation de Sp1 inhibe son interaction avec TAF 110 qui est indispensable à son activité transcriptionnelle *in vitro* (Yang *et al*. 2001), Golberg *et al*. (2006) ont montré que la GlcNAcylation de Sp1 augmente la transactivation du gène *PAI-1* dans des cellules mésangiales primaires de rat.

G. GlcNAcylation et pathologies

Bien que les rôles de la GlcNAcylation ne soient pas encore totalement élucidés, de nombreuses études démontrent son importance dans de nombreux processus physiologiques. Par ailleurs, des altérations des taux de *O*-GlcNAc sont associées à différentes pathologies telles que les maladies neurodégénératives, le diabète et certains cas de cancer.

a) La GlcNAcylation et le diabète de type II

Le diabète de type II est caractérisé par une hyperglycémie chronique. Cette maladie résulte en partie d'un phénomène de résistance à l'insuline des cellules pancréatiques, du foie, des adipocytes et des muscles squelettiques. Cette résistance est marquée par un défaut de transport du glucose dans ces cellules et donc, par une inhibition de la synthèse du glycogène. L'hyperglycémie sérique ainsi induite conduit à l'apparition de différents problèmes annexes dus à la toxicité du glucose. De nombreuses études relient le diabète de type II et la voie des hexosamines (voir pour revue : Marshall *et al.*, 1991). Comme les taux de *O*-GlcNAc sont dépendants de cette voie de biosynthèse, il a été proposé que cette modification post-traductionnelle puisse être un senseur des taux de glucose. Dans différents modèles cellulaires du diabète, des taux élevés de *O*-GlcNAc ont été retrouvés (McClain *et al.*, 2002). La GFAT, enzyme limitante de la voie de synthèse de l'UDP-GlcNAc, pourrait être un élément central du phénomène de résistance à l'insuline. En effet, les souris transgéniques, surexprimant la GFAT, sont résistantes à l'insuline (Hebert *et al.*, 1996). En outre, Marshall *et al* ont montré que l'azasérine (inhibiteur de la GFAT, Figure 18, page 48) réverse l'hyperglycémie tandis que la glucosamine potentialise la résistance à l'insuline (Marshall *et al.*, 1991). Il existe une étroite corrélation entre l'élévation des taux de *O*-GlcNAc induite par le PUGNAc (inhibiteur de l'OGA, Figure 18, page 48) et le phénomène de résistance à l'insuline. Suite à un traitement au PUGNAc, les adipocytes deviennent résistants à l'action de l'insuline (Vosseller *et al.*, 2002). De plus, la surexpression de l'OGT dans le muscle, le tissu adipeux ou dans le foie provoque cette même résistance à l'insuline (McClain *et al.*, 2002; Yang *et al.*, 2008) qui serait due à la GlcNAcylation de protéines clés de la voie de transduction de l'insuline (Whelan *et al.*, 2009). Dans des cultures primaires d'adipocytes de

rat, une diminution du taux de transporteur du glucose GLUT4 à la surface cellulaire est concomitante à la GlcNAcylation de IRS-1 et d'Akt2 (Park *et al.*, 2005). La glycosylation de ces deux protéines est induite par l'insuline suite au recrutement de l'OGT à la membrane plasmique par le biais du phosphatidylinositol 3,4,5 triphosphate (PIP3) et est responsable d'une inhibition de la transduction du signal (Yang *et al.*, 2008). La glycosylation d'Akt inhibe son activité par diminution de sa phosphorylation (Kang *et al.*, 2008; Soesanto *et al.*, 2008). Le facteur de transcription FoxO1 est l'une des cibles d'Akt. FoxO1 joue un rôle crucial dans l'expression des protéines impliquées dans la néoglucogénèse telles que la glucose 6-phosphatase et la phosphoenolpyruvate carboxykinase. Sa phosphorylation par Akt inhibe ses interactions avec l'ADN ce qui conduit à son inactivation (Issad et Kuo, 2008). Par ailleurs, l'activité transcriptionnelle de FoxO1 est augmentée lorsque celui-ci est GlcNAcylé (Kuo *et al.*, 2008). L'insuline active également la glycogène synthase ; l'activité de cette dernière étant inhibée par GlcNAcylation (Parker *et al.*, 2004; Parker *et al.*, 2003).

La GlcNAcylation est également impliquée dans la voie de synthèse de l'insuline. En effet, NeuroD1 et PDX-1, deux facteurs de transcription impliqués dans cette voie, sont GlcNAcylés (Andrali *et al.*, 2007; Gao *et al.*, 2003). La glycosylation de NeuroD1 induit sa relocalisation nucléaire (Andrali *et al.*, 2007) alors que celle de PDX-1 est associée à une augmentation de son interaction à l'ADN et à une surexpression de l'insuline (Gao *et al.*, 2003).

b) La GlcNAcylation et les maladies neurodégénératives

Les gènes de l'OGT et de l'OGA sont tous deux associés à des *loci* de prédisposition aux maladies neurodégénératives (Gao *et al.*, 2001; Heckel *et al.*, 1998; Myers *et al.*, 2000; Nolte et Muller, 2002; Nolte *et al.*, 2003; Shafi *et al.*, 2000). Par ailleurs, l'OGT et l'OGA sont particulièrement exprimées dans le cerveau (Akimoto *et al.*, 2003; Liu *et al.*, 2004b) et la modification post-traductionnelle par *O*-GlcNAc y est abondante principalement au niveau des protéines du cytosquelette (Dong *et al.*, 1996; Yao et Coleman, 1998a; Yao et Coleman, 1998b). Il a, de plus, été observé une diminution du taux de glucose intracellulaire chez les patients atteints de la maladie d'Alzheimer (Alexander *et al.*, 2002). Une dérégulation de la GlcNAcylation a été mise en évidence par la comparaison des taux de *O*-GlcNAc dans des cerveaux de personnes atteintes de cette maladie *versus* des cerveaux sains (Liu *et al.*, 2004a; Yao et Coleman, 1998b). Différentes de protéines GlcNAcylées sont connues pour être impliquées dans les maladies neurodégénératives mais les fonctions associées à leur glycosylation sont encore méconnues. Parmi elles, la protéine Tau pour « tubulin associated

unit » est la plus étudiée d'entre elles. Elle contrôle l'assemblage, la stabilité et l'orientation des microtubules axonaux. Dans la maladie d'Alzheimer, Tau est hyperphosphorylée ce qui induit son agrégation. Il a, par ailleurs, été observé une hypoGlcNAcylation de la protéine Tau dans des biopsies *post mortem* de patients atteints de la maladie d'Alzheimer (Liu *et al.*, 2009). La protéine Tau pourrait être hyperGlcNAcylée sur plus de 12 sites (Arnold *et al.*, 1996) et il a récemment été démontré *in vitro* que la GlcNAcylation du résidu sérine 356 ralentit l'agrégation de la protéine (Yu *et al.*, 2008). L'utilisation d'un inhibiteur pharmacologique a permis de montrer que l'hyperphosphorylation de Tau apparait parallèlement à une diminution de sa GlcNAcylation (Lefebvre *et al.*, 2003). A l'opposé, la surexpression de l'OGT induit l'hypophosphorylation de la protéine (Robertson *et al.*, 2004). De manière intéressante, il a été démontré que l'hypoglycosylation de Tau ainsi que son hyperphosphorylation résultante, sont corrélées à une diminution de l'entrée du glucose dans le cerveau (Li *et al.*, 2006).

c) La GlcNAcylation et le cancer

La littérature fournit de nombreuses données suggérant un lien entre la GlcNAcylation des protéines et la tumorigénèse. Des variations des taux de *O*-GlcNAc dues à une augmentation de l'activité de l'OGA ont été observées dans des carcinomes primaires de sein par rapport aux tissus sains environnants (Slawson *et al.*, 2001). L'inhibition de l'expression de la glutaminase est associée à une perte de 80 % de l'activité de la GFAT dans les cellules MCF7 (Donadio *et al.*, 2008). Les auteurs ont démontré que la perte d'activité de la GFAT conduit à une diminution de la GlcNAcylation de l'OGT, de Sp1 et de l'ATPase du protéasome. Par ailleurs, les cellules inhibées pour l'expression de la glutaminase présentent un phénotype plus différencié ainsi qu'une capacité proliférative moindre par rapport aux cellules contrôles (Donadio *et al.*, 2008). Les taux de *O*-GlcNAc varient au cours du cycle cellulaire et une inhibition de ces derniers induit un arrêt du cycle cellulaire en G1/S et conduit à une diminution de la durée de la phase S (Slawson *et al.*, 2005). A la transition G2/M, une augmentation globale de *O*-GlcNAc est observée sur les protéines extraites d'ovocytes de Xénope (Lefebvre *et al.*, 2004a). Cette augmentation de la GlcNAcylation semble indispensable à la poursuite du cycle puisque son inhibition ralentit la transition G2/M (Dehennaut *et al.*, 2007). Bien que la surexpression de l'OGT en ovocyte de Xénope induit l'entrée en phase M (Dehennaut *et al.*, 2007), elle conduit à un phénotype polyploïde et à un défaut de la cytocinèse dans les cellules de souris 3T3 (Slawson *et al.*, 2005). Ces différents travaux montrent que l'OGT et l'OGA sont impliquées dans la régulation et la progression du

cycle cellulaire mais pourraient également l'être dans le phénomène de remaniement chromosomique observé au cours de la tumorigénèse. Cependant, les acteurs moléculaires impliqués dans la cancérisation régulée ou dérégulée par GlcNAcylation restent encore peu connus. Certains oncogènes ou gènes suppresseurs de tumeur ont néanmoins été identifiés. En effet, la thréonine 58 de c-Myc est un site de GlcNAcylation mais également un « hot spot » de mutation dans les cas de lymphomes (Chou *et al.*, 1995b). La glycosylation de ce site préviendrait l'activation de la protéine (Chou *et al.*, 1995a). Par ailleurs, une étude récente a montré que la topoisomérase I, une enzyme impliquée dans la réparation de l'ADN, est GlcNAcylée dans une lignée cellulaire porcine (LLPCK-1) et, *in vivo*, dans les cellules de rein de souris. La modification de cette enzyme par GlcNAcylation conduit à une diminution de son activité (Noach *et al.*, 2007). HIC-1 (hypermethylated in cancer 1) possède trois sites majeurs modifiés par un résidu de GlcNAc qui pourraient contribuer à la dimérisation de ce répresseur de la transcription sous-exprimé dans les cas de cancer (Lefebvre *et al.*, 2004b). Le plus connu des suppresseurs de tumeur p53 est également GlcNAcylé. Sa GlcNAcylation permet de le stabiliser et de favoriser sa fixation sur ses promoteurs cibles (Yang *et al.*, 2006).

Travaux personnels

Les gènes cibles de la delta-lactoferrine

La régulation de la transcription des gènes est un élément central pour tout être vivant. En effet, toutes les fonctions d'un organisme, des plus essentielles aux plus complexes, nécessitent une régulation à la fois spécifique et temporelle, voire spaciale pour les organismes multicellulaires. Cette régulation s'effectue principalement en amont de la transcription des gènes. Au niveau nucléotidique, les promoteurs sont les régions régulatrices des gènes et comportent dans leurs séquences différents motifs de taille variable. Ces motifs sont de courtes séquences nucléotidiques conservées au sein d'un groupe de gènes ayant la même régulation. Ainsi, chez l'homme, les promoteurs des gènes dont la transcription est induite en réponse aux œstrogènes ont en commun le motif 5'-AGGTCAnnnTGACCT-3' (Rastinejad, 2001), alors que ceux répondant à une autre hormone stéroïdienne comme les glucocorticoïdes possèdent la séquence 5'-AGAACAnnnTCTTGT-3' (Beato et Klug, 2000; Horie-Inoue *et al.*, 2006). Ces courtes séquences d'ADN appelées éléments de réponse lient des éléments régulateurs spécifiques appelés facteurs de transcription (voir pour revue : Kondrakhin *et al.*, 1995). Ce sont ces protéines particulières qui régulent la transcription des gènes en recrutant différents co-facteurs mais également en recrutant le complexe transcriptionnel contenant l'ARN polymérase II. Les facteurs de transcription sont essentiels à l'expression des gènes assurant l'homéostasie cellulaire.

Contrairement à d'autres protéines intervenant dans la régulation des gènes comme les co-activateurs, les facteurs de transcription doivent posséder un ou plusieurs DBD capables d'interagir avec une ou plusieurs séquences nucléotidiques consensuelles qui leur sont propres. Il existe plus de 30 types de DBD permettant de classer les facteurs de transcription (voir pour revue : Garvie et Wolberger, 2001). L'analyse de la structure de la lactoferrine ne permet de pas d'identifier un DBD classique. L'étude de sa structure tridimensionnelle suggère, cependant, la présence de deux zones très basiques qui ont été proposées comme des domaines d'interaction à l'ADN potentiels (Baker *et al.*, 2002). En outre, de nombreux facteurs de transcription nécessitent, pour exercer leur activité transcriptionnelle, la reconnaissance d'un ligand ou des interactions protéine-protéine de type homo- ou hétérodimère (Marmorstein et Fitzgerald, 2003). La lactoferrine pourrait possèder cette caractéristique puisqu'il a été démontré qu'elle est capable de s'oligomériser *in vitro* (Semenov *et al.*, 1999). Les auteurs ont, de plus, démontré que cette oligomérisation de la

lactoferrine est inhibée par la fixation d'une molécule d'ATP sur le lobe C-terminal. La lactoferrine est capable de lier trois séquences nucléotidiques spécifiques et d'induire l'expression d'un gène rapporteur mis sous la dépendance de deux d'entre elles (He et Furmanski, 1995; Son *et al.*, 2002). Bien que la lactoferrine soit capable d'exercer son activité transcriptionnelle *in vitro via* l'une ou l'autre de ces séquences, une homodimérisation de la molécule pourrait nécessiter la présence des deux *in vivo*. Par ailleurs, sachant que la région inter-lobe se caractérise par une forte concentration de charges positives et pourrait correspondre à un site potentiel d'interaction à l'ADN (Baker *et al.*, 2002), nous pouvons nous demander si chacun des deux lobes de la lactoferrine pourrait faire office de monomère et si un changement de conformation ne pourrait pas conduire à un rapprochement de ces lobes favorisant ainsi la fixation à l'ADN.

I. *Skp1* et la régulation du cycle cellulaire

Le premier phénotype induit par l'expression de la delta-lactoferrine que nous avons décrit au laboratoire est un arrêt du cycle cellulaire observé dans un modèle inductible de cellules HEK 293. Le cycle cellulaire est le mécanisme essentiel qui permet la prolifération cellulaire. C'est une alternance de deux phases majeures : la phase S (Synthèse) de réplication du matériel génétique cellulaire et la phase M (Mitose) de division cellulaire. Au cours de la phase S, l'ADN est répliqué par un processus semi-conservatif pour produire deux chromatines identiques. Les histones sont également dupliquées lors de cette phase. En phase M, la cellule mère se divise en deux cellules filles identiques au cours de quatre étapes : la prophase, l'anaphase, la métaphase et la télophase. Les cellules filles peuvent ainsi recommencer un nouveau cycle. Chacune de ces phases est séparée par une phase intermédiaire de préparation à la suivante : la phase G1 (Gap 1) précède la phase S et permet la croissance cellulaire ; la phase G2 (Gap 2), précédant la phase M, permet de contrôler la bonne réplication de l'ADN et de préparer la division. Il existe une dernière phase qui n'entre pas à proprement parler dans le cycle cellulaire. En effet, il apparaît parfois que toutes les cellules ne se divisent pas et entrent alors dans une phase de latence dite G0 (Gap 0). Les cellules différenciées sont majoritairement en G0. Le passage d'une phase du cycle à la suivante est régulé au niveau des points de restriction. Une fois un point franchi, la cellule traversera la phase suivante jusqu'au prochain point de restriction sans nécessité de facteurs mitogènes. Il existe quatre points de restriction aux transitions G0-G1, G1-S, G2-M et au cours de la mitose (métaphase/anaphase). Chacun de ces points de restriction est extrêmement régulé et il existe des mécanismes de surveillance permettant le contrôle de l'intégrité de

l'ADN et autorisant ainsi ou non le passage à la phase suivante. La succession des différentes phases du cycle cellulaire est sous le contrôle de complexes protéiques spécifiques de type Cycline/Cdk (Figure 21). L'activité de ces complexes varie au cours du cycle grâce à deux mécanismes : le premier met en jeu un système de synthèse et de dégradation de la cycline et le second influe sur l'activité kinase du Cdk par le biais de régulation de la phosphorylation (voir pour revues : Blagosklonny et Pardee, 2002; Segurado et Tercero, 2009; van Leuken *et al.*, 2008). L'une des caractéristiques principales des cellules cancéreuses est la perte des régulations de ces points de contrôle (Dash et El-Deiry, 2004).

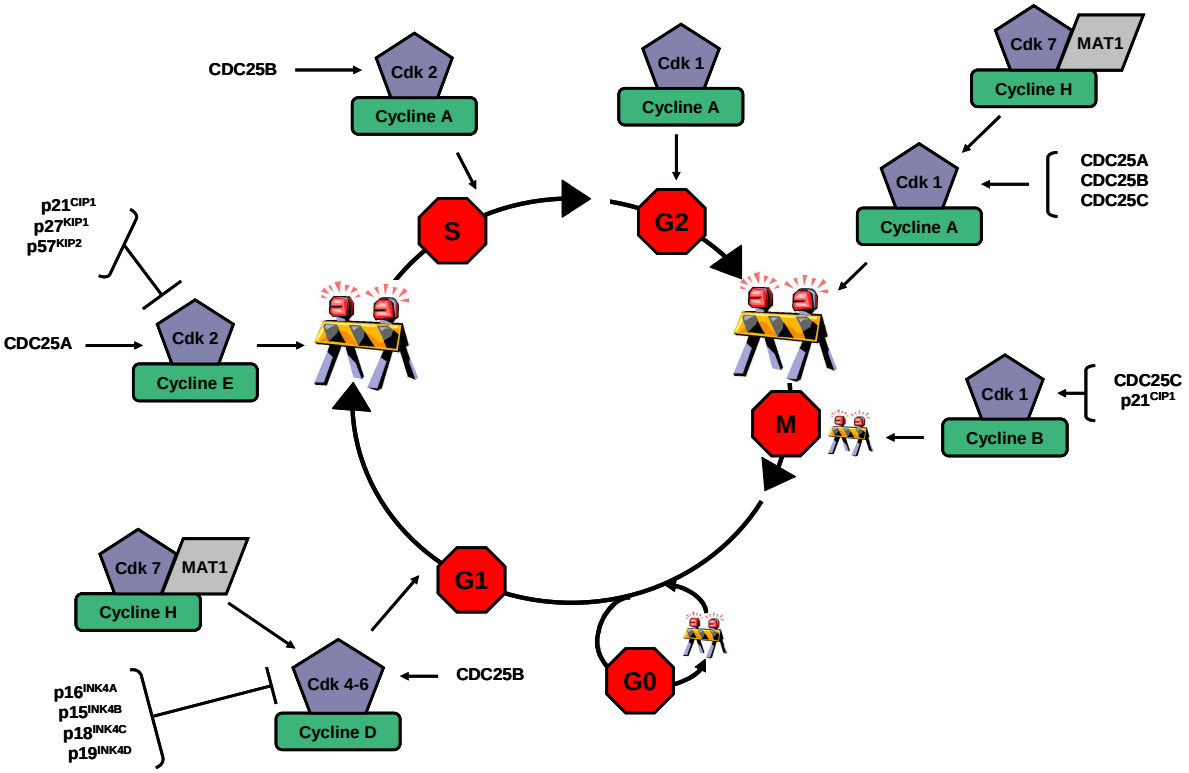


Figure 21 : Principales régulations mises en place lors des différentes phases du cycle cellulaire. Le franchissement des points de restriction nécessite l'activité des complexes Cycline/Cdk eux-mêmes régulés par d'autres complexes Cycline/Cdk ou par l'activité de kinases spécifiques.

Lors de sa découverte, la delta-lactoferrine a été décrite comme étant sous-exprimée dans les lignées cellulaires cancéreuses ce qui a été confirmé au laboratoire, quelques années plus tard, par l'étude de son expression dans 99 biopsies de tumeurs mammaires humaines (Benaïssa *et al.*, 2005). Au cours de cette étude, il a d'ailleurs été démontré une corrélation positive entre l'expression de la delta-lactoferrine et la survie des patientes. L'expression de la

delta-lactoferrine conduisant à l'arrêt du cycle cellulaire en phase S (Breton *et al.*, 2004), nous nous sommes en premier lieu intéressés à son action durant les phases G1 et S afin de mieux comprendre son implication dans la régulation du cycle.

Les résultats de cette étude ont fait l'objet de la publication suivante :

Mariller, C., Benaïssa, M., Hardivillé, S., Breton, M., Pradelle, G., Mazurier, J., et Pierce, A. « Human delta-lactoferrin is a transcription factor which enhances Skp1 (S-phase kinase associated protein) gene expression », 2007, *FEBS J.* 274 (8) : 2038-53

Cette étude nous a permis de démontrer, pour la première fois, que la delta-lactoferrine est effectivement un facteur de transcription et de définir sa ou ses cibles transcriptionnelles à la transition G1/S. Pour se faire, la technique de « macroarrays » a été utilisée, et l'expression de 23 ARNm de protéines clés de cette transition a été analysée. L'hybridation des ARNm de cellules HEK 293 exprimant ou non la delta-lactoferrine a permis de mettre en évidence une surexpression des ARNm de p19, p21, p27, p55, PCNA, Rb, Skp1 et Skp2. Parmi eux, Skp1 présente une augmentation d'un facteur deux à trois de ses taux d'ARNm et protéique. De plus, il a été retrouvé dans le promoteur du gène *Skp1* deux séquences, homologues à S1 et S2 (He et Furmanski, 1995), que nous avons pu définir comme des éléments de réponse fonctionnels à la delta-lactoferrine.

Skp1 est un des composants du système de dégradation des acteurs moléculaires de la transition G1/S appelé SCF (voir généralités page 36). Le complexe SCF ubiquitine-ligase permet la dégradation des inhibiteurs p27 et p21, des cyclines D1, D2, E, A (Yan et Mumby, 1999; Yew, 2001) et de la protéine Cdc6 essentielle pour la formation du complexe de pré-réplication (Figure 22 A, page 63). En outre, Skp1 est également présent lors de la transition G2/M puisqu'il a été montré que sa mutation chez la levure ou son « KO » (knock-out) chez la souris conduit à une mauvaise formation du fuseau mitotique suivie d'une aneuploïdie (Lehmann et Toda, 2004; Piva *et al.*, 2002). Chez la levure, Skp1 contribue à l'assemblage et la fonction normale du complexe CBF3 (kinétochore interne) *via* une Ctf13p (Rodrigo-Brenni *et al.*, 2004). Ctf13p est une F-box protéine qui constitue le cœur structural de CBF3 (Figure 22 B, page 63). Chez l'homme, il semble que Skp1 soit impliquée dans des fonctions similaires puisqu'elle est co-localisée au niveau des centrosomes durant la mitose.

L'interaction de Skp1 avec CENP-E (centromere-associated protein E) est requise pour le déroulement correct de la cytocinèse (Liu *et al.*, 2006).

Une variation des taux de Skp1 pourrait alors avoir d'importantes conséquences sur le déroulement du cycle cellulaire et sur la ségrégation des chromosomes lors de la mitose.

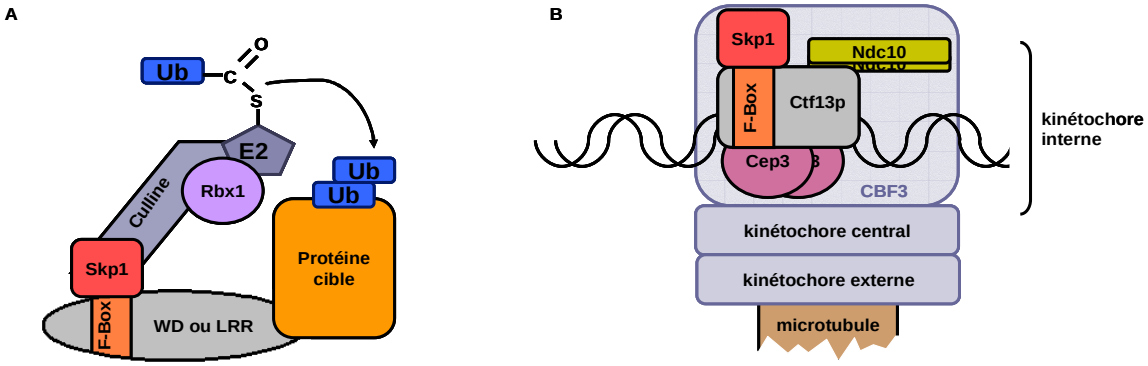


Figure 22 : Skp1 participe à la formation de deux complexes différents agissant à deux phases distinctes du cycle cellulaire (schéma inspiré de : Kitagawa et Hieter, 2001). (A) En phase G1-S, Skp1 participe au complexe SCF en recrutant les protéines à domaine F-Box, lesquelles permettent la sélectivité des substrats à ubiquitinyler. A ce stade du cycle cellulaire, le complexe SFC marque des protéines particulières telles que les inhibiteurs de kinases p27 ou p21 ainsi que les cyclines D, E et A ce qui conduit à leur dégradation et donc au passage des phases G1 et S. (B) Au cours de la mitose, Skp1 est requis pour l'assemblage du centrosome chez la levure.

the FEBS Journal

Human delta-lactoferrin is a transcription factor that enhances *Skp1* (S-phase kinase-associated protein) gene expression

Christophe Mariller, Monique Benaïssa, Stephan Hardivillé, Mathilde Breton, Guillaume Pradelle, Joël Mazurier and Annick Pierce

Unité de Glycobiologie Structurale et Fonctionnelle, Unité Mixte de Recherche 8576 CNRS-Université des Sciences et Technologies de Lille 1, Villeneuve d'Ascq, France

Keywords
cell cycle progression; delta-lactoferrin; proteasome; Skp1; transcription factor

Correspondence
A. Pierce, UGSF Unité Mixte de Recherche 8576 CNRS-Université des Sciences et Technologies de Lille 1, F-59655 Villeneuve d'Ascq cedex, France
Fax: +33 3 20 43 65 55
Tel: +33 3 20 33 72 38
E-mail: annick.pierce@univ-lille1.fr

(Received 3 October 2006, revised 29 January 2007, accepted 16 February 2007)

doi:10.1111/j.1742-4658.2007.05747.x

Delta-lactoferrin is a cytoplasmic lactoferrin isoform that can locate to the nucleus, provoking antiproliferative effects and cell cycle arrest in S phase. Using macroarrays, the expression of genes involved in the G_1/S transition was examined. Among these, *Skp1* showed 2–3-fold increased expression at both the mRNA and protein levels. Skp1 (S-phase kinase-associated protein) belongs to the Skp1/Cullin-1/F-box ubiquitin ligase complex responsible for the ubiquitination of cellular regulators leading to their proteolysis. *Skp1* overexpression was also found after delta-lactoferrin transient transfection in other cell lines (HeLa, MDA-MB-231, HEK 293) at comparable levels. Analysis of the *Skp1* promoter detected two sequences that were 90% identical to those previously known to interact with lactoferrin, the secretory isoform of delta-lactoferrin (GGCACTGTAC-S1^{Skp1}, located at −1067 bp, and TAGAAGTCAA-S2^{Skp1}, at −646 bp). Both gel shift and chromatin immunoprecipitation assays demonstrated that delta-lactoferrin interacts *in vitro* and *in vivo* specifically with these sequences. Reporter gene analysis confirmed that delta-lactoferrin recognizes both sequences within the *Skp1* promoter, with a higher activity on S1^{Skp1}. Deletion of both sequences totally abolished delta-lactoferrin transcriptional activity, identifying them as delta-lactoferrin-responsive elements. Delta-lactoferrin enters the nucleus via a short bipartite RRSDTSLTWNSVKGKK(417–432) nuclear localization signal sequence, which was demonstrated to be functional using mutants. Our results show that delta-lactoferrin binds to the *Skp1* promoter at two different sites, and that these interactions lead to its transcriptional activation. By increasing *Skp1* gene expression, delta-lactoferrin may regulate cell cycle progression via control of the proteasomal degradation of S-phase actors.

The ubiquitin–proteasome system controls the stability of numerous cell regulators, such as cyclins, cyclin inhibitors, transcription factors, tumor suppressor proteins, and oncoproteins [1–3]. Among the ligase complexes, the Skp1/Cullin-1/F-box ubiquitin ligase (SCF) complex is singled out in this work, as its temporal control of ubiquitin–proteasome-mediated protein degradation is critical for normal G_1- and S-phase progression. Here, we show that delta-lactoferrin (ΔLf), expression of which leads to cell cycle arrest in

Abbreviations
ChIP, chromatin immunoprecipitation; DBD, DNA-binding domain; ΔLf, delta-lactoferrin; ΔLfRE, delta-lactoferrin response element; Lf, lactoferrin; NLS, nuclear localization signal; SCF, Skp1/Cullin-1/F-box ubiquitin ligase; Skp1, S-phase kinase-associated protein 1.

S phase, upregulates the synthesis of *Skp1*, one of the SCF components.

ΔLf was first discovered as a transcript [4] that was found in normal cells and tissues but was downregulated in cancer cells and in breast cancer biopsy specimens [4,5]. Our recent investigations have shown that its expression level is of good prognostic value in human breast cancer, with high concentrations being associated with longer relapse-free and overall survival [5]. These findings suggest that ΔLf may play an important role in the regulation of normal cell growth, and demonstrated the need for better characterization of its role.

ΔLf transcription starts at the alternative promoter P2, present in the first intron of the lactoferrin (Lf) gene [6]. Translation of ΔLf starts at the first available AUG codon in-frame present in exon 2, as exon 1β contains a start codon immediately followed by a stop codon [4], and leads to the synthesis of a 73 kDa protein [7]. Thus, ΔLf is a protein devoid of the 45 first amino acid residues present in Lf, which include the leader sequence, implying that ΔLf is cytoplasmic. Moreover, a stretch of four arginine residues of Lf that has been identified as a nuclear localization signal (NLS) and as a putative DNA-binding domain (DBD) [8–10] is absent from ΔLf. However, this does not affect ΔLf nuclear targeting, as ΔLf and green fluorescent protein-tagged ΔLf have been observed in both the cytoplasm and the nucleus [6,7]. Concerning the putative DBD, a strong concentration of positive charges was found at the C-terminal end of the first helix (residues 27–30 in Lf and 2–5 in ΔLf) and at the interlobe region [11,12] that might create other potential DNA interaction sites. Lf is capable of binding DNA [13–16], and specific *in vitro* interactions between Lf and three DNA sequences have already been described [17]. Until now, only one of them had been found in a specific promoter [18].

Most of the previous studies concerning the function of the two isoforms refer to Lf, and do not discriminate between the two Lf isoforms. Whereas only Lf is involved in various aspects of host defense mechanisms [19,20], both Lf and ΔLf may possess antitumoral activities [21]. Whereas Lf acts exogeneously, either directly on tumor cell growth by modulating different transduction pathways [22–26], or via its immunomodulatory effects [20,27], ΔLf acts endogenously, its expression leading to cell cycle arrest in S phase and antiproliferative effects [7].

From these data, several questions arise concerning how ΔLf acts in cells and whether it could regulate cellular proliferation. As ΔLf is able to locate to the nucleus, it might behave as a transcription factor regulating cell cycle progression. We therefore investigated whether ΔLf induces regulation of cell cycle progression, and examined the impact of its expression on key genes involved in the G_1/S transition.

S phase kinase-associated protein (Skp1) is a highly conserved ubiquitous eukaryotic protein belonging to the SCF complex [28,29]. SCF has four components: Skp1, Cullin, and Rbx1, which form the core catalytic complex, and an F-box protein, which acts as a receptor for target proteins. Skp1 is an adaptor between one of the variable F-box proteins and Cullin. At the G_1/S transition, the F-box protein is Skp2, which begins to accumulate in late G_1, and is abundant during S and G_2 [30–32]. SCF is responsible for the ubiquitination of many cell cycle regulators, such as cyclins and cyclin-dependent kinase inhibitors, and at the G_1/S transition it is involved in the recruitment of cyclin E, cyclin A, p21 and p27, leading to their degradation by the proteasome [30,31,33]. At the G_2/M transition, Skp1 belongs to the CBF3 complex [34], which is crucial for kinetochore assembly. In yeast, Skp1 mutants showed increased rates of chromosome misaggregation [35]. In mice, *in vivo* interference with Skp1 function leads to genetic instability and neoplastic transformation [36]. Thus, Skp1 is essential for cell cycle progression at both the G_1/S and G_2/M transitions.

Our findings showed that ΔLf interacts directly with specific DNA sequences present in the *Skp1* promoter, and that these interactions lead to its transcriptional activation. Thus, by causing overexpression of Skp1, ΔLf may influence the proteasomal degradation of some S-phase actors.

Results

ΔLf upregulates *Skp1* expression

Lf expression leads to cell cycle arrest in S phase and antiproliferative effects. As the mechanism by which ΔLf acts in cells is unknown, macroarray analysis was initially performed. Membranes spotted with 23 different genes involved in the regulation of G_1/S phase progression were hybridized with biotin-labeled messengers isolated from 24 h doxycyclin-induced and noninduced ΔLf-HEK 293 cells. Densitometric data were normalized to the expression level of β-actin. The results, presented in Fig. 1A, are expressed as a percentage, where 100% represents the baseline level of each normalized mRNA expressed in the noninduced cells. Among the 23 genes screened (*cdk2*, *cdk4*, *cdk6*, *cyclin C*, *cyclin D2*, *cyclin D3*, *cyclin E1*, *DP1*, *DP2*, *EF*, *E2F-4*, *E2F5*, *p107*, *p130* (*RB2*), $p19^{Ink4d}$, $p21^{Waf1}$, $p27^{Kip1}$, $p55^{cdc}$, $p57^{Kip2}$, *PCNA*, *Rb*, *Skp1*, and *Skp2*),

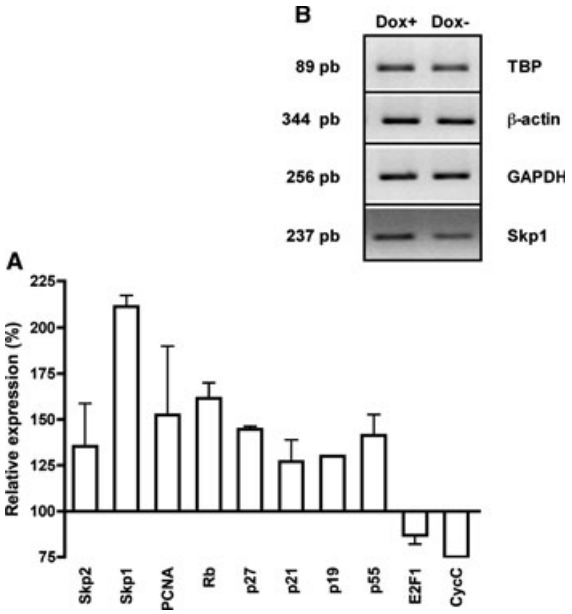


Fig. 1. ΔLf expression leads to *Skp1* upregulation. HEK 293 cells stably transfected with ΔLf (ΔLf-HEK 293) were induced or not with doxycyclin for 24 h. After harvesting, RNA was extracted, quantified, and biotin-labeled to generate separate probes. On each macroarray membrane, 23 genes involved in the G_1/S transition were spotted in duplicate, and two internal controls, *GAPDH* and *β-actin*, in triplicate. Each macroarray membrane was independently hybridized with probe overnight, washed, and exposed to film before densitometric quantification. Expression differences were calculated by the ratio of ΔLf-treated membrane intensity (of a specific gene spot) to its internal housekeeping gene and divided by the ratio of the control membrane intensity (same gene spot) to its internal housekeeping gene. *β-actin* was used to calculate response ratios. (A) The data summarized in the histogram are expressed as a percentage, where 100% represents the baseline level of each normalized mRNA expressed in the noninduced cells. Only significantly differentially expressed genes are presented. (B) Overexpression of *Skp1* in doxycyclin-induced cells was confirmed by RT-PCR using three different housekeeping genes.

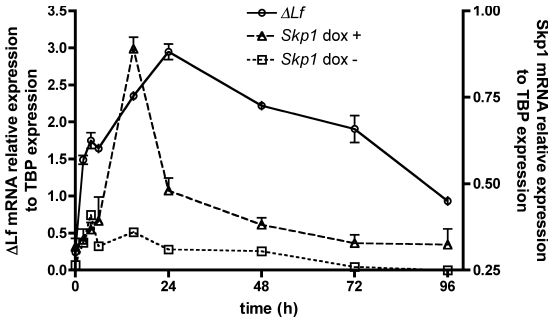


Fig. 2. *Skp1* overexpression is transient in ΔLf-HEK 293 cells. The expression levels of *ΔLf* and *Skp1* mRNA were measured by RT-PCR after induction or not by doxycyclin, and followed for 96 h. Total RNA from ΔLf-HEK 293 cells was harvested at different times, retrotranscribed, and amplified. PCR product signals were integrated using QUANTITY ONE software at cycles 35 for *ΔLf*, 30 for *Skp1*, and 25 for *TBP*. The expression of each transcript is normalized to *TBP* expression, and is expressed as the ratio of *Skp1* or *ΔLf* expression to *TBP* expression ($n = 3$).

few were significantly differentially expressed, and *Skp1* was the most affected by ΔLf overexpression, showing a two-fold increase. The increase of *Skp1* expression was confirmed by RT-PCR using the same RNA source. Whichever internal controls were used, a two-fold increase was observed (Fig. 1B). RT-PCR was also performed for the other genes, but the slight increases observed by macroarray analysis were not confirmed, apart for *Rb*, which was overexpressed 1.5-fold (data not shown).

Next, the upregulation of *Skp1* was followed after induction of ΔLf expression by doxycyclin for 4 days in ΔLf-HEK 293 cells (Fig. 2). ΔLf expression diminished only slightly after 48 h, due to the degradation of the doxycyclin. A very low level of *Skp1* was observed in uninduced ΔLf-HEK 293 cells. A peak of induction was visible, with a maximum 12 h after induction of ΔLf expression by doxycyclin. Therefore, *Skp1* upregulation follows induction of ΔLf expression, is transient, and corresponds to a 2–3-fold increase. These data suggest that this phenomenon might be strongly regulated.

In order to study the cell specificity of the process and to quantify putative ΔLf transcriptional activity, a transient transfection model was developed in parallel. Transient transfection was efficient, and also led to a 2–3-fold increased expression of *Skp1* (Fig. 3A). The maximum was observed with 2 μg of ΔLf plasmid for 10^6 cells (Fig. 3B). This overexpression was not specific to HEK 293 cells, but was also visible in HeLa and MDA-MB-231 cell lines at a comparable level.

As upregulation of gene expression is not always followed by overexpression of the protein, immunoblotting on HEK lysates transfected either with a 'null' plasmid or with increasing concentrations of pcDNA-ΔLf was performed. This showed that the amount of Skp1 protein increased in the lysate of the transfected HEK cells (Fig. 4A). The histogram corresponds to the compiled data from three independent experiments normalized to the cellular protein content. A maximum of 2–3-fold enhancement was obtained either with 1 μg or 2 μg of ΔLf-plasmid for 10^6 cells (Fig. 4B), suggesting that ΔLf concentration might be regulated either at the translational level or post-translationally by proteasomal degradation. Therefore, ΔLf expression leads to the upregulation of Skp1 at both the RNA and protein levels.

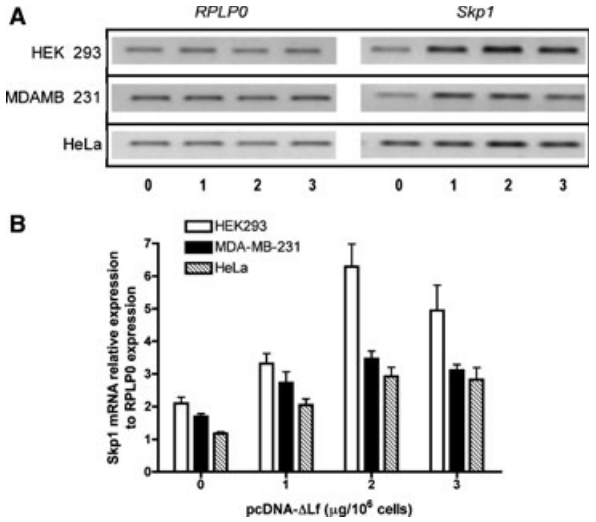


Fig. 3. Overexpression of *Skp1* is not cell-specific. (A) The expression pattern of *Skp1* transcripts in HEK 293, MDA-MB-231 and HeLa cells 24 h after transient transfection by increasing concentrations of pcDNA-ΔLf was followed by RT-PCR. (B) The expression of each transcript is normalized to *RPLP0* expression and is expressed as the ratio of *Skp1* expression to *RPLP0* expression ($n = 3$).

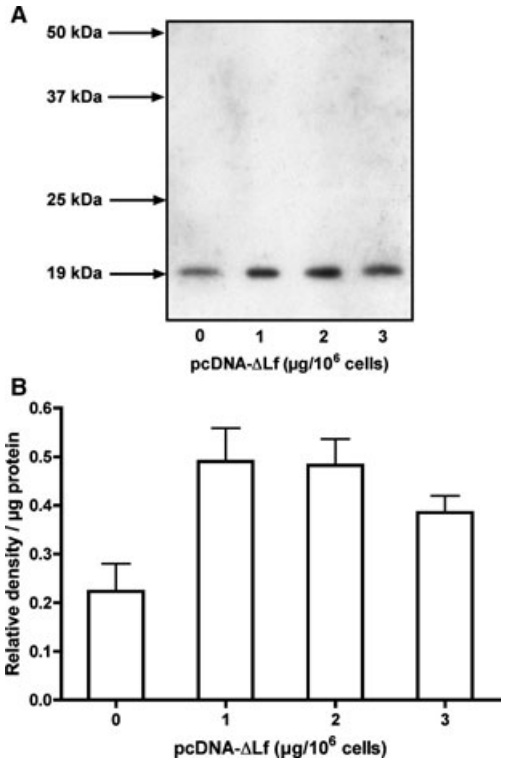


Fig. 4. Skp1 overexpression is visible at the protein level. HEK 293 cells were transfected by increasing concentrations of pcDNA-ΔLf. Twenty-four hours after transfection, total cell extracts were prepared from each transfected cell population. (A) Samples (15 μg of protein) were subjected to SDS/PAGE and immunoblotted with antibodies specific to Skp1. (B) The histogram represents the densitometric analysis of three independent experiments. The results are normalized to protein content, and are expressed in relative intensity per microgram of protein.

Presence of functional ΔLf response elements in the promoter *Skp1*

All the properties of ΔLf, such as nuclear targeting, antiproliferative effects, and Skp1 overexpression, argue in favor of ΔLf as a transcription factor. We therefore investigated the mechanism by which ΔLf potentiates *Skp1* transcription and whether it involves direct binding to DNA. Therefore, the human *Skp1* promoter was investigated. Screening of more than 3000 bases was done, and two sequences that were 90% identical to those already described were found. $S1^{Skp1}$ is the *Skp1* sequence homologous to the S1 sequence located at − 1067 bp, and $S2^{Skp1}$ is an *Skp1* sequence homologous to S2 at − 646 bp from the transcription initiation site (Fig. 5).

In order to determine whether these two sequences were ΔLf response elements (ΔLfREs), the *Skp1* promoter region was cloned using PCR. As *Skp1* is a single-copy gene, nested PCR was required. A 534 bp PCR product corresponding to the − 1164 bp to − 631 bp promoter region containing both the $S1^{Skp1}$ and $S2^{Skp1}$ sequences was cloned into the pGL3 promoter luciferase reporter vector. Next, a 132 bp product, which contains only the $S1^{Skp1}$ sequence (− 1164 bp to − 1033 bp), and a 138 bp product containing the $S2^{Skp1}$ sequence (− 768 bp to − 631 bp), were also cloned into pGL3 promoter luciferase reporter vectors. The constructs are shown in Fig. 6A.

Luciferase reporter assays were performed in HEK 293, MDAMB-231 and HeLa cells. As the results were comparable, only the data obtained with the HEK 293 cells are presented. The reporter luciferase vector was always used at the same concentration, and the ΔLf expression plasmid at increasing concentrations. Figure 6B shows that ΔLf was able to induce a marked increase in luciferase activity, whatever the reporter construct. The response of the reporter gene was dose-dependent up to 1 μg of pcDNA-ΔLf. Transactivation of $S1^{Skp1}$ in pGL3-$S1^{Skp1}$-Luc by ΔLf led to a 140-fold increase at the optimal concentration as compared to the basal expression level, and a 55-fold increase was observed for $S2^{Skp1}$ in pGL3-$S2^{Skp1}$-Luc. ΔLf therefore enhances transcription from the *Skp1* promoter, with both sequences responding to ΔLf, but $S1^{Skp1}$ responding at a higher level. The 534 promoter fragment is also transactivated by ΔLf, as the luciferase activity

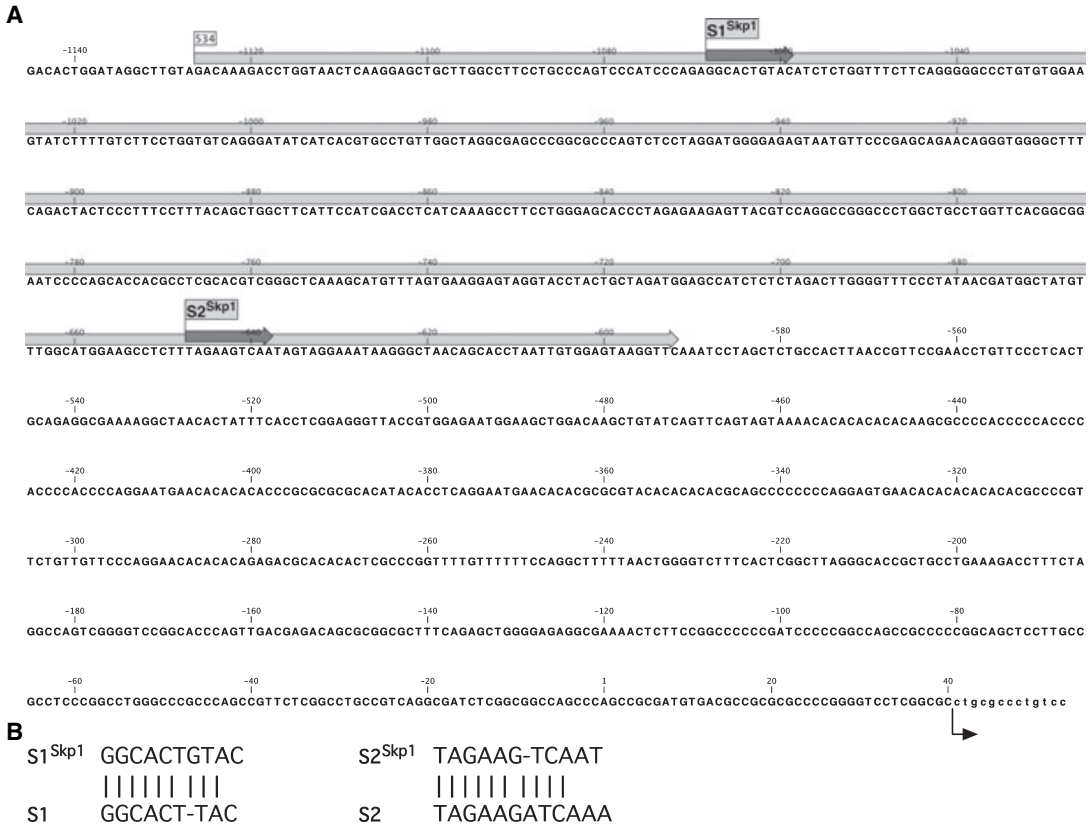


Fig. 5. ΔLfREs in the human *Skp1* promoter region. (A) The genomic sequence containing the human *Skp1* promoter was retrieved from the GenBank database (NC 00719). The 1.2 kbp range upstream of the mRNA start site was searched for possible ΔLfREs. The results showed that in the 5′-flanking region of the *Skp1* promoter, S1 and S2 ΔLf-like sequences are present and located at −1067 bp and −646 bp from the transcription start, respectively. (B) Comparison between these two sequences and those described by He & Furmanski [17].

corresponded to a 30-fold increase as compared to the basal expression level, but the presence of both ΔLfREs did not lead to a cumulative effect. This may be due to the presence of silencer elements in the intermediate region between the two response elements or to a limiting amount of ΔLf at each specific site.

In order to determine the contribution of each sequence to the overall activity of the native *Skp1* promoter, experiments with the 534 fragment construct, and constructs in which $S1^{Skp1}$ or $S2^{Skp1}$ had been deleted, were carried out. The sequence of the wild-type and deleted ΔLfREs within the reporter plasmids is shown in Fig. 7A. Six bases in the center of each of the sequences were deleted. Interestingly, deletion of the central core of either $S1^{Skp1}$ or $S2^{Skp1}$ strongly diminished ΔLf transcriptional activity (Fig. 7B). The percentage of inhibition measured at the optimal concentration of the expression plasmid was about 75% for $\Delta S1^{Skp1}$ and 85% for $\Delta S2^{Skp1}$ as compared to the wild-type promoter. These results therefore show that both sequences are ΔLfREs and are required for potentiating *Skp1* transcription.

We next investigated whether the homologous $S1^{Skp1}$ and $S2^{Skp1}$ sequences present in the *Skp1* promoter were also direct Lf targets. As we did not possess purified ΔLf, the gel shift assay was carried out using Lf. Shifted complexes were visible with *Skp1* probe sequences ($S'1^{Skp1}$ and S′2 Skp1) as well with S2 (Fig. 8A). Densitometric analysis of the interactions showed an equivalent interaction for $S1^{Skp1}$, $S2^{Skp1}$ and S2 as compared to a nonspecific probe (NS) (Fig. 8B). Binding to DNA occurs under stringent conditions (data not shown). The gel shift assay demonstrated that Lf interacts with these two sequences.

In order to demonstrate that ΔLf binds to the endogenous human *Skp1* promoter *in vivo*, we performed chromatin immunoprecipitation (ChIP) assays. Prior to the ChIP assay, ΔLf was N-terminus-tagged using the 3xFLAG epitope, in order to obtain the most reliable results, as shown in Fig. 9A. Comparison of the results of the immunoblots obtained either with

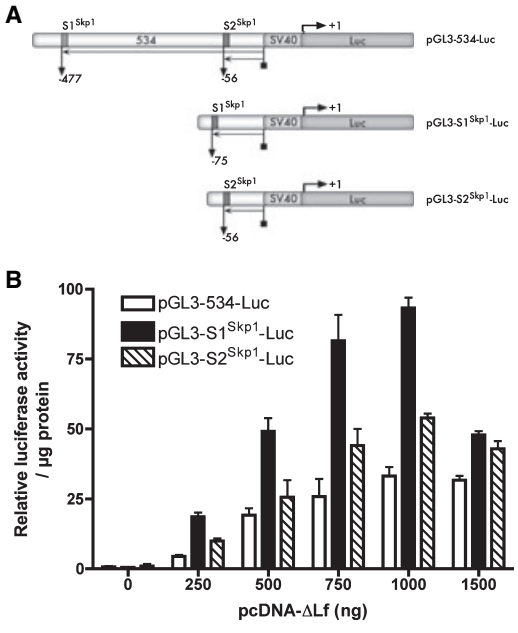


Fig. 6. ΔLf transactivates the *Skp1* promoter. (A) Diagrammatic presentation of the upstream promoter segments of the *Skp1* gene reporter constructs: pGL3-534-Luc, pGL3-S1^{Skp1}-Luc, and pGL3-S2^{Skp1}-Luc. (B) HEK 293 cells were cotransfected with these constructs (250 ng per well) and with a null plasmid or with pcDNA-ΔLf expression vector encoding ΔLf at increasing concentrations. Cells were lysed 24 h after transfection. Samples were assayed for protein content and luciferase activity. The relative luciferase activities reported were expressed as a ratio of the pGL3 reporter activity to protein content. Values represent the mean ± SD of triplicates from three independent measurements.

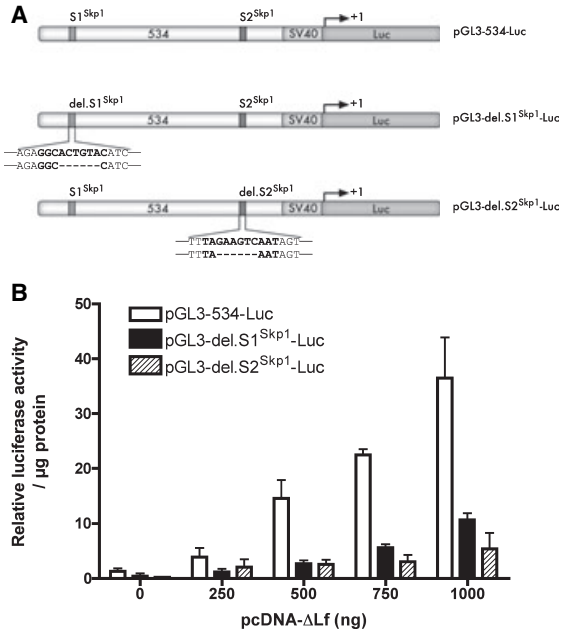


Fig. 7. Deletion mutation analyses of the human *Skp1* promoter. (A) Schematic diagram of the *Skp1* promoter showing the location of the S1^{Skp1} and S2^{Skp1} sequences as well as the deletion constructs. Mutated nucleotide sequences are emphasized by bold letters. A set of promoter constructs containing deleted S1^{Skp1} and S2^{Skp1} sequences was created by the protocol described in Experimental procedures. HEK 293 cells were transfected with wild-type 534 fragment or with the constructs of the del.S1^{Skp1} and del.S2^{Skp1} sequences at increasing concentrations. (B) Luciferase activities driven by the 534 bp fragment and mutated constructs. Twenty-four hours after the transfection, cells were lysed and luciferase activity was assayed. The relative luciferase activities reported were expressed as a ratio of the pGL3 reporter activity to protein content. The values represent the mean ± SE of three independent measurements.

antibodies to FLAG (M2) or antibodies to Lf (M90) showed that antibodies to FLAG could be used for the ChIP assay. Moreover, the tagged ΔLf was able to induce transcriptional activation of the luciferase reporter gene (Fig. 9B), indicating that FLAG-tagged ΔLf still bound to the *Skp1* promoter, validating the ChIP assay. The DNA purified from the sonicated chromatin was directly analyzed by PCR using *Skp1*-binding site-specific primers, which were used as an input control (lane 1). After immunoprecipitation by M2 antibodies, PCR amplification with the *Skp1*-specific primers revealed a product of the expected size (M2, lane 2, Fig. 9C). Control experiments involving nonspecific antibody (anti-rabbit IgG) showed only very slight amplification of the PCR product (IR, lane 4) and thus verified the results. The loading control, corresponding to the immunoprecipitation of chromatin with pure protein G Plus Sepharose (NS, lane 3), underlined the specificity of binding of ΔLf to the *Skp1* promoter. The PCR data shown in Fig. 9C corresponds to a significant experiment chosen among three independent assays. Densitometric analysis showed a four-fold higher level of amplification product for M2 *Skp1* promoter–ΔLf immunoprecipitate as compared to IR, and 10 times more compared to NS, after 36 cycles of amplification ($n = 3$) (Fig. 9D). Results correspond to the means of three separate experiments. The results show that antibodies to FLAG immunoprecipitate the ΔLf–*Skp1* promoter complex and demonstrate specific *in vivo* binding of ΔLf to *Skp1*. ΔLf is therefore a transcription factor.

These preliminary findings led us to examine the *Skp1* promoter sequences of other species. We compared the S1 and S2 DNA sequences of the response elements found in the human *Skp1* promoter with those of the chimpanzee, rat, and mouse, and compared them to those found in the interleukin-1β promoter [18] (Table 1). The comparisons showed that the chimpanzee *Skp1* promoter has one perfect copy of

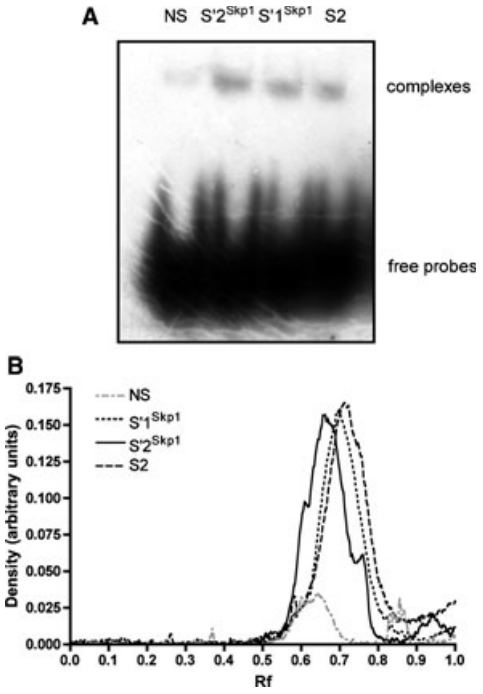


Fig. 8. Electrophoretic mobility shift assay of Lf with S′1^{Skp1} and S′2^{Skp1} elements. S′1^{Skp1} and S′2^{Skp1} correspond to 30-mer oligonucleotides containing one repeat of S1^{Skp1} or S2^{Skp1} placed in the center of the oligonucleotide and surrounded by their own native environment in the *Skp1* promoter. The NS oligonucleotide corresponds to a nonspecific DNA probe chosen within the *Skp1* promoter. As an internal control, the S2 sequence was chosen. As the DNA environment of S2 is unknown, it was placed in the same surrounding environment as S2^{Skp1}. All double-stranded oligonucleotides were labeled with ^{32}P and used as gel shift probes. Lf was used instead of ΔLf. The electrophoretic mobility shift assay was performed as described in Experimental procedures. (A) Retarded bands with S′1^{Skp1}, S′2^{Skp1} and S2 as probes were significantly induced in the presence of 25 ng of Lf (20 nM final) versus NS (nonspecific probe). (B) The densitometric profile of each retarded band shows specific interactions between Lf and S′1^{Skp1}, S′2^{Skp1}, and S2. All experiments were repeated three times, with comparable results.

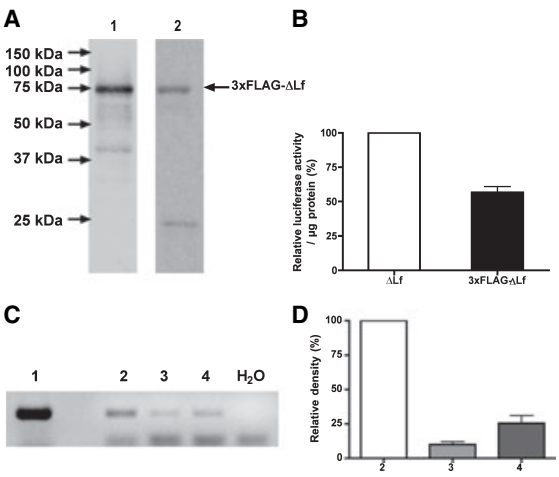


Fig. 9. ΔLf binds to the *Skp1* promoter *in vivo*. (A) HEK 293 cells were transiently transfected with p3xFLAG-CMV-10-ΔLf. Forty-eight hours after transfection, total cell extracts were prepared, and samples (15 μg of protein) were subjected to SDS/PAGE and immunoblotted with antibodies specific for the FLAG epitope (lane 1, anti-FLAG M2, 1 : 2000) or for Lf (lane 2, anti-hLf M90, 1 : 25 000). (B) The transcriptional activity of 3xFLAG-ΔLf as compared to ΔLf was examined using the luciferase reporter gene assay. HEK 293 cells were cotransfected with pcDNA-ΔLf or p3xFLAG-ΔLf constructs and pGL3-S1^{Skp1}-Luc plasmid. Cells were lysed 24 h after transfection. Values correspond to the mean ± SD of triplicates from two independent measurements. The data summarized in the histogram are expressed as a percentage, where 100% represents ΔLf transcriptional activity. (C) The binding of ΔLf to the *Skp1* promoter was examined in HEK 293 cells. ChIP was amplified by PCR using specific primers for the ΔLfRE of the *Skp1* promoter. Loading control (lane 1) corresponds to input (165 bp). ChIP assays were performed using anti-FLAG M2 (lane 2), and anti-rabbit IgG as nonspecific antibody control (lane 4). As a further control, the assay was performed without binding of an antibody to the protein G Plus Sepharose (lane 3). The results shown correspond to one experiment representative of the three performed. (D) Densitometric analysis of the ChIP assay (C, lanes 2–4). Results are expressed as a percentage, where 100% represents the signal obtained for the PCR product after immunoprecipitation with the anti-FLAG M2 (lane 2), and are the means of three separate experiments.

each ΔLfRE, whereas the mouse gene has two imperfect copies of each ΔLfRE-like sequence in a 3 kb region of the promoter. The rat gene has more divergent ΔLfRE-like sequences. Although the human promoter sequence has very limited identity overall with those of rodents, they all possess copies of ΔLfRE-like sequences in the 3 kb region of the promoter. The conservation of copies of ΔLfRE in *Skp1* promoters from these species might suggest an important role for ΔLf in regulating mammalian *Skp1* gene expression. Nevertheless, the location and sequence of the human ΔLfRE-like sequence are distinct from those of the cow and rodent species and more studies have to be done in order to confirm their function as ΔLfREs.

ΔLf possesses a functional bipartite NLS sequence

ΔLf, which lacks the GRRRR(1–5) pentapeptide present in Lf, which was identified as a functional nuclear import signal, was nevertheless observed in the nucleus. Among the other basic types of NLS, a short bipartite NLS sequence comprising two interdependent clusters of basic amino acids separated by a 10–12 amino acid spacer resembling the NLS of nucleoplasmin, Rb and interleukin-5 was found in ΔLf. This consensus sequence is conserved in Lfs from different species, such as the cow, mouse, pig, horse, and goat

Table 1. S1-like and S2-like sequences present in the *Skp1* promoter of different species compared to the S1-like sequences within the interleukin-1β promoter. ND, not determined.

Promoter	S1	Location[a]	S2	Location[a]	Accession number/Reference
He & Furmanski	GGCACTTA/GC		TAGA/GGATCAAA		[17]
Homo sapiens Skp1	GGCACTGTAC	− 1067 to − 1058	TAGAAGTCAATA	− 646 to − 637	AC007199
Mus musculus Skp1	GGCACTGAGC	− 2205 to − 2196	TAGAAGTCGGAT	− 2668 to − 2657	NT039267
	GGCACTGAGC		TGAAGTCACATA	− 496 to − 485	
Rattus norvegicus Skp1	GGCACTCTCAAC	− 104 to − 93	TGGAAGTCCC	− 213 to − 204	NM_001007608
Pan troglodytes Skp1	GGCACTGTAC	− 393 to − 384	TAGAAGTCAAT	+ 29 to + 37	NW_107077B
	GACACTGTAAC				
Homo sapiens IL-1β	GGCACTTGC	− 3202 to − 3193	ND		[18]
	GGAACTTGC	− 3137 to − 3129			
	GGAACTTGC	− 1052 to − 1043			
	GTCACGTGC	− 2384 to − 2376			
	GGCACTGTGC	− 1357 to − 1348			

[a] Location from the transcription start.

Table 2. Short bipartite NLSs in Lf from different species compared to those of nucleoplasmin, interleukin-5 (IL-5) and Rb.

Protein	Bipartite short-type NLSs[a]	Accession number/ reference
Xenopus nucleoplasmin	**KR**PAATKKAGQA**KKKK**	[48]
Human IL-5	**KK**YIDGQKKKCGEE**RRR**	[49]
Human Rb	**KR**SAEGSNPPKPL**KK**L**R**	[50]
Human Lf or ΔLf	**RR**SDTSLTWNSV**K**G**KK**	Q5EKS1
Bovine Lf	**KK**ANEGLTWNSL**K**D**KK**	P24627
Goat Lf	**KK**ANEGLTWNSL**K**G**KK**	Q29477
Mouse Lf	**RR**EDAGFTWSSL**R**G**KK**	P08071
Pig Lf	**RK**ANGGITWNSV**R**GT**K**	P14632
Horse Lf	**RK**SDADLTWNSLSG**KK**	077811

[a] The single-letter amino acid code is used; bold letters indicate the two arms of basic residues of the bipartite NLS.

(Table 2). In order to investigate whether this RRSDTSLTWNSVKGKK(417–432) NLS sequence may favor nuclear targeting, replacement of the arginine (417–418) and lysine (431–432) residues by alanine residues was performed, and the transcriptional activity of the $\Delta Lf^{del.RR}$, $\Delta Lf^{del.KK}$ and $\Delta Lf^{del.RRKK}$ mutants versus the wild-type (Fig. 10A) was assayed. Mutation of the KK residues leads to a 55% decrease in ΔLf transcriptional activation, whereas mutation of the RR residues leads to a larger decrease in ΔLf transcriptional activation of about 65%. The fact that $\Delta Lf^{del.KK432}$ retains a slightly higher nuclear import activity indicates that one part of the bipartite NLS (KK) may function individually as a weaker NLS. The double mutation RR-KK (75% inhibition) nearly completely abolishes the bipartite character of the NLS, abrogating its nuclear-targeting ability, as shown by a marked decrease in ΔLf transcriptional activation. The functionality of the short bipartite NLS was confirmed by comparing the subcellular distribution of the wild-type and mutated 3xFLAG-ΔLf fusion proteins. Immunohistochemistry was carried out using M2 murine antibody and goat anti-(mouse IgG) Alexa Fluor 488 in HEK 293 cells transiently transfected with expression plasmids encoding the FLAG epitope tag fused to the amino-ΔLf or the amino-$\Delta Lf^{del.RRKK}$ mutant. The wild-type and the $\Delta Lf^{del.RRKK}$ mutant fused to the FLAG epitope tag were similarly expressed (data not shown). The 3xFLAG-ΔLf fusion protein localized predominantly to the cytoplasm but was also present in the nucleus (Fig. 10B). In contrast, mutation of the NLS resulted in confinement of the mutated isoform to the cytoplasm (Fig. 10B). The double mutation RR-KK abolished the bipartite character of the NLS, as shown by the cytoplasmic retention of the mutated protein as compared to the wild-type.

Discussion

ΔLf is downregulated in cancer, and participates in the control of cell cycle progression, but the mechanism by which it exerts its antiproliferative properties is unknown. The data provided here show that ΔLf can locate to the nucleus and is involved in inducible gene expression. Transactivation by ΔLf targets the *Skp1* gene and, in particular, two specific DNA sequences located within the upstream promoter. Upregulation of *Skp1* is followed by a 2–3-fold increase at the protein level, and could explain in part the role of ΔLf in blocking cell cycle progression.

Skp1 is involved in a variety of crucial cellular functions. Modifications in its concentration may have

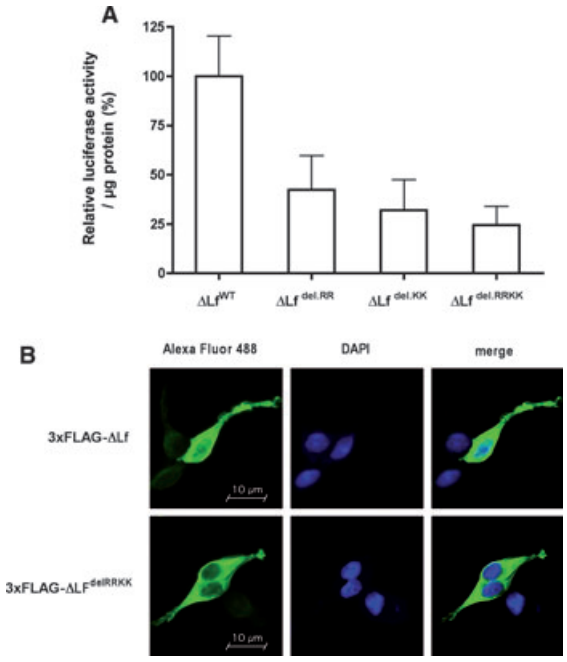


Fig. 10. Disruption of both basic amino acid sites in the bipartite NLS abolishes ΔLf transcriptional activity and nuclear traffic. (A) HEK 293 cells were transiently transfected with either the wild-type (ΔLf^{WT}) or the three mutated ΔLf-expressing plasmids, pcDNA-$\Delta Lf^{del.RR418}$, pcDNA-$\Delta Lf^{del.KK432}$, pcDNA-$\Delta Lf^{del.RRKK}$, corresponding, respectively, to the replacement by alanine residues of the sequences RR(417–418), KK(431–432) or both. The luciferase assay was performed 24 h after transfection. The relative luciferase activities reported were expressed as a ratio of the pGL3 reporter activity to protein content. The inhibition of the ΔLf transcriptional activity was expressed as a percentage of the relative luciferase activity of ΔLf-expressing mutants versus wild-type. The values represent the mean ± SD of three independent measurements. (B) Subcellular localization of 3xFLAG-ΔLf and 3xFLAG-$\Delta Lf^{delRRKK}$ isoforms using immunofluorescence microscopy. HEK 293 cells were transfected with ΔLf and $\Delta Lf^{delRRKK}$ tagged with 3xFLAG epitope, and examined after 24 h by fluorescent microscopy ($n = 3$). Nuclei were stained with DAPI. 3xFLAG-ΔLf and 3xFLAG-$\Delta Lf^{delRRKK}$ were stained using the M2 monoclonal antibody directed against the FLAG epitope and Alexa Fluor 488-conjugated goat anti-(mouse IgG). ΔLf is predominantly visible in the cytoplasm, but also enters the nucleus, as shown by the digital merge of the DAPI and Alexa Fluor 488 distributions. In contrast, $\Delta Lf^{delRRKK}$ was confined to the cytoplasm and excluded from the nucleus.

considerable consequences for cell cycle progression, leading, for example, to degradation of some cell cycle regulators before they could act. For instance, Skp2 is also a target of SCF [37], and its degradation would lead to cyclin accumulation and cell cycle arrest. On the other hand, Piva *et al*. [36], using Cul1 mutants able to sequestrate and inactivate Skp1, observed interference with the SCF degradation pathway and significant and specific increased expression of SCF substrates in cells expressing these mutants. They also observed the formation of multinucleated cells, centrosome and mitotic spindle abnormalities, and impaired chromosome segregation. They further generated Cul1 mutant transgenic mice in which Skp1 function was neutralized only in the T-cell lineage, leading to their death from T-cell lymphomas. Deregulation of the Cul1/Skp1 ratio affects the fidelity of chromosome transmission, and is directly responsible for neoplastic transformation. As Skp1 is required for the preservation of genetic stability and suppression of transformation, by upregulating its expression ΔLf might contribute to the control of cell division.

ΔLf is a transcription factor enhancing the *Skp1* promoter via two ΔLfREs: $S1^{Skp1}$ and $S2^{Skp1}$. Although $S1^{Skp1}$ was about three times more efficient than $S2^{Skp1}$, the different nucleotide environments of the two elements makes comparison difficult. However, our results are in agreement with those of He and Furmanski, in suggesting that the S1 sequence is the major transcriptional motif, whereas both $S1^{Skp1}$ and S^{2Skp1} (and S2) bind Lf equally efficiently. The role of $S2^{Skp1}$ as an independent *cis*-acting element was supported by mutational analysis of the promoter region containing both elements. In this case, deletion of the central core of either element led to a marked decrease in transactivation of the reporter gene, showing that in the native promoter, both motifs are required to mediate ΔLf transcriptional activity. Thus, the S1 sequence, when located near the initiation start point, efficiently led to *cis*-activation of transcription, whereas when located upstream in the promoter, it did not do so in the absence of $S2^{Skp1}$, as only 25% of the transcriptional activity remained. This suggests that multiple motifs or contact domains are required for ΔLf activity. Surprisingly, $S2^{Skp1}$ localized at the same place (– 56 bp upstream from the *Skp1* transcription initiation site) in the 137 bp and in 534 bp fragments when $S1^{Skp1}$ was deleted did not behave identically, as only 15% of the transcriptional activity was recovered in the latter case. This result might be explained by the presence of a silencer element that might not be strong enough to silence luciferase transcription when both ΔLfREs are present in the 534 bp fragment, whereas, when only one of them remains, silencing occurs. The intermediate region is currently under investigation.

The presence of two recognition sequences might contribute to transcriptional regulation. For example, the binding of ΔLf to the suboptimal S2 site prior to binding to the optimal S1 site, which may become accessible only under certain conditions determined by cell cycle signals, might serve as a pool for ΔLf. On the other hand, our results might suggest that the

distance between these two recognition elements and the initiation start point is crucial in order to promote the induction of transcription. For that, DNA bending might be necessary to lead to the juxtaposition of these two nonadjacent ΔLfREs, allowing ΔLf and regulatory proteins to interact together with the transcription apparatus. ΔLf function may require modification of the conformation of DNA at promoter sites by interaction with some cofactors such as cell cycle regulators.

The interaction of ΔLf with two response elements and the mutual dependency of both sites suggests that they are either bound by one ΔLf molecule via two DBDs, or that individual ΔLf molecules that are bound independently interact. A DBD has been located to the N-terminus [18], and the C-terminal end of the first helix might therefore represent a potent ΔLf DBD [11]. The interlobe region might also be a candidate [12], and using the ESCHER NG 1.0 docking software, we were able to observe that a DNA fragment could fit into the crevice between the two lobes (data not shown). More investigations need to be done in order to clarify these data.

The majority of sequence-specific DNA-binding proteins are multimers in solution, and multimerization is often necessary for high-affinity binding. Currently, nothing is known about the capability of Lf to undergo *in vivo* dimerization or multimerization. The data available concern only Lf and *in vitro* studies. Lf oligomerization usually occurs in solutions, depending on the ionic strength and/or the presence of calcium [38,39]. Lf and DNA complexes were also observed with a dependence on Lf concentration, with high concentrations favoring formation of large complexes [17]. Nevertheless, we do not know whether the *in vitro* oligomerization of Lf could have any physiologic relevance.

Our data show that the two basic amino acid clusters in the NLS contribute cooperatively to ΔLf nuclear import; disrupting one part of it reduced, but did not eliminate, ΔLf nuclear import, whereas disrupting both parts blocked ΔLf import, as shown by the loss of most its transcriptional activity and its cytoplasmic retention. This consensus sequence is conserved between Lf from different species. The remaining transcriptional activity observed with the double mutant may be due to an alternative NLS. Using the PSORT II server, the subprogram NUCDISC [40] has detected the KRKP(598–601) sequence as a putative NLS that could contribute to ΔLf nuclear import, but this sequence is not conserved in other species (data not shown), and might be irrelevant for Lf or ΔLf trafficking.

By causing the overexpression of *Skp1*, ΔLf may influence the proteasomal degradation of S-phase actors by controlling cell cycle progression or contribute to DNA preservation. Downregulation of transcription factors has been associated with pathologic states such as cancer. Therefore, the Lf gene was examined for structural alterations, and it was shown that the degree, as well as the pattern, of methylation were altered, notably in malignant breast cells [41–43]. Maintenance of a normal phenotype is the result of integrated effects of multiple tissue-specific transcriptional regulators, and ΔLf could be one of them. Nevertheless, two questions remain. What regulates ΔLf transcription, and is *Skp1* the only gene regulated by ΔLf during cell cycle progression? Lf and ΔLf promoters have been studied by Teng *et al.* [43], but nothing is known about the signaling pathways that drive ΔLf transcription. It will be interesting to study the kinetics of ΔLf synthesis in order to investigate whether it appears more specifically at the G_1/S transition. Collecting data on the regulation of ΔLf transcription may be important in developing strategies to enhance its expression in cancer cells. In order to answer the second question, *in silico* studies on other ΔLf target genes have been performed, and several genes involved in the control of cell cycle progression have been detected, the promoters of which are currently under investigation. Our preliminary studies on the *Rb* gene have shown that a sequence similar to S1 is present in its promoter region. The $S1^{Rb}$ sequence TGCACTTGTAT is located at −850 bp to −842 bp from the initiation start. Further investigations will be necessary to confirm its functionality.

Experimental procedures

Cell cultures

Human HEK 293 cells (ATCC CRL-1573) were kindly provided by J.-C. Dhalluin (INSERM U 524, Lille, France). HEK 293 stably transfected ΔLf (ΔLf-HEK 293) cells were obtained as previously described [7]. Human cervical cancer HeLa cells (ATCC CCL-2) were a kind gift from T. Lefebvre (UGSF, UMR 8576 CNRS, Villeneuve d'Ascq, France). Breast cancer MDA-MB-231 cell lines (ATCC HTB-26) were kindly provided by M. Mareel (Laboratory of Experimental Cancerology, University Hospital, Ghent, Belgium). All cell lines were routinely grown in monolayers as previously described [5,7,44]. Cell culture materials were obtained from Dutscher (Brumath, France), and culture media and additives from Cambrex Corporation (East Rutherford, NJ) and Invitrogen (Paisley, UK).

Plasmid construction

The translation-optimized ΔLf construct was generated by PCR using pBlueScript–ΔLf as template [7] and specific primer pairs (Table 3). The wild-type ΔLf and the $\Delta Lf^{del.RRKK}$ mutant with the 3xFLAG epitope tag fused in-frame at the N-terminus in the p3xFLAG expression vector (p3xFLAG-CMV-10; Sigma, St Louis, MO, USA) were generated by PCR using pcDNA-ΔLf or pcDNA-$\Delta Lf^{del.RRKK}$, respectively, as template and specific primer pairs (Table 3). The 534 bp *Skp1* promoter fragment (GenBank accession number: U33760) (from residues −630 to −1164 from the transcription initiation start) was amplified using nested PCR (primers are listed in Table 3) and purified genomic DNA as template. All PCR products were cloned in pCR BluntII-TOPO (Invitrogen), sequenced, and then transferred to the pcDNA3 vector (Clontech, Mountain View, CA), using *Kpn*I–*Xba*I for

Table 3. Oligonucleotides used for RT-PCR, ChIP, plasmid construction, site-directed mutagenesis and electrophoretic mobility shift assay (EMSA). S, sense; F, forward; NS, non-specific.

Method	Oligonucleotide* (5′– to 3′)	T_m (°C)	Cycle number	Amplicon size (bp)
RT-PCR				
ΔLf	S: TCCTCGTCCTGCTGTTCCTC	60	35	150
	F: GCTGTCTTTCGGTCCCGTAG			
Skp1	S: GTCTCCTTAACACCGA	55	30	522
	F: CACAACATTTCACTTCTC			
GAPDH	S: GTGGACCTGACCTGCCGTCTA	55	22	256
	F: CATGAGGTCCACCACCCTGTTGCTG			
RPLP0	S: GATGACCAGCCCAAAGGAGA	55	22	101
	F: GTGATGTGCAGCTGATCAAGACT			
TBP	S: CACGAACCACGGCACTGATT	60	25	89
	F: TTTTCTTGCTGCCAGTCTGGAC			
ChIP				
Skp1 promoter	S: GCTCAAAGCATGTTTAGTG	60	36	165
	F: GAACCTTACTCCACAATTAG			
Plasmid construction				
ΔLf	S: GGTACCGCCACCATGAGAAAAGTGCGTGGCCC			
	F: TCTAGATCTTCGGTTTTACTTCCTGAGGAATTC			
3xFLAG-CMV-10-ΔLf	S: AAGCTTATGAGAAAAGTGCGTGGCCC			
	F: TCTAGATCTTCGGTTTTACTTCCTGAG			
534 bp–*Skp1*	External S: GAGACTGGATAGGCTTGTAG			
	External F: GCGCCGAGGACCCCG			
	Internal S: ACAAAGACCTGGTAACTCA			
	Internal F: GAACCTTACTCCACAATTAG			
Site-directed mutagenesis				
$\Delta S1^{Skp1}$	S: CCCTGAAGAAACCAGAGATGGCCTCTGGGATGGGACTGGG			
	F: CCCAGTCCCATCCCAGAGGCCATCTCTGGTTTCTTCAGGG			
$\Delta S2^{Skp1}$	S: GTGCTGTTAGCCCTTATTTCCTACTATTAAAGAGGCTTCCATGCCAAACATAGCC			
	F: GGCTATGTTTGGCATGGAAGCCTCTTTAAATAGTAGGAAATAAGGGCTAACAGCAC			
$\Delta Lf^{del.RR}$	S: CTAGTGTCTGATGCTGCAACCACCGCCAC			
	F: GTGGCGGTGGTTGCAGCATCAGACACTAG			
$\Delta Lf^{del.KK}$	S: GTGTGGCAGGACGCTGCGCCTTTCACAG			
	F: CTGTGAAAGGCGCAGCGTCCTGCCACAC			
EMSA				
$S'1^{Skp1}$	S: TCCCAGAGGCACTGTACATCTCTG			
	F: CAGAGATGTACAGTGCCTCTGGGA			
$S'2^{Skp1}$	S: GCCTCTTTAGAAGTCAATAGTAGG			
	F: CCTACTATTGACTTCTAAAGAGGC			
S2	S: GCCTCTTTAGAAGATCAAAAGTAGG			
	F: CTACTTTTGATCTTCTAAAGAGGC			
NS*	S: TGGAGCCATCTCTCAGACTTGGG			
	F: CCCAAGTCTAGAGAGATGGCTCCA			

ΔLf, and to the pGL3-promoter-Luc vector (Promega, Madison, WI), using *Sac*I–*Xho*I for the 534 bp *Skp1* promoter fragment. Both PCR amplifications were carried out using the Proofstart DNA Polymerase (Qiagen, Hilden, Germany) for 40 cycles at 95 °C for 30 s, 55 °C for 90 s, and 72 °C for 120 s, with a final amplification step of 10 min at 72 °C. The pGL3-$S2^{Skp1}$-promoter-Luc vector was generated after removal of the *Kpn*I–*Kpn*I restriction fragment and religation of the pGL3-534-promoter-Luc vector used as template. The $S1^{Skp1}$ insert was obtained by removing an *Eco*RV–*Eco*RV fragment from the pCR-BluntII-TOPO-534 vector as template and religation. Then, the *Kpn*I–*Xho*I digest was isolated and cloned into the pGL3-promoter-Luc reporter vector, leading to the production of the pGL3-$S1^{Skp1}$-promoter-Luc vector. Ligations were performed using T4 DNA ligase (Invitrogen).

DNA and RNA isolation

Genomic DNA was extracted from HEK 293 cells as previously described [45], and purified using the Wizard Genomic DNA Purification kit (Promega), with yield being assessed by spectrophotometry. All plasmids were purified using the QIAprep Spin Miniprep Kit (Qiagen). Total RNA was extracted from cell cultures using the RNeasy Mini Kit (Qiagen) according to the manufacturer's specifications. The purity of the nucleic acid extracts were checked by measuring the ratio of the absorbance at 260 nm and 280 nm using a NanoDrop ND-1000-Spectrophotometer (Labtech International, Ringmer, UK), and their integrity was visualized on a BET-agarose gel.

GEArray

The human Cellcycle-2 GEArray kit was obtained from SuperArray Bioscience Corp. (Frederick, MD). It included reagents for probe generation and hybridization, and two identical gene arrays containing 23 marker genes for each array. The marker genes (*cdk2*, *cdk4*, *cdk6*, *cyclin C*, *cyclin D2*, *cyclin D3*, *cyclin E1*, *DP1*, *DP2*, *EF*, *E2F-4*, *E2F5*, *p107*, *p130* (*RB2*), $p19^{Ink4d}$, $p21^{Waf1}$, $p27^{Kip1}$, $p55^{cdc}$, $p57^{Kip2}$, *PCNA*, *Rb*, *Skp1*, and *Skp2*) were used to monitor the activation of genes involved in G_1/S progression. Negative and positive control genes were added (*pUC18*, *β-actin* and *GAPDH*). Prior to hybridization, expression of ΔLf in doxycyclin-induced HEK 293 cells was verified by RT-PCR. Hybridization was carried out with 16-dUTP-biotinylated RNA (10 μg) from induced and noninduced ΔLf-HEK 293 cells, according to the manufacturer's specifications. Experiments were performed in triplicate. Densitometric analyses were performed, and the average signal was obtained from duplicates of each gene. The normalized value for each gene was calculated from the ratio of averaged value of each gene divided by the average value of β-actin.

RT-PCR conditions

Primer pairs designed for the specific detection of target sequences such as *ΔLf*, *Skp1*, *TBP* (TATA box-binding protein), *GAPDH* and ribosomal protein large, P0 (*RPLP0*) are listed in Table 3. They were selected through computer analysis using PRIMER PREMIER Version 3.1 software (Biosoft International, Palo Alto, CA). Primer pairs are located on distinct exons to avoid amplification of contaminating genomic DNA. Primer pairs for *ΔLf*, *Skp1*, *GAPDH* and *RPLP0* were purchased from Eurogentec (Seraing, Belgium), and those for *TBP* from Genset SA (Paris, France).

Five micrograms of each RNA preparation were reverse transcribed into first-strand cDNA using oligo-dT primers and 200 units of Moloney murine leukemia virus (MMLV) reverse transcriptase (Promega). In order to minimize variations that could occur during retrotranscription, two first-strand cDNA batches were prepared as described above and mixed. Reverse transcriptase, oligo-dT primers and dNTPs were from PCR Nucleotide Mix (Promega). Silverstar polymerase (Eurogentec) was used. The first-strand cDNA preparation (2 μL) was then amplified by PCR as previously described [5,7]. Prior to the RT-PCR analysis, we first determined whether the PCR reactions detecting *RPLP0*, *TBP*, *Skp1* and *ΔLf* were optimal, that PCR products could all be visualized, and that the reactions remained within the exponential phase of amplification. Thus, the relative intensity of the various PCR signals reflects the initial abundance of the corresponding transcripts. RT-PCR assays were performed in triplicate. In all experiments, negative control reactions were done in which cDNA templates were replaced with sterile water to check for the absence of contaminants. The contamination of genomic DNA was excluded by performing 35 cycles of amplification without retrotranscription. RT-PCR conditions specific to each primer pair are summarized in Table 3. Amplification products were subcloned in either pGEM Easy-T (Promega) or pCR BluntII-TOPO (Invitrogen), and sequenced to confirm the specificity of the PCR.

Twenty-three microliters of each PCR reaction was loaded onto a 1.5% agarose gel stained with 0.5 $\mu g \cdot mL^{-1}$ ethidium bromide. Quantification was performed by UV transillumination using a Gel Doc1000 system (Bio-Rad, Hercules, CA) and densitometric analysis of the image using QUANTITY ONE v4.1 software (Bio-Rad). For each DNA sample, the level of *Skp1* or *ΔLf* expression was expressed as a ratio between mRNA expression and *RPLP0* or *TBP* expression, and was referred to as normalized expression.

Site-directed mutagenesis

All the mutants were generated using the QuikChange Site-directed Mutagenesis Kit (Stratagene, Garden Grove, CA), according to the manufacturer's instructions. The

pGL3-534-promoter-Luc construct was used as template for the single mutants $\Delta S1^{Skp1}$ and $\Delta S2^{Skp1}$. The oligonucleotides used are listed in Table 3. The right [RR(417–418] and the left [KK(431–432] parts of the bipartite NLS sequence were mutated, these four amino acid residues being replaced by alanine residues. The pcDNA-ΔLf construct was used as template for generating pcDNA3-$\Delta Lf^{del.RR}$ and the pcDNA-$\Delta Lf^{del.KK}$. pcDNA3-$\Delta Lf^{del.RR}$ was used to generate the double mutant pcDNA-$\Delta Lf^{del.RRKK}$. The two oligonucleotide pairs used are listed in Table 3. Following sequence verification, positive clones were used directly in transfection.

Transfection

Transfection studies were done using at least three independent plasmid preparations, and each transfection was repeated at least three times. All cell lines were cotransfected in triplicate with increasing concentrations of pcDNA-ΔLf. Transfections were performed using the transfection reagent Clonfectin (BD Biosciences, Franklin Lakes, NJ), according to the manufacturer's instructions. After incubation for 24 h, cells were washed with NaCl/P_i. Cells were then lysed in appropriate buffer, either for total RNA preparation or for protein extracts. Protease inhibitor (Pefablock; Roche, Basel) was added to protein extracts.

Reporter gene assays

HEK 293 cells were plated 1 day before transfection in 12-well plates at a density of 2×10^5 cells per well. The ΔLf transcriptional activity was assessed using pGL3-promoter-Luc reporter plasmids (pGL3-534; pGL3-$S1^{Skp1}$; pGL3-$S2^{Skp1}$; pGL3-534-$\Delta S1^{Skp1}$, and pGL3-534-$\Delta S2^{Skp1}$) (250 ng per well) and pcDNA-ΔLf in the range 250–1000 ng per well. The functionality of the ΔLf consensus NLS bipartite sequence was assessed using increasing concentrations of pcDNA-$\Delta Lf^{del.RR}$, pcDNA-$\Delta Lf^{del.KK}$ or pcDNA-$\Delta Lf^{del.RRKK}$ and pGL3-$S1^{Skp1}$-promoter-Luc reporter plasmid (250 ng per well). Each experiment represents at least three sets of independent triplicates. Twenty-four hours after the transfections, cells were lysed and assayed using a luciferase assay kit (Promega) in a Wallac Victor^2 1420 multilabel counter (Perkin Elmer, Boston, MA). For all experiments, protein content was used to normalize luciferase results. Protein concentrations of cell lysates were determined by a BCA assay, using BSA as standard. Absorbance measurements were carried out at 590 nm using a microplate reader (Model 550, Bio-Rad).

Electrophoretic mobility shift assays

Single-stranded oligonucleotides were end-labeled with [^{32}P]ATP[γP] (500 μCi·μL^{-1}; GE Healthcare Life Sciences, Little Chalfont, UK) and T4 polynucleotide kinase (Invitrogen) for 2 h at room temperature. The oligonucleotides ($S'1^{Skp1}$, $S'2^{Skp1}$, S2 and NS) used were are listed in Table 3. The sense and antisense strands were annealed at room temperature for 20 min and used as a probe. Labeled double-stranded probes were purified on a 20% acrylamide gel in Tris-borate-EDTA buffer. After autoradiography, probes were excised and eluted from the gel with 500 μL of sterile ultrapure water. Binding reactions were performed with 25–50 ng of Lf, 40 000–80 000 c.p.m. of radiolabeled probes, and 6.25 ng of poly(dI-dC) (GE Healthcare Life Sciences). Binding was performed at room temperature for 20 min. Loading buffer containing bromophenol blue and 50% glycerol was added to the DNA·protein complexes prior to separation on a 5% nondenaturing polyacrylamide gel in 0.25 × Tris/borate/EDTA, which was previously submitted to a 20 min prerun at 20 mA. Migration was carried out at 20 mA for 20 min in the same buffer. The gel was then dried and exposed (Hyperfilm+ GE Healthcare Life Sciences) for 24–48 h at − 80 °C.

ChIP assays

ChIP assays were conducted using the EZ ChIP Enzymatic kit (Upstate Biotech, Millipore, Billerica, MA) according to the manufacturer's instructions, with some modifications. At 24 h post-transfection, ΔLf-transfected HEK 293 cells were crosslinked with 1% formaldehyde for 10 min at room temperature. After the reaction had been stopped by the addition of 125 mM glycine for 5 min at room temperature, cells were washed in NaCl/P_i. Cells (10^7) were then incubated in 200 μL of lysis buffer [50 mM Tris, pH 8.1, 1% SDS, 10 mM EDTA, 1 mM Pefabloc (Roche)] for 10 min at 4 °C. After sonication and centrifugation (Heraeus, Biofuge $15R_1$, HFA 22.2 rotor, 12 000 ***g***, 15 min), lysates were diluted in the ChIP dilution buffer (1 : 10), precleared with 2 μL of mouse normal serum for 6 h at 4 °C under rotation, and precipitated with protein G Sepharose beads (GE Healthcare Life Sciences). The supernatant was further incubated with antibodies overnight at 4 °C or not incubated. An aliquot of untreated supernatant served as input control. An aliquot of supernatant was either incubated with M2 antibody (1 : 500, Sigma), or anti-(rabbit IgG) (1 : 1000, GE Healthcare Life Sciences) used as a non specific antibody control. An aliquot of supernatant not incubated with antibody was immunoprecipitated and used as a negative control. Complexes were precipitated for 2 h at 4 °C using protein G Sepharose beads (GE Healthcare Life Sciences). The captured immunocomplexes, containing bound transcriptional DNA fragments, were eluted overnight at 65 °C, and treated with 4 μL of ribonuclease A (20 mg·mL^{-1}; Sigma) and 2 μL of proteinase K (10 mg·mL^{-1}; Sigma). The DNA fragments were purified using a Quiagen DNA purification kit (Qiagen). Two

microliters of each supernatant was then used for PCR (36 cycles). Primer pairs specifically amplifying the *Skp1* promoter region are described in Table 3. PCR products were separated on a 2% agarose gel, and stained with ethidium bromide.

Western blotting and immunodetection

Cell extracts were prepared from frozen pellets of HEK cells transfected with pcDNA-ΔLf, p3xFLAG-ΔLf or pcDNA3 empty vector. Proteins were extracted in radioimmunoprecipitation assay (RIPA) buffer [46] for 20 min on ice. Cell debris were removed by centrifugation for 10 min at 12 000 ***g*** (Heraeus Biofuge $15R_1$ HFA 22.2 rotor), and the soluble material was submitted to SDS/PAGE and analyzed by western blotting. Blots were subsequently probed with primary antibodies (polyclonal goat anti-Skp1, 1 : 1000; murine anti-FLAG M2, 1 : 2000; rabbit anti-Lf, 1 : 25 000) for 1 h at 4 °C and secondary antibodies conjugated to peroxidase, before being detected by chemiluminescence (ECL+ GE Healthcare Life Sciences Biosciences). Primary antibodies against Skp1 were purchased from Santa Cruz Biotechnologies Inc. (Tébu-Bio, France), anti-FLAG M2 was purchased from Sigma-Aldrich (St Louis, MO), and secondary antibodies conjugated to horseradish peroxidase were purchased from GE Healthcare Life Sciences. Rabbit anti-human lactoferrin serum were prepared from purified hLf as in Fillebeen *et al.* [47].

Immunofluorescence microscopy

Twenty-four hours prior to transfection, cells were cultured onto glass slides pretreated with 50 $\mu g \cdot mL^{-1}$ collagen. Cells were processed for immunofluorescence 24 h after transfection, by washing in NaCl/P_i and fixing in 4% paraformaldehyde (pH 7.4) for 30 min. Cells were then rinsed two times in NaCl/P_i, permeabilized for 2 min in 0.15% Triton X-100 in NaCl/P_i, and washed again twice in NaCl/P_i. Glass slides were next placed in a solution containing 1% ethanolamine in NaCl/P_i for 20 min at 4 °C. After two washings in NaCl/P_i, cells were incubated overnight at 4 °C with the M2 primary antibodies (Sigma) diluted in blocking solution (1% BSA in NaCl/P_i) at 1 : 1000. Cells were then rinsed three times in NaCl/P_i, and incubated for 1 h at 37 °C in Alexa Fluor 488-conjugated goat (anti-mouse serum) (Invitrogen) diluted 1 : 2000 in blocking solution. Cells were then incubated for 30 min in 4^1, 6-diamidino-2-phenylindole (DAPI) (1 : 5000, Sigma) to label chromatin. Finally, glass slides were rinsed three times in NaCl/P_i and mounted with coverslips and Mowiol medium. Fluorescent microscopy images were obtained with a Zeiss Axioplan 2 imaging system (Carl-Zeiss S.A.S., Le Pecq, France) equipped with appropriate filter cubes using a 40× objective lens.

Densitometric analysis

The densitometric analysis of the Hyperfilm (GE Healthcare Life Sciences) obtained either by autoradiography or chemiluminescence was performed using the QUANTITY ONE v4.1 (Bio-Rad) acquisition software. Acquisition was carried out with a GelDoc camera (Bio-Rad) for the PCR products, or with a GS710-calibrated densitometer (Bio-Rad) for the films.

Acknowledgements

This investigation was supported in part by the CNRS Unité Mixte de Recherche 8576 (Unité de Glycobiologie Structurale et Fonctionnelle), the Institut Fédératif de Recherche no. 148, the Université des Sciences et Technologies de Lille 1, the Ministère de l'Education Nationale, the region Nord-Pas-de Calais (ARCir Signalization Cellulaire), the Ligue Nationale contre le Cancer and the Association de Recherche contre le Cancer (grant no. 5469). We are grateful to Professor M. M. Mareel (Laboratory of Experimental Cancerology, University Hospital, Ghent, Belgium) for providing us with the human breast cancer cell line MDA-MB-231, Dr M. Crépin (Institut d'Oncologie Cellulaire et Moléculaire Humaine, Bobigny, France) for the HBL 100 cells, Dr J.-C. Dhalluin (INSERM U 524, Lille, France) for the HEK 293 cells, and Dr T. Lefebvre (UGSF, UMR 8576-CNRS, Villeneuve d'Ascq, France) for the HeLa cells. We would like to thank INSERM U 547 (Director Professor M. Capron) for providing us with access to the Wallac Victor^2 1420 luminometer, and Dr R. J. Pierce (INSERM U 547, Institut Pasteur de Lille, France) for helping us with genomic analysis and reviewing this manuscript.

References

1 Hershko A & Ciechanover A (1998) The ubiquitin system. *Annu Rev Biochem* **67**, 425–479.
2 Willems AR, Schwab M & Tyers M (2004) A hitchhiker's guide to the cullin ubiquitin ligases: SCF and its kin. *Biochim Biophys Acta* **1695**, 133–170.
3 Ang XL & Wade Harper J (2005) SCF-mediated protein degradation and cell cycle control. *Oncogene* **24**, 2860–2870.
4 Siebert PD & Huang BC (1997) Identification of an alternative form of human lactoferrin mRNA that is expressed differentially in normal tissues and tumor-derived cell lines. *Proc Natl Acad Sci USA* **94**, 2198–2203.
5 Benaissa M, Peyrat JP, Hornez L, Mariller C, Mazurier J & Pierce A (2005) Expression and prognostic value of

lactoferrin mRNA isoforms in human breast cancer. *Int J Cancer* **114**, 299–306.
6 Liu D, Wang X, Zhang Z & Teng CT (2003) An intronic alternative promoter of the human lactoferrin gene is activated by Ets. *Biochem Biophys Res Commun* **301**, 472–479.
7 Breton M, Mariller C, Benaissa M, Caillaux K, Browaeys E, Masson M, Vilain JP, Mazurier J & Pierce A (2004) Expression of delta-lactoferrin induces cell cycle arrest. *Biometals* **17**, 325–329.
8 Penco S, Scarfi S, Giovine M, Damonte G, Millo E, Villaggio B, Passalacqua M, Pozzolini M, Garre C & Benatti U (2001) Identification of an import signal for, and the nuclear localization of, human lactoferrin. *Biotechnol Appl Biochem* **34**, 151–159.
9 van Berkel PH, Geerts ME, van Veen HA, Mericskay M, de Boer HA & Nuijens JH (1997) N-terminal stretch Arg2, Arg3, Arg4 and Arg5 of human lactoferrin is essential for binding to heparin, bacterial lipopolysaccharide, human lysozyme and DNA. *Biochem J* **328**(1), 145–151.
10 Garre C, Bianchi-Scarra G, Sirito M, Musso M & Ravazzolo R (1992) Lactoferrin binding sites and nuclear localization in K562(S) cells. *J Cell Physiol* **153**, 477–482.
11 Baker EN, Baker HM & Kidd RD (2002) Lactoferrin and transferrin: functional variations on a common structural framework. *Biochem Cell Biol* **80**, 27–34.
12 Baker EN & Baker HM (2005) Molecular structure, binding properties and dynamics of lactoferrin. *Cell Mol Life Sci* **62**, 2531–2539.
13 Hutchens TW, Magnuson JS & Yip TT (1989) Interaction of human lactoferrin with DNA: one-step purification by affinity chromatography on single-stranded DNA-agarose. *Pediatr Res* **26**, 618–622.
14 Hutchens TW, Henry JF, Yip TT, Hachey DL, Schanler RJ, Motil KJ & Garza C (1991) Origin of intact lactoferrin and its DNA-binding fragments found in the urine of human milk-fed preterm infants. Evaluation by stable isotopic enrichment. *Pediatr Res* **29**, 243–250.
15 Britigan BE, Lewis TS, Waldschmidt M, McCormick ML & Krieg AM (2001) Lactoferrin binds CpG-containing oligonucleotides and inhibits their immunostimulatory effects on human B cells. *J Immunol* **167**, 2921–2928.
16 Kanyshkova TG, Semenov DV, Buneva VN & Nevinsky GA (1999) Human milk lactoferrin binds two DNA molecules with different affinities. *FEBS Lett* **451**, 235–237.
17 He J & Furmanski P (1995) Sequence specificity and transcriptional activation in the binding of lactoferrin to DNA. *Nature* **373**, 721–724.
18 Son KN, Park J, Chung CK & Chung DK, Yu DY, Lee KK & Kim J (2002) Human lactoferrin activates transcription of IL-1beta gene in mammalian cells. *Biochem Biophys Res Commun* **290**, 236–241.
19 Legrand D, Elass E, Pierce A & Mazurier J (2004) Lactoferrin and host defence: an overview of its immunomodulating and anti-inflammatory properties. *Biometals* **17**, 225–229.
20 Legrand D, Elass E, Carpentier M & Mazurier J (2005) Lactoferrin: a modulator of immune and inflammatory responses. *Cell Mol Life Sci* **62**, 2549–2559.
21 Giancotti V (2006) Breast cancer markers. *Cancer Lett* **243**, 145–159.
22 Damiens E, El Yazidi I, Mazurier J, Duthille I, Spik G & Boilly-Marer Y (1999) Lactoferrin inhibits G1 cyclin-dependent kinases during growth arrest of human breast carcinoma cells. *J Cell Biochem* **74**, 486–498.
23 Xiao Y, Monitto CL, Minhas KM & Sidransky D (2004) Lactoferrin down-regulates G1 cyclin-dependent kinases during growth arrest of head and neck cancer cells. *Clin Cancer Res* **10**, 8683–8686.
24 Oh SM, Hahm DH, Kim IH & Choi SY (2001) Human neutrophil lactoferrin trans-activates the matrix metalloproteinase 1 gene through stress-activated MAPK signaling modules. *J Biol Chem* **276**, 42575–42579.
25 Fujita K, Matsuda E, Sekine K, Iigo M & Tsuda H (2004) Lactoferrin modifies apoptosis-related gene expression in the colon of the azoxymethane-treated rat. *Cancer Lett* **213**, 21–29.
26 Fujita K, Matsuda E, Sekine K, Iigo M & Tsuda H (2004) Lactoferrin enhances Fas expression and apoptosis in the colon mucosa of azoxymethane-treated rats. *Carcinogenesis* **25**, 1961–1966.
27 Ward PP, Paz E & Conneely OM (2005) Multifunctional roles of lactoferrin: a critical overview. *Cell Mol Life Sci* **62**, 2540–2548.
28 Bai C, Sen P, Hofmann K, Ma L, Goebl M, Harper JW & Elledge SJ (1996) SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. *Cell* **86**, 263–274.
29 Zheng N, Schulman BA, Song L, Miller JJ, Jeffrey PD, Wang P, Chu C, Koepp DM, Elledge S, Pagana M *et al.* (2002) Structure of the Cul1–Rbx1–Skp1–F boxSkp2 SCF ubiquitin ligase complex. *Nature* **416**, 703–709.
30 Zhang H, Kobayashi R, Galaktionov K & Beach D (1995) p19Skp1 and p45Skp2 are essential elements of the cyclin A-CDK2 S phase kinase. *Cell* **82**, 915–925.
31 Carrano AC, Eytan E, Hershko A & Pagano M (1999) SKP2 is required for ubiquitin-mediated degradation of the CDK inhibitor p27. *Nat Cell Biol* **1**, 193–199.
32 Marti A, Wirbelauer C, Scheffner M & Krek W (1999) Interaction between ubiquitin-protein ligase SCFSKP2 and E2F-1 underlies the regulation of E2F-1 degradation. *Nat Cell Biol* **1**, 14–19.

33 Tsvetkov LM, Yeh KH, Lee SJ, Sun H & Zhang H (1999) p27 (Kip1) ubiquitination and degradation is regulated by the SCF (Skp2) complex through phosphorylated Thr187 in p27. *Curr Biol* **9**, 661–664.

34 Kitagawa K & Hieter P (2001) Evolutionary conservation between budding yeast and human kinetochores. *Nat Rev Mol Cell Biol* **2**, 678–687.

35 Lehmann A & Toda T (2004) Fission yeast Skp1 is required for spindle morphology and nuclear membrane segregation at anaphase. *FEBS Lett* **566**, 77–82.

36 Piva R, Liu J, Chiarle R, Podda A, Pagano M & Inghirami G (2002) In vivo interference with Skp1 function leads to genetic instability and neoplastic transformation. *Mol Cell Biol* **22**, 8375–8387.

37 Wirbelauer C, Sutterluty H, Blondel M, Gstaiger M, Peter M, Reymond F & Krek W (2000) The F-box protein Skp2 is a ubiquitylation target of a Cul1-based core ubiquitin ligase complex: evidence for a role of Cul1 in the suppression of Skp2 expression in quiescent fibroblasts. *EMBO J* **19**, 5362–5375.

38 Mantel C, Miyazawa K & Broxmeyer HE (1994) Physical characteristics and polymerization during iron saturation of lactoferrin, a myelopoietic regulatory molecule with suppressor activity. *Adv Exp Med Biol* **357**, 121–132.

39 Bennett RM, Bagby GC & Davis J (1981) Calcium-dependent polymerization of lactoferrin. *Biochem Biophys Res Commun* **101**, 88–95.

40 Nakai K & Horton P (1999) PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem Sci* **24**, 34–36.

41 Panella TJ, Liu YH, Huang AT & Teng CT (1991) Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes, leukemic cells, and breast cancer. *Cancer Res* **51**, 3037–3043.

42 Liu LH, Gladwell W & Teng CT (2002) Detection of exon polymorphisms in the human lactoferrin gene. *Biochem Cell Biol* **80**, 17–22.

43 Teng C, Gladwell W, Raphiou I & Liu E (2004) Methylation and expression of the lactoferrin gene in human tissues and cancer cells. *Biometals* **17**, 317–323.

44 Lefebvre T, Alonso C, Mahboub S, Dupire MJ, Zanetta JP, Caillet-Boudin ML & Michalski JC (1999) Effect of okadaic acid on O-linked N-acetylglucosamine levels in a neuroblastoma cell line. *Biochim Biophys Acta* **1472**, 71–81.

45 Sambrook J, Fritsch EF & Maniatis T (1989) *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

46 Harlow E & Lane D (1988) *Antibodies: a Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

47 Fillebeen C, Mitchell V, Dexter D, Benaissa M, Beauvillain J, Spik G & Pierce A (1999) Lactoferrin is synthesized by mouse brain tissue and its expression is enhanced after MPTP treatment. *Brain Res Mol Brain Res* **72**, 183–194.

48 Robbins J, Dilworth SM, Laskey RA & Dingwall C (1991) Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell* **64**, 615–623.

49 Jans DA, Briggs LJ, Gustin SE, Jans P, Ford S & Young IG (1997) A functional bipartite nuclear localisation signal in the cytokine interleukin-5. *FEBS Lett* **406**, 315–320.

50 Efthymiadis A, Shao H, Hubner S & Jans DA (1997) Kinetic characterization of the human retinoblastoma protein bipartite nuclear localization sequence (NLS) in vivo and in vitro. A comparison with the SV40 large T-antigen NLS. *J Biol Chem* **272**, 22134–22139.

II. *Bax* et l'apoptose

Les facteurs de transcription ont souvent des cibles multiples et variées. La delta-lactoferrine ne fait pas exception à la règle puisque son implication dans le processus néoplasique ne se limite pas au contrôle du cycle cellulaire. En effet, si sa surexpression induit un arrêt du cycle cellulaire après 24 h, elle conduit à une augmentation significative de la mort cellulaire au bout de 48 h. A l'instar d'autres facteurs de transcription régulant le cycle cellulaire tels que p53, la delta-lactoferrine pourrait également induire l'apoptose.

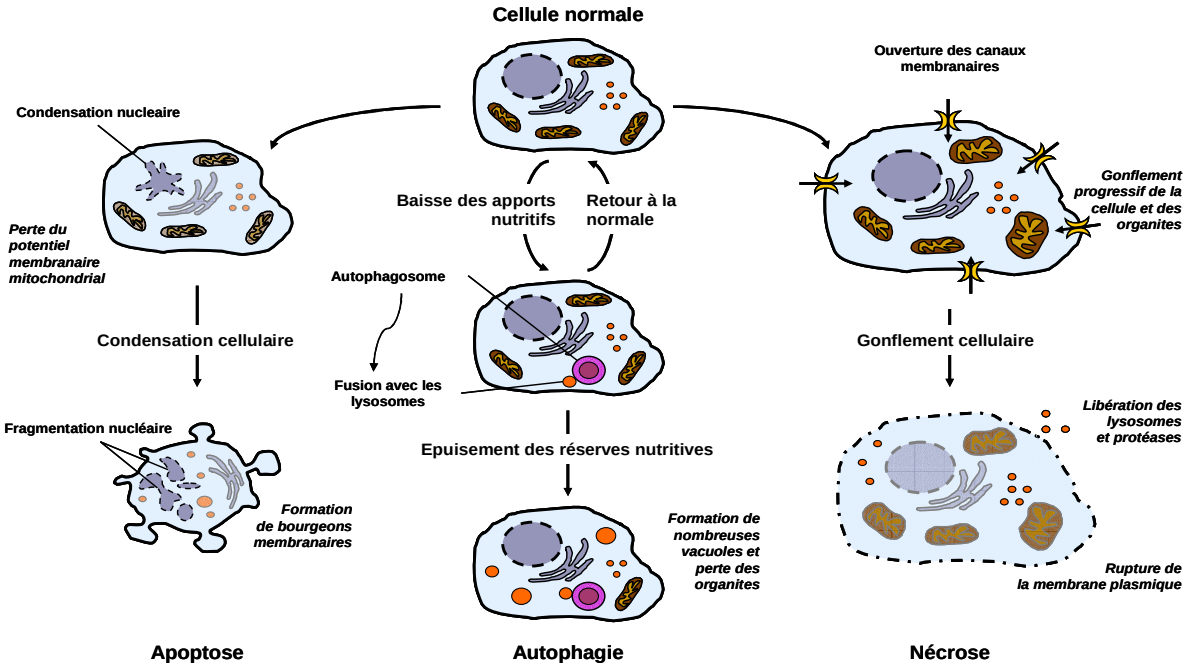


Figure 23 : Les principaux types de mort cellulaire : l'apoptose, l'autophagie et la nécrose. (Schéma inspiré de Hotchkiss et al., 2009).

Il existe trois types de mort possibles pour la cellule (Figure 23) : l'autophagie, la nécrose et l'apoptose (Hotchkiss *et al.*, 2009). L'autophagie est une réponse adaptative de la cellule, face à des conditions de stress comme la carence nutritionnelle, par laquelle elle produit de l'énergie par digestion de ses organites et macromolécules (Klionsky, 2007). L'autophagie permet à la cellule de « recycler » les macromolécules et les organites non essentiels, redondants ou non fonctionnels (Klionsky, 2007; Kroemer et Jaattela, 2005). La nécrose est considérée comme une mort cellulaire accidentelle. Elle résulte de l'effondrement des taux intracellulaires d'ATP et est fréquemment rencontrée lors de l'ischémie (Leist *et al.*, 1997; Malhi *et al.*, 2006). La rupture de la membrane plasmique qui en résulte conduit à la libération dans le milieu extérieur du contenu cytoplasmique ce qui induit la réponse

inflammatoire (Hotchkiss *et al.*, 2009). La mort cellulaire programmée ou apoptose est un processus cellulaire physiologique nécessaire à la survie des organismes pluricellulaires. C'est un mécanisme conservé sur le plan évolutif permettant d'éliminer les cellules vieillissantes, endommagées ou délétères pour l'organisme (Lawen, 2003; Ozoren et El-Deiry, 2003; Strasser *et al.*, 2000; Thorburn *et al.*, 2004). L'apoptose intervient dans de nombreux processus biologiques et dès l'embryogénèse. Elle intervient lors du développement des tissus normaux et de la réponse immunitaire (Vaux et Korsmeyer, 1999). Ainsi, au cours du développement humain, les cellules formant l'appendice caudal ou la palme entre les doigts entrent en apoptose. Pendant la sélection lymphocytaire, les lymphocytes B produisant des anticorps non réactifs ou auto-réactifs (attaquant le soi) sont éliminés de l'organisme par apoptose.

Bien que ne provoquant pas de réaction inflammatoire, l'apoptose s'accompagne, néanmoins, du recrutement de certaines cellules du système immunitaire telles que les macrophages. L'une des caractéristiques de l'apoptose est la translocation des résidus de phosphatidylsérine du feuillet interne au feuillet externe de la membrane plasmique, ce qui déclenche la phagocytose de ces cellules par les macrophages. L'apoptose est définie par d'autres phénotypes tels que la condensation de la chromatine, la fragmentation du noyau et la formation de corps apoptotiques (« bulles » de membrane plasmique). D'un point de vue biochimique, l'apoptose implique généralement l'activation par clivage protéique de caspases, protéines à activité cystéine protéase (Algeciras-Schimnich *et al.*, 2002; Ozoren et El-Deiry, 2003), responsables de la dégradation de substrats spécifiques conduisant à la fragmentation internucléosomique de l'ADN.

La transduction du signal de mort cellulaire implique deux voies distinctes et interconnectées : la voie extrinsèque, nécessitant l'interaction de récepteurs membranaires (appelés également récepteurs de mort) avec leurs ligands, et la voie intrinsèque faisant intervenir la mitochondrie (Figure 24, page 82).

Les récepteurs de mort appartiennent à la famille des récepteurs au TNF (Tumor Necrosis Factor) (Nagata, 1997). Dans cette famille, certains récepteurs sont spécialisés dans la survie cellulaire, d'autres dans l'induction de l'apoptose, d'autres encore ont une activité double puisque, selon le contexte cellulaire, ils induisent soit la survie soit la mort de la cellule. Les récepteurs spécialisés dans l'induction de l'apoptose, comme le récepteur Fas (appelé également CD95 ou APO-1) (Itoh *et al.*, 1991; Oehm *et al.*, 1992), TNF-R1 (Schall *et al.*, 1990), DR3 (Chinnaiyan *et al.*, 1996), DR4 (Chaudhary *et al.*, 1997), DR5 et DR6 (Pan *et al.*, 1998), ont en commun un domaine cytoplasmique de 80 acides aminés appelé domaine de

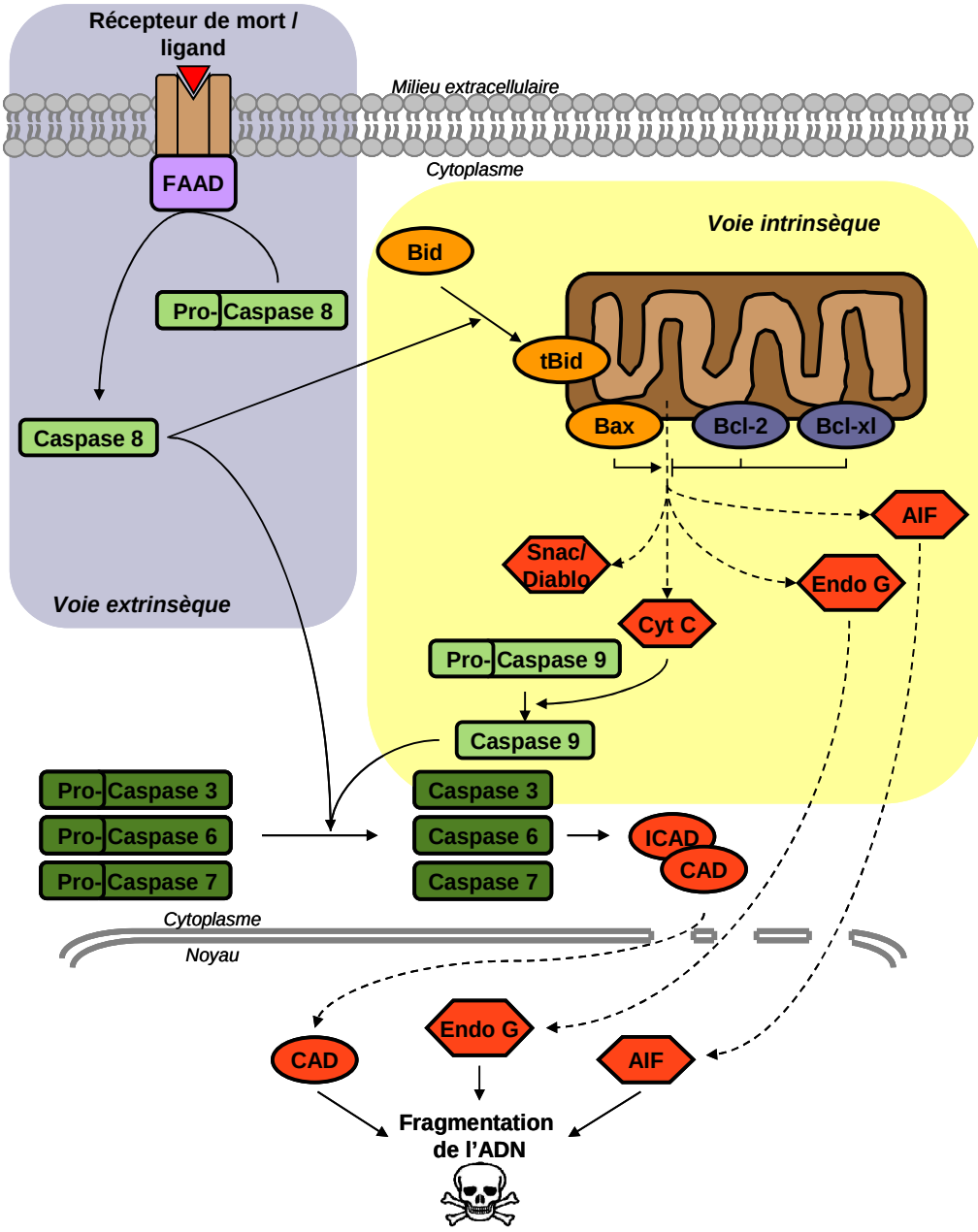


Figure 24 : Voies d'induction de l'apoptose. L'apoptose est classiquement induite par deux voies. Une voie dépendante de la reconnaissance d'un ligand par un récepteur membranaire et une voie faisant intervenir la libération de molécules mitochondriales initialement présentes dans l'espace intermembranaire. L'induction de l'une ou l'autre de ces voies aboutit à l'activation des caspases effectrices 3, 6 et 7 et à la dégradation de l'ADN nucléaire.

mort (Chaudhary *et al.*, 1997; Nagata, 1997; Tartaglia *et al.*, 1993). Le récepteur Fas est considéré comme l'archétype des récepteurs de mort. Il est exprimé à la surface de nombreux types cellulaires. L'activation de Fas par la liaison avec son substrat FasL conduit à son oligomérisation puis au recrutement de la protéine FADD (Fas Associated Death Domain)

(Chinnaiyan *et al.*, 1995). FADD interagit, alors, avec la pro-caspase 8. La formation de ce complexe induit le clivage activateur de la caspase 8. Une fois clivée, la caspase 8 se dimérise et peut alors cliver différentes caspases comme la caspase effectrice 3. Elle clive également la protéine Bid dont le fragment C-terminal conduit à l'activation de la voie mitochondriale (Li *et al.*, 1998; Luo *et al.*, 1998).

La mitochondrie joue un rôle primordial dans la régulation de l'apoptose par la voie intrinsèque (Brenner *et al.*, 1998; Kroemer *et al.*, 1997; Zamzami *et al.*, 1996). L'ouverture des pores de transition de perméabilité mitochondriaux conduit à l'induction de l'apoptose par libération, dans le cytoplasme, de molécules apoptogènes normalement séquestrées telles que le cytochrome c, le facteur AIF (apoptosis-inducing factor), l'endonucléase G, et les pro-caspases 2, 3 et 9. La libération des molécules précitées est sous le contrôle des protéines de la famille Bcl-2. Cette famille se caractérise par la présence de domaines BH conservés (Bcl-2 homology domain). La répartition de ces domaines permet de distinguer trois classes (Kelekar et Thompson, 1998). La première classe regroupe les membres à activité anti-apoptotique comme Bcl-2, Bcl-xl, Bcl-w, ou Mcl-1, lesquels présentent des homologies de séquence dans leurs domaines BH1, 2, 3 et 4. Cependant, ces membres ne possèdent pas tous de domaines BH4. La seconde classe comprend les protéines à activité pro-apoptotique telles que Bax, Bak, Bad, Bok ou Diva qui présentent des homologies de séquence pour leurs domaines BH1, 2 et 3 mais pas ou peu pour leur domaine BH4. Finalement, la dernière classe regroupe les membres ne possédant que le domaine BH3. Ce sont des protéines pro-apoptotiques comme Bik, Hrk, Bim, Blk, Bid et BNIP3 (Kelekar et Thompson, 1998). Tous les membres de la famille Bcl-2, hormis Bax et Bid, possèdent un domaine transmembranaire leur permettant d'être localisé à la membrane mitochondriale, au réticulum endoplasmique ou à la membrane nucléaire externe (Hsu *et al.*, 1997; Nguyen *et al.*, 1993). L'équilibre entre l'expression des membres de la famille Bcl-2 à activité pro-apoptotique, d'une part, et ceux à activité anti-apoptotique, d'autre part, permet de contrôler la libération des molécules séquestrées dans l'espace intermembranaire mitochondrial (Oltvai *et al.*, 1993). De part sa capacité à se lier à la membrane, la protéine Bcl-2 permet de bloquer la sortie du cytochrome c (Kluck *et al.*, 1997; Yang *et al.*, 1997) alors que la protéine Bax l'induit (Jurgensmeier *et al.*, 1998). Une autre caractéristique importante des protéines de la famille Bcl-2 est leur capacité à former des homo- ou hétérodimères grâce à leurs domaines BH. Ainsi, lorsque la protéine Bax est surexprimée par rapport à Bcl-2, elle s'homodimérise et se lie à la membrane mitochondriale (Bleicken *et al.*, 2009; George *et al.*) ce qui conduit à la libération du cytochrome c (Oltvai et Korsmeyer, 1994; Oltvai *et al.*, 1993; Sedlak *et al.*,

1995). Il semble, cependant, que la protéine Bcl-2 pourrait avoir une action anti-apoptotique sous sa forme monomérique (Conus *et al.*, 2000).

La prédominance de Bax sur Bcl-2 permet la perméabilisation de la membrane mitochondriale et la libération du cytochrome c, du facteur AIF et de l'endonucléase G. Ces deux derniers sont capables d'induire l'apoptose indépendamment de l'activation de caspases. Dans le cytoplasme, le cytochrome c s'associe à Apaf-1 (Apoptotic peptidase activating factor 1) pour former l'apoptosome (Acehan *et al.*, 2002) ce qui conduit au clivage de la caspase 9.

L'activation de la voie des récepteurs de mort ou de la voie mitochondriale aboutie au clivage de caspases effectrices telles que les caspases 3, 6 et/ou 7 (Cullen et Martin, 2009). Une fois activée, ces caspases permettent la libération de CAD (Caspase Activated DNase) de son inhibiteur ICAD (Inhibitor of Caspase Activated DNase) et clivent de nombreux substrats dont PARP (Poly (ADP-Ribose) Polymerase). Les protéines CAD, AIF et l'endonucléase G sont adressées au noyau et déclenchent la dégradation internucléosomique de l'ADN (formation de fragments d'environ 180 pb) responsable de la mort cellulaire (Li *et al.*, 2001; Varecha *et al.*, 2009).

Le but de cette étude a été de démontrer que la delta-lactoferrine est capable d'induire la mort cellulaire par apoptose mais également de déterminer le mécanisme d'induction mis en jeu.

Les résultats de cette étude ont été compilés et présentés sous la forme d'une publication :

Hardivillé, S., Hoedt, E., Mariller, C., Benaïssa, M., et Pierce, A., « Delta-lactoferrin induces cell death *via* the mitochondria death signaling pathway by upregulating Bax expression », en préparation pour *Eur. J. Cell. Biol. Chem.* (2010)

Au cours de cette étude nous avons pu démontrer que la delta-lactoferrine induit la mort cellulaire par apoptose grâce à son activité de facteur de transcription. En effet, nos résultats prouvent que l'expression de la delta-lactoferrine induit une augmentation des taux de la protéine Bax, ce qui conduit au clivage des caspases 9 et 7. L'utilisation des méthodes de gène rapporteur et d'immunoprécipitation de la chromatine couplée à la PCR temps réel, nous a permis de démontrer que la delta-lactoferrine reconnaît spécifiquement un élément de réponse présent dans les régions promotrices du gène *Bax* et induit l'expression de ce dernier.

DELTA-LACTOFERRIN INDUCES CELL DEATH *VIA* THE MITOCHONDRIA DEATH SIGNALING PATHWAY BY UPREGULATING BAX EXPRESSION

Stéphan Hardivillé[123], Esthelle Hoedt[123], Christophe Mariller[123], Monique Benaïssa[123] and Annick Pierce[123]*

[1]Univ Lille Nord de France, 59000 Lille,

[2] Université des Sciences et Technologie de Lille,

[3]Unité de Glycobiologie Structurale et Fonctionnelle, UMR 8576 CNRS, IFR-147, 59655 Villeneuve d'Ascq, France

Key words : Delta-lactoferrin, lactoferrin, transcription factor, apoptosis, Bax, tumor suppressor, intrinsic pathway

*Correspondance to: Pr. Annick Pierce, UGSF, UMR 8576 CNRS, USTL, 59655 Villeneuve d'Ascq, France.
Fax: +33 3 20 43 65 55; Tel: +33 3 20 33 72 38; E-mail: annick.pierce@univ-lille1.fr

Abbreviations : Lf, lactoferrin ; ΔLf, delta-lactoferrin ; ΔLfRE, delta-lactoferrin response element

Abstract

Delta-lactoferrin is a transcription factor belonging to the lactoferrin family the expression of which inhibits cell proliferation and leads to *Skp1* and *DcpS* gene transactivation. In this study, we showed that delta-lactoferrin expression also induces cell death and revealed the mechanisms by which it mediates apoptosis in HEK 293 and MCF7 cells using a cell viability assay and DNA fragmentation. Western blot analyses showed that apoptosis was caspase-9, 7 and 8 dependent. Proteolytic cleavage of the endonuclease PARP was significantly increased. The levels of expression of Bcl family members was detected by immunochemistry and showed that the Bcl-xl/Bax and Bcl-2/Bax protein ratios were decreased. Apoptosis induction by delta-lactoferrin triggers Bax up-regulation. Analysis of the *Bax* promoter region detected a ΔLfRE located at –155 bp from the transcription start point. Both luciferase reporter gene and chromatin immunoprecipitation assays confirmed that delta-lactoferrin interacts *in vitro* and *in vivo* specifically with this sequence. Its deletion, realized using directed mutagenesis, totally abolished delta-lactoferrin transcriptional activity, identifying it as a delta-lactoferrin-responsive element. These results indicate that the *Bax* gene is a novel delta-lactoferrin target and that delta-lactoferrin pro-apoptotic effects are mainly mediated by the activation of the mitochondria-dependent death-signaling pathway. Downregulated in the case of cancer, delta-lactoferrin, which exerts antiproliferative and proapoptotic effects, might be effective as an anti-tumoral agent and be considered as a tumor suppressor.

Introduction

Apoptosis is a universal cell suicide pathway characterized by membrane blebbing, nuclear breakdown, and DNA fragmentation. It is an indispensable process controling development, tissue homeostasis, and regulation of the immune system (Li et al., 2001) the dysregulation of which has been implicated in numerous pathogeneses. Two main induction pathways lead to cell death by apoptosis, the extrinsic or death receptor pathway and the intrinsic or mitochondrial-death pathway. Death receptors are activated through the binding of their ligands in response to extracellular stimuli (Nagata, 1997). Receptor stimulation induces formation of the death-induced signaling complex and leads to the recruitment to

and the cleavage of initiator caspases, like caspase 8. The mitochondrial pathway is characterized by the permeabilisation of the external mitochondrial membrane initiated by Bcl-2 family members (Kelekar and Thompson, 1998). Among them, anti-apoptotic members like Bcl-2 and Bcl-xl are responsible for preserving membrane integrity whereas apoptosis mediators such as Bax induce its permeabilisation (Bleicken et al., 2009) ; the balance between these two subgroups is critical and helps determine the susceptibility of cells to a death signal. Pro-apoptotic effectors trigger the release of apoptosis mediators such as cytochrome c, AIF or G endonuclease from the mitochondrial intermembrane space into the cytosol. These factors promote and amplify the apoptotic cascade directly or through the formation of the cytochrome c/Apaf-1/caspase-9-containing apoptosome complex (Newmeyer and Ferguson-Miller, 2003). Activation of both pathways lead to the clivage of downstream effector caspases such as the caspase 3, 6 and 7 (Cullen and Martin, 2009) which in turn are responsible for the release of the CAD nuclease from its inhibitor ICAD and to the cleavage of numerous proteins such as PARP. CAD, AIF and endonuclease G are then targeted to the nucleus to trigger DNA fragmentation (Li et al., 2001). Even if the extrinsic and intrinsic pathways appeared to be independent it is now clear that crosstalk may exist. For example, caspase 8 which is activated in extrinsic apoptosis also cleaves Bid to form tBid leading to cytochrome c release and higher apoptotic response (Li et al., 1998).

In this report, we show that delta-lactoferrin (ΔLf) is a transcription factor which exerts pro-apoptotic activities by upregulating the synthesis of Bax. ΔLf belongs to the lactoferrin family. Lfs exist as different variants due to post-translational modifications, gene polymorphisms and alternative splicing which lead to the production of lactoferrin isoforms that are all involved in host defence. The two main lactoferrin isoforms are the secretory lactoferrin (Lf) (Masson et al., 1969) and ΔLf, its intracellular counterpart (Liu et al., 2003, Mariller et al., 2007, Siebert and Huang, 1997). The 5'-regulatory region of the human Lf gene contains two promoters that are differentially transactivated leading to the synthesis of two transcripts (Hoedt et al., Liu et al., 2003). Transcription from P1 leads to the Lf mRNA and from P2 to ΔLf mRNA (Liu et al., 2003, Siebert and Huang, 1997). The N-terminus and their glycosylation pattern are the only features that discriminate the isoforms. ΔLf is a truncated Lf without the leader sequence and the first 26 aminoacid residues present

in Lf. Therefore, whereas Lf is secreted, ΔLf stays in the cytoplasm from which it enters the nucleus (Liu et al., 2003, Mariller et al., 2007). As a secreted protein, Lf is N-glycosylated (Spik et al., 1988) whereas ΔLf is *O*-GlcNAcylated (Hardivillé et al., 2010). The role of the Lf glycan moiety seems to be restricted to a decrease in the immunogenicity of the protein and its protection from proteolysis. On the other hand, *O*-GlcNAcylation which competes directly with phosphorylation (Hart et al., 1996), positively regulates ΔLf stability and negatively redulates its transcriptional activity (Hardivillé et al., 2010). Mutation of the *O*-GlcNAc sites in ΔLf leads to a constitutively active mutant with a higher activity than the wild type *O*-GlcNAcylated protein (Hardivillé et al., 2010, Mariller et al., 2009).

Both Lf isoforms display antitumoral properties and behave as tumor suppressors. Like tumor suppressor genes, Lf gene loses its function in cancer cells and expression of both Lf and ΔLf is downregulated or silenced in cancer cells or tissues (Benaissa et al., 2005, Penco et al., 1999). Genetic and epigenetic alterations have been found (Panella et al., 1991, Teng et al., 2004) that give a real proliferative advantage to cancer cells. In a study including 99 patients, we showed that a high level of expression of the two transcripts was correlated with a good prognosis (Benaissa et al., 2005). Both isoforms affect cell growth and lead to cell cycle arrest but whilst Lf antitumoral activity involved interactions with cell surface receptors and activation of different signaling pathways, the effects of ΔLf involved gene transactivation (Breton et al., 2004, Damiens et al., 1999, Son et al., 2006, Xiao et al., 2004). Lf exerts antitumoral and antimetastatic properties *in vitro* and *in vivo* although if it is mainly known as a regulator of the innate immune system of mammals (Bezault et al., 1994). Depending on the cellular models studied, Lf is described to inhibit cdk2 and cdk4 kinase activities, increase p21 cdk inhibitor expression, promote retinoblastoma protein-mediated growth arrest, increase p21 expression and upregulate the tumor suppressor p53 and its target genes (reviewed in Tsuda *et al*., 2005; Legrand *et al*., 2008; Pierce *et al*., 2009). Both Lf isoforms possess nuclear targeting sequences (Mariller et al., 2007) and a strong concentration of positive charges that might create potential DNA interaction sites (Baker, 2005). Moreover, they are both capable of binding DNA sequences, *in vitro* for Lf (He and Furmanski, 1995, Son et al., 2002) and *in vivo* for ΔLf (Mariller et al., 2007, Mariller et al., 2009).

Thus, while it is clear that ΔLf acts as a transcription factor *via* a functional ΔLfRE it is still unclear whether Lf possesses the same activity *in vivo*. ΔLf overexpression leads to cell cycle arrest in S phase (Breton et al., 2004) and to cell death. Its transcriptional activity targets *DcpS* and *Skp1* genes (Mariller et al., 2007, Mariller et al., 2009). DcpS is a scavenger decapping enzyme involved in cap nucleotide metabolism, the overexpression of which might be crucial for mRNA turnover and quality control of mRNA biogenesis (Nuss et al., 1975, van Dijk et al., 2003). Skp1, or S-phase kinase associated protein, is essential for cell cycle progression at both the G1/S and G2/M transitions. It belongs to the SCF complex at the G1/S transition (Ang and Wade Harper, 2005). This complex is responsible for the ubiquitinylation of many cell cycle regulators leading to their degradation by the proteasome (Bai et al., 1996). At the G2/M transition, it belongs to the CBF3 complex (Kitagawa and Hieter, 2001), which is critical for kinetochore assembly. Silencing of the *Skp1* gene in mice led to aneuploidy during mitosis (Piva et al., 2002). Therefore, modulation of Skp1concentration may have considerable consequences for cell cycle progression.

Both Lf isoforms promote apoptosis. Lf pro-apoptotic activity has been observed in human cell lines (Baumrucker et al., 2006, Valenti et al., 1999) and after administration to rodents (Chandra Mohan et al., 2006, Fujita et al., 2004b, Fujita et al., 2004a, Spagnuolo et al., 2007, Spagnuolo and Hoffman-Goetz, 2008). This process seems involve the death receptor Fas and activation of caspase-3 and -8 (Lee et al., 2006, Fujita et al., 2006). Like Lf, ΔLf is also at the crossroads between cell survival and cell death. Our findings showed that ΔLf interacts directly with a ΔLfRE in both the *Fas* and *Bax* promoters, and that these interactions lead to their transcriptional activation with a much higher response for *Bax*. Thus, by causing overexpression of Bax, ΔLf triggers apoptosis through the intrinsic mitochondrial-death pathway.

Material et Methods

Cell culture, reagents and transfections

Human kidney HEK293 (ATC CRL-1573) cells, human breast cancer MDA-MB-231 (ATCC HTB-26) and MCF7 (ATCC HTB-22) cell lines were grown in monolayers in DMEM (« Dulbecco's Modified Eagle's Medium ») or in EMEM (« Eagle's Minimum Essential Medium ») medium only for MCF7, at 37°C and under 5% CO_2. The media were both complemented with

10% FCS and 2mM L-glutamine. Human recombinant insulin (0.01 mg/ml) was added only to MCF7 medium. Both cells were transfected as described (Hardivillé *et al.*, 2010) using DreamFect™ (OZ Biosciences) at the rate of 1µg of pcDNA-ΔLf or its mutant constructs to $1x10^6$ cells. The amount of ΔLf expression vectors was adjusted to keep the ΔLf amount similar to that found in normal NBEC cells (Benaïssa *et al.*, 2005). Transfections wee done in triplicate ($n \geq 3$). The stable MDA-MD-231 expressing ΔLf cell line under doxycyclin induction (MDA-MB-231-ΔLf) was obtained as previously described (Breton *et al.*, 2004) using the tetracyclin inducible Tet-on system (Clontech). The presence of ΔLf messengers was confirmed by TaqMan qPCR after RNA extraction (Hoedt *et al.*, 2010). Cell culture reagents were from Dutscher, Cambrex Corporation and Invitrogen. Other reagents were from Sigma. Antibodies against 3xFLAG epitope (mouse monoclonal anti-FLAG M2 antibody) were from Sigma. Rabbit polyclonal antibodies against caspase 7, caspase 9, PARP, PARP (Asp 214), Bcl-2, Bcl-xl and Bax were from Cell Signaling. Antibodies against caspase 8 (rabbit polyclonal anti-caspase 8) and actin (goat polyclonal antibodies I-19) were from Santa Cruz Biotechnologies.

Cell viability

Cell viability of ΔLf expressing cells ($1x10^6$) was assessed by counting the cells 24 h and 48 h after transfection using Trypan blue 0.4% (v/v) (Cambrex). Dead and viable cells were discriminated by ATP-dependent exclusion of the colorant.

DNA fragmentation

For DNA gels, 5×10^5 ΔLf-expressing cells were lysed 20 min on ice, 24 h after transfection without vortexing or pipetting in 20 µl DNA extraction buffer (2 mM EDTA, 0.8% SDS, 100 mM Tris-HCl pH 8). The lysat was first incubated 30 min at 37°C with 2 µl of 50 mg/ml RNase A and then incubated 90 min at 50°C with 10 µl of 20 mg/ml proteinase K. The DNA pellet was then subjected to electrophoresis on a 1.8 % agarose gel (Invitrogen) containing ethidium bromide at 25 V and visualized under ultraviolet light.

DNA purification

Genomic DNA was purified using the QIAprep Spin Miniprep Kit (Qiagen). Plasmid extraction was performed using QIAprep® Spin Miniprep Kit (Qiagen). The purity and integrity of each extract were checked using the nanodrop ND-1000 spectrophotometer (Labtech International).

Vector constructs and site-directed mutagenesis

pcDNA-ΔLf, pcDNA-ΔLf^{M4} and p3xFLAG-CMV10-ΔLf were constructed as described (Mariller et al., 2007).

Promoter regions containing the S1 element of *Bax* (139 bp) and *Fas* (137 bp) were amplified with specific primer pairs as follows : sense 5'-TGGATAAATGAAGGCATTA-3', antisense 5'-TGACCGCACCTGCCTCGCT-3' for *Bax*, and sense 5'-CTATTAGATGCTCAGAGTG-3', antisense 5'-CATCTGTCACTGCACTTAC-3' for *Fas*. Amplification products were further cloned in the basic pGL3-Luc reporter vector as described (Mariller et al., 2007) and sequenced. Mutants were generated using the QuikChange® Sitedirected Mutagenesis Kit (Stratagene) with pGL3-S1^{Bax}-Luc and pGL3-S1^{Fas}-Luc as templates and primer pairs: sense 5'-CGCCACTGCTGTATCGGGAGATGC-3', antisense 5'-AGCATCTCCCGATACAGCAGTGGCG-3' (pGL3-delS1^{Bax}-Luc) and sense 5'-GGTCTTCCTCATGAACAGTCTACTG-3', antisense 5'-CAGTAGACTGTTCATGAGGAAGACC-3' (pGL3-delS1^{Fas}-Luc). Following sequencing, the inserts were cloned into fresh pGL3-Luc for reporter gene assays.

Reporter gene assay

Reporter gene assay was performed using pGL3-S1^{Bax}-Luc, pGL3-delS1^{Bax}-Luc, pGL3-S1^{Fas}-Luc and pGL3-delS1^{Fas}-Luc reporter vectors and the pcDNA-ΔLf construct as described in (Mariller et al., 2009). Cell lysates were assayed using a luciferase assay kit (Promega) in a Tristar multimode microplate reader LB 941 (Berthold Technologies). Relative luciferase activities correspond to the luciferase activity normalized to protein content as in (Mariller et al., 2007). Each experiment represents at least three sets of independent triplicates.

***In vivo* DNA binding assays and Q-PCR**

ChIP assay was performed as described (Mariller et al., 2009, Metivier et al., 2003). For ChIP assay, cells were transfected with p3xFLAG-CMV10-ΔLf in order to produce a 3xFLAG-tagged-ΔLf (Mariller et al., 2007, Hardivillé et al., 2010). Briefly, ChIP complexes (8 x 10^6 cells) were immunoprecipitated with M2 (anti 3xFLAG-tag) used at 1:250 or rabbit anti-IgG (GE Healthcare Life Sciences) or without antibodies. Input corresponds to DNA not submitted to immunoprecipitation. Samples were then decrosslinked overnight at 65 °C and DNA purified using the Qiagen DNA purification Kit (Qiagen). Amplification was carried out in triplicate (n=3) in the presence of 2 µl of purifed DNA, primer pairs used to amplify the *Bax* promoter fragment and Brilliant SYBER Green QPCR Master Mix (Stratagene) according to the manufacturer's instructions. Samples were then submitted to 40 cycles of

amplification (denaturation : 30 s at 90 °C, hybridation : 30 s at 55 °C, elongation : 30 s at 72 °C) in a thermocycler Mx4000 (Stratagene). Data presented in Figure 9 are expressed as a percentage of input.

Immunoblotting

Proteins were extracted from frozen cell pellets in 500 µL RIPA buffer as described (Mariller et al., 2007). Protein content was assayed using microBCA (Pierce) according to the manufacturer's instructions. For direct immunoblotting, samples mixed with 4x Laemmli buffer were boiled for 5 min. 20 µg of protein from each sample were submitted to 7.5% SDS-PAGE and immunoblotted. Blots were saturated by 3% SAB in TBS-Tween 0.1%, 2h at room temperature. Blots were subsequently probed in the same solution with primary antibodies (anti-actin, 1:10000; anti-caspase 8, 1:1000 and anti-caspase 7, anti-caspase 9, anti-PARP, anti-PARP (Asp 214), anti-Bcl-2, anti-Bcl-xl and anti-Bax 1:500) overnight at 4°C and with secondary antibodies conjugated to horseradish peroxidase (GE Healthcare Life Sciences) at room temperature for 1 h before being detected by chemiluminescence (ECL advance, GE Healthcare Life Sciences). Each result in which immunoblots are presented corresponds to one representative experiment of at least three conducted.

Results and discussion

ΔLf induces programmed cell death

Our previous studies have shown that ΔLf expression induces cell cycle arrest at S phase (Breton et al., 2004) and is correlated with a good prognostis in the case of breast cancer (Benaissa et al., 2005) suggesting that ΔLf may exerts antitumoral activities. However, the activity of ΔLf on cell homeostasis is still poorly understood. To better characterize its function, we transiently expressed ΔLf and its constitutively active isoform (M4) (Hardivillé et al., 2010) in two different cell lines and performed cell viability assays. HEK 293 and MCF7 were chosen since they express none or an extremely low level of ΔLf transcripts (Benaissa et al., 2005, Hoedt et al.). Figure 1 shows cell viability at 24 h and 48 h after ΔLf expression. The cell viability of HEK 293 cells was not affected 24 h after ΔLf or M4 expression compared to cells transfected with empty pcDNA vector whereas it decreased 48 h after transfection. The percentage of dead cells corresponded to approximately 30% of total cells at 48 h. The total number of cells did not really change significantly between 24 h to 48 h, possibly due to the doubling time of these cells which is around 36 h to 48 h. Therefore, the death of HEK 293 cells appears to be ΔLf or M4 mutant dependent.

Cell death was visible earlier for MCF7 cells. The total number of cells decreased slightly at 24 h and strongly at 48 h for ΔLf expressing cells compared to their respective controls. The total number of dead MCF7 cells was approximately 10 % at 24 h whether cells were treated or not and 30 % at 48 h suggesting that MCF7 cells are more susceptible than HEK 293 cells to cell death. Nevertheless, the total number of viable cells decreased for both ΔLf isoform expressing cells by around 30 % at 24 h and 70 % at 48 h compared to their respective controls suggesting that both ΔLf isoforms have a strong impact either on cell cycle progression or cell death. Our data show that both ΔLf isoforms have a higher impact on MCF7 cell viability than on HEK 293 cells, the effects of the expression of the M4 mutant are slightly more visible than those of the wild type at 48 h.

In order to determine whether cell death is due to apoptosis or necrosis we next studied the genomic DNA degradation pattern in ΔLf expressing cells. Since apoptosis is characterized by the activation of endogenous endonucleases such as CAD and endonuclease G (Li et al., 2001) with subsequent cleavage of chromatin DNA into internucleosomal fragments of 180 bp, whereas necrosis results in non-enzymatic degradation of DNA which forms a "smear" on agarose gels (Shiokawa and Tanuma, 2004), we determined the electrophoretic behaviour of the genomic DNA extracted from these cells. Although the detection of DNA fragmentation was not quantitative, each well was loaded with the same amount of DNA and comparison to controls allows a relative estimation of the numbers of cells undergoing apoptosis. The presence of a "ladder" pattern at 180 bp intervals for HEK 293 cells that overexpressed ΔLf or its constitutively activate isoform is in favor of an apoptotic DNA fragmentation (Figure 2A, lanes 3 and 4). This fragmentation was strongly visible with the later. MCF7 cells show a basal level of apoptotis as evidenced by the presence of a " DNA scale" in both control lanes ie NT and pEGFP representing either untransfected cells or GFP expressing cells, respectively (Figure 2B, lanes 1 and 2). This pattern suggests that a small subpopulation of MCF7 cells was already in apoptosis. This feature has already been reported (Chen et al., 2009, Kagawa et al., 2001, Luzhna and Kovalchuk, 2009) and is consistent with our cell viability results (Figure 1B). The signals observed for the tracks corresponding to MCF7 cells which express ΔLf or its M4 mutant have a higher intensity than that observed for the control tracks (Figure 2B, lanes 3 and 4). These results allows us to conclude that a larger population of MCF7 cells is entering into

apoptosis when ΔLf isoforms are expressed. Like its secretory counterpart (Baumrucker et al., 2006), ΔLf is also able to induce programmed cell death in cancerous mammary gland cell lines. Therefore ΔLf, with its restricted expression and its antiproliferative and proapoptotic activities is a good candidate as a tumor suppressor. Thus, further work will be done to collect data on the regulation of ΔLf transcription in order to develop strategies to enhance its expression in cancer cells.

O-GlcNAcylation is a dynamic and reversible post-translational modification mainly found in cytoplasm and nucleus (Hart et al., 1989). There are proteins in almost every functional class that are subjected to this modification and about one-quarter of the identified *O*-GlcNAcylated proteome is composed of proteins involved in transcription or translation (Hart et al., 2007). ΔLf belongs to this group of *O*-GlcNAcylated members and since *O*-GlcNAcylation negatively controls ΔLf transcriptional activity we produced a mutant in which the *O*-GlcNAcylation sites were mutated (Hardivillé et al., 2010). The higher DNA cleavage in the presence of the ΔLf constitutively active M4 mutant confirms that this isoform is a much more efficient proapoptotic agent than the wild type protein. ΔLf and M4 proapoptotic activities are currently under study using a prototype helper-dependent adenovirus gene delivery system (Long et al., 2009).

ΔLf triggers the mitochondrial death pathway

ΔLf induces apoptosis in both HEK 293 and MCF7 cells. Since normal mammary gland epithelial cells express ΔLf whereas breast cancer cells such as MCF7 do not, we next studied the impact of « re-introducing » a pro-apoptotic transcription factor like ΔLf and investigated whether the death receptor or the mitochondrial pathway wss involved. DNA fragmentation would occur whatever the pathway induced. Indeed, the release of endonuclease G from the intermembrane space of mitochondria is sufficient in itself for the fragmentation of genomic DNA (Li et al., 2001) but the release of the CAD nuclease by effector caspases may also be involved (Shiokawa and Tanuma, 2004). Figure 3 shows the time course of the cleavage of PARP, a substrate of these effector caspases, in ΔLf-expressing MCF7 cells. In the absence of ΔLf, full length PARP and a basal level of cleaved PARP are visible using antibodies directed against both isoforms (Figure 3, lane 1, top panel). However, when ΔLf is expressed the amount of cleaved PARP is markedly increased (Figure 3, lanes 2-5, top panel).

The cleavage of PARP was confirmed using cleaved PARP (Asp214) antibody which specifically detects endogenous levels of the large fragment (85 kDa) of human PARP produced by caspase cleavage. This antibody recognizes neither full length PARP nor other PARP isoforms (Figure 3, lanes 2-5, medium panel). Actin was used as loading control (Figure 3, bottom panel).

In order to determine which pathway was induced following ΔLf expression, we next studied the cleavage of different caspases (Figure 4) except caspase 3 which is not functional in MCF7 cells (Fazi et al., 2008, Janicke et al., 1998). Since the cleavage of PARP can also be triggered by the activation of caspase 7 (Walsh et al., 2008), we further investigated cleavage of this caspase. The immunorevelation of the time course of the cleavage of caspase 7 is presented in Figure 4A, top panel. A fragment of 20 kDa corresponding to the cleaved caspase 7 appeared about 18 h post transfection while the signal corresponding to the 35 kDa full length caspase 7 decreased concomittantly 24 h after transfection. This result demonstrates that the activation of caspase 7 is dependent on ΔLf expression. Actin was used as loading control (Figure 4A, bottom panel). The cleavage of caspase 7 follows that of PARP (Figure 3).

Caspase 7 is itself cleaved by caspases known as initiator caspases such as caspases 8 and 9. Caspases 8 and 9 are involved upstream either the extrinsic or the intrinsic pathway, respectively. The formation of the complex death receptor/ligand/FADD will therefore trigger the activation by cleavage of caspase 8 whereas the release of cytochrome c from mitochondria and the formation of the apoptosome will trigger the cleavage of caspase 9 (Cullen and Martin, 2009). Figures 4B and 4C show the cleavage of these two caspases. In both cases, immunodetection points out the cleavage of both caspases. The signals at 37 and 35 kDa corresponding to the cleaved forms of caspase 9 are visible from 18 h post transfection and increased until 48 h (Figure 4 B, top panel). Similarly, the appearance of signals at 26/28 kDa and 18 kDa, reflecting the activation of the cleavage of caspase 8 is also visible (Figure 4 C, top panel). Although both caspase-dependent pathways seem induced by ΔLf expression, we further investigated which of these pathways might be the principal one involved and started by studying the intrinsic pathway.

The release of molecules sequestered in the mitochondrial intermembrane space, such as cytochrome c, is under the control of Bcl-2 family members and notably of the Bax/Bcl-2 or

the Bax/Bcl-xl ratio (Oltvai et al., 1993). Figure 5 shows the variation of the expression of the pro-apoptotic Bax and anti-apoptotic Bcl-2 and Bcl-xl proteins during ΔLf expression. The production of ΔLf by MCF7 cells leads to the overexpression of the Bax protein. The amount of Bax increased from 18 h and peaked at 24 h. In parallel, no alteration in the level of expression of Bcl-2 was detected, suggesting that the presence of ΔLf does not influence its expression. The immunodetection of Bcl-xl shows that the amount of this protein decreases 18 h after transfection (Figure 5, lane 2) and strongly diminushes after 24 h (Figure 5, lanes 3 and 4) compared to the control (Figure 5, lane 1). Therefore in both cases the Bax/Bcl-xl and Bax/Bcl-2 ratios are altered. The highest level was reached 24 h after transfection and corresponded to the maximal expression level of ΔLf as previously shown (Mariller et al., 2007, Mariller et al., 2009). Our data demonstrate that ΔLf modifies the balance between Bcl-2 members at the mitochondrial membrane leading to cytochrome c release (Sedlak et al., 1995) and in turn to the cleavage of caspase-9 (Figure 4) confirming the proapoptotic effect of ΔLf and the activation of the intrinsic mitochondria-death pathway. In parallel, we investigated the possibility that ΔLf might also induce the extrinsic pathway. The cleavage of caspase 8 suggests that, besides the mitochondrial pathway, the extrinsic pathway may also be involved. But our preliminary results failed to demonstrate that the expression of the Fas receptor is upregulated. Nevertheless, activation of the other death receptors or crosstalk between the extrinsic and the intrinsic pathways might be responsible for caspase 8 cleavage. Recently it was shown that, in addition to complement caspase 3 in cells where it is inactive, caspase 7 is able to cleave caspases 2 and 6 which in turn target caspase 8. Therefore, the cleavage of caspase 8 may occur independently of the activation of death receptors (Inoue et al., 2009) suggesting that it may no longer be a critical criterion for the activation of the death receptor pathway. Further investigations will be necessary to clarify this point.

Bax is a new ΔLf target gene

Since ΔLf is a transcription factor, we next investigated the putative presence of a ΔLf response element (ΔLfRE) in the *Bax* promoter region. The *in silico* study highlighted the presence of a putative response element at -155 bp from the transcription start site (Figure 6). Ths sequence is homologous to the S1 sequence described as an Lf binding

sequence by He and Furmanski (1995) and found functional in the promoter of the *Skp1* and *DcpS* genes (Mariller et al., 2007, Mariller et al., 2009). In this study, we also examined the promoter region of death receptor genes. Although we were unable to demonstrate an upregulation of the Fas receptor in response to ΔLf expression, *in silico* analyses of the *Fas* promoter region revealed the presence of a putative ΔLfRE at -1013 bp from the transcription start site (Figure 7).

In order to demonstrate the functionality of these two response elements, the sequences surrounding each response element were amplified using PCR. A 139 bp PCR product corresponding to the – 216 bp to - 77 bp of the *Bax* promoter region and a 137 bp PCR product corresponding to the - 1068 bp to - 931 bp of the *Fas* promoter region were cloned into pGL3 promoter luciferase reporter vectors. Luciferase reporter assays were performed in HEK 293 cells and verified in the inducible ΔLf-MDA-MB-231 cell line. As the results were comparable, only the data obtained with the MCF7 cells are presented. The results are summarized in Figure 8A for the *Bax* promoter and Figure 8B for the *Fas* promoters. Transactivation of *Bax* by ΔLf led to a 82-fold increase as compared to the basal expression level whereas under the same conditions, transactivation of *Fas* led to a 10-fold increase. ΔLf therefore enhances transcription from both ΔLfRE sequences but at a much higher level for *Bax*. In order to demonstrate the functionality of both ΔLfREs, the GGCACT bases in each sequences were deleted. The results obtained with the deleted ΔLfREs within the pGL3 luciferase reporter plasmids are shown in Figure 8. Interestingly, deletion of part of each ΔLfRE strongly diminished ΔLf transcriptional activity of either $S1^{Bax}$ or $S1^{Fas}$ (Fig. 7B). The percentage of inhibition measured was about 80 % for $delS1^{Bax}$ and 100 % for $delS1^{Fas}$ as compared to the wild-type promoters. Nevertheless, even if both cis-acting elements are functional, *Bax* ΔLfRE is much more efficient for potentiating transcription and confirmed that ΔLf preferentially overexpressed Bax.

The alignment of ΔLfREs present in different ΔLf target genes (Mariller et al., 2007, Mariller et al., 2009) with the S1 sequence described by He and Furmanski (1995) shown in Table 1 points out a core of five conserved nucleotides (GCACT). Among the four ΔLfREs, $S1^{Skp1}$ and $S1^{DcpS}$ are the most efficient with a higher score for $S1^{Skp1}$ while $S1^{Fas}$ is the least efficient (data not shown). The biggest difference that may account for the lower efficiency

of $S1^{Fas}$ is the presence of an adenylic residue instead of a conserved thymidylic residue at position 8. Nevertheless, these four ΔLfREs are not present in the same nucleotidic environment and the presence of negative cis-regulatory elements can not be excluded. Therefore, the different ΔLfREs should be assayed for their relative efficiency when located in the same nucleotidic environment.

In order to determine whether ΔLf interacts *in vivo* with the *Bax* ΔLfRE, we performed chromatin immunoprecipitation assays and the quantification of *Bax* promoter region specifically linked to ΔLf was done using qPCR. The qChIP results are shown in Figure 9. To this end, a 3XFLAG-Nterminus-tagged ΔLf was used to obtain the most reliable results (Mariller et al., 2007, Mariller et al., 2009). The DNA purified from the sonicated chromatin immunoprecipitated with anti-FLAG (M2) or not was analyzed by qPCR using *Bax*-promoter-specific primer pairs. Figure 9 shows the percentage of immunoprecipitated *Bax* promoter sequence containing a ΔLfRE bound to ΔLf compared to the whole pool of *Bax* promoter sequence containing a ΔLfRE present in the same cell sample prior immunoprecipitation. A weak signal was detected in control conditions (NIP and IR) and corresponded to the "background" inherent to the method. The immunoprecipitation of ΔLf with the M2 antibody allows the amplification of the promoter region of the *Bax* promoter the signal of which is significantly different from that observed in control conditions. The input corresponding to the positive control, represents all the ΔLfRE present in the sample before immunoprecipitation. The upstream sequence of the albumin promoter devoid of ΔLf response elements, was also amplified and served as a negative control. As expected, this promoter region was not enriched by the M2 antibodies (data not shown). The signal corresponding to the M2 immunoprecipitated material represents approximately 0.7% of the input, meaning that 0.7% of the response elements of the *Bax* promoter are physically occupied by ΔLf. This value is extremely low but coherent with bibliographic data described for other transcription factors (Nagaoka et al., 2005, Yochum et al., 2007) and indicates that ΔLf interacts *in vivo* with its response element in the *Bax* promoter.

Acknowlegments

This investigation was supported in part by the CNRS UMR 8576 (Unité de Glycobiologie Structurale et Fonctionnelle), the Institut Fédératif de Recherche n° 147, the Université des Sciences et Technologies de Lille I, the Comité du Nord de la Ligue Nationale contre le Cancer and the Association pour la Recherche sur le Cancer (« *Etude du rôle de la delta-lactoferrine, des gangliosides et des neurotrophines dans le développement et la progression du cancer du sein*"). We would like to thank Dr. R. J. Pierce (CIIL, Institut Pasteur de Lille, France) for critical reading of this manuscript.

References

Ang, X. L. & Wade Harper, J. (2005). SCF-mediated protein degradation and cell cycle control. *Oncogene, 24***,** 2860-2870.

Bai, C., Sen, P., Hofmann, K., Ma, L., Goebl, M., Harper, J. W. & Elledge, S. J. (1996). SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. *Cell, 86***,** 263-274.

Baker, E. N. (2005). Lactoferrin: a multi-tasking protein par excellence. *Cell Mol Life Sci, 62***,** 2529-2530.

Baumrucker, C. R., Schanbacher, F., Shang, Y. & Green, M. H. (2006). Lactoferrin interaction with retinoid signaling: cell growth and apoptosis in mammary cells. *Domest Anim Endocrinol, 30***,** 289-303.

Benaissa, M., Peyrat, J. P., Hornez, L., Mariller, C., Mazurier, J. & Pierce, A. (2005). Expression and prognostic value of lactoferrin mRNA isoforms in human breast cancer. *Int J Cancer, 114***,** 299-306.

Bezault, J., Bhimani, R., Wiprovnick, J. & Furmanski, P. (1994). Human lactoferrin inhibits growth of solid tumors and development of experimental metastases in mice. *Cancer Res, 54***,** 2310-2312.

Bleicken, S., Classen, M., Padmavathi, P. V., Ishikawa, T., Zeth, K., Steinhoff, H. J. & Bordignon, E. (2009). Molecular details of Bax activation, oligomerization and membrane insertion. *J Biol Chem*.

Breton, M., Mariller, C., Benaissa, M., Caillaux, K., Browaeys, E., Masson, M., Vilain, J. P., Mazurier, J. & Pierce, A. (2004). Expression of delta-lactoferrin induces cell cycle arrest. *Biometals, 17***,** 325-329.

Chandra Mohan, K. V., Devaraj, H., Prathiba, D., Hara, Y. & Nagini, S. (2006). Antiproliferative and apoptosis inducing effect of lactoferrin and black tea polyphenol combination on hamster buccal pouch carcinogenesis. *Biochim Biophys Acta, 1760***,** 1536-1544.

Chen, T., Wong, Y. S. & Zheng, W. (2009). Induction of G1 cell cycle arrest and mitochondria-mediated apoptosis in MCF-7 human breast carcinoma cells by selenium-enriched Spirulina extract. *Biomed Pharmacother*.

Cullen, S. P. & Martin, S. J. (2009). Caspase activation pathways: some recent progress. *Cell Death Differ, 16***,** 935-938.

Damiens, E., El Yazidi, I., Mazurier, J., Duthille, I., Spik, G. & Boilly-Marer, Y. (1999). Lactoferrin inhibits G1 cyclin-dependent kinases during growth arrest of human breast carcinoma cells. *J Cell Biochem, 74***,** 486-498.

Fazi, B., Bursch, W., Fimia, G. M., Nardacci, R., Piacentini, M., Di Sano, F. & Piredda, L. (2008). Fenretinide induces autophagic cell death in caspase-defective breast cancer cells. *Autophagy, 4***,** 435-441.

Fujita, K., Kato, T., Shibayama, K., Imada, H., Yamauchi, M., Yoshimoto, N., Miyachi, E. & Nagata, Y. (2006). Protective effect against 17beta-estradiol on neuronal apoptosis in hippocampus tissue following transient ischemia/recirculation in mongolian gerbils via down-regulation of tissue transglutaminase activity. *Neurochem Res, 31***,** 1059-1068.

Fujita, K., Matsuda, E., Sekine, K., Iigo, M. & Tsuda, H. (2004a). Lactoferrin enhances Fas expression and apoptosis in the colon mucosa of azoxymethane-treated rats. *Carcinogenesis, 25***,** 1961-1966.

Fujita, K., Matsuda, E., Sekine, K., Iigo, M. & Tsuda, H. (2004b). Lactoferrin modifies apoptosis-related gene expression in the colon of the azoxymethane-treated rat. *Cancer Lett, 213***,** 21-29.

Hardivillé, S., Hoedt, E., mariller, C., Benaissa, M. & Pierce, A. (2010). *O*-GlcNAcylation/Phosphorylation cycling at Ser 10 controles both transcriptional activity and stability of delta-lactoferrin. *J Biol Chem*.

Hart, G. W., Haltiwanger, R. S., Holt, G. D. & Kelly, W. G. (1989). Nucleoplasmic and cytoplasmic glycoproteins. *Ciba Found Symp, 145***,** 102-112, discussion 112-108.

Hart, G. W., Housley, M. P. & Slawson, C. (2007). Cycling of O-linked beta-N-acetylglucosamine on nucleocytoplasmic proteins. *Nature, 446***,** 1017-1022.

Hart, G. W., Kreppel, L. K., Comer, F. I., Arnold, C. S., Snow, D. M., Ye, Z., Cheng, X., DellaManna, D., Caine, D. S., Earles, B. J., Akimoto, Y., Cole, R. N. & Hayes, B. K. (1996). O-GlcNAcylation of key nuclear and cytoskeletal proteins: reciprocity with O-phosphorylation and putative roles in protein multimerization. *Glycobiology, 6***,** 711-716.

He, J. & Furmanski, P. (1995). Sequence specificity and transcriptional activation in the binding of lactoferrin to DNA. *Nature, 373***,** 721-724.

Hoedt, E., Hardiville, S., Mariller, C., Elass, E., Perraudin, J. P. & Pierce, A. Discrimination and evaluation of lactoferrin and delta-lactoferrin gene expression levels in cancer cells and under inflammatory stimuli using TaqMan real-time PCR. *Biometals*.

Inoue, S., Browne, G., Melino, G. & Cohen, G. M. (2009). Ordering of caspases in cells undergoing apoptosis by the intrinsic pathway. *Cell Death Differ, 16***,** 1053-1061.

Janicke, R. U., Sprengart, M. L., Wati, M. R. & Porter, A. G. (1998). Caspase-3 is required for DNA fragmentation and morphological changes associated with apoptosis. *J Biol Chem, 273***,** 9357-9360.

Kagawa, S., Gu, J., Honda, T., McDonnell, T. J., Swisher, S. G., Roth, J. A. & Fang, B. (2001). Deficiency of caspase-3 in MCF7 cells blocks Bax-mediated nuclear fragmentation but not cell death. *Clin Cancer Res, 7***,** 1474-1480.

Kelekar, A. & Thompson, C. B. (1998). Bcl-2-family proteins: the role of the BH3 domain in apoptosis. *Trends Cell Biol, 8***,** 324-330.

Kitagawa, K. & Hieter, P. (2001). Evolutionary conservation between budding yeast and human kinetochores. *Nat Rev Mol Cell Biol, 2***,** 678-687.

Lee, D. H., Park, T. & Kim, H. W. (2006). Induction of apoptosis by disturbing mitochondrial-membrane potential and cleaving PARP in Jurkat T cells through treatment with acetoxyscirpenol mycotoxins. *Biol Pharm Bull, 29***,** 648-654.

Li, H., Zhu, H., Xu, C. J. & Yuan, J. (1998). Cleavage of BID by caspase 8 mediates the mitochondrial damage in the Fas pathway of apoptosis. *Cell, 94,* 491-501.

Li, L. Y., Luo, X. & Wang, X. (2001). Endonuclease G is an apoptotic DNase when released from mitochondria. *Nature, 412,* 95-99.

Liu, D., Wang, X., Zhang, Z. & Teng, C. T. (2003). An intronic alternative promoter of the human lactoferrin gene is activated by Ets. *Biochem Biophys Res Commun, 301,* 472-479.

Long, N., Hardiville, S., Pierce, A., Fukamachi, K., Futakuchi, M., Alexander, D. B., Shirai, T. & Tsuda, H. (2009). Construction of a Multi-Functional Helper-Dependent Adenovirus Based System for Cancer Gene Therapy. *Asian Pac J Cancer Prev, 10,* 939-960.

Luzhna, L. & Kovalchuk, O. (2009). Modulation of DNA methylation levels sensitizes doxorubicin-resistant breast adenocarcinoma cells to radiation-induced apoptosis. *Biochem Biophys Res Commun.*

Mariller, C., Benaissa, M., Hardiville, S., Breton, M., Pradelle, G., Mazurier, J. & Pierce, A. (2007). Human delta-lactoferrin is a transcription factor that enhances Skp1 (S-phase kinase-associated protein) gene expression. *FEBS J, 274,* 2038-2053.

Mariller, C., Hardiville, S., Hoedt, E., Benaissa, M., Mazurier, J. & Pierce, A. (2009). Proteomic approach to the identification of novel delta-lactoferrin target genes: Characterization of DcpS, an mRNA scavenger decapping enzyme. *Biochimie, 91,* 109-122.

Masson, P. L., Heremans, J. F. & Schonne, E. (1969). Lactoferrin, an iron-binding protein in neutrophilic leukocytes. *J Exp Med, 130,* 643-658.

Metivier, R., Penot, G., Hubner, M. R., Reid, G., Brand, H., Kos, M. & Gannon, F. (2003). Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell, 115,* 751-763.

Nagaoka, H., Ito, S., Muramatsu, M., Nakata, M. & Honjo, T. (2005). DNA cleavage in immunoglobulin somatic hypermutation depends on de novo protein synthesis but not on uracil DNA glycosylase. *Proc Natl Acad Sci U S A, 102,* 2022-2027.

Nagata, S. (1997). Apoptosis by death factor. *Cell, 88,* 355-365.

Newmeyer, D. D. & Ferguson-Miller, S. (2003). Mitochondria: releasing power for life and unleashing the machineries of death. *Cell, 112,* 481-490.

Nuss, D. L., Furuichi, Y., Koch, G. & Shatkin, A. J. (1975). Detection in HeLa cell extracts of a 7-methyl guanosine specific enzyme activity that cleaves m7GpppNm. *Cell, 6,* 21-27.

Oltvai, Z. N., Milliman, C. L. & Korsmeyer, S. J. (1993). Bcl-2 heterodimerizes in vivo with a conserved homolog, Bax, that accelerates programmed cell death. *Cell, 74,* 609-619.

Panella, T. J., Liu, Y. H., Huang, A. T. & Teng, C. T. (1991). Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes, leukemic cells, and breast cancer. *Cancer Res, 51,* 3037-3043.

Penco, S., Caligo, M. A., Cipollini, G., Bevilacqua, G. & Garre, C. (1999). Lactoferrin expression in human breast cancer. *Cancer Biochem Biophys, 17,* 163-178.

Piva, R., Liu, J., Chiarle, R., Podda, A., Pagano, M. & Inghirami, G. (2002). In vivo interference with Skp1 function leads to genetic instability and neoplastic transformation. *Mol Cell Biol, 22,* 8375-8387.

Sedlak, T. W., Oltvai, Z. N., Yang, E., Wang, K., Boise, L. H., Thompson, C. B. & Korsmeyer, S. J. (1995). Multiple Bcl-2 family members demonstrate selective dimerizations with Bax. *Proc Natl Acad Sci U S A, 92,* 7834-7838.

Shiokawa, D. & Tanuma, S. (2004). Differential DNases are selectively used in neuronal apoptosis depending on the differentiation state. *Cell Death Differ, 11,* 1112-1120.

Siebert, P. D. & Huang, B. C. (1997). Identification of an alternative form of human lactoferrin mRNA that is expressed differentially in normal tissues and tumor-derived cell lines. *Proc Natl Acad Sci U S A, 94***,** 2198-2203.

Son, H. J., Lee, S. H. & Choi, S. Y. (2006). Human lactoferrin controls the level of retinoblastoma protein and its activity. *Biochem Cell Biol, 84***,** 345-350.

Son, K. N., Park, J., Chung, C. K., Chung, D. K., Yu, D. Y., Lee, K. K. & Kim, J. (2002). Human lactoferrin activates transcription of IL-1beta gene in mammalian cells. *Biochem Biophys Res Commun, 290***,** 236-241.

Spagnuolo, P. A., Bird, R. P. & Hoffman-Goetz, L. (2007). Effect of short-term dietary intake of bovine lactoferrin on intestinal lymphocyte apoptosis in healthy mice. *Nutrition, 23***,** 812-817.

Spagnuolo, P. A. & Hoffman-Goetz, L. (2008). Lactoferrin effect on lymphocyte cytokines and apoptosis is independent of exercise. *Med Sci Sports Exerc, 40***,** 1013-1021.

Spik, G., Coddeville, B. & Montreuil, J. (1988). Comparative study of the primary structures of sero-, lacto- and ovotransferrin glycans from different species. *Biochimie, 70***,** 1459-1469.

Teng, C., Gladwell, W., Raphiou, I. & Liu, E. (2004). Methylation and expression of the lactoferrin gene in human tissues and cancer cells. *Biometals, 17***,** 317-323.

Valenti, P., Greco, R., Pitari, G., Rossi, P., Ajello, M., Melino, G. & Antonini, G. (1999). Apoptosis of Caco-2 intestinal cells invaded by Listeria monocytogenes: protective effect of lactoferrin. *Exp Cell Res, 250***,** 197-202.

van Dijk, E., Le Hir, H. & Seraphin, B. (2003). DcpS can act in the 5'-3' mRNA decay pathway in addition to the 3'-5' pathway. *Proc Natl Acad Sci U S A, 100***,** 12081-12086.

Walsh, J. G., Cullen, S. P., Sheridan, C., Luthi, A. U., Gerner, C. & Martin, S. J. (2008). Executioner caspase-3 and caspase-7 are functionally distinct proteases. *Proc Natl Acad Sci U S A, 105***,** 12815-12819.

Xiao, Y., Monitto, C. L., Minhas, K. M. & Sidransky, D. (2004). Lactoferrin down-regulates G1 cyclin-dependent kinases during growth arrest of head and neck cancer cells. *Clin Cancer Res, 10***,** 8683-8686.

Yochum, G. S., McWeeney, S., Rajaraman, V., Cleland, R., Peters, S. & Goodman, R. H. (2007). Serial analysis of chromatin occupancy identifies beta-catenin target genes in colorectal carcinoma cells. *Proc Natl Acad Sci U S A, 104***,** 3324-3329.

Table 1: ΔLfRE sequences found in the *Fas* and *Bax* promoters are compared to the S1 sequence described by He and Furmanski (1995) and to the ΔLfRE found in the *DcpS* and *Skp1* promoters.

Promoter	ΔLfRE										Location[a]	Accession number
	1	2	3	4	5	6	7	8	9	10		
consensus	G	G	C	A	C	T	-	T	A/G	C/G		
S1	G	G	C	A	C	T	-	T	A/G	C		
Skp1	G	G	C	A	C	T	G	T	A	C	-1067 to -1058	AC007199
DcpS	A	G	C	A	C	T	-	T	G	G	-2228 to -2220	NT_033899.7 (Mariller et al., 2009)
Bax	G	G	C	A	C	T	-	T	A	T	-155 to -144	NW_001838497
Fas	G	G	C	A	C	T	-	A	A	C	-1013 to -1004	NT_030059

[a] Location from the transcription start

Legends to figures

Figure 1: ΔLf induces cell death. Cell viability of HEK 293 (A) and MCF7 cells (B) was assayed 24h and 48h post-transfection using Trypan blue. Prior to counting, cells were transfected with pcDNA-ΔLf, pcDNA-ΔLf ^{M4} or empty vector. Data shown as means ± SEM ($n \geq 5$).

Figure 2: DNA fragmentation of cells expressing either ΔLf or its constitutively active mutant. HEK 293 (A) and MCF7 (B) were transfected 48 h before experiment with empty vector or pEGFP as negative control and pcDNA-ΔLf or pcDNA-ΔLf ^{M4} expression vectors. Genomic DNA was extracted from the same quantity of cells for each transfection and submitted to BET agarose gel electrophoresis. Molecular sizes are expressed in bp.

Figure 3: ΔLf induces PARP cleavage. MCF 7 cells were transfected with pcDNA-ΔLf expression vector and lysed at the indicated times post transfection. 20 µg of total protein extract were submitted to SDS-PAGE and immunoblotted with anti-PARP, anti-cleaved PARP (Asp214) or anti-actin antibodies.

Figure 4: ΔLf induces cleavage of caspase 7 (A), caspase 9 (B) and caspase 8 (C). MCF 7 cells were transfected with pcDNA-ΔLf expression vector and lysed at the indicated times post transfection. 20 µg of total protein extract were submitted to SDS-PAGE and immunoblotted with anti-caspase 7, anti-caspase 9, anti-caspase 8 or anti-actin antibodies.

Figure 5: The expression of Bcl family members is modified under ΔLf expression. MCF 7 cells were transfected with pcDNA-ΔLf expression vector and lysed at the indicated times. 20 µg of total protein extract were submitted to SDS-PAGE and immunoblotted with anti-Bax, anti-Bcl-2, anti-Bcl-xl and anti-actin antibodies.

Figure 6: A ΔLfRE is present in the *Bax* promoter. The genomic sequence containing the human *Bax* promoter was retrieved from the Genbank database (NW_001838497). A 1 kbp sequence upstream the transcription start site was searched for the presence of a putative ΔLfRE. The S1 sequence is in bold letters and boxed. The transcription start site is indicated in bold characters (+1).

Figure 7: A ΔLfRE is present in the *Fas* promoter. The genomic sequence containing the human *Fas* promoter was retrieved from the Genbank database (NT_030059). A 1.2 kbp sequence upstream the transcription start site was searched for putative ΔLfRE. The S1 sequence is in bold letters and boxed. The transcription start site is indicated in bold characters (+1).

Figure 8: ΔLf transactivates both the *Bax* (A) and *Fas* promoters (B). HEK 293 cells were co-transfected with pGL3-S1^{Bax}-Luc, pGL3-S1^{DelBax}-Luc, pGL3-S1^{Fas}-Luc or pGL3-S1^{DelFas}-Luc vector and with null plasmid or pcDNA-ΔLf expression vector. Cells were lysed 24 h after transfection. Samples were assayed for protein content and luciferase activity. The relative luciferase activity is expressed as a ratio of the luciferase activity to the protein content. Values represent the mean ± SD of triplicates from three independent measurements.

Figure 9: ΔLf binds to the *Bax* promoter *in vivo*. HEK 293 cells were transfected with the p3X FLAG-CMV-10-ΔLf. 24h after transfection, a ChIP assay was performed. ChIP product was then amplified by real time PCR using specific primer pairs targeting the ΔLf RE containing fragment of the *Bax* promoter. Total ΔLfRE of the *Bax* promoter corresponds to input. ChIP assays were performed using an anti-FLAG (M2) and anti-rabbit IgG as nonspecific antibody control (IR). As a further control, the assay was performed without binding of an antibody to the protein G plus Sepharose (NIP). The results are expressed as a percentage, where 100 % represent the signal obtained for the Q-PCR product of the input, and are the mean ± SD of triplicates from three independent assays.

Figure 1

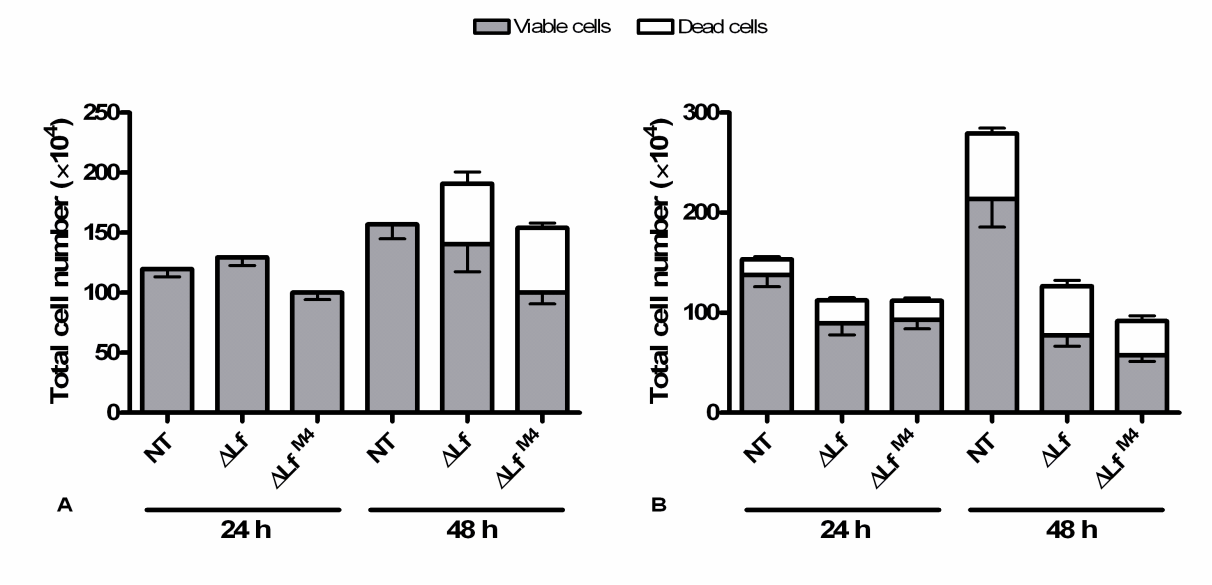


Figure 2

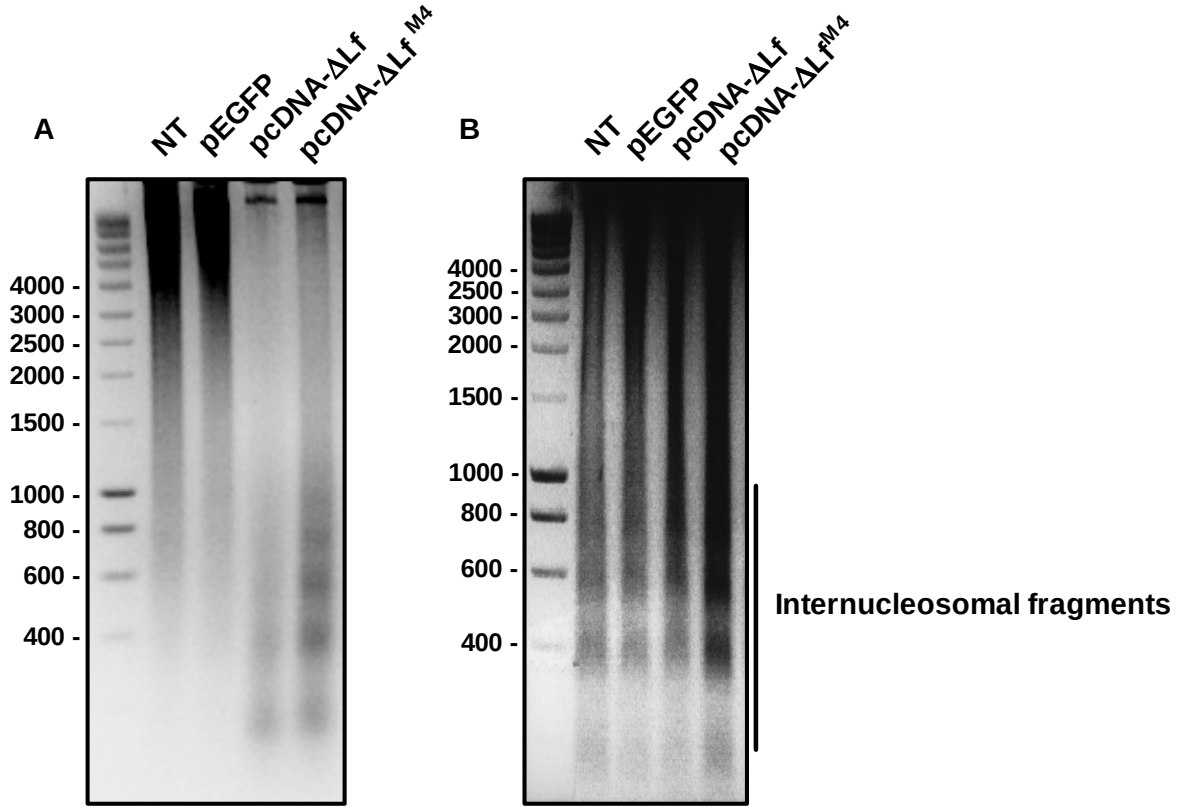


Figure 3

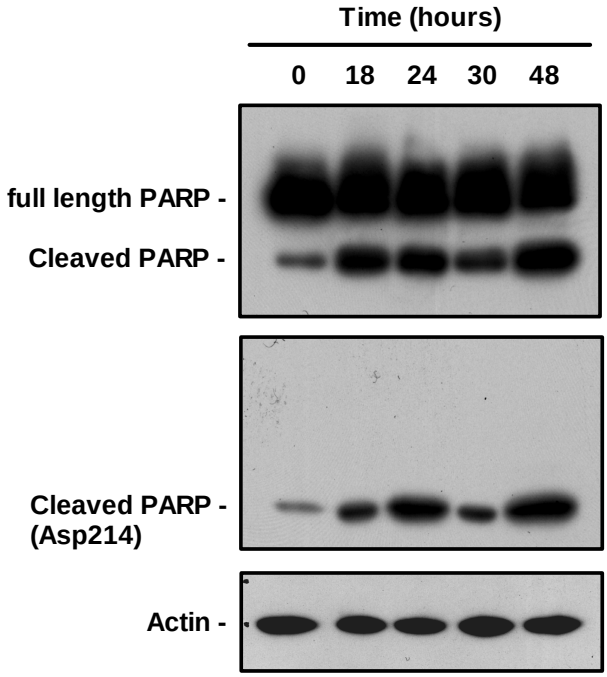


Figure 4

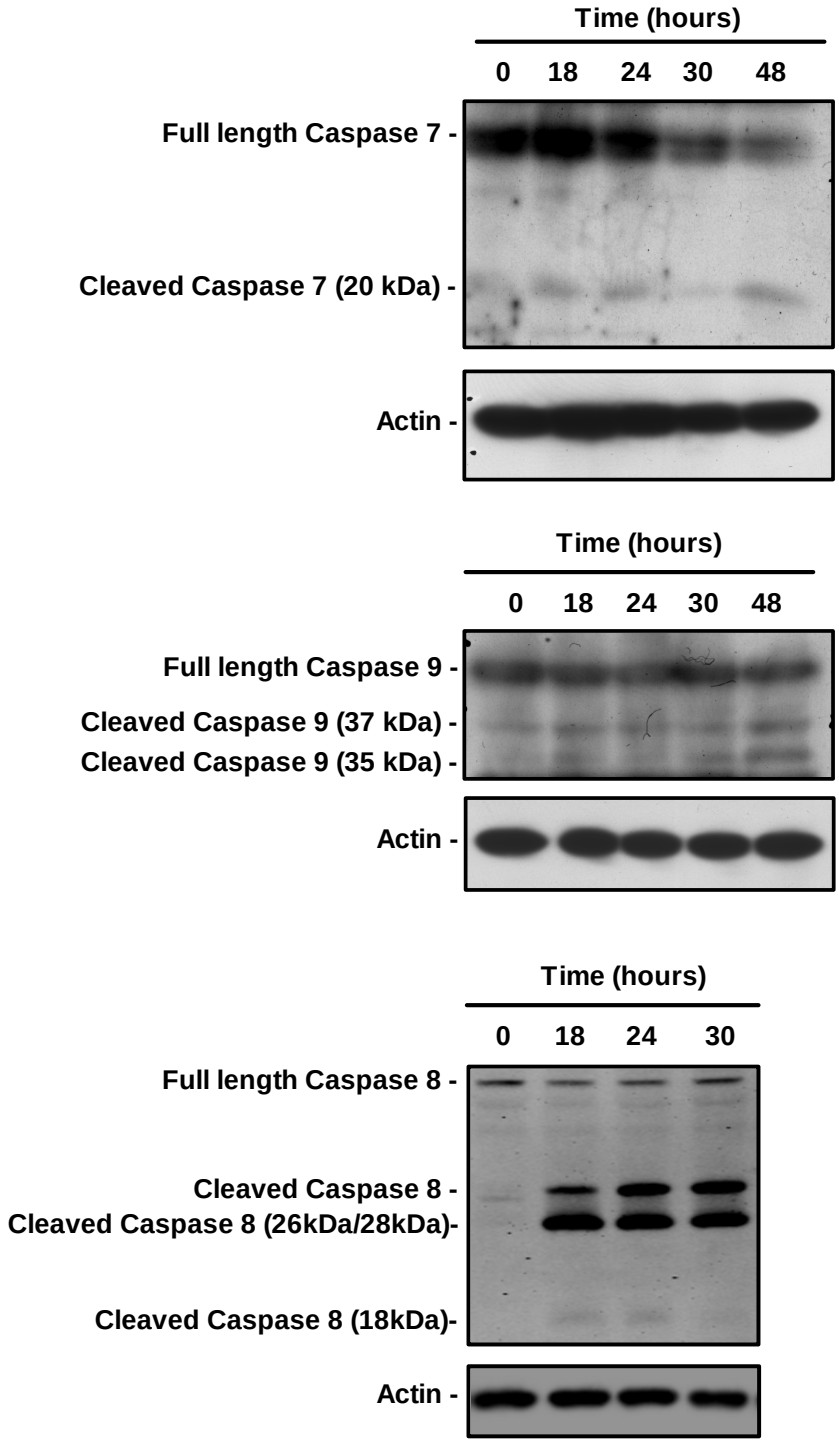


Figure 5

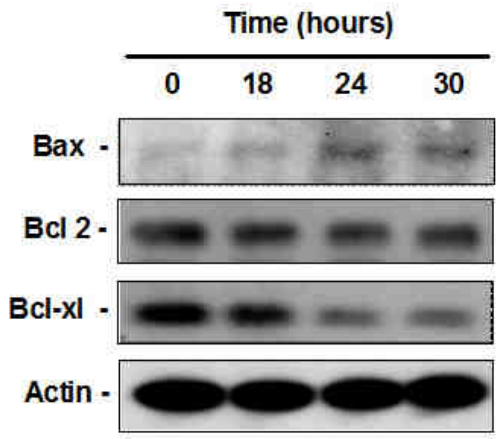


Figure 6

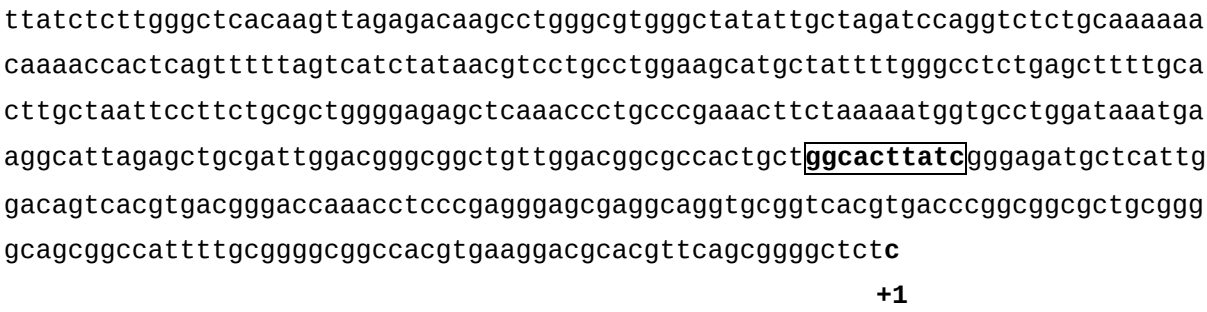


Figure 7

Figure 8

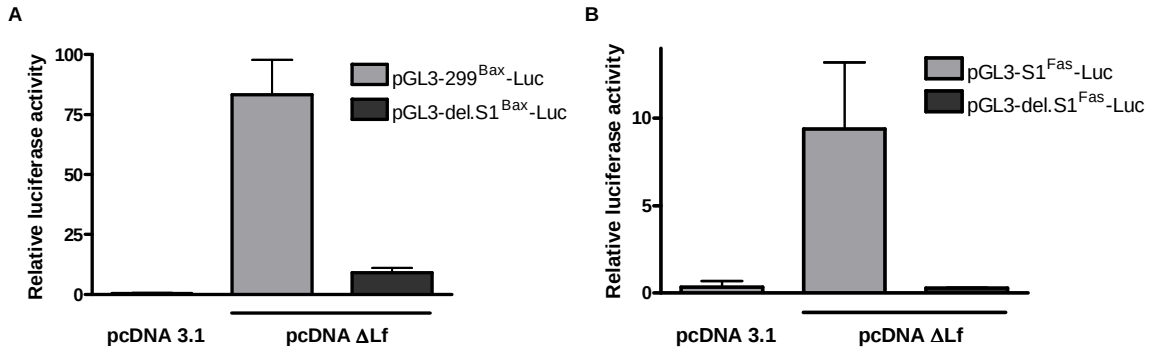
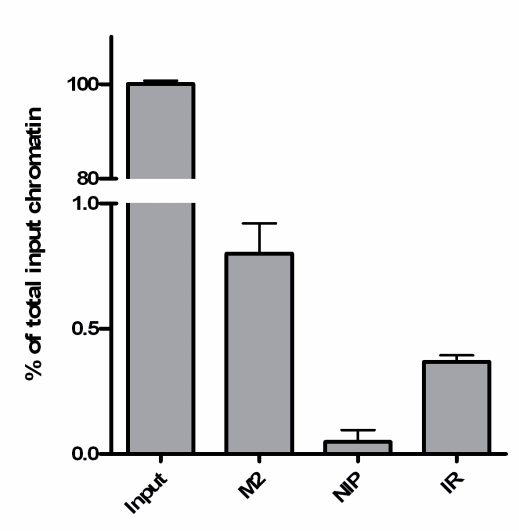


Figure 9



III. *DcpS* et la dégradation des ARNm

Afin de mettre en évidence de nouvelles cibles de la delta-lactoferrine, nous avons entrepris une étude comparative des protéomes des cellules exprimant ou non la delta-lactoferrine. En parallèle, nous avons regardé l'impact de l'expression d'une isoforme constitutivement activée de la delta-lactoferrine appelé mutant M4.

L'ensemble des résultats de cette étude a fait l'objet de la publication suivante :

Mariller, C., Hardivillé, S., Hoedt, E., Benaïssa, M., Mazurier, J., et Pierce, A. « Proteomic approach to the identification of novel delta-lactoferrin target genes : characterization of DcpS, an mRNA scavenger decapping enzyme», 2009, *Biochimie* 91(01) : 109-22

Au cours de cette étude nous avons comparé les variations du protéome des cellules HEK 293 soumis à quatre conditions différentes : non traitées, transfectées avec un vecteur d'expression vide, transfectées avec le vecteur d'expression de la delta-lactoferrine ou celui du mutant M4. L'électrophorèse bidimensionnelle des lysats cellulaires a permis de mettre en évidence après coloration au nitrate d'argent, environ 500 protéines dont 214 ont pu être retrouvées dans les quatre conditions. L'analyse informatique du signal des 214 protéines a révélé une surexpression significativement différente de 12 d'entre elles lorsque la delta-lactoferrine ou son isoforme M4 était exprimée. Parmi ces 12 protéines, 7 présentent une augmentation de quantité supérieure lorsque le mutant M4 est exprimé par rapport à l'isoforme sauvage confirmant le caractère constitutivement activé de cette isoforme.

L'analyse des digestions trypsiques en MALDI-TOF ou en LC-MS/MS a permis d'identifier 8 protéines parmi celles surexprimées. Il s'agit des protéines *(1)* PSB3 qui participe à l'architecture du protéasome (Tanahashi *et al.*, 1993), *(2)* CCT2 qui est une protéine chaperonne permettant la bonne conformation des protéines (Abe *et al.*, 2009; Zebol *et al.*, 2009), *(3)* PDIA3, une disulfide isomérase qui est impliquée dans la formation des ponts disulfure (Jensen, 2007), *(4)* NSF qui est une ATPase participant au désassemblage des complexes SNARE (Soluble N-ethylmaleimide-sensitive fusion attachment protein receptor) (Littleton *et al.*, 2001), *(5)* PRP19 et *(6)* HNRPL qui sont impliquées dans la maturation des ARNm (Lee *et al.*, 2009; Makarova *et al.*, 2004; Soderberg *et al.*, 2007), *(7)* PCBP2 qui permet la traduction des ARNm à partir de séquences IRES (internal ribosomal entry sequence) (Emerald *et al.*, 2007; Sean *et al.*, 2009), et de *(8)* DcpS qui intervient dans la dégradation des ARNm (Garneau *et al.*, 2007). Cependant, seules DcpS et CCT2 présentent

une augmentation des quantités de leurs transcrits d'un facteur deux par rapport à la condition contrôle. L'analyse de leurs promoteurs a permis de mettre en évidence un élément de réponse à la delta-lactoferrine dans le promoteur du gène *DcpS*. Nous avons ensuite démontré que cet élément de réponse est fonctionnel, confirmant que DcpS est une nouvelle cible de la delta-lactoferrine.

DcpS est l'enzyme intervenant à la fin de la voie de dégradation des ARNm. Les processus de synthèse, de maturation et de dégradation des ARN sont des mécanismes hautement conservés par l'évolution et indispensables au maintien de la vie cellulaire. La dégradation des ces mêmes ARN occupe une place primordiale dans le contrôle de l'expression génique, le « contrôle qualité » des ARNm et la défense contre les infections virales. Chez les eucaryotes, la dégradation des ARNm se fait par deux voies principales débutantes toutes deux par l'hydrolyse de la coiffe 7 méthylguanosine ou 5'-cap (Figure 25, page 117) (Liu et Kiledjian, 2006). Ce nucléotide particulier, ajouté de manière post-transcriptionnelle sur les ARNm, permet leur protection face aux ribonucléases et favorise leur traduction *via* le recrutement du ribosome. La dégradation récurrente des ARNm dans le sens 3'-5' par l'exosome libère la coiffe sous la forme du dinucléotide m7GpppN. La dégradation dans le sens 5'-3' fait intervenir une première enzyme, Dcp2, qui hydrolyse la coiffe sous la forme m7GDP et libère l'extrémité 5'-phosphate des ARNm. Il s'en suit alors une dégradation des ARNm de 5' en 3' effectuée par la ribonucléase Xrn1 (Coller et Parker, 2004; Liu et Kiledjian, 2006). Les nucléotides m7GpppN et m7GDP sont alors pris en charge par DcpS et libérés sous la forme de m7GMP qui sera recyclé (van Dijk *et al.*, 2003).

DcpS pourrait, ainsi, être impliquée indirectement dans la régulation de l'épissage ainsi que dans la traduction des ARNm. En effet, ces deux mécanismes nécessitent respectivement l'interaction du complexe CBC (Cap-binding complex) (Lewis et Izaurralde, 1997) et d'eIF4E (eukaryotic translation initiation factor 4E) (Gingras *et al.*, 1999) avec la coiffe des ARNm. Les coiffes libérées lors de la dégradation des ARNm, non converties en 7mGMP par DcpS, entrent en compétition avec l'extrémité coiffée des ARNm pour lier les complexes CBC et eIF4E, interférant ainsi avec les mécanismes d'épissage et de traduction (Bail et Kiledjian, 2008). La surexpression de DcpS permettrait alors de diminuer les quantités de 7mGpppN et 7mGDP et, ainsi, d'augmenter l'efficacité de l'épissage et la traduction des ARNm.

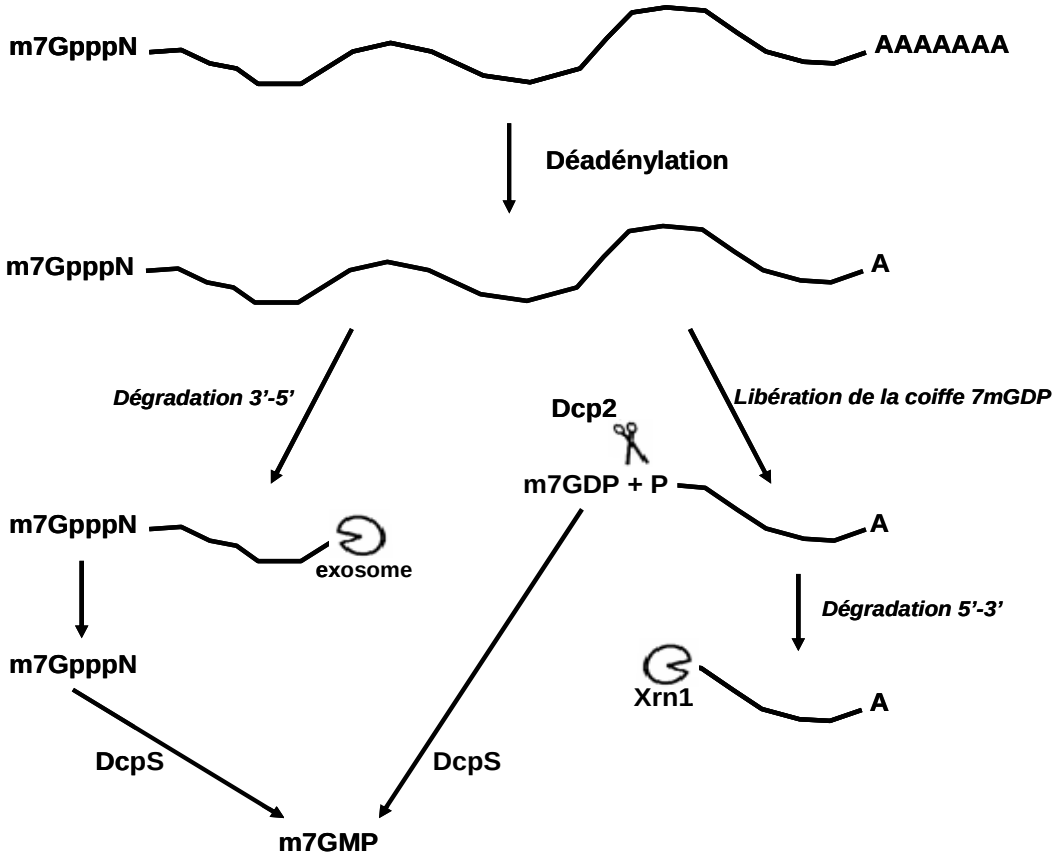


Figure 25 : Modèle de la dégradation des ARNm. Après la déadénylation, les ARNm sont dégradés par l'exosome de 3' en 5' libérant ainsi du m7GpppN. L'autre voie de dégradation fait intervenir Dcp2 pour l'hydrolyse de la coiffe libérant du 7mGDP. La dégradation des ARNm de 5' en 3' peut alors être effectuée par Xrn1. Les deux produits de la dégradation issus de la coiffe des ARNm sont alors convertis par DcpS en m7GMP qui peut alors être recyclé ou éliminé de la cellule. (Adapté de van Dijk et al., 2003).

Available online at www.sciencedirect.com
ScienceDirect
Biochimie 91 (2009) 109–122
BIOCHIMIE
www.elsevier.com/locate/biochi

Research paper

Proteomic approach to the identification of novel delta-lactoferrin target genes: Characterization of DcpS, an mRNA scavenger decapping enzyme

Christophe Mariller*, Stephan Hardivillé, Esthelle Hoedt, Monique Benaïssa, Joël Mazurier, Annick Pierce

Université des Sciences et Technologies de Lille, CNRS, UMR 8576, Unité de Glycobiologie Structurale et Fonctionnelle, F-59655 Villeneuve d'Ascq cedex, France

Received 18 March 2008; accepted 30 July 2008
Available online 7 August 2008

Abstract

The expression of the transcription factor ΔLf is deregulated in cancer cells. Its overexpression provokes cell cycle arrest along with antiproliferative effects and we recently showed that the *Skp1* gene promoter was a target of ΔLf. Skp1 belongs to the Skp1/Cullin-1/F-box ubiquitin ligase complex responsible for the ubiquitination and the proteosomal degradation of numerous cellular regulators. The transcriptional activity of ΔLf is highly controlled and negatively regulated by *O*-GlcNAc, a dynamic post-translational modification known to regulate the functions of many intracellular proteins. We, therefore, constructed a ΔLf-M4 mutant corresponding to a constitutively active ΔLf isoform in which all the glycosylation sites were mutated. In order to discover novel targets of ΔLf transcriptional activity and to investigate the impact of the *O*-GlcNAc regulation on this activity *in situ* we compared the proteome profiles of ΔLf- and ΔLf-M4-expressing HEK293 cells versus null plasmid transfected cells. A total of 14 differentially expressed proteins were visualized by 2D electrophoresis and silver staining and eight proteins were identified by mass spectrometry analyses (MALDI-TOF; LC-MS/MS), all of which were upregulated. The identified proteins are involved in several processes such as mRNA maturation and stability, cell viability, proteasomal degradation, protein and mRNA quality control. Among these proteins, only DcpS and TCPB were also upregulated at the mRNA level. Analysis of their respective promoters led to the detection of a *cis*-regulating element in the *DcpS* promoter. The $S1^{DcpS}$ is 80% identical to the S1 sequence previously described by He and Furmanski [Sequence specificity and transcriptional activation in the binding of lactoferrin to DNA, Nature 373 (1995) 721–724]. Reporter gene analyses and ChIP assays demonstrated that ΔLf interacts specifically with the *DcpS* promoter *in vivo*. These data established that DcpS, a key enzyme in mRNA decay, is a new target of ΔLf transcriptional activity.

Keywords: Delta-lactoferrin; Transcription factor; Scavenger decapping enzyme DcpS; mRNA turnover; *O*-GlcNAc/P interplay; Proteomics

1. Introduction

The process of mRNA turnover is a critical mechanism for the regulation of gene expression, quality control of mRNA biogenesis and antiviral defenses. The major mRNA degradation pathways involve shortening the poly(A) tail, exonucleolytic decay and decapping (reviewed in refs. [1,2]). Among the decapping enzymes, DcpS, also known as the scavenger decapping enzyme, is involved in cap nucleotide metabolism [3,4]. Here, we show that delta-lactoferrin (ΔLf), a transcription factor involved in the regulation of cell cycle progression at the G1/S transition, enhances *DcpS* transcription.

First discovered as a transcript in normal tissues [5], ΔLf was found to be downregulated in cancer cells and in breast cancer biopsies [6]. Its expression level was of good prognosis value in human breast cancer with high concentrations

Abbreviations: ΔLf, delta-lactoferrin; DcpS, scavenger decapping enzyme; hDcp2, human decapping enzyme 2; Skp1, S-phase kinase 1; RPLP0, ribosomal protein, large, P0; ΔLfRE, ΔLf response element; NLS, nuclear localization signal; *O*-GlcNAc/P, *O*-glycosylation/phosphorylation.

* Corresponding author. Tel.: +33 3 2033 7238; fax: +33 3 2043 6555.
E-mail address: christophe.mariller@univ-lille1.fr (C. Mariller).

0300-9084/$ - see front matter © 2008 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.biochi.2008.07.009

being associated with longer relapse-free and overall survival [6]. ΔLf is a lactoferrin isoform, the transcription of which starts at the alternative promoter P2 present in the first intron of the Lf gene [7]. The comparison of the two enhancer/promoter regions revealed that this gene is differentially transactivated [5,7]. The deregulation of Lf gene expression that occurs in tumors is mainly due to genetic and epigenetic changes [8–11]. The alternative selection of promoters produces an alternative N-terminal domain. Thus, compared to Lf, ΔLf is a protein devoid of the 45 first amino acid residues including the leader sequence, implying that it is a 73 kDa cytoplasmic isoform [12]. However, ΔLf was also observed in the nucleus [12,13] and a short bipartite nuclear localization signal at the C-terminus conserved in Lfs from different species, has been identified [14]. Recently, we showed that ΔLf is an efficient transcription factor interacting *in vivo* with a ΔLf response element (ΔLfRE) found in the *Skp1* promoter [14]. This specific GGCACTTGC sequence had previously been described [15] and found to be responsible for IL-1β transactivation by Lf [16]. Studies of the three-dimensional (3D) Lf structure indicated two putative DNA-binding domains (DBDs) located either at the N-terminus (residues 27–30 in Lf and 2–5 in ΔLf) and/or at the interlobe region [17,18].

ΔLf expression provokes antiproliferative effects, cell cycle arrest in S phase [12] and Skp1 upregulation [14]. At the G1-S transition, Skp1 (S phase kinase associated protein) belongs to the SCF (Skp1/Cullin-1/F-box ubiquitin ligase) complex responsible for the ubiquitination of cellular regulators such as cyclins and cyclin-dependent kinase (CDK) inhibitors leading to their degradation by the proteasome [19,20]. At the G2/M transition Skp1 belongs to the CBF3 complex involved in the preservation of genetic stability [21,22]. By upregulating Skp1 gene expression, ΔLf may survey cell cycle progression *via* the control of the proteasomal degradation of S phase actors. Thus, ΔLf transcriptional activity should be strongly controlled. The presence of putative *O*-*N*-acetylglucosaminylation (*O*-GlcNAc) sites (YinOYang 1.2 server, http://www.cbs.dtu.dk/services/YinOYang/) could imply a control of the transcriptional activity or the half-life of ΔLf *via* the balance between *O*-GlcNAc and phosphorylation as already described for other factors [23]. Four *O*-GlcNAc sites are present (Ser10, Ser227, Ser472 and Thr559) the mutation of which produces the constitutively active ΔLf-M4 mutant with a 2.5-fold increased transcriptional activity compared to wild type.

In order to identify factors that are differentially expressed in response to ΔLf or the ΔLf-M4 mutant isoform, we have undertaken a differential proteomic approach using 2D gel electrophoresis combined with mass spectrometry. Among the eight differentially expressed proteins described here, we identified DcpS as a new ΔLf target gene. DcpS is a member of the HIT family of pyrophosphatases which performs catalysis of the 5′ cap structure [24]. This cap, that has to be removed during mRNA decay, is involved in a variety of functions such as pre-mRNA splicing, export, stability and efficient translation [25–27]. Eukaryotic mRNA degradation proceeds through two main pathways. In the 3′–5′ mRNA decay pathway, degradation generates free m7GpppN that is hydrolyzed by DcpS, the scavenger decapping enzyme, generating m7GMP. In the 5′–3′ pathway, the cleavage of the cap of deadenylated mRNAs is performed by the hDcp2 decapping enzyme producing 5′-phosphorylated mRNA and m7GDP which is then converted to m7GMP by DcpS [3,24,28–30].

Our findings showed that ΔLf and its constitutively active mutant modulate the expression of proteins involved in the cell cycle, cell survival and mRNA turnover. Among them, DcpS was shown to be a new target of ΔLf transcriptional activity.

2. Materials and methods

2.1. Cell culture

Human HEK 293 cells (ATC CRL-1573) were kindly provided by Dr. J.-C. Dhalluin (INSERM U 524, Lille, France). Human cervical cancer HeLa cells (ATCC CCL-2) were a kind gift from Dr. T. Lefebvre (UGSF, UM5 8576 CNRS, Villeneuve d'Ascq, France). The Breast Cancer MDA-MB231 (ATCC HTB-26) cell line was kindly provided by M. Mareel (Laboratory of Experimental Cancerology, University Hospital, Ghent, Belgium). Cells were routinely grown in monolayers as previously described [12,14]. MDA-MB-231 cells stably transfected with ΔLf (MDA-MB-231-ΔLf) were produced as in [12,14]. Expression of ΔLf is induced by doxycycline (2 μg/ml) [12] in these ΔLf-expressing cell lines. Cell culture materials were obtained from Dutscher (France), and culture media and additives from Cambrex Corporation (NJ, USA) and Invitrogen (UK).

2.2. DNA and RNA isolation

Genomic DNA was extracted from HEK 293 cells as previously described [33] and purified using QIAprep Spin Miniprep Kit (Qiagen, Germany). Total RNA was extracted from cell cultures using the RNeasy Mini Kit (Qiagen) according to the manufacturer's specifications. The purity of the extracts was checked by measuring the ratio of the absorbance at 260 nm and 280 nm using a nanodrop ND-1000 spectrophotometer (Labtech International, UK) and their integrity was visualized on a BET-agarose gel.

2.3. Transfection

Transfections were performed using the Dreamfect® reagent (OZ Biosciences, France), according to the manufacturer's instructions. After incubation for 24 h, cells were washed with NaCl/Pi. They were then lysed in appropriate buffer, either for total RNA preparation or for protein extracts. Protease inhibitor (Pefablock SC and Complete, Roche, Switzerland) was added to protein extracts. Except for transfections dedicated to the proteomic analysis, each transfection

was repeated at least three times and all cells were transfected in triplicate.

2.4. Plasmid construction

pGL3-S1^{Skp1}-Luc, pcDNA-ΔLf and p3XFLAG-CMV10-ΔLf were constructed as described in [Mariller et al., 2007]. pGL3-S1^{DcpS}-Luc was obtained as in ref. [14] except that the 136 bp *DcpS* promoter fragment (GenBank accession number: NT_033899.7) was amplified with the primer pair listed in Table 1 and cloned into the pcDNA3.1 vector (Clontech, CA, USA), sequenced and transferred using KpnI–XhoI to the pGL3-promoter-Luc vector (Promega).

2.5. Site-directed mutagenesis

Mutants were generated using the QuikChange® Site-directed Mutagenesis Kit (Stratagene, Garden Grove, CA), according to the manufacturer's instructions. The oligonucleotides used are listed in Table 1. The pcDNA-ΔLf-M4 was generated using pcDNA-ΔLf as template. The four mutations were carried out sequentially with Ser10, Ser27, Ser472 and Thr559 being replaced by alanine residues. Following sequence verification, positive clones were used directly in transfection.

2.6. Western blotting and immunodetection

Cell extracts were prepared from frozen pellets of HEK cells transfected with p3XFLAG-CMV-10-ΔLf, p3XFLAF-CMV-10-ΔLf-M4, pcDNA, pcDNA-ΔLf or pcDNA-ΔLf-M4 vector. Proteins were extracted in RIPA buffer [50 mM Tris–HCl (pH 7.4) containing 1% (vol/vol) NP-40 (Roche), 1% (wt/wt) Na-deoxycholate (Sigma-Aldrich, MO, USA), 0.1% (wt/wt) SDS, 0.15 M NaCl and Complete® (according to manufacturer's instructions, Roche)] for 20 min on ice. Cell debris were removed by centrifugation for 10 min at 12,000 × *g*. Protein concentration was determined using the BCA assay (Pierce, MA, USA). Samples were mixed with 4× Laemmli sample buffer [250 mM Tris–HCl (pH 6.8) containing 20% (vol/vol) β-mercaptoethanol, 6% (wt/wt) SDS, 40% (vol/vol) and 0.04% (wt/wt) bromophenol blue] and boiled for 5 min. A total of 30 μg of protein in each sample was submitted to SDS-PAGE and analyzed by Western blotting. Blots were subsequently probed with primary antibodies (monoclonal murine anti-FLAG M2 antibody, 1:2000; monoclonal murine anti-DcpS antibody, 1:200 and polyclonal rabbit anti-actin antibody, 1:1000) for 2 h at room temperature and secondary antibodies conjugated to peroxidase, before being detected by chemiluminescence (ECL+, GE Healthcare Life Sciences, UK) on a ChemiGenius 2 (Syngene, UK) imaging system. Primary antibodies against DcpS were purchased from Abcam (UK), against actin from Santa Cruz Biotechnologies Inc. (CA, USA), anti-FLAG M2 from Sigma-Aldrich and secondary antibodies conjugated to horseradish peroxidase from GE Healthcare Life Sciences. The densitometric analysis was performed using the Quantity One v4.1 software (Bio-Rad, CA, USA).

2.7. Sample preparation for proteomic analysis

After medium removal, 2×10^6 cells were rinsed three times in ice-cold Tris-sucrose [50 mM Tris–HCl (pH 7.4) containing 8.5% (wt/vol) sucrose]. Each pellet was lysed on ice with lysis buffer [50 mM Tris–HCl (pH 7.4) containing 3% (wt/vol) SDS and 0.3% (wt/vol) dithiotreitol (DTT, Bio-Rad)] for 15 min and then heated at 100 °C for 5 min. The concentration of proteins in the samples was assayed using the BCA Protein Assay (Pierce). Samples corresponding to 200 μg or 500 μg were precipitated with three volumes of acetone for 60 min at −20 °C and centrifuged at 10,000 × *g* for 15 min at +4 °C. The pellets were resuspended in 200 μl of resolubilization buffer [Tris 0.04 M, DTT 0.01 M, urea 8 M, 4% (wt/vol) CHAPS] and 300 μl of rehydration buffer [Tris 0.04 M, DTT 0.01 M, urea 8 M, 4% (wt/vol) CHAPS, 0.7% (vol/vol) ampholytes (Bio-Rad), 0.1% (wt/vol) bromophenol blue].

2.8. 2D gel analysis

Proteins were separated by 2D gel electrophoresis. In the first dimension, isoelectric focusing (IEF), samples prepared as above (200 μg for analytical gels or 500 μg for preparative gels) were applied onto 18 cm IPG strips (Bio-Rad) with a linear pH gradient from 3 to 10 prior to rehydration for at least 12 h at 4 °C in a reswelling tray (Bio-Rad). After complete sample uptake onto the strip, IEF was performed using a Protean IEF Cell System (Bio-Rad) at room temperature with a current limit of 50 μA/strip. An in-gel incorporation step under low voltage (50 V for 6 h) was used to enhance protein uptake. Using rapid ramping, the voltage was increased to 250 V for 1 h, 1000 V for 2 h and 8000 V for 10 h.

For SDS-PAGE, IPG strips were incubated for 15 min in equilibration buffer [50 mM Tris–HCl (pH 8.6) containing 6 M urea, 1% (wt/vol) SDS, 65 mM DTT, 30% (vol/vol) glycerol] followed by a 20 min incubation in equilibration buffer supplemented with 50 mM iodoacetamide. The equilibrated strips were applied on 10% acrylamide gels (20 × 18.5 cm) and proteins were resolved using a Protean II XL system (Bio-Rad) in the Laemmli buffer system. Electrophoresis was carried out at 60 V for 1 h and 120 V for 12–16 h. Gels for analytical purposes were silver-stained as described previously by [31]. The preparative gels were stained as described by ref. [32].

2.9. Image acquisition and data analyses

2D gels were digitized with a precision of 42.3 μm (600 dpi) using a GS-800 Bio-Rad calibrated densitometer and the Quantity One v4.1.0 software (Bio-Rad). Differential analysis was performed using ImageMaster 2D Platinum v6 (Genebio, Switzerland). Four classes of gels were defined

Table 1
Oligonucleotides used for RT-PCR, qRT-PCR, ChIP, mutagenesis and plasmid construction

Method	Oligonucleotide (5′–3′)	Tm (°C)	Cycle number	Amplicon size (pb)
RT-PCR				
DcpS	S: 5′-TACCTGTGCTCCTGTCC-3′ F: 5′-TGTAGTCATCTCCCGTCT-3′	60	25	371
hnRNPL	S: 5′-GGCTCAAGACTGACAACG-3′ F: 5′-TGTAAAGAACATCCGTGGT-3′	60	35	159
PDIA3	S: 5′-TCCAACCCTGAAGATA-3′ F: 5′-TGAAGGACGAAATAAGA-3′	50	35	325
Prp19	S: 5′-TCCCTAATCTGCTCCAT-3′ F: 5′-CGCACCCACAACACTT-3′	60	30	464
NSF	S: 5′-ATTGACTCCAACCCT-3′ F: 5′-CATGTTTACAACCCATC-3′	50	35	435
PCBP2	S: 5′-CGGAAAGAAAGGAGAA-3′ F: 5′-GTTGAGTTGGGTAGCATA-3′	50	30	333
TCPB	S: 5′-CTTTAAGGCAGGAGCT-3′ F: 5′-AACGGTAACAGAGGTAGTG-3′	50	35	280
PSB3	S: 5′-TCAGGTTCTCCTCGTGC-3′ F: 5′-GTTTAATGGCTTTCTTGCT-3′	45	40	393
RPLP0	S: 5′-GATGACCAGCCCAAAGGAGA-3′ F: 5′-GTGATGTGCAGCTGATCAAGACT-3′	55	22	101
qRT-PCR				
DcpS	S: 5′-AGACGGGAGATGACTACAGG-3′ F: 5′-GTGGGTGTATAACATTCTCGAC-3′	55	40	86
RPLP0	S: 5′-GATGACCAGCCCAAAGGAGA-3′ F: 5′-GTGATGTGCAGCTGATCAAGACT-3′	55	40	101
ChIP				
DcpS promoter	S: 5′-CCTATCCCCAGGCTCTTTCAC-3′ F: 5′-GGTCTCAATTTCCTGACCTCG-3′	60	36	136
Albumin promoter	S: 5′-GTGGCTCTGATTGGCTTTCTG-3′ F: 5′-ACTCATGGGAGCTGCTGGTTC-3′	60	36	139
Site-directed mutagenesis				
ΔLf-M4:	Ser10: S: 5′-CGTGGCCCTCCTGTCGCCTGCATAAAGAGAGA-3′ F: 5′-TCTCTCTTTATGCAGGCGACAGGAGGGCCACG-3′ Ser227: S: 5′-CCCGGGTCCCTGCTCATGCCGTTG-3′ F: 5′-CAACGGCATGAGCAGGGACCCGGG-3′ Ser472: S: 5′-GTGCCCCTGGGGCTGACCCGAGAT-3′ F: 5′-ATCTCGGGTCAGCCCCAGGGGCAC-3′ Thr559: S: 5′-ACGGAAGCCTGTGGCTGAGGCTAGAAGC-3′ F: 5′-GCTTCTAGCCTCAGCCACAGGCTTCCGT-3′		manufacturer's instructions	
Plasmid construction				
DcpS	S: 5′-CCTATCCCCAGGCTCTTTCAC-3′ F: 5′-GGTCTCAATTTCCTGACCTCG-3′	60	36	136

corresponding to non-transfected cells, cells transfected with the pcDNA3.1 plasmid alone, pcDNA-ΔLf expression vector or pcDNA-ΔLf-M4 expression vector respectively, namely NT class, pcDNA class, pcDNA-ΔLf class and pcDNA-ΔLf-M4 class. Scatter plots were used to discard eventual experimental disparities in stain intensities or sample loading. Spot detection was realized independently for each gel with a visual check to avoid false negatives or positives. In order to optimize spot detection and matching, a master gel was designated in each class to ascribe to it groups which consist of spots that are present on at least three independent gels of the same class. Automatic matching was realized using the four master gels. The quantification of expression was expressed as percent volume % vol = ($vol_{S(n)}/\sum$vol of all spots resolved in the gel) where $vol_{S(n)}$ is the volume of spot S in a gel containing n spots. Inter-class differences in protein expression were analyzed using the Kolmogorov–Smirnov statistical test with the median as the central tendency and the mean absolute deviation as the dispersion to allow a more robust analysis of the outliers.

2.10. Protein identification and mass spectrometry

For identification of proteins by mass fingerprinting analysis, spots from preparative gels were cut from the gel and washed three times with 500 µl of a 50 mM ammonium

carbonate/acetonitrile 1:1 (vol/vol) solution. Polyacrylamide fragments were dehydrated three times in acetonitrile and dried with a vacuum concentrator (Concentrator 5301, Eppendorf, Germany). Protein reduction and alkylation were performed by reswelling polyacrylamide fragments in 50 mM ammonium carbonate solution containing 0.15 mg/ml DTT at 56 °C. This solution was replaced by a 50 mM ammonium carbonate solution containing 10 mg/ml 2-iodoacetamide (Bio-Rad) for 45 min at room temperature and in the dark. Fragments were dried as previously and tryptic cleavage was initiated by reswelling the gel in 50 µl of a 50 mM ammonium carbonate solution containing 20 µg/ml of trypsin (Promega, WI, USA) for 20 min. The solution was then replaced by 40 µl of 50 mM ammonium carbonate solution and digestion carried out overnight at 30 °C. Tryptic peptides were extracted at 30 °C for 45 min with a 45% (vol/vol) acetonitrile and 10% (vol/vol) acetic acid solution. A second extraction was performed at room temperature for 15 min with a 45% (vol/vol) acetonitrile and 5% (vol/vol) acetic acid solution. Both eluates were pooled, dried under vacuum and reconstituted with 10 µl of a 0.01% (vol/vol) trifluoroacetic acid solution.

Protein identification by mass spectrometry was performed by MALDI-TOF and by LC-MS/MS. For MALDI-TOF analyses, 0.5 µl of reconstituted extract was mixed directly onto the target with 1 µl of a matrix solution [CH_3OH/H_2O 7:3 (vol/vol) containing 10 mg/ml 2,5-dihydroxybenzoic acid] and air-dried. Mass spectra were recorded in positive ion mode at an accelerating voltage of 21 kV with a 200 ns extraction delay and a 750–3500 mass gate on a Voyager-DE™ STR Biospectrometry Workstation (Applied Biosystems, MA, USA). Four hundred shots were accumulated to obtain a final spectrum. Internal calibration was performed using the average mass of the two autolytic trypsin fragment ions at m/z 843.014 and 2212.425. Nano-LC-nano-ESI-MS/MS analysis of the trypsin digests was performed on an ion trap mass spectrometer (Thermo Fisher Scientific, MA, USA) equipped with a nanoelectrospray ion source coupled with a nano-high pressure liquid chromatography system (LC Packings Dionex, CA, USA). Tryptic digests were resuspended in 10 µl of 0.1% HCOOH, and 1 µl was injected into the mass spectrometer using a Famos autosampler (LC Packings Dionex). The samples were first desalted and then concentrated on a reverse phase precolumn of 5 mm × 0.3 mm inner diameter (Dionex) by solvent A (H_2O/acetonitrile 95:5 (vol/vol) containing 0.1% (vol/vol) HCOOH) delivered by the Switchos pumping device (LC Packings Dionex) at a flow rate of 10 µl/min for 3 min. Peptides were separated on a 15 cm × 75 µm-inner diameter C_{18} PepMap column (Dionex). The flow rate was set at 200 nl/min. Peptides were eluted using a 5–100% linear gradient of solvent B (H_2O/acetonitrile 20:80 (vol/vol) containing 0.08% (vol/vol) HCOOH) for 45 min. Coated nanoelectrospray needles were obtained from New Objective (MA, USA). Spray voltage was set at 1.5 kV, and the capillary temperature was set at 170 °C. The mass spectrometer was operated in positive ion mode. Data acquisition was performed in a data-dependent mode consisting of, alternatively in a single run, full-scan MS over the range m/z 500–2000 and full MS/MS of the ion selected in an dynamic exclusion mode (the most intense ion is selected and excluded for further selection for a duration of 3 min). MS/MS data were acquired using a 2 m/z unit ion isolation window and 35% relative collision energy. MS/MS raw data files were transformed with Bioworks 3.1 software (Thermo Fischer Scientific) and neutral mass of the precursor and sequence information were used to identify proteins.

2.11. Bioinformatic data analyses

Resulting peptide masses were used to perform MASCOT search (MASCOT Peptide Mass Fingerprint: http://www.matrixscience.com/cgi/search_form.pl?FORMVER=2&SEARCH=PMF) against the *Homo sapiens* proteome in the MSDB database (mass spectrometry protein sequence database, release 20063108, 09/31/2006). Searches were set for a mass accuracy of 100 ppm, one missed cleavage of trypsin in matching peptides, oxidation of methionine, carbamidomethylation of cysteine and removal of the autolysis trypsin fragment ions in the case of MALDI-TOF spectra. No restrictions on experimental isoelectric point and mass were applied but a probability-base Mowse score (score is $-10 * \log_{10}(P)$ where P is the probability that the observed match is a random event) of 60 was adopted as the strict minimum for protein identification. Protein annotations followed those of UniprotKB/Swiss-Prot and UniProtKB/trEMBL databases (http://www.expasy.org/sprot/).

2.12. RT-PCR and qRT-PCR conditions

Primer pairs designed for the specific detection of target sequences (NSF, hnRNPL, TCPB, PDIA3, Prp19, DcpS, PSB3, PCBP2, RPLP0) are listed in Table 1. They were selected through computer analysis using Primer Premier v3.1 software (Biosoft International, CA, USA). All primer pairs were purchased from Eurogentec (Belgium).

RT-PCR analyses were performed mainly as described in ref. [15]. Briefly, 5 µg of each RNA preparation were reverse transcribed into first-strand cDNA using oligo-dT primers and 200 units of Moloney murine leukemia virus (MMLV) reverse transcriptase. Reverse transcriptase, oligo-dT primers and dNTPs were from Promega and the Silverstar polymerase from Eurogentec. The first-strand cDNA preparation and the determination of the optimal PCR conditions were as in ref. [7]. RT-PCR assays were performed in triplicate. Negative control reactions were performed using sterile water instead of the cDNA template. Contamination by genomic DNA was excluded by performing 35 cycles of amplification without retrotranscription. RT-PCR conditions specific to each primer pair are reported in Table 1. PCR products were separated on a 1.5% agarose gel stained with ethidium bromide and image acquisition was performed by UV transillumination using a Gel Doc 1000 system (Bio-Rad). Amplification products were subcloned in either pGEM Easy-T (Promega) or pCR BluntII-TOPO (Invitrogen), and sequenced to confirm the specificity of the PCR.

Real-time PCR (qRT-PCR) was performed as using an Mx4000 thermal cycler system and Brilliant SYBER Green QPCR Master Mix (Stratagene, The Netherlands). The qRT-PCR conditions consisted of one denaturing cycle at 90 °C for 30 s, annealing at 55 °C for 60 s and elongation at 72 °C for 60 s. At the end of the PCR, the samples were subjected to a melting curve analysis. To control for any variations due to efficiencies of the reverse transcription and PCR, PRLP0 was used as internal control. The PCR efficiency of DcpS and RPLP0 was determined by making seven tenfold serial dilutions in triplicate of their cDNAs to plot, after amplification, the C_T versus log cDNA dilution and to determine the slope of the line. The PCR efficiency was then calculated as $m = -(1/\log E)$ where m is the slope and E is the efficiency. This step was used to ensure that the PCR efficiency of DcpS and RPLP0 were within 10% of each other. The comparative DcpS expression was calculated with the comparative C_T method where $2^{-\Delta\Delta C_T}$ is [(DcpS $C_{T,\text{transfected}}$ − RPLP0 $C_{T,\text{transfected}}$) − (DcpS $C_{T,\text{null vector}}$ − RPLP0 $C_{T,\text{null vector}}$)] [34]. $2^{-\Delta\Delta C_T}$ is reported as arbitrary units representing the fold change in DcpS expression. All PCR runs were performed in triplicate. qRT-PCR primer sequences are provided in Table 1.

2.13. Reporter gene assay

HEK 293 cells were plated the day before transfection in 12-well plates at a density of 2×10^5 cells per well. The ΔLf transcriptional activity was assessed using pGL3-promoter-Luc reporter plasmids (pGL3-S1^{Skp1}-Luc and pGL3-S1^{DcpS}-Luc) (50 ng per well) and pcDNA-ΔLf or pcDNA-ΔLf-M4 (200 ng per well). Each experiment represents three sets of independent triplicates. Twenty-four hours after transfection, cells were lysed and assayed using a luciferase assay kit (Promega) in a Wallac Victor2 1420 multilabel counter (Perkin Elmer, MA, USA). For all experiments, protein content was used to normalize luciferase results. Protein concentrations of cell lysates were determined by a BCA® assay, using bovine serum albumin as standard. Absorbance measurements were carried out at 590 nm using a microplate reader (Model 550, Bio-Rad).

2.14. ChIP assays

ChIP assay was conducted using the EZ ChIP Enzymatic kit (Upstate Biotech, Millipore, MA, USA) according to the manufacturer's instructions, with some modifications. At 24 h post-transfection, ΔLf transfected cells were cross-linked with 1% formaldehyde for 10 min at room temperature. After the reaction had been stopped by the addition of 125 mM of glycine for 5 min at room temperature, cells were washed in NaCl/Pi. Cells (10^7) and were then incubated in 200 µl of lysis buffer [50 mM Tris–HCl (pH 8.1) containing 1% (vol/vol) SDS, 10 mM EDTA, 1 mM Pefablock] for 10 min at 4 °C. After sonication and centrifugation (12,000 × g, 15 min), lysates were diluted in the ChIP dilution buffer (1:10), pre-cleared with 2 µl of mouse normal serum for 6 h at 4 °C under rotation, and precipitated with protein G Sepharose beads (GE Healthcare Life Sciences). The supernatant was further incubated with antibodies overnight at 4 °C or not incubated. An aliquot of untreated supernatant served as input control. An aliquot of supernatant was either incubated with monoclonal anti-FLAG M2 antibody (1:500, Sigma) or anti-rabbit IgG (1:1000, GE Healthcare Life Sciences, UK) used as non-specific antibody control. An aliquot of supernatant not incubated with antibody was immunoprecipitated and used as a negative control. Complexes were precipitated for 2 h at 4 °C using protein G Sepharose beads (GE Healthcare Life Sciences). The captured immunocomplexes, containing bound DNA fragments, were eluted overnight at 65 °C, and treated with 4 µl of ribonuclease A (20 mg/ml Sigma) and 2 µl of proteinase K (10 mg/ml Sigma). The DNA fragments were purified using a Qiagen DNA purification kit (Qiagen). After a 1:50 dilution, 2 µl of each supernatant were then used for PCR. Primer pairs specifically amplifying the *DcpS* or the *albumin* promoter region are described in Table 1. PCR products were separated on a 2% agarose gel, and stained with ethidium bromide.

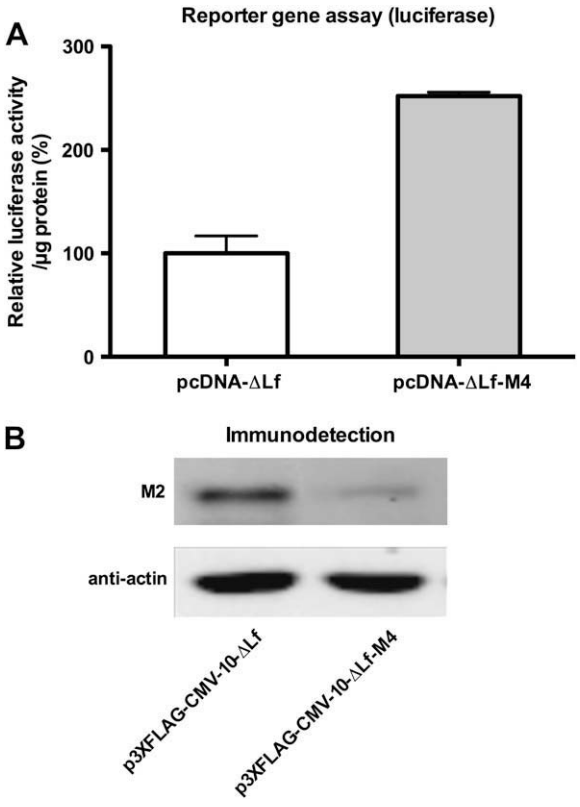


Fig. 1. ΔLf and ΔLf-M4 transactivate transcription from the *Skp1* promoter. Panel A. HEK 293 cells were co-transfected with pGL3-S1^{Skp1} and with either pcDNA-ΔLf or pcDNA-ΔLf-M4 expression vector (200 ng/well) encoding ΔLf or ΔLf-M4 respectively. Cells were lysed 24 h after transfection and samples were assayed for protein content and luciferase activity. The relative luciferase activity reported is expressed as the fold increase of the ratio of the pGL3 reporter activity to protein content. Values represent the mean ± SD of triplicates from three independent measurements. Panel B. HEK 293 cells were transfected with p3XFLAG-CMV-10-ΔLf or p3XFLAF-CMV-10-ΔLf-M4 vectors. Cells were lysed 24 h after transfection and samples (20 µg of protein) were subjected to SDS-PAGE and immunoblotted with antibodies specific to 3XFLAG tag (M2) or actin antibodies (anti-actin).

3. Results

3.1. Differential proteomic analysis

A proteomic approach was used to determine the changes in protein expression resulting from the expression of ΔLf, either native or deleted for *O*-glycosylation sites, in transiently transfected HEK 293 cells. HEK 293 cells were chosen since they do not express detectable levels of endogeneous ΔLf. Since *O*-GlcNAc negatively regulates ΔLf activity, mutagenesis of the four *O*-glycosylation sites was performed and the resulting mutant named ΔLf-M4. The transcriptional activity and the level of expression of ΔLf-M4 were analyzed and compared to the wild type protein. In Fig. 1A, the luciferase reporter assay showed that the mutant exhibits a higher transcriptional activity corresponding to a 2.5-fold increase compared to native ΔLf. The evaluation of the expression of both ΔLf isoforms at the protein level in HEK 293 cells was carried out using tagged isoforms of ΔLf and antibodies against the tag peptide. Fig. 1B showed that 3XFLAG-N-terminus-tagged ΔLf is expressed at a higher level than 3XFLAG-N-terminus-tagged ΔLf-M4. Taken together, these two results suggested that *O*-GlcNAc negatively regulates ΔLf transcriptional activity since mutation of the four glycosylation sites lead to a more efficient ΔLf isoform, but also that the presence of *O*-GlcNAc probably increases ΔLf stability.

Proteomic analysis was then performed. Four classes of experiments were used to highlight the impacts of ΔLf or ΔLf-M4 expression on the proteomic profile. Two control classes were used, one corresponding to cells that were not transfected (NT class) and the other corresponding to HEK 293 cells transfected with the pcDNA empty plasmid, in order to eliminate from the analyses all protein expression variations due to the transfection itself.

After selection to reject the most uneven gels, three to nine replicates of the 2D gels (controls, ΔLf, ΔLf-M4) were used. For analytical purposes, silver-stained 2D gels were able to resolve more than 500 protein spots each. In order to avoid inconsistent matches, only 214 protein spots were simultaneously assigned on the 2D gels from the four classes. The intensity of each spot was calculated as mean ± SD and expressed as percentage of volume. To minimize the impact of a non-normal distribution of the sample data, the Kolmogorov–Smirnov non-parametric statistical test was chosen to highlight protein spots of which the expression level differed significantly ($p < 0.05$). This analysis evidenced 12 protein spots with differential protein levels and a statistically reproducible difference over both control groups (Fig. 2A). One example of a differentially expressed protein, spot #1742, was enlarged from one gel in each class (Fig. 2B–E). The intensity of this protein spot was clearly increased in the gel corresponding to the ΔLf-M4 class by comparison with the control and ΔLf classes. Fig. 3 summarizes the levels of the differentially

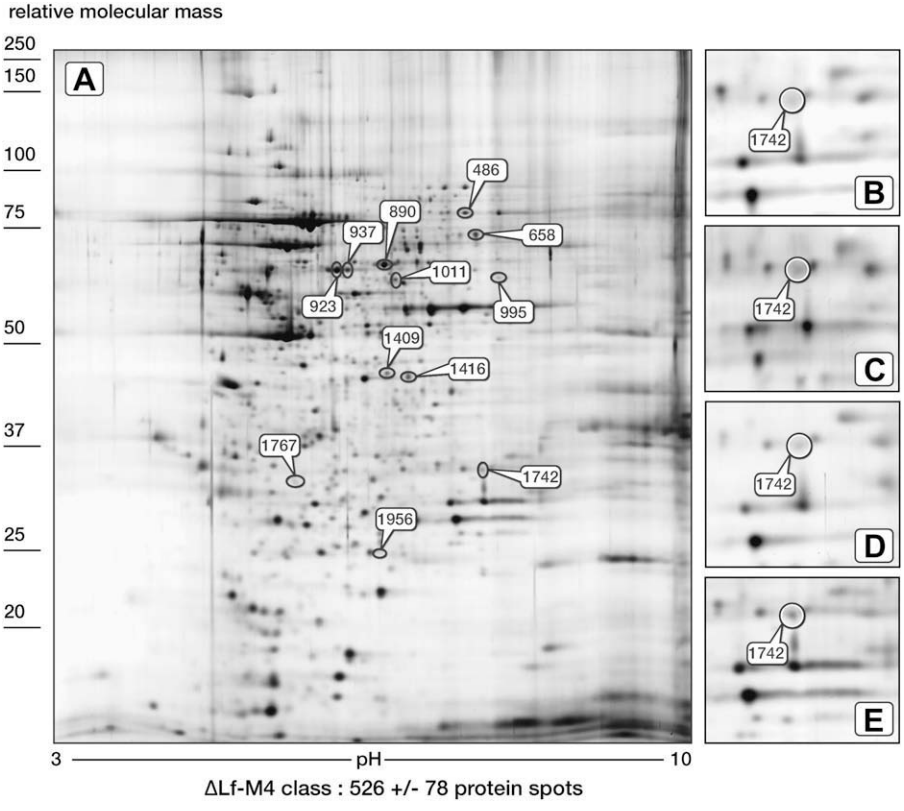


Fig. 2. Proteomic profiling of HEK 293 cells transfected with ΔLf constructs. Panel A: representative 2D image of HEK 293 proteins after transfection by pcDNA-ΔLf-M4. Panels B–E: 2D silver-stained portions encompassing DcpS protein from HEK 293 cells. B: not transfected; C: transfected with pcDNA only; D: transfected with pcDNA-ΔLf; E: transfected with pcDNA-ΔLf-M4.

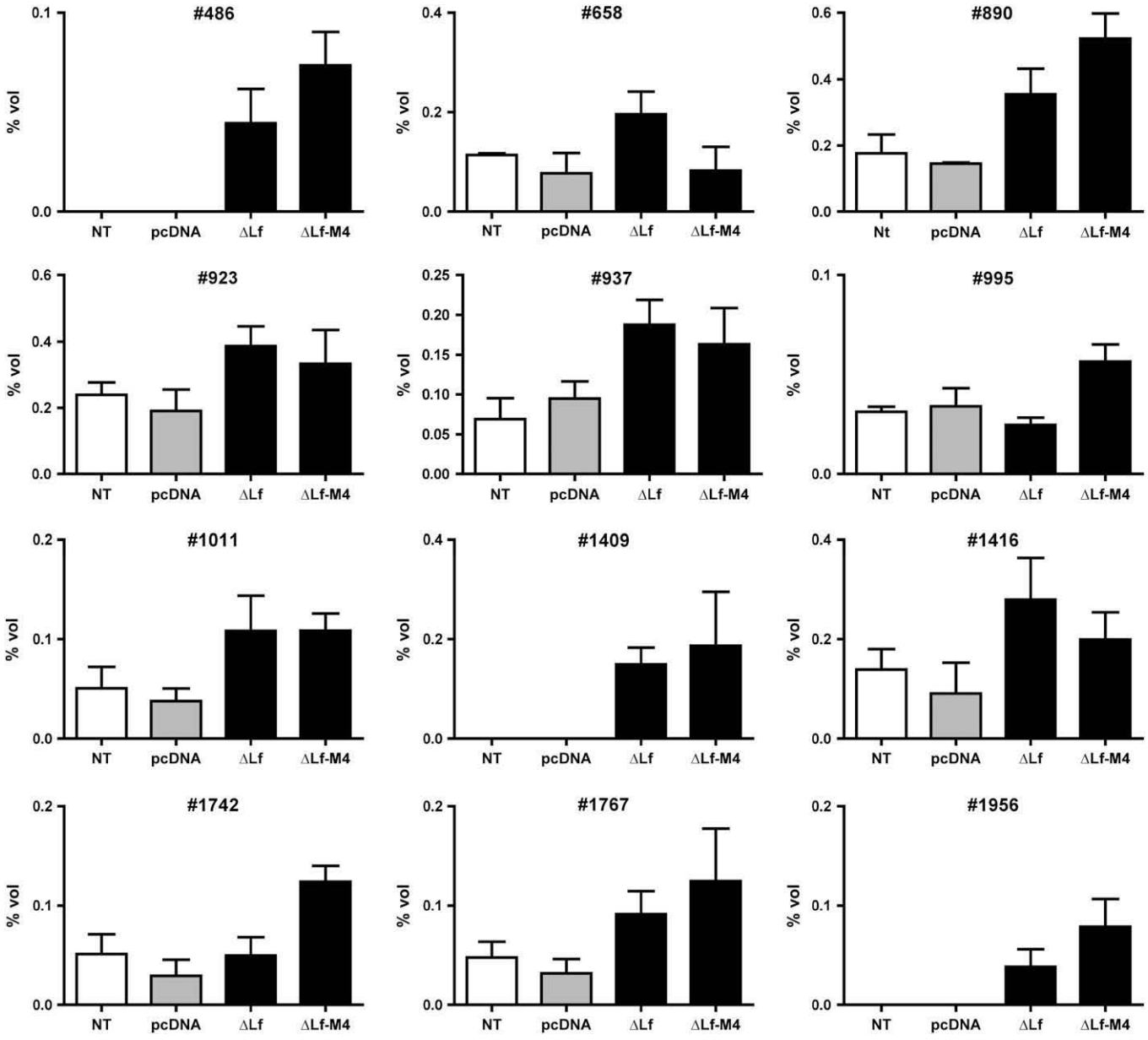


Fig. 3. Modifications in protein levels in ΔLf or ΔLf-M4-expressing-cells. HEK 293 cells were not transfected (NT) or transfected with pcDNA3.1 alone, pcDNA-ΔLf (ΔLf) or pcDNA-ΔLf-M4 (ΔLf-M4). Their total protein content was separated by 2D electrophoresis. Twelve protein spots, of which the intensity varied significantly ($p < 0.05$) were selected and plotted as mean ± SD of their % vol.

detected proteins in each group. Nine proteins (spots 486, 890, 923, 937, 1011, 1409, 1416, 1767 and 1956) were upregulated when cells were transfected with pcDNA-ΔLf or pcDNA-ΔLf-M4, whereas two proteins (spots 995 and 1742) were strongly upregulated when transfection occurred with pcDNA-ΔLf-M4 and one protein (spot 658) was upregulated only in the presence of pcDNA-ΔLf expression vector. For spots 486, 1409 and 1956, no matching spot was detected in the control classes, probably due to the fact that their expression level is below the silver staining sensitivity threshold. The feeble over-expression of proteins induced in the presence of M4-ΔLf compared to wild type might be due to the higher transcriptional activity of the mutant being compensated by its shorter half-life when *O*-GlcNAc sites are removed.

3.2. Mass spectrum identification of differentially expressed proteins

For identification purposes, 2D gels were performed with 500 μg of sample and stained with colloidal Coomassie brilliant blue instead of silver staining to avoid interference with mass spectrometry. The spots corresponding to the chosen proteins were excised, subjected to trypsin digestion and mass spectrometry identification by MALDI-TOF or LC-MS/MS. Among the 12 differentially expressed proteins, four were not identified (spots #937, #995, #1409 and #1767) either by MALDI-TOF or LC-MS/MS. The heterogeneous nuclear ribonucleoprotein L (spot #658), T-complex 1 subunit β (spot #890), protein disulfide isomerase A3 (spot #923) and

Table 2
Protein identification by mass spectrometry

Spot number	Accession number in SWISSPROT database	Description	UniProtKB/Swiss-Prot entry *Gene name*	Sequence coverage (%)	Number of peptides
MALDI-TOF					
#486	P46459	Vesicle-fusing ATPase	NSF_HUMAN *NSF*	29	12
#658	P14866	Heterogeneous nuclear ribonucleoprotein L	HRNPL_HUMAN *HNRNPL*	25	9
#890	P78371	T-complex 1 subunit β	TCPB_HUMAN *CCT2*	43	20
#923	P30101	Protein disulfide isomerase A3	PDIA3_HUMAN *PDIA3*	30	16
#1011	Q9UMS4	Pre-mRNA processing factor 19	PRP19_HUMAN *PRPF19*	24	10
#1742	Q96C86	Scavenger mRNA-decapping enzyme DcpS	DCPS_HUMAN *DCPS*	21	10
#1956	P49720	Proteasome subunit beta type 3	PSB3_HUMAN *PSMB3*	41	12
LC-MS/MS					
#658	P14866	Heterogeneous nuclear ribonucleoprotein L	HRNPL_HUMAN *HNRNPL*	19	8
#890	P78371	T-complex 1 subunit β	TCPB_HUMAN *CCT2*	26	17
#923	P30101	Protein disulfide isomerase A3	PDIA3_HUMAN *PDI3*	27	15
#1011	Q9UMS4	Pre-mRNA processing factor 19	PRP19_HUMAN *PRPF19*	21	10
#1416	Q15366	Poly(rC)-binding protein 2	PCBP2_HUMAN *PCBP2*	12	4

pre-mRNA-processing factor 19 (spot #1011) were all identified by MALDI-TOF and MS/MS while the poly(rC)-binding protein 2 (spot #1416) was identified only by LC-MS/MS and the vesicle-fusing ATPase (spot #486), the scavenger mRNA-decapping enzyme DcpS (spot #1742) and the proteasome subunit beta type 3 (spot #1956) only by MALDI-TOF (Table 2). As an example of peptide mass fingerprinting, the MALDI-TOF spectrum of spot #1742 obtained after trypsin digestion is shown in Fig. 4A. Its analysis allowed the characterization of the scavenger mRNA-decapping enzyme DcpS (Fig. 4B).

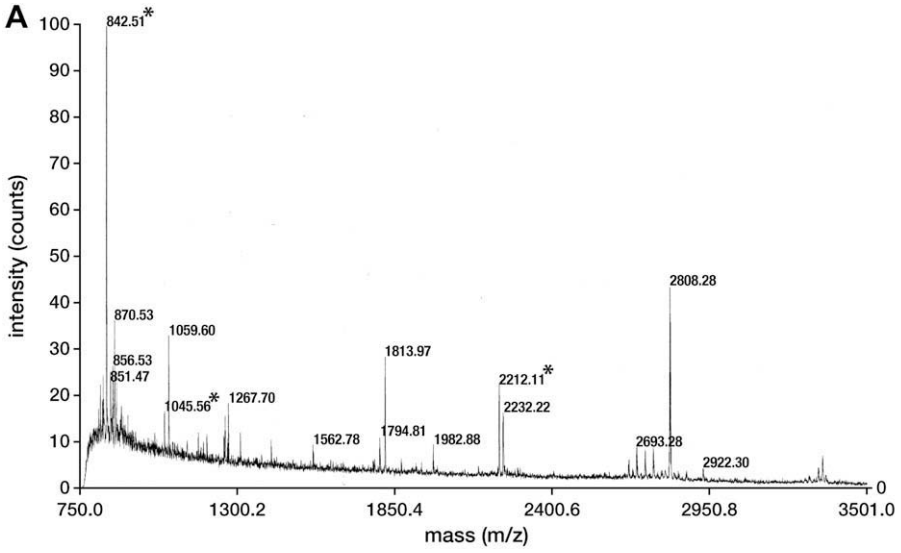
3.3. DcpS is a target of ΔLf transcriptional activity

In order to determine whether our identified proteins were also over-expressed at the mRNA level we first performed RT-PCR. As shown in Fig. 5A, a nearly 2-fold increase was observed for *DcpS* and *TCPB* transcripts but not for the others suggesting that the increased expression of these latter identified proteins may be due to translational events. We next investigated the potential existence of a ΔLfRE in the promoter region of *TCPB* and *DcpS* genes and found one in the promoter region of *DcpS* (Table 3). DcpS over-expression was then confirmed by qRT-PCR (Fig. 5B1). A higher level of transactivation was observed when ΔLf-M4 was used instead of native ΔLf. In order to study, the cell specificity of this process, the level of expression of DcpS mRNA was assayed using qRT-PCR in an HeLa cell line in which ΔLf was transiently expressed and in a stable ΔLf-expressing MDA-MB-231 cell line under doxycyclin induction. In all three cellular models, ΔLf expression led to a 2–3.5-fold increase in DcpS mRNA expression as shown in Fig. 5B. This overexpression was not specific to HEK 293 cells (Fig. 5B1), but was also visible in HeLa (Fig. 5B2) and MDA-MB-231-ΔLf (Fig. 5B3) cell lines at a comparable level using either a transient or a stable cellular expression model. Moreover, the response was dose-dependent with a maximum of 3.5-fold enhancement obtained with 2 μg of pcDNA-ΔLf plasmid for 10^6 cells in the HeLa cells.

To investigate whether overexpression of DcpS was also visible at the protein level, immunoblotting on HEK 293 lysates transfected with a 'null' plasmid, pcDNA-ΔLf or pcDNA-ΔLf-M4, was performed. This showed that the amount of DcpS protein increased in the transfected cells (Fig. 6A), consistent with the results obtained either by RT-PCR or qRT-PCR and confirming that ΔLf or ΔLf-M4 expression leads to the overexpression of DcpS at both the RNA and protein levels.

To monitor the efficiency of ΔLf transcriptional activity on the *DcpS* promoter, luciferase reporter assays were performed in HEK 293 cells. Fig. 7A shows that ΔLf induces a marked increase in luciferase activity after binding to the *DcpS* enhancer/promoter region compared to the *Skp1* promoter region used as positive control. Gene transactivation by ΔLf *via* $S1^{Skp1}$ or $S1^{DcpS}$ led to a 40-fold increase over the basal expression level. These results suggest that the central part of the core element is sufficient to allow this transactivation since these two *cis*-acting elements only differ in their external bases (Table 3). In order to demonstrate that ΔLf binds to the human *DcpS* promoter *in vivo*, we performed a chromatin

A



B

Observed	Mr(expt)	Mr(calc)	Delta	Start	End	Miss	Peptide
821.42	820.41	820.48	-0.07	316	322	0	TLTFALR
827.47	826.46	826.51	-0.04	61	67	0	IIFLHGK
830.45	829.44	829.48	-0.04	52	58	1	VLRESAR
856.53	855.52	855.44	0.08	258	264	1	MKGDHLR
870.54	869.53	869.46	0.07	2	10	0	ADAAPQLGK
916.52	915.51	915.47	0.04	330	337	0	LLQEAQQS
1668.92	1667.92	1667.87	0.04	323	337	1	ADDPLLKLLQEAQQS
1813.97	1812.96	1812.85	0.11	13	28	1	RELDVEEAHAASTEEK

Fig. 4. Panel A. MALDI-TOF peptide-mass fingerprint spectrum of the tryptic digest of spot 1742 protein. *, the trypsin autolytic fragments. Panel B. Proposed peptides were obtained by MASCOT search (MASCOT Peptide Mass Fingerprint: http://www.matrixscience.com/cgi/search_form.pl?FORMVER=2&SEARCH=PMF) against the *Homo sapiens* proteome in the MSDB database (mass spectrometry protein sequence database, release 20063108, 09/31/2006).

immunoprecipitation assay (ChIP). To this end, a 3XFLAG-N-terminus-tagged ΔLf was used to obtain the most reliable results. The DNA purified from the sonicated chromatin was directly analyzed by PCR using a *DcpS*-promoter-specific primer pair and was used as an input control (lane 1, Fig. 7B). After immunoprecipitation by M2 antibodies, PCR amplification with the *DcpS*-specific primers revealed a product of the expected size (lane 2, Fig. 7B). Control experiments involving non-specific antibodies showed only a slight amplification of the PCR product (lane 4, Fig. 7B) confirmed specificity of the results and this was reinforced by the loading control, corresponding to the immunoprecipitation of chromatin with pure protein G Plus Sepharose (lane 3, Fig. 7B). The PCR data shown in Fig. 7B correspond to a significant experiment chosen among three independent assays. The upstream sequence of the albumin promoter devoid of ΔLf response elements, was also amplified and served as a negative control. As expected, this promoter region was not enriched by the M2 antibodies (lane 2, Fig. 7B).

4. Discussion

The reduced expression of ΔLf in cases of cancer, and its role as a transcription factor regulating cell cycle progression suggests that it may act as a tumor suppressor gene. The present study was designed to identify, using proteomic tools, proteins that are differentially expressed in the presence of ΔLf and might be targets of its transcriptional activity.

Since ΔLf acts in cell surveillance its activity should be highly regulated. We previously showed that the *O*-GlcNAc/phosphorylation interplay is involved in this regulation. *O*-GlcNAc modification of transcription factors, cytoskeletal proteins, kinases, and nuclear pore proteins is a crucial regulatory post-translational modification akin to phosphorylation (reviewed in refs. [35,36]). This modification controls numerous processes ranging from nutrient sensing to the regulation of proteosomal degradation, cell signaling, cell cycle progression and gene expression [35,37–40]. Thus, this modification might play a key role in tumorigenesis since many oncogene and tumor suppressor gene products are modified by *O*-GlcNAc [39,41] and alterations of the *O*-GlcNAc status are observed in breast cancer [38]. In order to investigate the possibility that ΔLf might be used in therapeutic approaches and to avoid a negative control of its transcriptional activity by *O*-GlcNAc in cancer cells, we produced a ΔLf isoform in which all the *O*-GlcNAc sites were mutated. The ΔLf-M4 mutant is, thus, a constitutively active ΔLf isoform of which the transcriptional activity is independent of regulation by *O*-GlcNAc. We therefore added this glycosylation mutant in our study in order to confirm that it behaves like ΔLf *in vivo*. Our data showed that nearly all the same proteins were overexpressed in ΔLf- and ΔLf-M4-expressing HEK293 cells with a higher expression in the latter.

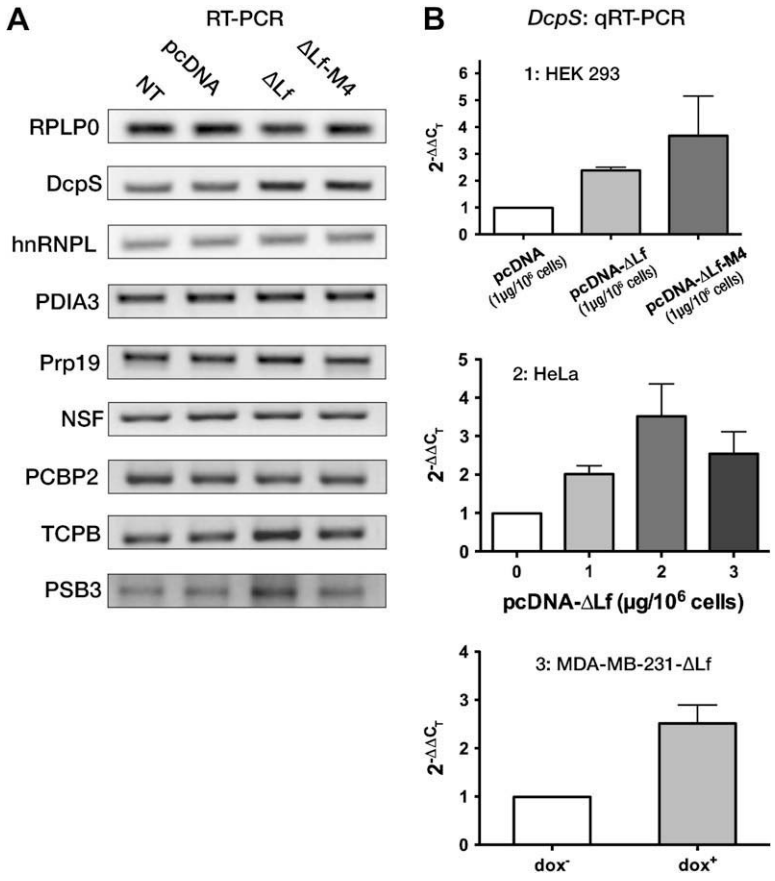


Fig. 5. Panel A. Upregulation of *DcpS* and *TCPB* mRNA in ΔLf and ΔLf-M4-expressing-HEK 293 cells. HEK 293 cells were either not transfected (NT) or transfected with pcDNA3.1 only (pcDNA), pcDNA-ΔLf (ΔLf) or pcDNA-ΔLf-M4 (ΔLf-M4). Among the eight identified genes, overexpression was confirmed for DcpS and TCPB by RT-PCR using RPLP0 as internal control. RPLP0, ribosomal protein, large, P0; DcpS, scavenger mRNA-decapping enzyme DcpS; hnRNPL, heterogeneous nuclear ribonucleoprotein L; TCPB, T-complex 1 subunit β; PDIA3, protein disulfide isomerase A3; Prp19, pre-mRNA processing factor 19; NSF, vesicle-fusing ATPase; PCBP2, Poly(rC)-binding protein 2; PSB3, proteasome subunit beta type 3. Panel B. Overexpression of DcpS is not cell specific. The expression pattern of DcpS transcripts in HEK 293, HeLa and MDA-MB-231-ΔLf 24 h after transient transfection or 24 h after doxycyclin induction was followed by qRT-PCR. The expression of each transcript is normalized to RPLP0 expression and is expressed as $2^{-\Delta\Delta CT}$ of DcpS expression compared to RPLP0 expression ($n = 3$). Panel B1. HEK 293 cells transiently transfected with pcDNA, pcDNA-ΔLf or pcDNA-ΔLf-M4. Panel B2. HeLa cells transiently transfected with increasing concentrations of pcDNA-ΔLf. Panel B3. MDA-MB-231-ΔLf cells stably expressing ΔLf 24 h after doxycyclin induction.

Among the identified proteins, some involved in proteasomal degradation were found. Degradation of ubiquitinylated substrates by the proteasome is highly regulated to maintain normal cell growth and we previously showed that ΔLf expression may lead to the control of S phase cell cycle actor degradation via the overexpression of Skp1, a component of the SCF complex (Skp1/Cullin-1/F-box ubiquitin ligase) [15]. Therefore the increased expression of proteins such as PSB3, the proteasome subunit beta type 3, one of the ten proteolytically-active beta subunits of the 20S core complex within the 26S proteosome [42], and the pre-mRNA-processing factor 19 (Prp19) that interestingly possesses an active conserved U-box domain present in E3 ubiquitin ligases and of which the activity has been demonstrated *in vitro* [43], reinforces our previous work showing the involvement of ΔLf in controlling S phase compound degradation. Increased synthesis of proteins in the ΔLf-expressing cells seems to be followed by a higher activity of proteins that act as chaperones/mediators of protein folding such as the protein disulfide isomerase A3 (ERp57), an essential folding catalyst of the ER [44], and the cytoplasmic chaperonin T-complex protein 1 subunit beta TCPB [45].

Among the overexpressed proteins, three are known to be critical for cell growth and viability. Thus, the poly(rC)-binding protein 2 (PCBP2 also called hnRNPE2) of which overexpression leads to a decrease in cellular viability and an

Table 3
S1-like sequences found in the *Skp1* and *DcpS* promoters compared to the S1 sequence

Name	Sequence									Location	
S1	G	G	C	A	C	T	T	A/G	C		[15]
$S1^{Skp1}$	G	G	C	A	C	T	G–T	A	C	−1067 to −1058	[14]
$S1^{DcpS}$	A	G	C	A	C	T	T	G	G	−2228 to −2220	

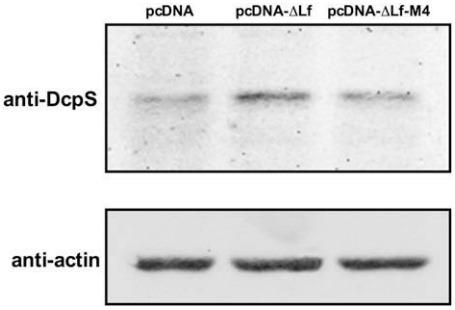


Fig. 6. Immunodetection of DcpS. HEK 293 cells were transfected with pcDNA, pcDNA-ΔLf or pcDNA-ΔLf-M4 expression vector (1 μg/10^6 cells). 24 h after transfection, total cell extracts were prepared, and samples (30 μg of protein) were loaded on 10% SDS-PAGE. Western blot detection was done using anti-DcpS (upper panel) or anti-actin (lower panel) antibodies developed by ECL.

increase in caspase 3 and 9 activity, possesses pro-apoptotic activity [46]. Unlike PCB2, it is the downregulation of TCPB, a protein overexpressed at the G1/S transition and interacting with cyclin E [47], that leads to cell cycle arrest and apoptosis [48]. Over-expression of Prp19 involved in DNA repair also

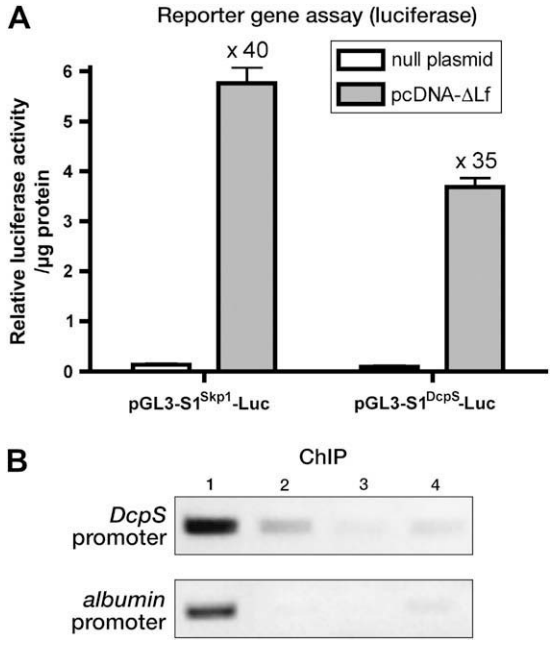


Fig. 7. ΔLf transactivates transcription from the *DcpS* promoter. Panel A. HEK 293 cells were cotransfected with pGL3-S1^{DcpS}-Luc (A) or pGL3-S1^{Skp1}-Luc (B) constructs (50 ng/well) and with a null plasmid or with pcDNA-ΔLf expression vector (200 ng/well) encoding ΔLf, respectively. Cells were lysed 24 h after transfection. Samples were assayed for protein content and luciferase activity. The relative luciferase activity reported is expressed as the fold increase of the ratio of the pGL3 reporter activity to protein content. Values represent the mean ± SD of triplicates from three independent measurements. Panel B. ΔLf binds to the *DcpS* promoter *in vivo*. 3XFLAG-ΔLf/DNA complexes were captured with M2 antibodies and amplified using specific primers for the *DcpS* promoter or for the *albumin* promoter. PCR products representing the *DcpS* and the *albumin* promoter (negative control) sequences are noted: lane 1, loading control corresponding to input; lane 2, ChIP using anti-FLAG M2 antibody; lane 3, control without M2 antibody; lane 4, control with an anti-rabbit IgG as non-specific antibody.

favors cellular viability [49]. Therefore, these three proteins might participate with ΔLf in the maintenance of cell integrity.

Another group of proteins that function in biogenesis, stability and maturation of mRNAs such as Prp19, the heterogeneous nuclear ribonucleoprotein L (hnRNP L), the RNA-binding protein PCBP2 and the scavenger decapping enzyme (DcpS) were also upregulated. Thus, Prp19 is involved in DNA repair and pre-mRNA splicing [50,51]. It belongs to a complex required for activation and stability of the splicesome [50] and its U-box domain is essential for mRNA maturation *in vivo* [43]. The heterogenous nuclear ribonucleoproteins, hnRNPs, have been thought to regulate the export, stability, and translation of specific mRNA [52–55]. The PCBP proteins are involved in mRNA stabilization, transcriptional regulation, translational control and apoptotic program activation (reviewed by ref. [56]). DcpS, the scavenger decapping enzyme also acts as a translation, transport, and/or mRNA quality control regulator. Data generated in this proteomic study provide a starting point for future studies aimed at examining the relationship between ΔLf and mRNA biosynthesis and turnover.

The combined examination of both protein and mRNA expression data and promoter analysis provided a novel insight into the cellular processes regulated by ΔLf. Of the eight upregulated proteins only two were also upregulated at the transcription level. For those that were not, this may be due to a translational level of control such as differential mRNA stability. Furthermore of the two upregulated genes, only *DcpS* had an identifiable ΔLfRE in its promoter. Therefore the upregulation of the TCPB gene may be due to an indirect effect on transcription. The functionality of the *DcpS* promoter ΔLfRE was demonstrated both using ChIP and a reporter gene assay. DcpS, the scavenger decapping enzyme, is a nucleocytoplasmic shuttling protein which is predominantly observed in the nucleus [27]. It catalyzes the release of m^7GMP ensuring that no excess unhydrolyzed cap accumulates. Thus, DcpS is a modulator of cap binding proteins. It positively modulates translation and cap-proximal intron splicing by controlling eIF4E [26] and Cbp20 [27] availability, respectively. It may also perform the cap removal of aberrant nuclear transcripts and thereby may improve the quality control of mRNA biogenesis [57,58]. Since this enzyme possesses numerous functions, modifications in DcpS concentration may have considerable consequences for cell survival. Our result raises the question of whether ΔLf, by overexpressing DcpS, may modulate activities such as mRNA turnover and quality control of mRNA synthesis, or the prevention of eIF4E and Cbp20 sequestration. Further investigations will be necessary to determine whether all or only some of these activities effectively take place in ΔLf-expressing cells.

In conclusion, the proteome profiling technique provided a broad-based and effective approach to identify protein changes induced by ΔLf and has allowed us to pinpoint DcpS as a new ΔLf target gene. Identification and characterization of functionally modulated proteins involved in the cell cycle, cell viability and mRNA biogenesis and degradation events may lead to a better understanding of the effects of ΔLf and its

constitutionally active isoform at the molecular level and may contribute to the future development of novel cancer therapeutic approaches.

Acknowledgments

This investigation was supported in part by the CNRS Unité Mixte de Recherche 8576 (Unité de Glycobiologie Structurale et Fonctionnelle, director Dr. J.C. Michalski), the Institut Fédératif de Recherche n° 147, the Université des Sciences et Technologies de Lille I, the Ministère de l'Education Nationale and the comité du Nord de la Ligue Nationale contre le Cancer (*Etude du rôle de la delta-lactoferrine*, *des gangliosides et des neurotrophines dans le développement et la progression du cancer du sein*) and the Association pour la Recherche sur le Cancer (contrat N° 7936). We would like to thank the INSERM U 547 (Dir. Prof. M Capron) for providing us with access to the Wallac Victor2 1420 luminometer and Dr. R.J. Pierce (INSERM U 547, Institut Pasteur de Lille, F) for reviewing this manuscript.

References

[1] J. Coller, R. Parker, Eukaryotic mRNA decapping, Annu. Rev. Biochem. 73 (2004) 861−890.
[2] H. Liu, M. Kiledjian, Decapping the message: a beginning or an end, Biochem. Soc. Trans. 34 (2006) 35−38.
[3] E. van Dijk, H.L. Hir, B. Séraphin, DcpS can act in the 5′−3′ mRNA decay pathway in addition to the 3′−5′ pathway, Proc. Natl. Acad. Sci. U.S.A. 100 (2003) 12081−12086.
[4] D.L. Nuss, Y. Furuichi, G. Koch, A.J. Shatkin, Detection in HeLa cell extracts of a 7-methyl guanosine specific enzyme activity that cleaves m7GpppNm, Cell 6 (1975) 21−27.
[5] P.D. Siebert, B.C. Huang, Identification of an alternative form of human lactoferrin mRNA that is expressed differentially in normal tissues and tumor-derived cell lines, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 2198−2203.
[6] M. Benaïssa, J. Peyrat, L. Hornez, C. Mariller, J. Mazurier, A. Pierce, Expression and prognostic value of lactoferrin mRNA isoforms in human breast cancer, Int. J. Cancer. 114 (2004) 299−306.
[7] D. Liu, X. Wang, Z. Zhang, C.T. Teng, An intronic alternative promoter of the human lactoferrin gene is activated by Ets, Biochem. Biophys. Res. Commun. 301 (2003) 472−479.
[8] L.E. Liu, W. Gladwell, C.T. Teng, Detection of exon polymorphisms in the human lactoferrin gene, Biochem. Cell Biol. 80 (2002) 17−22.
[9] C. Teng, W. Gladwell, I. Raphiou, E. Liu, Methylation and expression of the lactoferrin gene in human tissues and cancer cells, Biometals 17 (2004) 317−323.
[10] Y. Yang, J. Li, A. Szeles, M.P. Imreh, M. Kost-Alimova, H. Kiss, I. Kholodnyuk, L. Fedorova, E. Darai, G. Klein, S. Imreh, Consistent downregulation of human lactoferrin gene, in the common eliminated region 1 on 3p21.3, following tumor growth in severe combined immunodeficient (SCID) mice, Cancer Lett. 191 (2003) 155−164.
[11] H. Yi, H. Li, D. Peng, H. Zhang, L. Wang, M. Zhao, K. Yao, C. Ren, Genetic and epigenetic alterations of LTF at 3p21.3 in nasopharyngeal carcinoma, Oncol. Res. 16 (2006) 261−272.
[12] M. Breton, C. Mariller, M. Benaïssa, K. Caillaux, E. Browaeys, M. Masson, J. Vilain, J. Mazurier, A. Pierce, Expression of delta-lactoferrin induces cell cycle arrest, Biometals 17 (2004) 325−329.
[13] G.S. Goldberg, T. Kunimoto, D.B. Alexander, K. Suenaga, F. Ishidate, K. Miyamoto, T. Ushijima, C.T. Teng, J. Yokota, T. Ohta, H. Tsuda, Full length and delta lactoferrin display differential cell localization dynamics, but do not act as tumor markers or significantly affect the expression of other genes, Med. Chem. 1 (2006) 57−64.
[14] C. Mariller, M. Benaïssa, S. Hardivillé, M. Breton, G. Pradelle, J. Mazurier, A. Pierce, Human delta-lactoferrin is a transcription factor that enhances Skp1 (S-phase kinase-associated protein) gene expression, FEBS J 274 (2007) 2038−2053.
[15] J. He, P. Furmanski, Sequence specificity and transcriptional activation in the binding of lactoferrin to DNA, Nature 373 (1995) 721−724.
[16] K. Son, J. Park, C. Chung, D.K. Chung, D. Yu, K. Lee, J. Kim, Human lactoferrin activates transcription of IL-1beta gene in mammalian cells, Biochem. Biophys. Res. Commun. 290 (2002) 236−241.
[17] E.N. Baker, H.M. Baker, Molecular structure, binding properties and dynamics of lactoferrin, Cell. Mol. Life Sci. 62 (2005) 2531−2539.
[18] E.N. Baker, H.M. Baker, R.D. Kidd, Lactoferrin and transferrin: functional variations on a common structural framework, Biochem. Cell Biol. 80 (2002) 27−34.
[19] X.L. Ang, J.W. Harper, Interwoven ubiquitination oscillators and control of cell cycle transitions, Sci. STKE 2004 (2004) pe31.
[20] W. Krek, Proteolysis and the G1-S transition: the SCF connection, Curr. Opin. Genet. Dev. 8 (1998) 36−42.
[21] A. Lehmann, T. Toda, Fission yeast Skp1 is required for spindle morphology and nuclear membrane segregation at anaphase, FEBS Lett. 566 (2004) 77−82.
[22] R. Piva, J. Liu, R. Chiarle, A. Podda, M. Pagano, G. Inghirami, In vivo interference with Skp1 function leads to genetic instability and neoplastic transformation, Mol. Cell. Biol. 22 (2002) 8375−8387.
[23] G.W. Hart, L.K. Kreppel, F.I. Comer, C.S. Arnold, D.M. Snow, Z. Ye, X. Cheng, D. DellaManna, D.S. Caine, B.J. Earles, Y. Akimoto, R.N. Cole, B.K. Hayes, *O*-GlcNAcylation of key nuclear and cytoskeletal proteins: reciprocity with *O*-phosphorylation and putative roles in protein multimerization, Glycobiology 6 (1996) 711−716.
[24] H. Liu, N.D. Rodgers, X. Jiao, M. Kiledjian, The scavenger mRNA decapping enzyme DcpS is a member of the HIT family of pyrophosphatases, EMBO J 21 (2002) 4699−4708.
[25] Y. Furuichi, A.J. Shatkin, Viral and cellular mRNA capping: past and prospects, Adv. Virus Res. 55 (2000) 135−184.
[26] S. Liu, X. Jiao, H. Liu, M. Gu, C.D. Lima, M. Kiledjian, Functional analysis of mRNA scavenger decapping enzymes, RNA 10 (2004) 1412−1422.
[27] V. Shen, H. Liu, S. Liu, X. Jiao, M. Kiledjian, DcpS scavenger decapping enzyme can modulate pre-mRNA splicing, RNA 14 (2008) 1132−1142.
[28] E. van Dijk, N. Cougot, S. Meyer, S. Babajko, E. Wahle, B. Séraphin, Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures, EMBO J 21 (2002) 6915−6924.
[29] N.D. Rodgers, Z. Wang, M. Kiledjian, Regulated alpha-globin mRNA decay is a cytoplasmic event proceeding through 3′-to-5′ exosome-dependent decapping, RNA 8 (2002) 1526−1537.
[30] Z. Wang, M. Kiledjian, Functional link between the mammalian exosome and mRNA decapping, Cell 107 (2001) 751−762.
[31] M. Chevallet, S. Luche, T. Rabilloud, Silver staining of proteins in polyacrylamide gels, Nat. Protocols 1 (2006) 1852−1858.
[32] V. Neuhoff, N. Arold, D. Taube, W. Ehrhardt, Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250, Electrophoresis 9 (1988) 255−262.
[33] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, New York, 1989.
[34] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2(−Delta Delta *C*(T)) method, Methods 25 (2001) 402−408.
[35] N.E. Zachara, G.W. Hart, Cell signaling, the essential role of O-GlcNAc!, Biochim. Biophys. Acta 1761 (2006) 599−617.
[36] N.E. Zachara, G.W. Hart, *O*-GlcNAc a sensor of cellular state: the role of nucleocytoplasmic glycosylation in modulating cellular function in response to nutrition and stress, Biochim. Biophys. Acta 1673 (2004) 13−28.
[37] G.W. Hart, M.P. Housley, C. Slawson, Cycling of *O*-linked beta-*N*-acetylglucosamine on nucleocytoplasmic proteins, Nature 446 (2007) 1017−1022.

[38] C. Slawson, J. Pidala, R. Potter, Increased *N*-acetyl-beta-glucosaminidase activity in primary breast carcinomas corresponds to a decrease in *N*-acetylglucosamine containing proteins, Biochim. Biophys. Acta 1537 (2001) 147–157.
[39] W.H. Yang, J.E. Kim, H.W. Nam, J.W. Ju, H.S. Kim, Y.S. Kim, J.W. Cho, Modification of p53 with *O*-linked *N*-acetylglucosamine regulates p53 activity and stability, Nat. Cell Biol. 8 (2006) 1074–1083.
[40] X. Yang, F. Zhang, J.E. Kudlow, Recruitment of *O*-GlcNAc transferase to promoters by corepressor mSin3A: coupling protein *O*-GlcNAcylation to transcriptional repression, Cell 110 (2002) 69–80.
[41] T.Y. Chou, C.V. Dang, G.W. Hart, Glycosylation of the c-Myc transactivation domain, Proc. Natl. Acad. Sci. U.S.A. 92 (1995) 4417–4421.
[42] N. Tanahashi, C. Tsurumi, T. Tamura, K. Tanaka, Molecular structure of 20S and 26S proteasomes, Enzyme Protein 47 (1993) 241–251.
[43] C.W.V. Kooi, M.D. Ohi, J.A. Rosenberg, M.L. Oldham, M.E. Newcomer, K.L. Gould, W.J. Chazin, The Prp19 U-box crystal structure suggests a common dimeric architecture for a class of oligomeric E3 ubiquitin ligases, Biochemistry 45 (2006) 121–130.
[44] K. Bedard, E. Szabo, M. Michalak, M. Opas, Cellular functions of endoplasmic reticulum chaperones calreticulin, calnexin, and ERp57, Int. Rev. Cytol 245 (2005) 91–121.
[45] A.Y. Dunn, M.W. Melville, J. Frydman, Review: cellular substrates of the eukaryotic chaperonin TRiC/CCT, J. Struct. Biol. 135 (2001) 176–184.
[46] P. Roychoudhury, R.R. Paul, R. Chowdhury, K. Chaudhuri, HnRNP E2 is downregulated in human oral cancer cells and the overexpression of hnRNP E2 induces apoptosis, Mol. Carcinog 46 (2007) 198–207.
[47] J. Grantham, K.I. Brackley, K.R. Willison, Substantial CCT activity is required for cell cycle progression and cytoskeletal organization in mammalian cells, Exp. Cell Res. 312 (2006) 2309–2324.
[48] C. Coghlin, B. Carpenter, S.R. Dundas, L.C. Lawrie, C. Telfer, G.I. Murray, Characterization and over-expression of chaperonin t-complex proteins in colorectal cancer, J. Pathol 210 (2006) 351–357.
[49] X. Lu, R.J. Legerski, The Prp19/Pso4 core complex undergoes ubiquitylation and structural alterations in response to DNA damage, Biochem. Biophys. Res. Commun. 354 (2007) 968–974.
[50] S. Chan, S. Cheng, The Prp19-associated complex is required for specifying interactions of U5 and U6 with pre-mRNA during spliceosome activation, J. Biol. Chem. 280 (2005) 31190–31199.
[51] M. Grey, A. Düsterhöft, J.A. Henriques, M. Brendel, Allelism of PSO4 and PRP19 links pre-mRNA processing with recombination and error-prone DNA repair in *Saccharomyces cerevisiae*, Nucleic Acids Res. 24 (1996) 4009–4014.
[52] B. Hahm, O.H. Cho, J.E. Kim, Y.K. Kim, J.H. Kim, Y.L. Oh, S.K. Jang, Polypyrimidine tract-binding protein interacts with HnRNP L, FEBS Lett. 425 (1998) 401–406.
[53] X. Liu, J.E. Mertz, HnRNP L binds a cis-acting RNA sequence element that enables intron-dependent gene expression, Genes Dev. 9 (1995) 1766–1780.
[54] D. Perrotti, B. Calabretta, Post-transcriptional mechanisms in BCR/ABL leukemogenesis: role of shuttling RNA-binding proteins, Oncogene 21 (2002) 8577–8583.
[55] S.C. Shih, K.P. Claffey, Regulation of human vascular endothelial growth factor mRNA stability in hypoxia by heterogeneous nuclear ribonucleoprotein L, J. Biol. Chem. 274 (1999) 1359–1365.
[56] A.V. Makeyev, S.A. Liebhaber, The poly(C)-binding proteins: a multiplicity of functions and a search for mechanisms, RNA 8 (2002) 265–278.
[57] P.A. Frischmeyer, A. van Hoof, K. O'Donnell, A.L. Guerrerio, R. Parker, H.C. Dietz, An mRNA surveillance mechanism that eliminates transcripts lacking termination codons, Science 295 (2002) 2258–2261.
[58] A. van Hoof, P.A. Frischmeyer, H.C. Dietz, R. Parker, Exosome-mediated recognition and degradation of mRNAs lacking a termination codon, Science 295 (2002) 2262–2264.

La régulation de l'activité de la delta-lactoferrine

Comme nous venons de le voir, l'expression de la delta-lactoferrine induit un arrêt du cycle cellulaire en phase S (Breton *et al.*, 2004) ainsi que la mort cellulaire par apoptose (article 2). Son activité semble donc primordiale dans la défense de l'organisme face au cancer mais serait délétère dans des conditions non pathologiques. En outre, la delta-lactoferrine est un facteur de transcription possédant des gènes cibles variés : *Skp1* (article 1), *Bax* (article 2) et *DcpS* (article 3). La participation de ces protéines dans des processus cellulaires cruciaux suggère que l'activité de la delta-lactoferrine doit être hautement régulée.

Les modifications post-traductionnelles sont un élément clé de la régulation des facteurs de transcription. En effet, la régulation de l'activité de p53 en est un exemple. Dans un contexte cellulaire normal, p53 est constamment exprimée et constamment dégradée à cause de son interaction avec la protéine mdm2 qui conduit à son ubiquitinylation (Lavin et Gueven, 2006; Yang *et al.*, 2004). En condition de stress cellulaire, p53 est phosphorylée par des kinases spécifiques ce qui inhibe son interaction avec mdm2 et la stabilise (Yang *et al.*, 2004). La stabilisation de la protéine dans la cellule lui permet alors d'exercer son activité. De plus, Yang *et al*. ont montré que la GlcNAcylation de p53 conduit à l'inhibition de son ubiquitinylation et à sa stabilisation (Yang *et al.*, 2006). Un premier niveau de régulation de l'activation de la protéine p53 se fait donc par ces modifications post-traductionnelles. D'autres modifications post-traductionnelles peuvent également entrer en jeu et vont lui conférer une activité transcriptionnelle sélective pour certains gènes cibles. Ainsi, l'acétylation de la lysine 320 de p53 inhibe la phosphorylation en N-terminal et conduit à la transactivation de ses gènes cibles possédant un élément de réponse dit de haute affinité (retrouvé dans le promoteur du gène *p21*) alors que l'acétylation de la lysine 373 conduit à l'hyperphosphorylation de son extrémité N-terminale et permet la transactivation de gènes possédant un élément de réponse de faible affinité (retrouvé dans le promoteur du gène *Bax*) (Knights *et al.*, 2006). Cet exemple montre que l'activité des facteurs de transcription est régulée de manière complexe et fine par les modifications post-traductionnelles.

La localisation cytoplasmique et nucléaire de la delta-lactoferrine fait qu'elle pourrait être la cible de ces mêmes modifications post-traductionnelles. Cependant, les modifications post-traductionnelles connues pour être responsables de la régulation d'une activité transcriptionnelle sont restreintes. En effet, ces modifications post-traductionnelles doivent être dynamiques et réversibles. Nous nous sommes intéressés dans un premier temps à la

phosphorylation et la GlcNAcylation. Les résultats de l'analyse de la séquence primaire de la delta-lactoferrine par les serveurs Ying*O*Yang 1.2, NetPhosK 1.0 et NetPhos 2.0, sont présentés dans la Figure 26. Ils montrent que la delta-lactoferrine possède de nombreux sites de phosphorylation putatifs et quatre sites de GlcNAcylation putatifs en position 10, 227, 472 et 559. Certains de ces sites sont situés au niveau de domaines particuliers (Figure 7, page 20) : la sérine 10 et la thréonine 559 se trouvent à proximité des domaines putatifs d'interaction à l'ADN, sur le lobe N-terminal pour le site S10 et, de part la conformation tridimentionnelle de la molécule, dans l'environnement de la région inter-lobe pour le site T559. Le site T559 est également situé à proximité de la séquence NLS (article 1). La sérine 227, quant à elle, est située à l'intérieur du site de fixation du fer qui ne devrait pas être fonctionnel, l'environnement oxydant régnant dans le cytosol étant incompatible avec une fixation du fer. La sérine 472 est aussi située près de la séquence NLS et d'une séquence PEST putative. La difficulté à isoler la delta-lactoferrine même lorsqu'elle est produite par des systèmes de surexpression (*Pichia pastoris*, Baculovirus) suggère que cette séquence PEST putative pourrait être pleinement fonctionnelle.

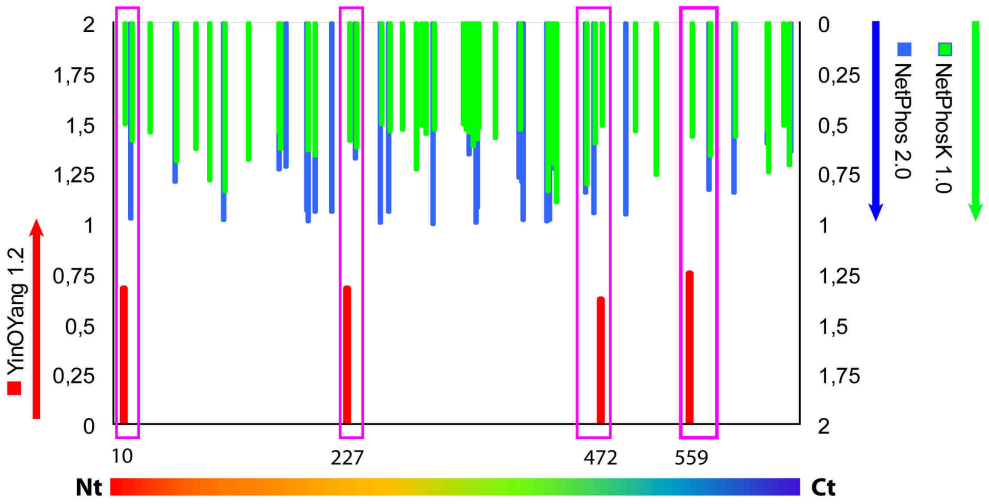


Figure 26 : Sites potentiels de GlcNAcylation et de phosphorylation de la delta-lactoferrine. L'utilisation du serveur YinOYang 1.2 permet de cibler les acides aminés pouvant être modifiés par un résidu de O-GlcNAc et celle des serveurs NetPhosK 1.O et NetPhos 2.0 met en évidence les acides aminés susceptibles d'être modifiés par un résidu de phosphate. Les quatre sites de GlcNAcylation prédits sont également des sites putatifs de phosphorylation suggérant que la réciprocité GlcNAcylation/phosphorylation pourrait s'exercer sur la delta-lactoferrine (Etude réalisée par le Dr. Christophe Mariller).

A mon arrivée au laboratoire, la caractérisation de ces sites n'avait pas encore pu être menée à bien à cause de la difficulté à visualiser et/ou isoler la protéine. Cependant, une étude préliminaire utilisant la delta-lactoferrine traduite *in vitro* puis une reconnaissance par la lectine WGA (« wheat germ agglutinin ») ou les anticorps RL2 (spécifique de la GlcNAcylation) avait permis de montrer que cette protéine pouvait être GlcNAcylée. Une partie de mes travaux de thèse a consisté à confirmer l'existence de ces sites modifiés par GlcNAcylation et d'étudier l'effet de cette ou de ces modification(s) sur l'activité transcriptionnelle et la stabilité de la delta-lactoferrine.

Les résultats de cette étude ont été rassemblés et font l'objet de la publication suivante actuellement en révision :

Hardivillé, S., Hoedt, E., Mariller, C., Benaïssa, M., et Pierce, A. « A cycling between *O*-GlcNAcylation and phosphorylation at Ser10 controlled both delta-lactoferrin transcriptional activity and stability », *J. Biol. Chem*. en révision (2010)

Pour cette étude, nous avons du mettre au point un système nous permettant de visualiser la protéine. Son absence dans les lignées en culture ainsi que l'absence d'anticorps efficaces dirigés contre la delta-lactoferrine, nous a conduit à utiliser une construction permettant la fusion d'une séquence immunoréactive à la séquence polypeptidique de la delta-lactoferrine. Notre choix s'est porté sur le vecteur p3xFLAG-CMV et l'étiquette FLAG. En effet, les anticorps dirigés contre le FLAG sont hautement réactifs et permettent la détection de quelques dizaines de fentomoles de molécules de fusion. Grâce à cette construction nous avons pu mettre en évidence pour la première fois la delta-lactoferrine. L'utilisation de la protéine « FLAGée » nous a permis de confirmer que la delta-lactoferrine est GlcNAcylée mais qu'en plus elle est phosphorylée et ubiquitinylée. Il existe une relation étroite entre ces trois modifications. En effet, GlcNAcylation et phosphorylation peuvent être antagonistes (voir le chapitre « Modifications post-traductionnelles », page 48). Par ailleurs, la phosphorylation, notamment des séquences PEST, induit l'ubiquitinylation des protéines concernées (Rechsteiner et Rogers, 1996). Enfin, il a récemment pu être démontré au laboratoire que l'ubiquitinylation des protéines est aussi régulée par la GlcNAcylation (Guinez *et al.*, 2008). Nos travaux montrent que ces trois modifications post-traductionnelles sont étroitement liées et qu'elles gèrent de concert l'activité transcriptionnelle et la stabilité de la delta-lactoferrine.

En effet, nous montrons que l'augmentation des taux de GlcNAcylation intracellulaires inhibe l'activité transcriptionnelle de la delta-lactoferrine alors que l'augmentation de la phosphorylation antagoniste a l'effet inverse. Nous avons pu également démontrer grâce à différentes constructions mutantes pour les sites de GlcNAcylation que seuls les résidus de sérine 10, sérine 227 et sérine 472 sont la cible de l'OGT. L'analyse de l'activité transcriptionnelle de ces mutants de glycosylation démontre que la modification du résidu de sérine 10 par un résidu de GlcNAc est responsable de l'inhibition de l'activité transcriptionnelle de la delta-lactoferrine. Par ailleurs, l'activité transcriptionnelle de l'isoforme mimant une phosphorylation sur la sérine 10 (mutant S10D) est supérieure d'un facteur 4,5 à celle de l'isoforme sauvage. Ces résultats montrent que la réciprocité GlcNAcylation/phosphorylation contrôle l'activité transcriptionnelle de la delta-lactoferrine. Parallèlement à ces travaux, les résultats concernant l'étude de la fonctionnalité de la séquence PEST de la delta-lactoferrine ont montré que cette séquence est pleinement fonctionnelle. En effet, la mutation de la séquence PEST conduit à une augmentation de la durée de demi-vie de la protéine ainsi qu'à l'inhibition de la polyubiquitinylation de la delta-lactoferrine. La mutation des deux résidus de lysine proches de la séquence PEST, conservés phylogénétiquement parmi les lactoferrines, a permis de montrer que le résidu de lysine 379 est la cible majeure de la polyubiquitinylation. Par ailleurs, nos résultats montrent que la GlcNAcylation inhibe la polyubiquitinylation de la delta-lactoferrine. Les isoformes mutantes pour les sites de GlcNAcylation n'ont pas toutes la même stabilité. En effet, l'isoforme pour laquelle la sérine 10 est conservée a une durée de demi-vie comparable à celle de l'isoforme sauvage. Nous avons pu montrer que les modifications post-traductionnelles sur ce site sont au cœur des régulations de l'activité transcriptionnelle et de la stabilité de la delta-lactoferrine. En effet, nos résultats démontrent que la GlcNAcylation du résidu de sérine 10 inhibe la polyubiquitinylation de la delta-lactoferrine et la protège de la dégradation protéasomale ; la phosphorylation de ce site conduisant à sa dégradation. Finalement, des expériences d'immunoprécipitation de la chromatine ont confirmé nos résultats car elles montrent que l'isoforme GlcNAcylée de la delta-lactoferrine n'interagit pas avec l'ADN. En outre, la purification du complexe transcriptionnel contenant la delta-lactoferrine suggère que cette dernière est fixée à son élément de réponse sous une forme phosphorylée et ubiquitinylée. L'ensemble des travaux de cette étude nous a permis de mieux comprendre le mode de régulation de l'activité transcriptionnelle de la delta-lactoferrine et de proposer un modèle de régulation (Figure 27, page 136).

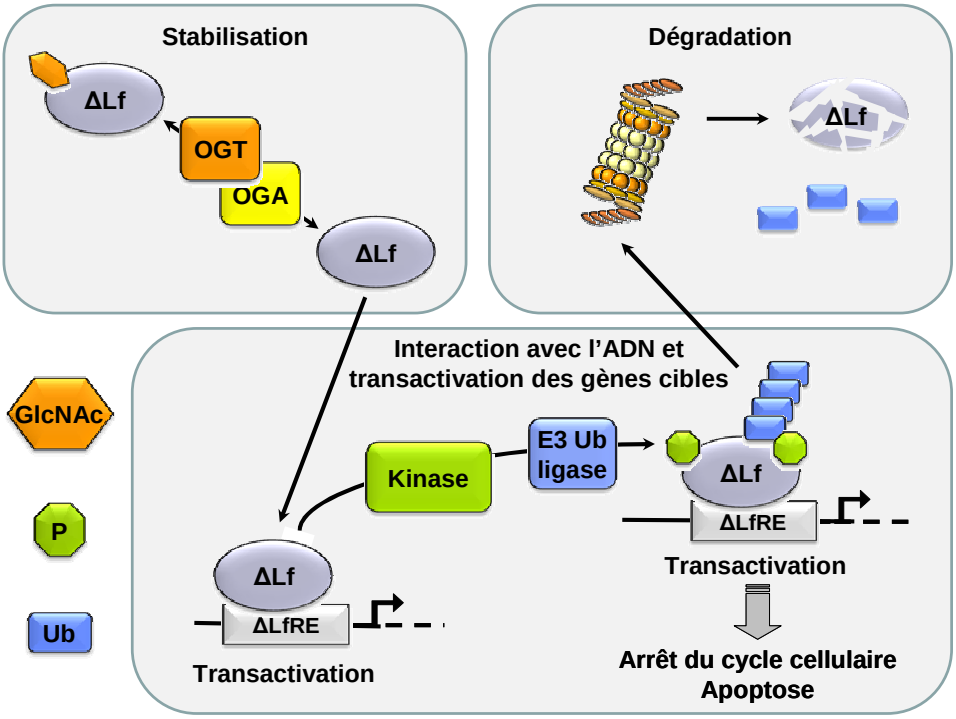


Figure 27 : Modèle de la régulation de la delta-lactoferrine. La delta-lactoferrine existerait dans la cellule sous une forme glycosylée stable mais inactive. Certains stimuli cellulaires induiraient sa déGlcNAcylation lui conférant son activité transcriptionnelle. Elle deviendrait alors la cible de kinases. La phosphorylation de la sérine 10 conduirait à la phosphorylation de la séquence PEST et au déclenchement de la polyubiquitinylation de la delta-lactoferrine. Cette dernière serait alors dégradée par le protéasome.

O-GLCNACYLATION/PHOSPHORYLATION CYCLING AT SER 10 CONTROLS BOTH TRANSCRIPTIONAL ACTIVITY AND STABILITY OF DELTA-LACTOFERRIN*

Stéphan Hardivillé, Esthelle Hoedt, Christophe Mariller, Monique Benaïssa, and Annick Pierce

From the Unité de Glycobiologie Structurale et Fonctionnelle, Unité Mixte de Recherche 8576 CNRS, Université des Sciences et Technologies de Lille, IFR 147, 59655 Villeneuve d'Ascq, France

Running head: *O*-GlcNAc/P interplay regulates ΔLf stability and activity

Address correspondance to : Annick Pierce, Pr, UGSF, UMR 8576 CNRS, UST Lille1, 59655 Villeneuve d'Ascq, France. Fax: +33 3 20 43 65 55; Tel: +33 3 20 33 72 38; E-mail: annick.pierce@univ-lille1.fr

Delta-lactoferrin (ΔLf) is a transcription factor that upregulates *DcpS*, *Skp1* and *Bax* genes provoking cell cycle arrest and apoptosis. It is posttranslationally modified either by *O*-GlcNAc or phosphate but the effects of the *O*-GlcNAc/P interplay on ΔLf function are not yet understood. Here, using a series of glycosylation mutants, we showed that Ser 10 (S^{10}) is *O*-GlcNAcylated and that this modification is associated with increased ΔLf stability, achieved by blocking ubiquitin-dependent proteolysis, demonstrating that *O*-GlcNAcylation protects against polyubiquitination. We highlighted the ^{391}KSQQSSDPDPNCVD^{404} sequence as a functional PEST motif responsible for ΔLf degradation and defined K^{379} as the main polyubiquitin acceptor site. We next investigated the control of ΔLf transcriptional activity by the *O*-GlcNAc/P interplay. Reporter gene analyses using the *Skp1* promoter fragment containing a ΔLf response element showed that *O*-GlcNAcylation at S^{10} negatively regulates ΔLf transcriptional activity whereas phosphorylation activates it. Using a ChIP assay, we showed that *O*-GlcNAcylation inhibits DNA binding. Deglycosylation leads to DNA binding and transactivation of the *Skp1* promoter at a basal level. Basal transactivation was markedly enhanced by 2-3 fold when phosphorylation was mimicked at S^{10} by aspartate. Moreover, using Re-ChIP assays, we showed that the ΔLf transcriptional complex binds to the ΔLf response element and is phosphorylated and/or ubiquitinated suggesting that ΔLf transcriptional activity and degradation are concomitant events. Collectively, our results indicate that reciprocal occupancy of S^{10} by either *O*-phosphate or *O*-GlcNAc coordinately regulates ΔLf stability and transcriptional activity.

O-GlcNAcylation is a ubiquitous post-translational modification consisting of a single N-acetylglucosamine moiety linked to Ser or Thr residues (1). It is a dynamic and reversible process mediated by the combined actions of *O*-GlcNAc transferase (OGT) and *O*-GlcNAcase (OGA). Disruption of β-*O*-linked *N*-acetylglucosamine (*O*-GlcNAc) cycling through inhibitors or gene manipulations results in cellular defects (2,3) and alterations of the *O*-GlcNAc status are associated with type-2 diabetes, neurological disorders and cancer (4).

Since numerous proteins such as transcription factors, signaling components and metabolic enzymes are modified, *O*-GlcNAcylation is therefore critical to normal cell homeostasis and gene regulation (5). It notably modulates gene expression depending on the promoter and its associated transcription initiation complexes. For instance, the carboxy-terminal domain of RNA polymerase II and a subset of general transcription factors are *O*-GlcNAcylated at transcription initiation (6). Gene silencing may be effected *via* the recruitment of OGT onto promoters by transcriptional corepressors. It then catalyzes the *O*-GlcNAcylation of specific transcription actors, modulating their activity. For instance, the association of OGT with the co-repressor mSin3A leads to the recruitment of histone deacetylase, thereby increasing transcriptional downregulation (7,8). OGA may favour gene transcription, not only by reducing the level of glycosylation of but also *via* its intrinsic histone acetyltransferase domain (9). *O*-GlcNAcylation may also modulate the activity of transcription factors *via* the regulation of their trafficking, binding affinity either to protein partners or DNA and/or turnover (8,10-13).

Increasing evidence links *O*-GlcNAcylation to the proteosome pathway. It has been shown that *O*-GlcNAcylation is associated with lower proteasomal susceptibility of transcription factors

such as Sp1 (14,15), p53 (16) and the estrogen receptor β (ER-β) (17). Most of these proteins have high PEST scores and phosphorylation of their PEST (Pro-Glu-Ser-Thr) domain targets them for polyubiquitination (18) and subsequent degradation by the proteasome, whereas *O*-GlcNAcylation prolongs their half-lives. The proteasome is itself regulated through *O*-GlcNAcylation of both its regulatory and catalytic subunits (19,20) as well as the ubiquitin (Ub)-activating enzyme E1 (21). Reduced degradation of *O*-GlcNAcylated proteins might also be due to their specific interaction with chaperones such as Hsp70 family members that display lectin activity towards the *O*-GlcNAc motif, protecting them from proteolysis (22).

In many *O*-GlcNAcylated proteins, a phosphate group can alternatively occupy the same or adjacent sites (16,17,23,24). This *O*-GlcNAc/P interplay, which leads to a rapid-response mechanism, high molecular diversity and fine-tunes protein interactions and functions, may also target delta-lactoferrin (ΔLf) and regulate its transcriptional activity and stability. ΔLf is a transcription factor that was first discovered as a transcript, the expression of which was observed only in normal cells and tissues (25). Its absence from cancer cells (25,26) is due to genetic and epigenetic alterations (27,28). ΔLf messenger is therefore a healthy tissue marker and we previously showed that its expression level is correlated with a good prognosis in human breast cancer, high concentrations being associated with longer overall survival (26). ΔLf is transcribed from the alternative promoter P2 in the Lf gene (29) and the use of the first available AUG codon in frame produces an alternative N-terminus. Thus, compared to lactoferrin (Lf), its secretory counterpart, ΔLf is a 73 kDa cytoplasmic protein able to enter the nucleus (30). Potential DBDs have been suggested for Lf implicating the strong concentration of positive charges at the C-terminal end of the first helix, which is truncated in ΔLf, and at the interlobe region (31,32).

ΔLf expression provokes anti-proliferative effects and induces cell cycle arrest in S phase (33). It is a transcription factor interacting *via* a ΔLf response element (ΔLfRE) found in the *Skp1* and *DcpS* promoters (30,34). ΔLf is also at the crossroads between cell survival and cell death since we recently linked ΔLf overexpression to upregulation of the *Bax* promoter and apoptosis (35). Since ΔLf has several crucial target genes and may act as a tumor suppressor, modifications in its activity or concentration may have marked effects on cell survival and its transcriptional activity should be strongly controlled. Results of screening ΔLf for *O*-GlcNAcylation and phosphorylation sites showed that the protein potentially undergoes both post-translational modifications. Four putative Ying Yang sites were found at S^{10}, S^{227}, S^{472} and T^{559} and their mutation led to a constitutively active ΔLf^{M4} mutant (34). Here, we map the major *O*-GlcNAc/P site to S^{10}, the PEST sequence (aa 391-404) and the main polyUb site to K^{379}. We also report that *O*-GlcNAcylation at S^{10} down-regulates ΔLf transcriptional activity and upregulates its stability by abrogating Ub-mediated proteolysis, whereas phosphorylation activates both transcription and degradation.

Experimental procedures

Cell culture, reagents, and transfection - HEK293 cells (ATC CRL-1573) were grown in monolayers and transfected as described (2 µg of DNA for 2 x 10^6 cells) (30) using DreamFect™ (OZ Biosciences). The amounts of ΔLf expression vectors were adjusted to maintain ΔLf amounts similar to those found in normal NBEC cells (26). Transfections were done in triplicate ($n \geq 4$). Cell viability was assessed by counting using Trypan blue 0.4% (v/v). Cell culture reagents were from Dutscher, Cambrex Corporation and Invitrogen. Other reagents were from Sigma. Antibodies against the 3xFLAG epitope (mouse monoclonal anti-FLAG M2 antibody, Sigma), HA epitope (goat HA-Tag polyclonal antibodies, BD Biosciences; mouse monoclonal HA.11 antibody 16B12, Covance Research Products), *O*-GlcNAc motif (mouse monoclonal RL2 antibody, Affinity Bioreagents; mouse monoclonal CTD110.6, Covance Research Products), P-ser motif (rabbit polyclonal antibodies, Millipore) and actin (goat polyclonal antibodies I-19, Santa Cruz Biotechnologies) were used for immunofluorescence, immunoprecipitation and/or immunoblotting.

Immunofluorescence and microscopy- HEK 293 cells were transfected by ΔLf C-terminal

fused GFP expression vector 24h prior the DAPI (Sigma) staining. The pΔLf-N-EGFP vector was kindly provided by Dr C. Teng (NIH, Research Triangle Park, NC). Immunofluorescence and microscopy were performed as in (30). Fluorescent microscopy images were obtained at room temperature with a Zeiss Axioplan 2 imaging system (Carl-Zeiss S.A.S., Le Pecq, France) equipped with appropriate filter cubes using a 40X objective lens.

Purification of DNA, RNA and poly(A)+ RNA - Genomic DNA was purified using the QIAprep Spin Miniprep Kit (Qiagen), total RNA using the RNeasy Mini Kit (Qiagen) and poly(A)+ RNA using the polyATrack® mRNA isolation system (Promega). The purity and integrity of each extract were checked using the nanodrop ND-1000 spectrophotometer (Labtech International) and the Bioanalyzer 2100 (Agilent Technologies).

qPCR conditions- qPCR analyses were performed as in (30). The primer pairs used for the detection of ΔLf (forward 5'-AAGCCAGTGGACAAGTTCA-3', reverse 5'-GCTTTGTTGGGATTTGTAGTT-3', annealing temperature 55°C), RPLP0 (34) and HPRT (forward 5'-GACCAGTCAACAGGGGACAT-3', reverse 5'-AACACTTCGTGGGGTCCTTTTC-3', annealing temperature 55°C) were purchased from Eurogentec.

Plasmid construction and site-directed mutagenesis- pGL3-S1^{Skp1}-Luc, pcDNA-ΔLf and p3xFLAG-CMV10-ΔLf were constructed as described (30). p3xFLAG-CMV10 (Sigma) and pcDNA3.1 (Invitrogen) were used as null vectors. The ubiquitin-hemagglutinin A (Ub-HA) expression vector was a gift from Dr. C. Couturier (IBL, Lille, France). The pcDNA-OGT expression vector was constructed using OGT cDNA isolated from the pShuttle-OGT vector (36) (a kind gift of Dr. J. Hart, John Hopkins University School of Medicine, Baltimore, USA) and further cloned into the pcDNA3.1 vector. Mutants were generated using the QuikChange® Site-directed Mutagenesis Kit (Stratagene) with pcDNA-ΔLf as template and primer pairs listed in Table I. The constructs in which several sites were mutated were done sequentially. Following sequencing, the *Hind* III-*Not* I digests were cloned either into pcDNA3.1 for reporter gene assays or into p3xFLAG-CMV10 for protein experiments.

Reporter gene assay- Reporter gene assays were performed using the pGL3-S1^{Skp1}-Luc reporter vector and the different pcDNA-ΔLf mutant constructs or a null vector as described in (34). Cell lysates were assayed using a luciferase assay kit (Promega) in a Tristar multimode microplate reader LB 941 (Berthold Technologies). Relative luciferase activities were normalized to basal luciferase expression and protein content as in (30) and expressed as a percentage; 100 % corresponds to the relative luciferase activity of ΔLf^{WT}. Basal luciferase expression was assayed using a null vector and was determined for each condition (OGT, OA, $GlcNH_2$) at each concentration. Each experiment represents at least three sets of independent triplicates.

In vivo DNA binding assays- ChIP and re-ChIP assays were performed as described (34,37) with some modifications introduced for re-ChIP. Briefly, ChIP complexes (8 x 10^6 cells) were immunoprecipitated with M2, RL2, HA-tag or anti-P-ser antibodies all used at 1:250 and twice eluted with 100 µL of 1 mM DTT for 30 min at 37°C. After centrifugation, pooled eluted fractions were diluted 40 times to reduce the DTT concentration to 25 µM with ChIP dilution buffer and then immunoprecipitated with either M2 or rabbit anti-IgG (GE Healthcare Life Sciences) or without antibodies. Amplification conditions of *Skp1* and *albumin* promoters are described in (34). ChIP or re-ChIP results presented in Fig 5 correspond to one representative experiment among three. qPCR was performed only for the ChIP assay. Amplification was carried out in triplicate (n=3) in the presence of 2 µl of purifed DNA, primer pairs used to amplify the *Skp1* promoter fragment (34) and Brilliant SYBER Green QPCR Master Mix (Stratagene) according to the manufacturer's instructions. Samples were then submitted to 40 cycles of amplification (denaturation: 30 s at 90 °C, hybridation: 30 s at 55 °C, elongation: 30 s at 72 °C) in a thermocycler Mx4000 (Stratagene). Data presented in Figure 5 D are expressed as a percentage of input.

Proteasomal degradation assay- Proteasomal activity assay was performed according to the assay instructions (Chemicon International) on HEK 293 cell lysates. Lactacystin was used as a 20S proteasome inhibitor. Fluorescence data were collected using a Tristar multimode microplate

reader LB 941 (Berthold Technologies) using a 380 nm excitation and 460 nm emission filters.

Immunoblotting and immunoprecipitation- Proteins were extracted from frozen cell pellets in RIPA buffer as described (30). For direct immunoblotting, samples mixed with 4x Laemmli buffer were boiled for 5 min. 20 µg of protein from each sample were submitted to 7.5% SDS-PAGE and immunoblotted. For immunoprecipitation experiments, 1 mg of total protein was preabsorbed with protein G Sepharose 4 Fast Flow (GE Healthcare). RL2 (1/250), M2 (1/500) or anti-HA polyclonal (1/100) antibodies were mixed with Protein G Sepharose beads for 1 h prior to an overnight incubation with the preabsorbed lysate supernatant at 4°C. The beads were then washed five times with lysis buffer. Proteins bound to the beads were eluted with 4X Laemmli buffer and analyzed by immunoblotting as above. Blots were probed at room temperature with primary antibodies (M2, 1/2000; CTD110.6, 1/3000; HA.11, 1/4000; anti-P-ser, 1/500; RL2, 1/2000 and anti-actin, 1/10000) for 2 h and with either secondary anti-IgG antibodies conjugated to horseradish peroxidase (GE Healthcare Life Sciences) or secondary anti-IgM antibodies conjugated to HRP (Chemicon International) for 1 h before detection by chemiluminescence (ECL advance, GE Healthcare Life Sciences). Each result in which immunoblots are presented corresponds to one representative experiment among at least three.

Densitometric and statistical analyses- The densitometric analyses were performed using Quantity One v4.1 software (Bio-Rad) and acquisition carried out with a GS710-calibrated densitometer (Bio-Rad). M2 densitometric values were normalized to actin and expressed as $R = D^{M2}/D^{act}$. The fold stability (X) is expressed as this ratio related to the wild type ratio and to the t_0 value as follows for ΔLf^{PEST} $X = [{}^{PEST}R_t/{}^{PEST}R_{t0}/{}^{WT}R_t/{}^{WT}R_{t0}]$. All the statistical analyses were done using Origin® 7 software (OriginLab Corporation). Means were statistically analysed using the t-test or ANOVA and differences assessed at $p<0.05$ (*) or $p<0.01$ (**).

RESULTS

Impact of the O-GlcNAc/P interplay on ΔLf transcriptional activity and stability. Investigation of the *O*-GlcNAc function has mainly relied on the manipulation of the hexosamine biosynthesis pathway *via* an increased production of UDP-GlcNAc, the substrate for OGT (38). Thus, cells exposed to increased concentrations of glucosamine ($GlcNH_2$) or overexpressing OGT exhibit enhanced levels of protein *O*-GlcNAcylation (39). On the other hand, the use of okadaic acid (OA), an inhibitor of PP2A and PP1 phosphatases, is a valuable tool for inducing protein hyperphosphorylation (40,41).

Prior to investigating whether ΔLf transcriptional activity is regulated *via* *O*-GlcNAc/P interplay, we first established that HEK293 cells possess rapid, inducible *O*-GlcNAc/P mechanisms at the OGT, $GlcNH_2$ and OA concentrations employed (42). First of all, we verified that cell viability was not perturbed (Fig. 1A). At the concentrations usually used in the literature such as 40 mM GlcNH2 and 50 nM OA, cell viability was markedly decreased in HEK 293 cells. For this reason we used lower concentrations such as 10 mM of $GlcNH_2$ and 10 nM of OA that did not affect cell viability but at which modulation of the *O*-GlcNAc/P status was visible (Fig 1B). Co-transfection of ΔLf (1 µg DNA/10^6 cells) and OGT (2.5 µg DNA/10^6 cells) expression vectors did not significantly affect cell viability (Fig. 1A). Figure 1B shows that ΔLf is indeed sensitive to OA and $GlcNH_2$ or OGT but with opposite effects. Treatment with OA led to decreased ΔLf glycosylation whereas treatment with $GlcNH_2$ or co-transfection with OGT increased it. The same GlcNAcylation pattern was observed using either the RL2 or the CDT110.6 antibody. This result demonstrates clearly that ΔLf possesses *O*-GlcNAc site(s). OA treatment, which favours phosphorylation, decreases the ΔLf glycosylation level, suggesting that glycosylation site(s) may exist in balance with phosphorylation site(s). Since RNA polymerase II activity is also controlled by this interplay we next verified that transcription of ΔLf, RPLP0 or HPRT was indeed not altered under OGT, $GlcNH_2$ or OA treatment (Fig. 1C). Our control experiments showed that modulation of the *O*-GlcNAc content does not impair cell functions at the concentration of OA, $GlcNH_2$ or OGT we used.

We then investigated ΔLf transcriptional activity using reporter gene assays and a *Skp1* promoter fragment containing the ΔLfRE known

to be highly transactivated by ΔLf (30). ΔLf transcriptional activity increased in line with OA concentration (Fig 1D) whereas it decreased in a dose dependent manner in the presence of $GlcNH_2$ (Fig. 1E). Thus, when phosphorylation was augmented, transactivation was increased 6-7 fold compared to controls whereas when *O*-GlcNAcylation was increased in ΔLf-expressing cells, transactivation of the *Skp1* promoter was strongly reduced. ΔLf transcriptional activity also decreased when cells overexpressed OGT, but at a lower level (Fig 1F).

Since, as for many transcription factors, ΔLf is rapidly degraded, we next investigated whether its turnover is dependent on both the Ub-proteasome pathway and *O*-GlcNAc/P interplay. We first verified that treatment with OA, $GlcNH_2$ or OGT did not disturb the proteasome pathway. As shown in Fig. 1G, these treatments did not alter or exarcerbate proteasomal degradation compared to the untreated and to the lactacystin treated conditions. Figure 1H shows that a ladder of polyubiquitinated ΔLf forms is visible (upper panel, lanes 3 and 6). We next evaluated whether *O*-GlcNAcylation regulates ΔLf degradation and the intensity of polyubiquitination was indeed decreased in a dose dependent manner after $GlcNH_2$ treatment (Fig 1H, upper panel, lanes 4 and 5) and after OGT overexpression (Fig 1H, upper panel, lane 7). Equivalent loadings of Ub-HA protein (Fig 1H, medium panel) and ΔLf (Fig 1H, lower panel) were confirmed by immunoblotting. These data demonstrated that ΔLf is more stable in an environment favouring *O*-GlcNAcylation.

We further investigated whether ΔLf traffic might be altered, leading to an exclusive nuclear targeting of the phosphoform. As previously described (29,30), a ΔLf-GFP fused protein localizes predominantly to the cytoplasm but also to the nucleus (Fig. 1I, panel 2). Here, we showed that the subcellular localization of ΔLf-GFP was not modified with either $GlcNH_2$ or OA (Fig. 1I, panels 3 and 4, respectively) suggesting that ΔLf traffic is not regulated by the *O*-GlcNAc/P interplay.

Mapping the key O-GlcNAc site to S^{10}. The low abundance of ΔLf, the necessity for producing a 3xFLAG-tagged protein in order to detect it and the inherent limitation of the sensitivity of tritium labelling render the detection of carbohydrate moieties on ΔLf and the subsequent mapping of its glycosylated sites extremely difficult. Therefore, in order to confirm the presence of the *O*-GlcNAc sites and characterize their roles, we made a series of glycosylation mutants in which only one *O*-GlcNAc site is preserved, named ΔLf^{S10+}, ΔLf^{S227+}, ΔLf^{S472+} and ΔLf^{T559+} respectively (Fig 2A). ΔLf^{M4} (34) and ΔLf^{WT} were used as controls. ΔLf and its glycosylation mutants were then expressed in HEK 293 cells and their levels of expression compared. Figure 2B shows that the ΔLf^{S10+} mutant was expressed at the same level as ΔLf^{WT} (short exposure time) in contrast to the other mutants (long exposure time). ΔLf^{S227+} and ΔLf^{S472+} were slightly more expressed than were ΔLf^{M4} and ΔLf^{T559+}, which were both feebly expressed. These data suggest that the post-translational modifications present on S^{10} may participate in ΔLf stability and that its absence from the other mutants leads to their rapid turnover.

O-GlcNAcylation was then investigated on the ΔLf isoforms. Since ΔLf mutants are feebly produced we first immunoprecipitated ΔLf-expressing cell lysates with RL2 in order to accumulate enough *O*-GlcNAcylated material (Fig. 2C). A reverse immunoprecipitation was then performed using the M2 antibody in order to specifically immunoprecipitate ΔLf or its glycovariants (Fig. 2D). Figure 2C shows that ΔLf was effectively glycosylated whereas ΔLf^{M4} was not, confirming that no other *O*-GlcNAc sites are present on the protein. ΔLf^{S10+}, ΔLf^{S227+} and ΔLf^{S472+} mutants were glycosylated whereas ΔLf^{T559+} was not (Fig. 2C). The reverse immunoprecipitation of the cell lysates with M2 antibody followed by *O*-GlcNAc immunodetection with the CTD 110.6 antibody (Fig. 2D) confirmed that ΔLf and its ΔLf^{S10+} mutant were glycosylated whereas ΔLf^{M4} and ΔLf^{T559+} were not. The *O*-GlcNAcylated signals corresponding to ΔLf^{S227+} and ΔLf^{S472+} mutants that were effectively visible when RL2 antibody was used, were poorly visible for ΔLf^{S227+} and not visible for ΔLf^{S472+} when the CTD 110.6 antibody was used. RL2 (43) and CTD110.6 (44) are the two most commonly used antibodies with CTD110.6 described as being the most specific. Therefore, this divergent result may be due to the fact that both isoforms were too

poorly expressed to be detected or that the S472 site was not fully glycosylated. However, we cannot exclude that performing the immunoprecipitation first with RL2 antibody may favour immunoprecipitation of *O*-GlcNAcylated ΔLf partners. Nevertheless, both immunoprecipitations confirmed without ambiguity that ΔLf^{S10+}, ΔLf^{S227+} and ΔLf^{WT} mutants were glycosylated whereas ΔLf^{T559+} and ΔLf^{M4} were not.

We next assayed the transcriptional activity of the mutants compared to wild type (WT) (Fig. 2E). The strong transcriptional activity of the glycosylation-null mutant confirmed that *O*-GlcNAcylation negatively regulates ΔLf activity as previously indicated (Figs. 1E and 1F). We compared the activity of mutants in which only one glycosylation site was preserved to that of ΔLf^{M4} in order to evaluate the impact of adding only one regulatory site at a time. The ΔLf^{S227+} and ΔLf^{S472+} mutants showed transcriptional activities nearly 2-fold greater than WT and close to that of ΔLf^{M4} (Fig. 2E) suggesting that the presence of either *O*-GlcNAc or phosphate on these two sites does not crucially regulate ΔLf transcriptional activity. These sites might be priming sites as described for phosphorylation, necessary to target OGT or specific kinases to the other sites and may only be transiently *O*-GlcNAcylated. In contrast, the transcriptional activities of ΔLf^{S10+} and ΔLf^{T559+} were strongly inhibited compared to ΔLf^{M4} suggesting that these two sites are primordial for regulation. The absence of response of the ΔLf^{T559+} mutant might be due to the feeble expression or rapid degradation of this non glycosylated and transcriptionally inactive mutant. However, this may not be the only explanation since the removal of all four glycosylation sites in ΔLf^{M4} leads to a constitutively active isoform. Preliminary results on the ΔLf^{T559+} mutant showed that it may compete with WT for DNA binding (data not shown) and further work will be necessary to understand the intrinsic role of T559. *O*-GlcNAc/P modifications on the S^{10} site led to a 5-fold inhibition of ΔLf transcriptional activity compared to ΔLf^{M4} and by 2-fold compared to ΔLf^{WT} (Fig. 2E). This site seems therefore to be a crucial target for the regulation of both ΔLf stability and transcriptional activity. We therefore focused our initial attention on S^{10} and first investigated the impact of *O*-GlcNAc/P modifications on ΔLf turnover by studying their relationship with the proteasome pathway.

ΔLf turnover is driven through a PEST motif (aa 384-404) and K^{379}. Short intracellular half-life proteins frequently have a short hydrophilic stretch of amino acids termed a PEST motif. Phosphorylation of the Ser and/or Thr residues and ubiquitination, often of the flanking Lys residues, trigger degradation. Analysis of the ΔLf sequence did not allow identification of a PEST motif in the S^{10} environment but indicated one at the C-terminus with three nearly contiguous Ser (S^{392}, S^{395} and S^{396}) and two flanking Lys (K^{379} and K^{391}) residues as potential targets either for kinase/OGT or Ub ligase, respectively. Alignement of Lf sequences from other species to this PEST motif shows that the locus is conserved (Table II).

We evaluated the functionality of the PEST sequence using a ΔLf^{PEST} mutant in which the three Ser residues were replaced by Ala and showed that this mutation leads to a slight increase in ΔLf content of about 40 % compared to WT (Figs. 3A and 3B). To measure the ΔLf turnover rate indirectly, we performed incubations (0-150 min) with cycloheximide (CHX), a potent inhibitor of *de novo* protein synthesis (45,46). The ΔLf content of HEK cells transfected with either ΔLf^{WT} or ΔLf^{PEST} constructs was analyzed following addition of CHX (Fig 3C). Differences in the steady state levels of ΔLf were readily apparent after 30 min, which may correspond to the delay necessary for observing the first effects of CHX treatment (Fig 3C, panel 1). Mutation of the Ser residues in the PEST sequence conferred stability on ΔLf (panel 5). $GlcNH_2$ treatment of HEK cells transfected with either ΔLf^{WT} or ΔLf^{PEST} constructs was also performed (Fig. 3C, panels 3 and 7, respectively), and OGT was coexpressed with ΔLf^{WT} (Fig 3C, panel 9). Actin, which is stable under the same experimental conditions, was used as an internal control (panels 2, 4, 6, 8 and 10). Densitometric data are expressed as fold stability as described in Materials and Methods (Fig. 3D). Invalidation of the PEST sequence led to a 5-6 fold gain in stability and confirmed that this sequence is determinant for ΔLf degradation. $GlcNH_2$ treatment or OGT overexpression led to a 2-fold or 3-fold increase in ΔLf^{WT} stability respectively compared to controls, visible at 90 min, confirming that increasing *O*-GlcNAcylation

protects ΔLf from degradation. However, when ΔLf^{PEST}-expressing cells were submitted to the $GlcNH_2$ treatment the stability of ΔLf^{PEST} was not significantly different in the presence or absence of $GlcNH_2$ suggesting that mutation of the PEST sequence is sufficient to confer stability on ΔLf. Moreover, these results also suggest that these two events could be linked since, if they were independent, a greater stability of the ΔLf^{PEST} isoform should be visible in the presence of $GlcNH_2$.

We next studied the invalidation of the PEST sequence on the ΔLf^{M4} mutant. This mutant is detected at low levels in transfected cells, indicating that it is either feebly expressed or rapidly degraded (Fig 2B). Figure 3E shows that this invalidation increased ΔLf^{M4} stability and rendered this mutant more resistant to proteasomal proteolysis. We further investigated whether a particular Ser within the PEST motif was involved in this process using a series of single Ser mutants (Table II). Whatever the Ser mutated, ΔLf expression was identical, suggesting that the three Ser residues were equivalent phosphorylation targets due to their proximity (Fig 3E). Moreover, prediction results for putative phosphorylation sites using the NetPhos 2.0 server (CBS.TDU; http://www.cbs.dtu.dk/services/NetPhos/) also emphasized these three Ser residues as kinase targets, albeit with a higher score for S^{396} (S^{392} : 0.766 ; S^{395} : 0.789 ; S^{396} : 0.977).

We then studied whether ΔLf-mediated ubiquitination occurs predominantly through K^{379}- or K^{391}-linked chains by constructing a series of mutants in which residues 379 or 391 were mutated to Ala either in ΔLf or ΔLf^{M4}. The K379A mutation led to a slightly increased expression level of ΔLf and completely restored stability to ΔLf^{M4} compared to controls, whereas the K391A mutation had no effect on the ΔLf expression level and only slightly increased expression of ΔLf^{M4} (Fig 3F). This result confirms that the flanking K^{379}, which is highly conserved among species (Table II), is involved in ΔLf turnover and suggests that it is the major polyUb acceptor site. We next verified that ubiquitination of ΔLf is indeed K^{379} linked. Figure 3G shows that polyubiquitination was strongly visible on ΔLf^{WT} (upper panel, lane 3), was lower on ΔLf^{PEST} (lane 4) and ΔLf^{K391A} mutants (lane 6), poorly visible on ΔLf^{K379A} (lane 5) and not at all visible on the double mutant ΔLf^{KK} (lane 7) in which both Lys residues were mutated. Control levels of ubiquitination and ΔLf expression are shown in the middle and lower panels, respectively (Fig 3G). These data confirmed that the K^{379} residue corresponds to the main Ub ligase target and K^{391} to a minor site. We next investigated which type of relationship may exist between the functional PEST sequence at the C-terminus and the *O*-GlcNAc/P site at the N-terminus.

O-GlcNAcylation of S^{10} protects ΔLf from polyubiquitination. To determine whether ΔLf protein stability was controlled *via* an *O*-GlcNAc/P switch on S^{10}, other mutants were constructed (Fig 4A) such as the ΔLf^{S10A} mutant that has only the S^{10} residue mutated and the ΔLf^{S10D} mutant in which an Asp residue was introduced in place of Ser in order to mimic constitutive phosphorylation as previously described (47,48). Immunoblotting of the different mutants with M2 antibodies is presented in Fig 4B. ΔLf^{S10A} and ΔLf^{S10+} had similar expression levels to WT whereas ΔLf^{S10D} had an extremely short half-life (Fig. 4C), suggesting that this mutant is an interesting tool for studying degradation of the ΔLf phosphoform. Due to the absence of S^{10}, ΔLf^{S10A} was expected, like the ΔLf^{M4}, ΔLf^{T559+}, ΔLf^{S227+} and ΔLf^{S472+} mutants (Fig. 2B), to be less stable than WT. In ΔLf^{S10A}, only S^{10} is mutated. Therefore, the stability of ΔLf^{S10A} might be due to the other sites, which could be used as « protecting sites » in the absence of S^{10}.

The turnover of these different S^{10} mutants compared to WT and actin (internal control) is shown in Fig 4C. Differences in the steady state levels of ΔLf and ΔLf^{S10+} mutant were readily apparent around 30-60 min and strongly visible after 90 min (Fig 4C, panels 1 and 5). Invalidation of the S^{10} site in ΔLf^{S10A} resulted in a markedly prolonged half-life (panel 3). Comparable results were obtained when cells expressing ΔLf^{S10+} were cultured in the presence of $GlcNH_2$, confirming the crucial role of *O*-GlcNAcylation in ΔLf stability (panel 7). Interestingly, ΔLf^{S10D} had a faster turnover rate compared to WT (panel 9) indicating that mimicking phosphorylation at this locus triggers degradation. Figure 4D summarizes the densitometric data of the ΔLf immunoblots

expressed as fold stability as described in Materials and Methods. ΔLf^{S10+}, which could either be phosphorylated or glycosylated, was slightly more stable than WT (1.5-fold) whereas the same mutant expressed in hyper-*O*-GlcNAcylation conditions was 4-fold more stable than WT. Interestingly, the mutation of S^{10} to Ala also led to ΔLf stability (3.5-fold compared to WT) which suggests that stability is not due to the presence of the *O*-GlcNAc moiety but to the absence of the phosphate group. Mimicking phosphate at S^{10} in the ΔLf^{S10D} mutant shortened its half-life. From these experiments we conclude that phosphorylation at S^{10} accelerates ΔLf degradation whereas *O*-GlcNAcylation at S^{10} controls its stability, confirming the existence of a strong link between the *O*-GlcNAc/P interplay and the Ub degradation pathway.

We next investigated whether ubiquitination of ΔLf is linked to S^{10} phosphorylation. Figure 4E shows that polyubiquitination was marked on ΔLf^{WT} and ΔLf^{S10+} (upper panel, lanes 3 and 4) whereas it was reduced on ΔLf^{S10A} (lane 5). Control levels of ubiquitination and ΔLf expression are shown in the middle and lower panels (Fig 4E). Unfortunately, the high turnover rate of the ΔLf^{S10D} mutant or of ΔLf^{WT} under OA treatment precluded the observation of a polyubiquitination signal (data not shown).

In conclusion our data showed that ΔLf turnover is driven through a PEST sequence located at the C-terminus with polyubiquitination occurring mainly at K^{379}. We also demonstrated that the degradation process is regulated *via* the *O*-GlcNAc/P interplay which targets S^{10}. As a glycoform, ΔLf is stable whereas as a phosphoform it is sensitive to degradation. Since proteasomal degradation is triggered by phosphorylation we suggest that phosphorylation of S^{10} favours phosphorylation of the PEST sequence whereas *O*-GlcNAcylation of S^{10} prevents it.

Phosphorylation at S^{10} controls ΔLf transcriptional activity. Since OA treatment increases ΔLf transcriptional activity, we next questioned whether the phosphoform might be responsible for gene transactivation. Using immunoprecipitation with the M2 antibody and probing the resulting blot with an anti-P-Ser antibody, we studied the phosphorylation status of ΔLf (Fig. 5A, left panel). Phosphatase treatment markedly abrogated the phosphorylation signal confirming antibody specificity (right panel). Immunoblotting (Fig. 5A) showed that ΔLf and its S^{10} mutants exist as phosphoforms. The decreased phosphorylation signals observed under $GlcNH_2$ treatment confirm that phosphorylation and *O*-GlcNAcylation may alternate on some of the sites. Therefore the weaker phosphorylation signal observed with the hyperglycosylated ΔLf^{S10+} isoform (lane 9) compared to control (lane 5) strongly suggests that the *O*-GlcNAc/phosphate interplay targets the S10 site. However, since ΔLf^{M4} is phosphorylated, ΔLf is also phosphorylated on sites different from the *O*-GlcNAc/P interplay sites.

We next performed gene reporter analyses as described above and investigated whether phosphorylation at S^{10} controls ΔLf transcriptional activity. Figure 5B shows that, compared to ΔLf^{WT}, ΔLf^{S10+} transcriptional activity was inhibited 2-fold as in Fig 2E, whereas the transcriptional activity of ΔLf^{S10A} was increased 1.5-2 fold and that of ΔLf^{S10D} 4.5-5 fold. The prevention of glycosylation of S^{10} favoured transcription suggesting that *O*-GlcNAcylation at this site inhibits ΔLf transcriptional activity. Mimicking phosphorylation at S^{10} rendered ΔLf even more active than ΔLf^{M4} (Fig. 2E) and strongly suggests that the presence of a phosphate group on this site favours transactivation (Fig. 5B). This result reinforces the status of ΔLf^{S10D} as a constitutive phosphorylated mutant. Since S^{10} is present in a basic environment (^{1}MRKVRGPPVSCIKR^{14}) within a putative truncated DBD we constructed a $\Delta Lf^{\Delta 1\text{-}14}$ mutant in which the first fourteen amino acid residues were deleted. Surprisingly, this deletion did not affect ΔLf transcriptional activity (Fig. 5B) suggesting that the ΔLf DBD must be located at the hinge region (31,32). Since *O*-GlcNAcylation and phosphorylation might occur on neighbouring sites we screened the vicinity of S^{10} and identified S^{16} that might be used as a replacement target by kinases. We therefore constructed a ΔLf^{S16D} mutant in order to mimic phosphorylation at this site. Expression of this mutant led to a basal expression level of the reporter gene (Fig 5B) showing that constitutive phosphorylation at this locus does not lead to increased transactivation as

for ΔLf^{S10D} and does not take over when the major acceptor site is invalidated, confirming the key role of S^{10}.

Since ΔLf transcriptional activity is altered by *O*-GlcNAcylation at S^{10} and that an OGT/ OGA complex has been described in the vicinity of transcription factors bound to their response elements (8,9) we next considered whether glycosylated ΔLf binds DNA. Using a chromatin immunoprecipitation (ChIP) assay we investigated the binding of the different S^{10} mutants compared to WT. As shown in figure 5C, specific ChIP PCR products were detected for each mutant. It is interesting to note that the PCR product signals for ΔLf^{WT} and ΔLf^{S10+} were equivalent whereas treatment with $GlcNH_2$ led to a weaker signal for both, suggesting that fewer promoter sites were occupied. Since ΔLf^{WT} and ΔLf^{S10+} were equivalently expressed (Fig. 4B) even under $GlcNH_2$ treatment (Figs. 1B and 4C, respectively) we suggest that glycosylation inhibits binding to DNA and that amongst the ΔLf intracellular pool, only the S^{10}-phosphoforms bind ΔLfRE. These results were confirmed by the detection of a PCR product signal comparable to that of WT for ΔLf^{S10D}, which was poorly expressed but extremely active (Fig. 4B), suggesting that a large proportion of ΔLf^{S10D} binds ΔLfRE (Fig. 5C). The detection of a weaker signal for ΔLf^{S10A}, which was expressed similarly to WT, shows that without phosphorylation and glycosylation at S^{10} ΔLf still binds DNA but its capacity to occupy promoter sites is reduced. Real time PCR was next performed to quantify promoter site occupancy (Fig. 5D). The qPCR data confirmed the PCR results except that promoter site occupancy for ΔLf^{S10+} and ΔLf^{S10D} was twice as high as that of WT. Treatment with $GlcNH_2$ led to a 0.5-fold promoter site occupancy compared to WT, confirming that favouring GlcNAcylation prevents DNA binding.

In addition, we performed a double chromatin immunoprecipitation (re-ChIP) assay to investigate whether ΔLf or a ΔLf-associated transcriptional complex binds to the endogenous human *Skp1* promoter *in vivo* as a phosphoform. Moreover, since the half-life of ΔLf is short as a phosphoform we studied the possibility that ΔLf also exists as an ubiquitinated isoform on DNA. Using a re-ChIP assay we showed that phosphorylated and ubiquitinated but not *O*-GlcNAc ΔLf-complexes were specifically co-localized on the *Skp1* promoter fragment (Fig. 5E). The slight amplification observed in panel 1 (NIP and IR) might be due to the fact that the two immunoprecipitations were performed with the same antibody, increasing the background level. Our results clearly demonstrate that phosphorylated and/or ubiquitinated ΔLf or ΔLf associated with phosphorylated and/or ubiquitinated proteins specifically bind the *Skp1* promoter segment with close proximity *in vivo* whereas glycosylated ΔLf or ΔLf associated with glycosylated proteins do not. Since ΔLf is ubiquitinated at K^{379} and phosphorylated at S^{10} we suggest that these two PTMs might be concomittantly present on ΔLf bound to DNA and may both be determinant in its activity. Further work will have to be done to demonstrate such a partnership and for that, specific antibodies against the phospho-S^{10} or the Ub-K^{379} or polyUb-K^{379} will be obtained.

DISCUSSION

O-GlcNAc/P modification of transcription factors modulates their transcriptional activity by regulating their turnover, traffic, binding to DNA or cofactor recruitment. ΔLf, is a transcription factor controlling the expression of key molecular actors, and as such should be highly regulated. In this study, we demonstrated that it is alternatively *O*-GlcNAcylated or *O*-phosphorylated at S^{10} and that these two alternative modifications play distinct roles in modulating its turnover and transcriptional activity.

The concentration of transcription activators and the rate of their degradation are under the control of the proteasome and there is direct evidence that a switch between *O*-GlcNAcylation and phosphorylation regulates the process. Phosphorylation drives proteins to degradation *via* the capping of PEST hydroxyl groups whereas *O*-GlcNAcylation hinders it mainly by competing for and masking these hydroxyl groups from kinases. Numerous proteins, such as the transcription factor Sp1 (14), the estrogen receptor (49), the eukaryotic initiation factor (eIF)2a–p67 (50) or p53 (16) are protected from proteasomal degradation by *O*-GlcNAcylation. Here, we show

that ΔLf has a short half-life compatible with its function and is stabilized when S^{10} is *O*-GlcNAcylated. Moreover, we showed that S^{10} is not present within a phosphodegron, which is a recognition signal for Ub ligases. The ΔLf degradation motif (^{391}KSQQSSDPDPNCVD^{404}) is conserved in Lf from different species and the mutation of all three Ser residues led to increased stability of the protein, clearly confirming the functionality of this motif. Mutation of each Ser separately indicated that they behave similarly, suggesting that they are equivalent targets of kinases due to their proximity, but we do not know whether they are also substituted with GlcNAc moities. Ying Yang server predictions indicated S^{292} and S^{395} as OGT targets but with low scores. The ΔLf^{M4} isoform is not glycosylated, suggesting that no further glycosylation sites are present but we cannot exclude glycosylation of the PEST motif only when S^{10} is glycosylated or that the ΔLf^{M4} isoform, which is extremely unstable, exists only as a phosphorylated-PEST isoform.

We next investigated Ub targets by mutating lysine residues neighbouring the PEST motif and demonstrated that ΔLf ubiquitination occurs on K^{379} and K^{391} with K^{379} as the main target. The ΔLf^{KK} double mutant was devoid of Ub confirming that only these two residues are involved. The formation of Ub-ladders observed with ΔLf^{K379} and ΔLf^{K391} also revealed that, despite the possibility of its multi-monoubiquitination, ΔLf undergoes polyubiquitination. Unexpectedly, the ΔLf^{PEST} mutant was still ubiquitinated suggesting the existence of other degradation motifs. ΔLf is involved in S phase control and should be ubiquitinated *via* the SCF complex but it is possible that another complex such as APC/C might be involved. Interestingly, ΔLf possesses a ^{475}RSNLCAL^{491} sequence which may behave as a potential RxxLxx[LIVM] D-box motif (ELM D-box entry), the target of APC/C (51). The presence of two degradation motifs suggests that ΔLf may be degraded throughout the cell cycle. Nevertheless, further work has to be done in order to prove the functionality of this D-Box.

The relationship linking *O*-GlcNAcylation and the Ub pathway has not yet been elucidated. Whilst Yang et al (16) demonstrated that *O*-GlcNAcylation inhibits ubiquitination of p53, a recent study by Guinez et al (21) shows that *O*-GlcNAc and Ub can coexist on the same protein and suggests that the Ub/*O*-GlcNAc ratio may send proteins either to destruction or repair. Here, we demonstrated that enhancement of the *O*-GlcNAc status within the cells inhibited ΔLf ubiquitination and the absence of S^{10} as in the ΔLf^{S10A} mutant was accompanied by a decrease in polyubiquitination, suggesting that this modification of ΔLf^{S10+} and ΔLf^{WT} only occurs on the phosphoforms. Phosphorylation at S^{10}, by acting through the creation of a negatively charged region and/or the triggering of transient conformational changes may lead to phosphorylation at the PEST locus, conferring a priming site role on S^{10}.

The *O*-GlcNAc/P interplay also modulates transcriptional activity. *O*-GlcNAcylation directly activates FoxO1 (52), p53 (53,54) and NF-κB (55) and Sp1 indirectly *via* cofactors (14,15), whereas it inhibits c-myc (24) and mER-β (49). In this work we have demonstrated that GlcNAcylation inhibits ΔLf transcriptional activity whereas phosphorylation activates it and that S^{10} is central to this regulation. An absence of modification at S^{10} leads to gene transactivation whereas phosphomimetism increases it, confirming the inhibitory role of glycosylation. Since the expression of ΔLf^{S10+} is much greater than that of ΔLf^{S10D} we suggest that ΔLf exists normally in the cell as a pool of stable but inactive glycoforms that, under appropriate stimuli become activated by phosphorylation and sensitive to degradation. However, another explanation is that only the phosphoform is present in the nucleus. Nucleocytoplasmic traffic may be regulated *via* *O*-GlcNAcylation since the modification of the *O*-GlcNAc status leads to a change in the cellular distribution of Tau (56), Alpha4 and Sp1 (57) but does not influence Stat5a traffic (58). Here, we showed that ΔLf-GFP traffic was not affected by $GlcNH_2$ or OA treatment. But even if the nucleocytoplasmic traffic is not governed by the *O*-GlcNAc/P interplay, since the OGT/OGA complex and kinases are present within both compartments (59,60), nuclear ΔLf might exist only as a phosphoform.

Phosphorylated transcription factors are usually more competent to bind DNA and activate transcription than their non-phosphorylated counterparts but there is direct evidence for the involvement of *O*-GlcNAcylation. PDX-1 *O*-

GlcNAcylation increases its ability to bind DNA (61) and enhances p53 DNA binding by hiding an inhibitory domain at the N-terminus (54). *O*-GlcNAcylation of HIC1 does not affect its specific DNA binding (62) and whatever the modifications present on Stat5, it binds its response element similarly (58). However *O*-GlcNAcylation at the C-terminus of Sp1 abolishes homopolymerization and dramatically affects its function (15). In this study, we demonstrated that *in vivo* ΔLf binding to ΔLfRE occured with the unmodified or the phosphomimetic S^{10} isoform but decreased when *O*-GlcNAcylation was increased, suggesting that the glycoform is unable to bind DNA. ΔLf^{S10A} bound DNA and transactivated transcription at a basal level, but given the dynamic nature of the *O*-GlcNAc/P interplay, it is doubtful whether an unmodified ΔLf isoform exists. Nevertheless, we can infer that transactivation by ΔLf is a two step process, starting at a basal level and increasing with phosphorylation as depicted in Fig 6. Moreover, using the Re-ChIP assay, we were able to show that the ΔLf transcriptional complex linked to ΔLfRE is phosphorylated. Our data demonstrate that *O*-GlcNAcylation at S^{10} inhibits DNA binding whereas phosphorylation favours it and promotes transactivation.

The *O*-GlcNAc/P content fluctuates during cell cycle progression. A recent study showed that increasing *O*-GlcNAc levels induce a slowing down of both S and G_2/M phases whereas a reduced *O*-GlcNAc level impairs the G1/S checkpoint transition (63). Since temporal control of Ub-proteasome-mediated protein degradation is critical for normal G_1 and S phase progression, ΔLf modifications may switch between glycosylation and phosphorylation depending on the cell cycle phase. Progression of the cell cycle requires degradation of cyclins and cyclin inhibitors. At the G_1/S check point Skp1, one of the targets of ΔLf, is involved in the process when associated with the SCF complex (64,65). Thus, *O*-GlcNAcylation of ΔLf, by downregulating Skp1 expression, may alter SCF activity whereas phosphorylation of ΔLf may increase it. Regulation of the transcriptional activity of ΔLf by the *O*-GlcNAc/P interplay may therefore modulate the Ub-protesome-mediated degradation of cell cycle regulators. Furthermore, we demonstrated that ΔLf is itself ubiquitinated, thus its turnover could be regulated by feedback control *via* overexpression of Skp1. On the other hand, ubiquitination also occurs on the ΔLf-DNA complex. Modification by Ub is not only a destruction signal but also determines membrane receptor internalization, sorting at the endosomal compartment, activation of DNA repair or transactivation of transcription factors such as c-myc and SRC-3 (66-68). As an example, SRC-3 is first activated by multi-(mono) ubiquitination and then polyubiquitinated prior to degradation. Therefore, ΔLf might require concomitant pre-ubiquitination and phosphorylation as a transcriptional activation signal before being degraded as a polyubiquitinated isoform.

REFERENCES

1. Hart, G. W., Haltiwanger, R. S., Holt, G. D., and Kelly, W. G. (1989) *Ciba Found Symp* **145**, 102-112, discussion 112-108
2. Hanover, J. A. (2001) *FASEB J* **15**, 1865-1876
3. O'Donnell, N., Zachara, N. E., Hart, G. W., and Marth, J. D. (2004) *Mol Cell Biol* **24**, 1680-1690
4. Lefebvre, T., Dehennaut, V., Guinez, C., Olivier, S., Drougat, L., Mir, A. M., Mortuaire, M., Vercoutter-Edouart, A. S., and Michalski, J. C. (2009) *Biochim Biophys Acta*
5. Comer, F. I., and Hart, G. W. (1999) *Biochim Biophys Acta* **1473**, 161-171
6. Hart, G. W. (1997) *Annu Rev Biochem* **66**, 315-335
7. Roos, M. D., and Hanover, J. A. (2000) *Biochem Biophys Res Commun* **271**, 275-280
8. Yang, X., Zhang, F., and Kudlow, J. E. (2002) *Cell* **110**, 69-80
9. Toleman, C., Paterson, A. J., Whisenhunt, T. R., and Kudlow, J. E. (2004) *J Biol Chem* **279**, 53665-53673

10. Wells, L., Kreppel, L. K., Comer, F. I., Wadzinski, B. E., and Hart, G. W. (2004) *J Biol Chem* **279**, 38466-38470
11. Hart, G. W., Housley, M. P., and Slawson, C. (2007) *Nature* **446**, 1017-1022
12. Slawson, C., Housley, M. P., and Hart, G. W. (2006) *J Cell Biochem* **97**, 71-83
13. Zachara, N. E., and Hart, G. W. (2004) *Trends Cell Biol* **14**, 218-221
14. Han, I., and Kudlow, J. E. (1997) *Mol Cell Biol* **17**, 2550-2558
15. Yang, X., Su, K., Roos, M. D., Chang, Q., Paterson, A. J., and Kudlow, J. E. (2001) *Proc Natl Acad Sci U S A* **98**, 6611-6616
16. Yang, W. H., Kim, J. E., Nam, H. W., Ju, J. W., Kim, H. S., Kim, Y. S., and Cho, J. W. (2006) *Nat Cell Biol* **8**, 1074-1083
17. Cheng, X., Cole, R. N., Zaia, J., and Hart, G. W. (2000) *Biochemistry* **39**, 11609-11620
18. Rechsteiner, M., and Rogers, S. W. (1996) *Trends Biochem Sci* **21**, 267-271
19. Sumegi, M., Hunyadi-Gulyas, E., Medzihradszky, K. F., and Udvardy, A. (2003) *Biochem Biophys Res Commun* **312**, 1284-1289
20. Zhang, F., Su, K., Yang, X., Bowe, D. B., Paterson, A. J., and Kudlow, J. E. (2003) *Cell* **115**, 715-725
21. Guinez, C., Mir, A. M., Dehennaut, V., Cacan, R., Harduin-Lepers, A., Michalski, J. C., and Lefebvre, T. (2008) *FASEB J* **22**, 2901-2911
22. Guinez, C., Lemoine, J., Michalski, J. C., and Lefebvre, T. (2004) *Biochem Biophys Res Commun* **319**, 21-26
23. Comer, F. I., and Hart, G. W. (2001) *Biochemistry* **40**, 7845-7852
24. Kamemura, K., and Hart, G. W. (2003) *Prog Nucleic Acid Res Mol Biol* **73**, 107-136
25. Siebert, P. D., and Huang, B. C. (1997) *Proc Natl Acad Sci U S A* **94**, 2198-2203
26. Benaïssa, M., Peyrat, J. P., Hornez, L., Mariller, C., Mazurier, J., and Pierce, A. (2005) *Int J Cancer* **114**, 299-306
27. Teng, C., Gladwell, W., Raphiou, I., and Liu, E. (2004) *Biometals* **17**, 317-323
28. Yang, Y., Li, J., Szeles, A., Imreh, M. P., Kost-Alimova, M., Kiss, H., Kholodnyuk, I., Fedorova, L., Darai, E., Klein, G., and Imreh, S. (2003) *Cancer Lett* **191**, 155-164
29. Liu, D., Wang, X., Zhang, Z., and Teng, C. T. (2003) *Biochem Biophys Res Commun* **301**, 472-479
30. Mariller, C., Benaïssa, M., Hardivillé, S., Breton, M., Pradelle, G., Mazurier, J., and Pierce, A. (2007) *FEBS J* **274**, 2038-2053
31. Baker, E. N., and Baker, H. M. (2005) *Cell Mol Life Sci* **62**, 2531-2539
32. Baker, E. N., Baker, H. M., and Kidd, R. D. (2002) *Biochem Cell Biol* **80**, 27-34
33. Breton, M., Mariller, C., Benaïssa, M., Caillaux, K., Browaeys, E., Masson, M., Vilain, J. P., Mazurier, J., and Pierce, A. (2004) *Biometals* **17**, 325-329
34. Mariller, C., Hardivillé, S., Hoedt, E., Benaïssa, M., Mazurier, J., and Pierce, A. (2009) *Biochimie* **91**, 109-122
35. Hardivillé, S., Hoedt, E., Mariller, C., Benaïssa, M. and Pierce. A. (2010) *Eur. J. Cell Biol. Chem.*
36. Zachara, N. E., O'Donnell, N., Cheung, W. D., Mercer, J. J., Marth, J. D., and Hart, G. W. (2004) *J Biol Chem* **279**, 30133-30142
37. Metivier, R., Penot, G., Hubner, M. R., Reid, G., Brand, H., Kos, M., and Gannon, F. (2003) *Cell* **115**, 751-763
38. Marshall, S., Bacote, V., and Traxinger, R. R. (1991) *J Biol Chem* **266**, 4706-4712
39. Kreppel, L. K., and Hart, G. W. (1999) *J Biol Chem* **274**, 32015-32022
40. Haystead, T. A., Sim, A. T., Carling, D., Honnor, R. C., Tsukitani, Y., Cohen, P., and Hardie, D. G. (1989) *Nature* **337**, 78-81
41. Lefebvre, T., Alonso, C., Mahboub, S., Dupire, M. J., Zanetta, J. P., Caillet-Boudin, M. L., and Michalski, J. C. (1999) *Biochim Biophys Acta* **1472**, 71-81
42. Huynh-Delerme, C., Fessard, V., Kiefer-Biasizzo, H., and Puiseux-Dao, S. (2003) *Environ Toxicol* **18**, 383-394

43. Holt, G. D., Snow, C. M., Senior, A., Haltiwanger, R. S., Gerace, L., and Hart, G. W. (1987) *J Cell Biol* **104**, 1157-1164
44. Comer, F. I., Vosseller, K., Wells, L., Accavitti, M. A., and Hart, G. W. (2001) *Anal Biochem* **293**, 169-177
45. Fu, Y., Fang, Z., Liang, Y., Zhu, X., Prins, P., Li, Z., Wang, L., Sun, L., Jin, J., Yang, Y., and Zha, X. (2007) *J Cell Biochem* **102**, 704-718
46. Yonashiro, R., Sugiura, A., Miyachi, M., Fukuda, T., Matsushita, N., Inatome, R., Ogata, Y., Suzuki, T., Dohmae, N., and Yanagi, S. (2009) *Mol Biol Cell*
47. Egelhoff, T. T., Lee, R. J., and Spudich, J. A. (1993) *Cell* **75**, 363-371
48. Huang, W., and Erikson, R. L. (1994) *Proc Natl Acad Sci U S A* **91**, 8960-8963
49. Cheng, X., and Hart, G. W. (2001) *J Biol Chem* **276**, 10570-10575
50. Datta, R., Choudhury, P., Ghosh, A., and Datta, B. (2003) *Biochemistry* **42**, 5453-5460
51. Peters, J. M. (2006) *Nat Rev Mol Cell Biol* **7**, 644-656
52. Kuo, M., Zilberfarb, V., Gangneux, N., Christeff, N., and Issad, T. (2008) *Biochimie* **90**, 679-685
53. Fiordaliso, F., Leri, A., Cesselli, D., Limana, F., Safai, B., Nadal-Ginard, B., Anversa, P., and Kajstura, J. (2001) *Diabetes* **50**, 2363-2375
54. Shaw, P., Freeman, J., Bovey, R., and Iggo, R. (1996) *Oncogene* **12**, 921-930
55. James, L. R., Tang, D., Ingram, A., Ly, H., Thai, K., Cai, L., and Scholey, J. W. (2002) *Diabetes* **51**, 1146-1156
56. Lefebvre, T., Ferreira, S., Dupont-Wallois, L., Bussiere, T., Dupire, M. J., Delacourte, A., Michalski, J. C., and Caillet-Boudin, M. L. (2003) *Biochim Biophys Acta* **1619**, 167-176
57. Dauphinee, S. M., Ma, M., and Too, C. K. (2005) *J Cell Biochem* **96**, 579-588
58. Gewinner, C., Hart, G., Zachara, N., Cole, R., Beisenherz-Huss, C., and Groner, B. (2004) *J Biol Chem* **279**, 3563-3572
59. Kreppel, L. K., Blomberg, M. A., and Hart, G. W. (1997) *J Biol Chem* **272**, 9308-9315
60. Wells, L., Gao, Y., Mahoney, J. A., Vosseller, K., Chen, C., Rosen, A., and Hart, G. W. (2002) *J Biol Chem* **277**, 1755-1761
61. Gao, Y., Miyazaki, J., and Hart, G. W. (2003) *Arch Biochem Biophys* **415**, 155-163
62. Lefebvre, T., Pinte, S., Guerardel, C., Deltour, S., Martin-Soudant, N., Slomianny, M. C., Michalski, J. C., and Leprince, D. (2004) *Eur J Biochem* **271**, 3843-3854
63. Slawson, C., Zachara, N. E., Vosseller, K., Cheung, W. D., Lane, M. D., and Hart, G. W. (2005) *J Biol Chem* **280**, 32944-32956
64. Ang, X. L., and Wade Harper, J. (2005) *Oncogene* **24**, 2860-2870
65. Krek, W. (1998) *Curr Opin Genet Dev* **8**, 36-42
66. Kim, S. Y., Herbst, A., Tworkowski, K. A., Salghetti, S. E., and Tansey, W. P. (2003) *Mol Cell* **11**, 1177-1188
67. Schnell, J. D., and Hicke, L. (2003) *J Biol Chem* **278**, 35857-35860
68. Wu, R. C., Feng, Q., Lonard, D. M., and O'Malley, B. W. (2007) *Cell* **129**, 1125-1140
69. Querfurth, H. W., Jiang, J., Geiger, J. D., and Selkoe, D. J. (1997) *J Neurochem* **69**, 1580-1591

FOOTNOTES

*This investigation was supported in part by the CNRS UMR 8576 (Unité de Glycobiologie Structurale et Fonctionnelle), the Institut Fédératif de Recherche n° 147, the Université des Sciences et Technologies de Lille I, the Comité du Nord de la Ligue Nationale contre le Cancer and the Association pour la Recherche sur le Cancer (« *Etude du rôle de la delta-lactoferrine, des gangliosides et des neurotrophines dans le développement et la progression du cancer du sein*"). We would like to thank Dr. R. J. Pierce (CIIL, INSERM U1019, CNRS UMR 8024, Institut Pasteur de Lille, F) and Pr. T. Lefebvre (UMR 8576 CNRS, USTL, F) for reviewing this manuscript.

[1]The abbreviations used are: ΔLf, delta-lactoferrin; Lf, lactoferrin; OGA, *O*-GlcNAc hydrolase; OGT, *O*-GlcNAc transferase; Ub, ubiquitin; HPRT, Hypoxanthine-Guanine-Phospho-Ribosyl-Transferase; RPLP0, ribosomal protein, large, P0; ΔLfRE, ΔLf response element, OA, okadaic acid; $GlcNH_2$, glucosamine; *O*-GlcNAc, β-*O*-linked *N*-acetylglucosamine; P, phosphate; ChIP, chromatin immunoprecipitation; WT, wild type, NT, not transfected.

KEYWORDS

O-GlcNAcylation, *O*-GlcNAc, delta-lactoferrin, transcriptional activity, proteasome.

FIGURE LEGENDS

Fig. 1. *O*-GlcNAc/P interplay regulates ΔLf transcriptional activity. HEK 293 cells were incubated with glucosamine ($GlcNH_2$), okadaic acid (OA) or transfected with an OGT- construct (OGT) to assess the impact of the *O*-GlcNAc/P interplay on ΔLf. (A) Cell viability. Cell viability of 10^4 HEK 293 was assayed 24 h after $GlcNH_2$ or OA treatment or after transfection with pcDNA-OGT at 2.5 or 5 µg DNA/10^6 cells (n=9). (B) ΔLf *O*-GlcNAcylation status. Treated and untreated 3xFLAG-ΔLf-expressing HEK293 cell extracts were M2 immunoprecipitated prior to SDS-PAGE and CTD110.6 or RL2 immunodetection. Input was used as the loading control. n=3 (C) Gene expression is not altered under $GlcNH_2$ and OA treatment or OGT overexpression. Poly(A)+ RNA was purified from total RNA of ΔLf-expressing cells treated with OGT, $GlcNH_2$ or OA 24 h after transfection and assayed using real time PCR. RPLPO and HPRT are internal controls (n=3). (D-F) ΔLf transcriptional activity is modulated by OGT, $GlcNH_2$ or OA treatment. Cells were co-transfected with pcDNA-ΔLf and pGL3-$S1^{Skp1}$-Luc and incubated with $GlcNH_2$ or OA or co-transfected with pcDNA-ΔLf, pcDNA-OGT and pGL3-$S1^{Skp1}$-Luc for 24h prior to lysis. Relative luciferase activities are expressed as described in Materials and Methods (n≥9). (G) The endoproteolytic activity of the proteasome is not altered under OGT, $GlcNH_2$ and OA treatment. The histograms represent the proteasomal activity assayed by following the fluorescence emitted during the degradation of a synthetic fluorescent peptide (69). Lactacystin is a proteasome inhibitor (n=3). (H) Ub-dependent degradation of ΔLf is $GlcNH_2$ sensitive. Cells were co-transfected with or without 3xFLAG-tagged ΔLf (ΔLf-3xFLAG), and HA-tagged-ubiquitin (Ub-HA) vectors, treated or not with $GlcNH_2$, or transfected or not with pcDNA-OGT. Cells were incubated 2 h with 10 µM MG132 before lysis in order to inhibit proteasomal degradation. Total cell extracts were immunoprecipitated with anti-HA polyclonal antibodies or used as input. Samples were immunoblotted with M2 (upper and lower panels) or HA.11 (middle panel) antibodies. (I) ΔLf traffic is not affected by $GlcNH_2$ or OA treatment. Cells were transfected with pEGFP empty or pEGFP-ΔLf vector and incubated with $GlcNH_2$ or OA. Fluorescent microscopy was performed after DAPI staining.

Fig. 2. Post-translational modifications of S^{10} modulate ΔLf transcriptional activity and stability. (A) Schematic representation of ΔLf glycosylated mutant constructs. (B) Expression of 3xFLAG-ΔLf^{WT} and its *O*-GlcNAcylation mutants. ΔLf^{WT} and mutant constructs were transfected for 24 h prior to lysis. Whole cell extract was immunoblotted with either M2 or anti-actin antibodies used as loading control. Development at two exposure times is shown for M2. (C-D) Mapping of ΔLf *O*-GlcNAcylated sites. Cells were transfected by the above constructs and lysed 24 h later. (C) Lysates were immunoprecipitated with RL2 and immunoblotted with M2. (D) Lysates were immunoprecipitated with M2 and immunoblotted with CTD110.6. (E) Relative luciferase activity of ΔLf and its mutants. Cells were co-transfected with pGL3-$S1^{Skp1}$-Luc reporter vector and pcDNA-ΔLf (ΔLf^{WT}) vector or the *O*-GlcNAc mutant constructs. Relative luciferase activities are expressed as described in Materials and Methods (n≥9; $p < 0.01$ (**)).

Fig. 3. Ub-dependent ΔLf degradation is mediated through a PEST sequence at the C-terminus and K^{379} and is inhibited by *O*-GlcNAcylation. (A-B) Deletion of the PEST sequence slightly increases ΔLf stability. HEK 293 cells were transfected with either ΔLf^{WT} or ΔLf^{PEST} constructs for 24 h. Total protein extracts were immunoblotted with M2. (C) Modulation of ΔLf half life by *O*-GlcNAcylation. Cells were transfected with either ΔLf^{WT} or ΔLf^{PEST} and $GlcNH_2$ treated or cotransfected with the OGT-construct or not and then incubated with fresh medium supplemented by 10 µg.ml^{-1} CHX for the indicated time 24 h after transfection. Total protein extracts were immunoblotted with either M2 or anti-actin antibodies. (D) Data are expressed as fold stability as described in Materials and Methods. $p < 0.05$ (*). (E) The three Ser residues of the PEST sequence are equivalent. Mutation of Ser residues was done on the 3xFLAG-ΔLf^{M4} construct as template. Cells were transfected by the different constructs and 24 h after transfection, total protein extracts were immunoblotted with either M2 or anti-actin antibodies. (F) Mutation of K^{379} rather

than of K^{391} inhibits degradation. 3xFLAG-ΔLf and 3xFLAG-ΔLf^{M4} constructs were used as template to obtain K^{379} and K^{391} mutants. Cells were transfected by the different constructs and 24 h after transfection, total protein extracts were immunoblotted with either M2 or anti-actin antibodies. (G) K^{379} is the main Ub-ligase target. HEK 293 cells were co-transfected with or without the 3xFLAG-ΔLf constructs and the Ub-HA-expressing vector for 24 h and then incubated with 10 µM of the proteasomal inhibitor MG132 for 2 h prior to lysis. Total cell extracts were immunoprecipitated with anti-HA polyclonal antibodies or used as input. Samples were immunoblotted with M2 (upper and lower panel) or with HA.11 (middle panel) antibodies.

Fig. 4. S^{10} posttranslational modifications control ΔLf turnover: *O*-GlcNAc confers stability whereas phosphorylation triggers degradation. (A) Schematic representation of the S^{10} mutant constructs. (B) Total protein extract of HEK293 cells expressing 3xFLAG-ΔLf or its S^{10} mutant constructs were immunoblotted with M2 or anti-actin antibodies. (C) *O*-GlcNAc posttranslational modification at S^{10} increases ΔLf stability. Transfected cells were incubated 24 h posttransfection with fresh media supplemented with 10 $\mu g.ml^{-1}$ CHX. ΔLf^{WT} and ΔLf^{S10+} transfected cells were incubated with or without $GlcNH_2$. Cell lysates were immunoblotted with M2 or anti-actin antibodies. (D) The graph data are expressed as fold stability as described in Materials and Methods (n=5; $p < 0.05$ (*)). (E) Invalidation of S^{10} protects against Ub-dependent degradation. Cells were transiently co-transfected with or without the 3xFLAG-tagged ΔLf or S^{10} mutant contructs and the Ub-HA expression vector for 24 h and then incubated with 10 µM MG132 for 2 h prior to lysis. Total cell extracts were immunoprecipitated with anti-HA polyclonal antibodies or used as input. Samples were immunoblotted with M2 (upper and lower panel) or with HA.11 antibodies (medium panel).

Fig. 5. *O*-GlcNAcylation at S^{10} negatively controls DNA binding and ΔLf transcriptional activity. (A) ΔLf is a phosphorylated protein. HEK293 cells were transfected by different ΔLf constructs in the presence or not of $GlcNH_2$ for 24 h prior to lysis. M2 immunoprecipitates were immunoblotted with anti-P-Ser and as a loading control, input was immunoblotted with M2 (left panel). Phosphatase treatment (1U/IP) confirms anti-P antibody specificity (AP, right panel) (B) Relative luciferase activity of ΔLf and its *O*-GlcNAc mutants. HEK 293 cells were co-transfected with pGL3-$S1^{Skp1}$-Luc vector and pcDNA3.1-ΔLf^{WT} or S^{10} mutant constructs. Relative luciferase activities are expressed as in Materials and Methods ($n \geq 9$; $p < 0.01$ (**)). (C-D) *O*-GlcNAcylation inhibits DNA binding. The *in vivo* binding of ΔLf and its S^{10} mutants to the *Skp1* promoter fragment was examined in HEK 293 cells treated or not with $GlcNH_2$ (n=3). Cross-linked DNA-ΔLf complexes were immunoprecipitated and precipitated DNA fragments were PCR-amplified (C) or real time PCR amplified (D) with specific primers covering the ΔLfRE present in the *Skp1* promoter. The PCR-amplified DNA purified from the sonicated chromatin was used as input and loading control. ChIP assays were performed using M2, anti-rabbit IgG antibodies as a nonspecific control (IR) and without antibody (NIP). Amplification of the albumin promoter region was used as a negative control. (E) ΔLf transactivation complex is not *O*-GlcNAcylated. Re-ChIP was performed as above for the ChIP assay with some modifications. The first immunoprecipitation was performed using M2, RL2, anti-P-Ser or anti-HA antibodies. Then, prior to reversal of protein–DNA cross-linking, the chromatin fragments were subjected to re-precipitation using M2, irrelevant (IR) or no antibodies (NIP). (n=3).

Fig. 6. Illustration of the regulation of ΔLf activity and stability by the *O*-GlcNAc/P interplay. ΔLf exists as an inactivated and stable S^{10}-*O*-GlcNAcylated isoform pool in cells. Depending on cellular events such as transcription, translation, cell cycle progression, cell signaling, or in the case of cell stress/injury, OGA will trigger ΔLf deglycosylation unmasking the S^{10} hydroxyl group. Deglycosylated ΔLf can bind DNA and transactivate basal transcription. Phosphorylation at S^{10} by an appropriate kinase leads to a strong amplification of the process and mediates PEST phosphorylation, polyUb at K^{379} and ΔLf degradation.

Table I: Names of mutants, amino acid modification, location and oligonucleotides used for mutagenesis.

Name of mutants	Mutated amino acid(s)[a]	Site-directed mutagenesis oligonucleotides
ΔLf ^{S10A}	S10A	F: 5'-CGTGGCCCTCCTGTCGCCTGCATAAAGAGAGA-3' R: 5'-TCTCTCTTTATGCAGGCGACAGGAGGGCCACG-3'
ΔLf ^{S227A}	S227A	F: 5'-CCCGGGTCCCTGCTCATGCCGTTG-3' R: 5'-CAACGGCATGAGCAGGGACCCGGG-3'
ΔLf ^{S472A}	S472A	F: 5'-GTGCCCCTGGGGCTGACCCGAGAT-3' R: 5'-ATCTCGGGTCAGCCCCAGGGGCAC-3'
ΔLf ^{S559A}	T559A	F: 5'-ACGGAAGCCTGTGGCTGAGGCTAGAAGC-3' R: 5'-GCTTCTAGCCTCAGCCACAGGCTTCCGT-3'
ΔLf ^{S10D}	S10D	F: 5'-GTGGCCCTCCTGTCGACTGCATAAAGAGAG-3' R: 5'-CTCTCTTTATGCAGTCGACAGGAGGGCCAC-3'
ΔLf ^{S16D}	S16D	F: 5'-CTGCATAAAGAGAGACGACCCCATCCAGTGTATC-3' R: 5'-GATACACTGGATGGGGTCGTCTCTCTTTATGCAG-3'
ΔLf ^{M4}	S10A, S227A, S472A, T559A	
ΔLf $^{S10+}$	S227A, S472A, T559A	
ΔLf $^{S227+}$	S10A, S472A, T559A	
ΔLf $^{S472+}$	S10A, S227A, T559A	
ΔLf $^{S559+}$	S10A, S227A, S472A	
ΔLf ^{S392A}	S392A	F: 5'-CAGAGAACTACAAAGCCCAACAAAGCAGTG-3' R: 5'-CACTGCTTTGTTGGGCTTTGTAGTTCTCTG-3'
ΔLf ^{S395A}	S395A	F: 5'-CTACAAATCCCAACAAGCCAGTGACCCTGATCC-3' R: 5'-GGATCAGGGTCACTGGCTTGTTGGGATTTGTAG-3'
ΔLf ^{S396A}	S396A	F: 5'-CAAATCCCAACAAAGCGCTGACCCTGATCCTAAC-3' R: 5'-GTTAGGATCAGGGTCAGCGCTTTGTTGGGATTTG-3'
ΔLf ^{PEST}	S392A, S395A, S396A	
ΔLf $^{M4\text{-}S392A}$	S10A, S227A, S392A, S472A, T559A	
ΔLf $^{M4\text{-}S395A}$	S10A, S227A, S395A, S472A, T559A	
ΔLf $^{M4\text{-}S396A}$	S10A, S227A, S396A, S472A, T559A	
ΔLF $^{M4\text{-}PEST}$	S10A, S227A, S392A, S395A, S396A, S472A, T559A	
ΔLF ^{K379A}	K379A	F: 5'-CTGGCAGAGAACTACGCATCCCAACAAAGCAG-3' R: 5'-CTGCTTTGTTGGGATGCGTAGTTCTCTGCCAG-3'
ΔLF ^{K391A}	K391A	F: 5'-GTGTACACTGCATGCGCATGTGGTTTGGTGC-3' R: 5'-GCACCAAACCACATGCGCATGCAGTGTACAC-3'
ΔLf ^{KK}	K379A, K391A	
ΔLf $^{M4\text{-}K379A}$	S10A, S227A, K379A, S472A, T559A	
ΔLf $^{M4\text{-}K379A}$	S10A, S227A, K391A, S472A, T559A	
ΔLf $^{\Delta 1\text{-}14}$	Deletion of amino acids 1 to 14	F: 5'-AAGCTTATGGACTCCCCCATCCAGTGTATCC-3' R: 5'-GGATACACTGGATGGGGGAGTCCATAAGCTT-3'

[a] The single-letter amino acid code is used

Table II: Alignement of ΔLf PEST sequence to lactoferrin sequences from different species

Species	Sequences[a]	Accession number / ref
	391 392 395 396	
Delta-lactoferrin		
Homo sapiens	379 K̲CGLVPVLAENY**KSQQSSDPDPNCVD**RP 406	(25)
Lactoferrin		
Homo sapiens	423 KCGLVPVLAENYK***S***QQ***SS***DPDPNCVDRP 450	Q5EK51
Sus scrofa	419 KCGLVPVLAENQK***S***RQ**SSSS**D--CVHRP 444	P14632
Mus musculus	422 KCGLVPVLAENQK***SS***K***S***NGLD--CVNRP 447	P08071
Equus caballus	410 KCGLVPVLAENQK***S***QN***S***NAPD--CVHRP 435	O77811
Bubalus bubalis	423 KCGLVPVLAENRK***SS***KH***SS***LD--CVLRP 448	O77698
Bos taurus	423 KCGLVPVLAENRK***SS***KH***SS***LD--CVLRP 448	P24627
Capra hircus	423 KCGLVPVMAENRK***SS***KY***SS***LD--CVLRP 448	Q29477
Ovis aries	423 KCGLVPVMAENRE***SS***KY***SS***LD--CVLRP 448	NP_001020033
Rattus norvegicus	422 KCGLVPVLAEIQK***S***PN***S***NG***S***D--CVDRP 447	NP_001100334
Camelus dromedarius	423 KCGLVPVLAESQQ***S***PE***SS***GLD--CVHRP 448	Q9TUM0

[a] The single-letter amino acid code is used ; the PEST sequence of ΔLf predicted using the PESTfind software (EMBnet Austria; https://emb1.bcc.univie.ac.at/toolbox/pestfind/) is in bold type; italic bold letters indicate the Ser residues within the putative PEST sequence in Lf from different species; the Ub-targeted Lys residue is underlined in the ΔLf PEST sequence.

Figure 1

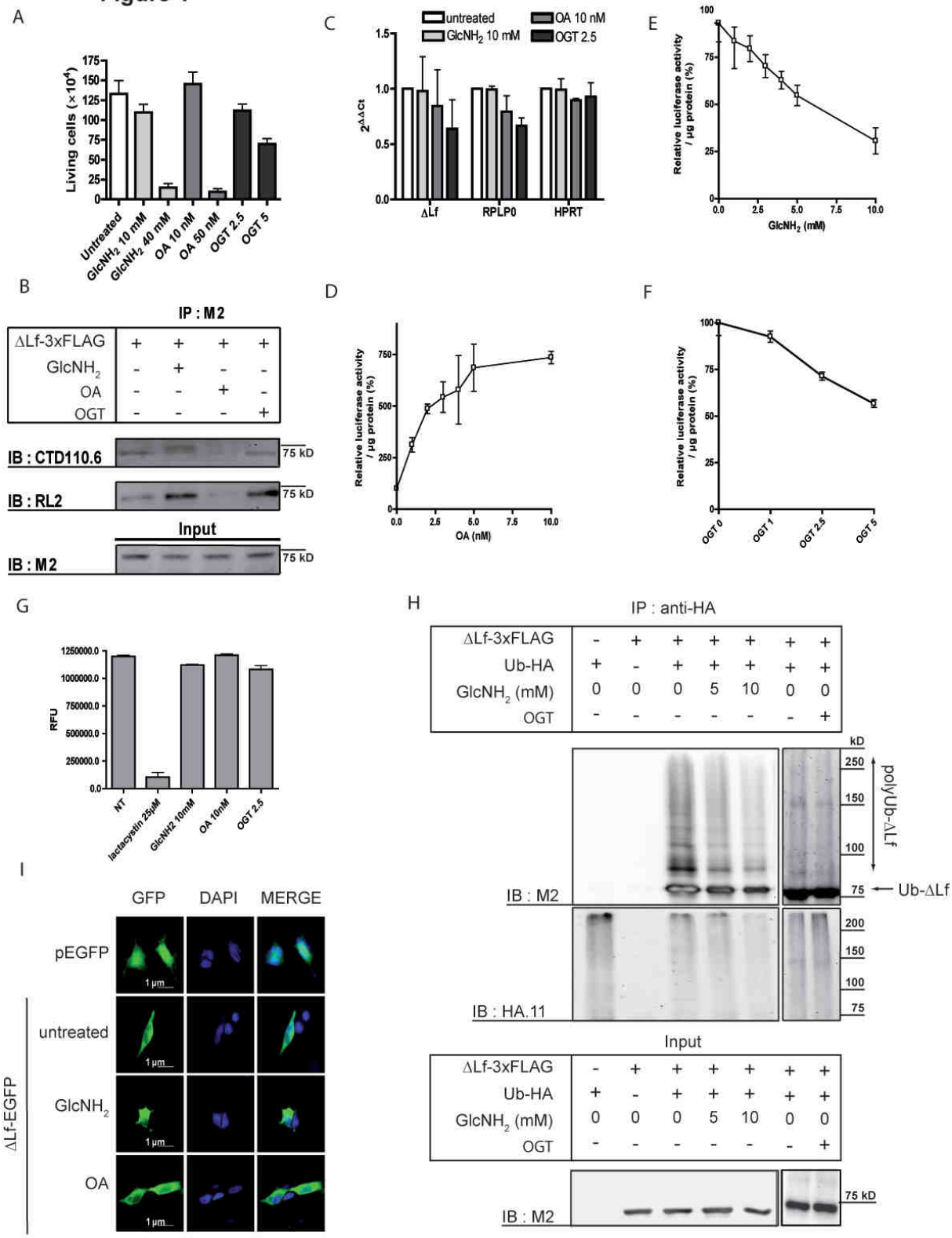


Figure 2

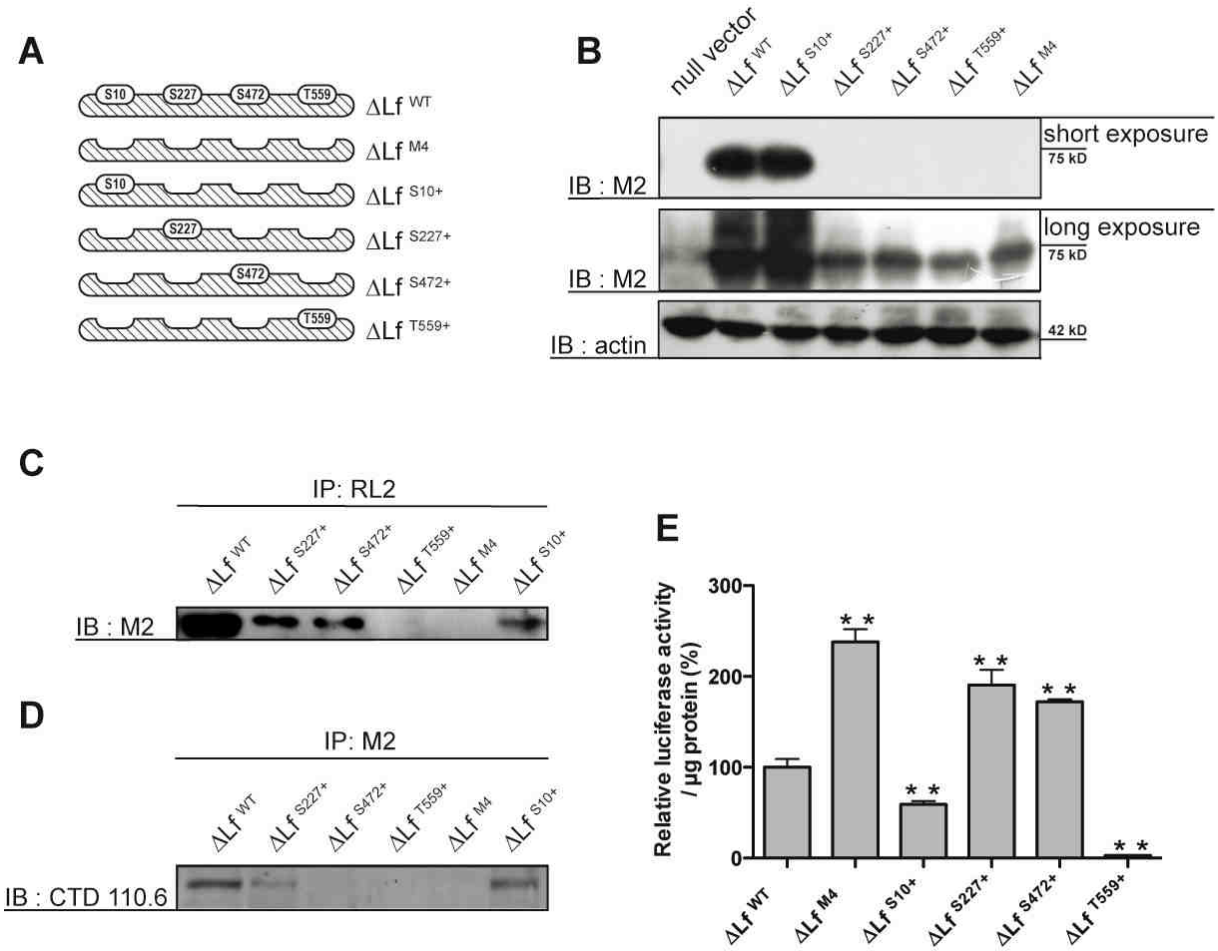
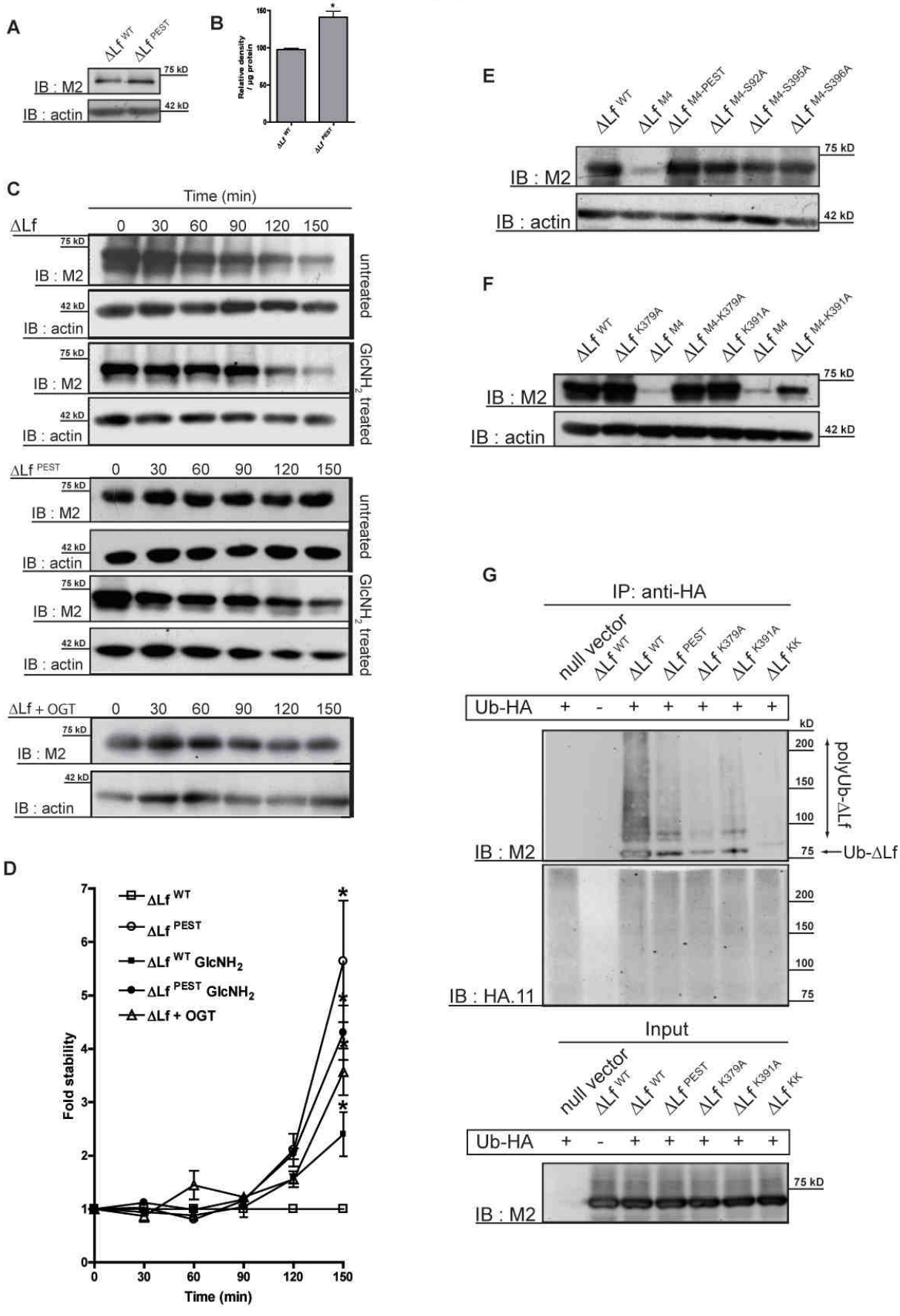


Figure 3



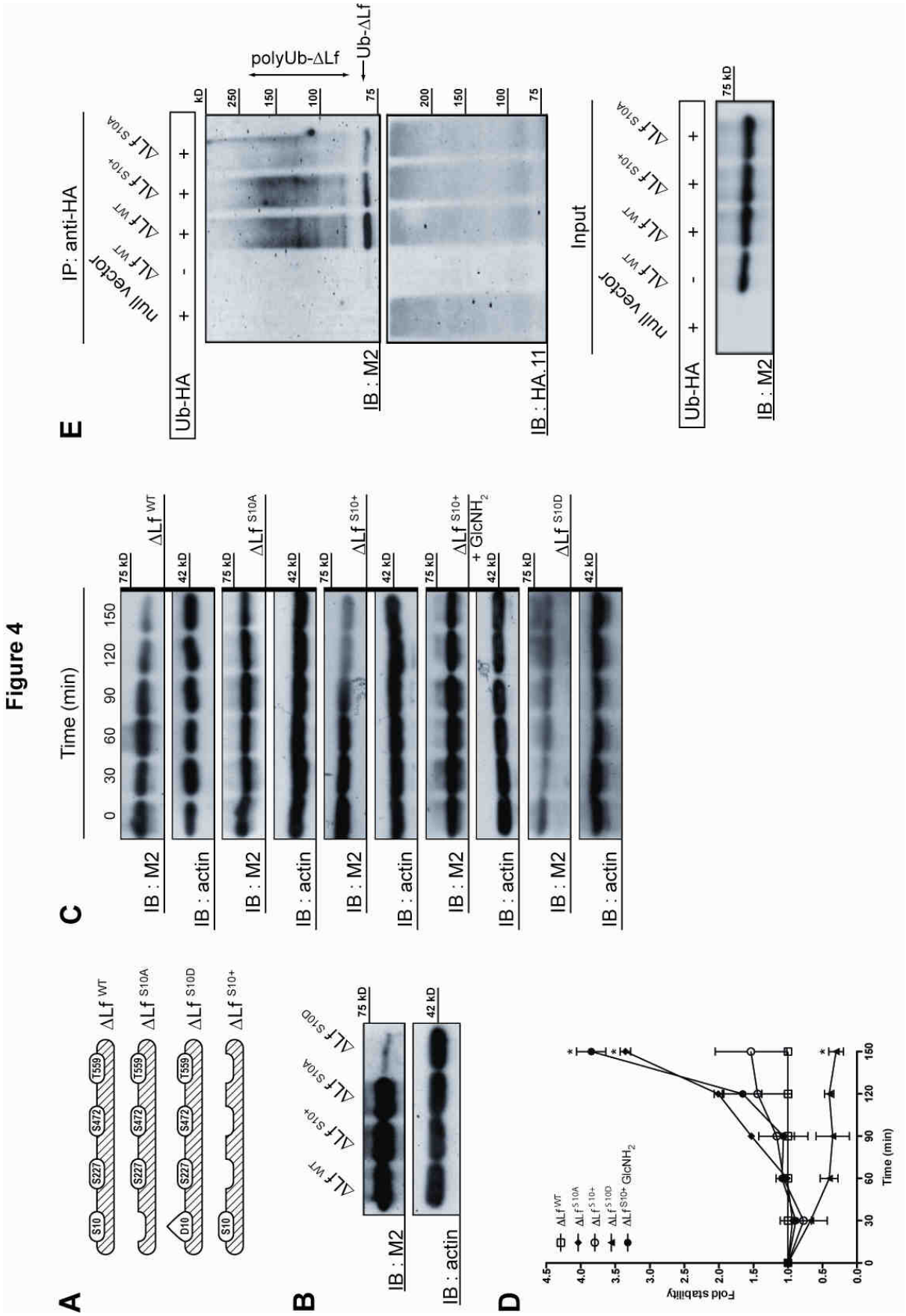


Figure 5

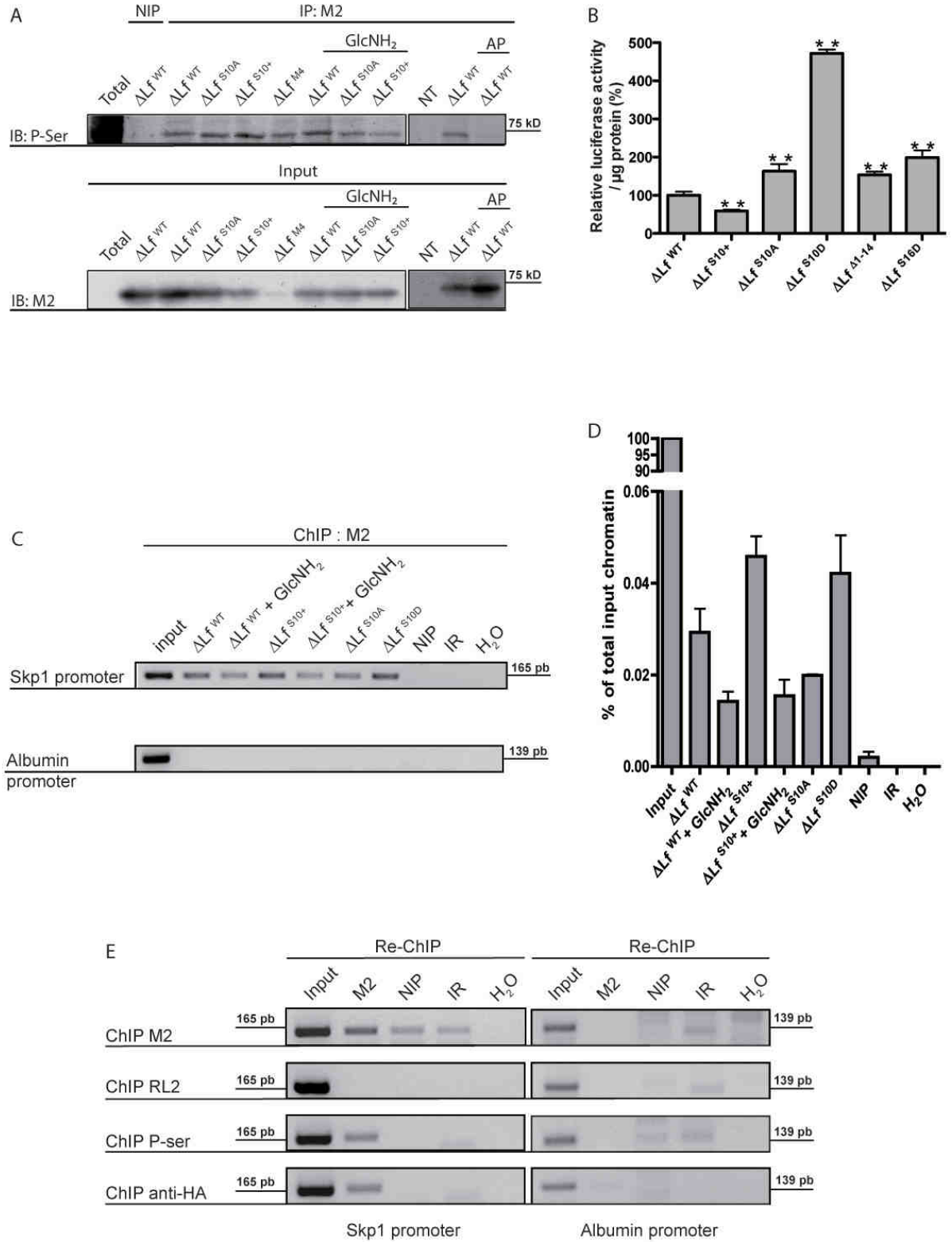
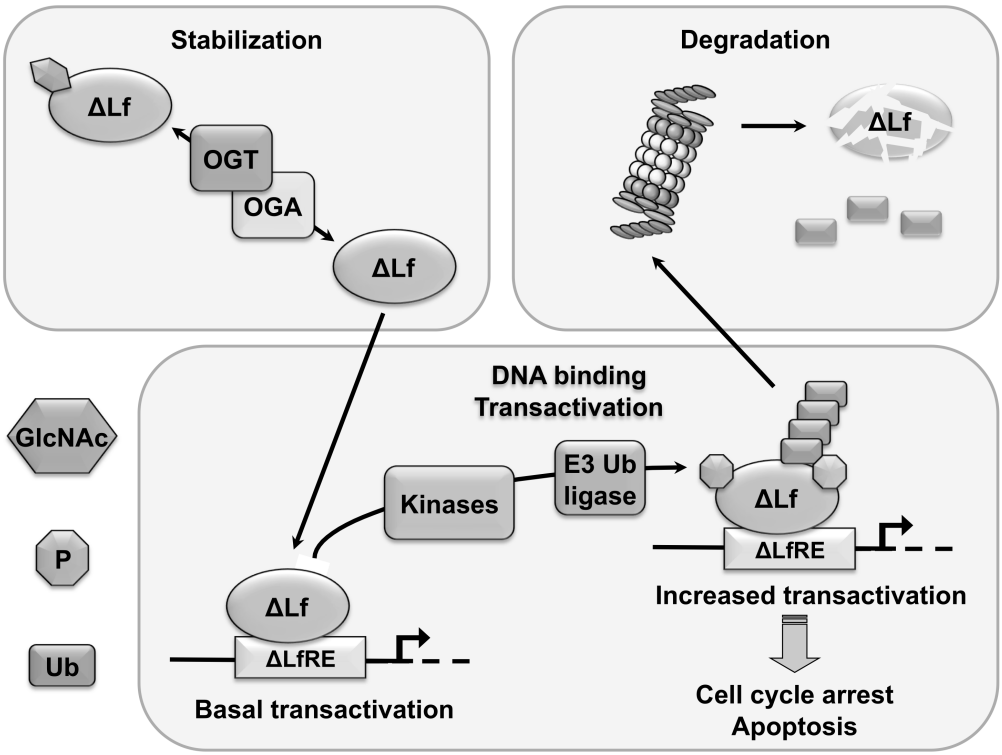


Figure 6



Discussion et Conclusion

A mon arrivée au laboratoire, la delta-lactoferrine n'avait été décrite que sous la forme de transcrits et peu de données étaient présentes dans la littérature à son propos. Durant mes travaux de thèse, nous avons pu déterminer plusieurs caractéristiques structurales de la delta-lactoferrine (articles 1 et 4) et démontrer que c'est un facteur de transcription ciblant différents gènes cibles (articles 1, 2 et 3) dont la régulation fait appel à plusieurs modifications post-traductionelles (article 4).

Les domaines fonctionnels de la delta-lactoferrine

Les tentatives de production de la delta-lactoferrine purifiée dans le but de comparer sa structure à celle de son isoforme sécrétée ont été un échec. Contrairement à la lactoferrine, la delta-lactoferrine est peu représentée en termes de quantité. De plus, les deux isoformes ont la même séquence ce qui ne permet pas d'avoir des anticorps spécifiquement dirigés contre la delta-lactoferrine. Par ailleurs, la delta-lactoferrine doit être produite « en intracellulaire » afin de conserver sa glycosylation atypique. Tout ceci rend sa production et sa purification difficiles et nous oblige à la produire dans un système recombinant où la lactoferrine n'est pas présente ou à produire une delta-lactoferrine étiquetée. A mon arrivée au laboratoire, en Master II, plusieurs systèmes d'expression avaient été testés. L'expression en cellules de mammifères d'une delta-lactoferrine de fusion avec un polypeptide poly-histidine n'a pas permis sa mise en évidence. En outre, le Dr. Christophe Mariller, en collaboration avec le Dr. Patricia Nagnan-Le Meillour, a testé un système d'expression en *Pichia Pastoris*, mais là encore la protéine n'a pas pu être détectée. Seule son expression en ovocytes de Xénope (Breton *et al.*, 2004) ou *in vitro* en lysat de réticulocytes de lapin a permis de visualiser la protéine après marquage à la méthionine S^{35}. Cependant, les quantités de protéine ainsi produites ne permettent pas d'envisager l'étude de sa structure tridimensionnelle. Au cours de mes travaux de thèse, j'ai réussi à produire et mettre en évidence la delta-lactoferrine fusionnée à l'étiquette 3xFLAG en cellules humaines. Avec cette construction, j'ai également tenté sa purification après expression en système Baculovirus/cellules d'insecte en collaboration avec M^{me} Marie-Christine Slomianny et le Dr. Dominique Legrand. Malheureusement, ce système n'a pas non plus permis de produire suffisamment de protéines. La faible quantité de la delta-lactoferrine dans les lysats protéiques malgré sa surexpression suggère qu'elle est peu stable quelque soit le système de production utilisé. Comme les protéines cytosoliques sont dégradées par le protéasome après leur polyubiquitinylation, nous nous sommes donc intéressés à l'ubiquitinylation de la delta-lactoferrine. Nous avons pu montrer que la delta-lactoferrine est polyubiquitinylée sur les résidus de lysine positionnés en

379 et 391 (article 4). La mutation de ces deux résidus abolit complètement l'ubiquitinylation de la protéine alors que cette polyubiquitinylation subsiste partiellement lorsqu'un seul de ces deux résidus de lysine est conservé. Nos résultats montrent, néanmoins, que la lysine 379 est la cible préférentielle de l'ubiquitinylation.

La polyubiquitinylation des protéines est sous le contrôle de séquences de dégradation rapide. La delta-lactoferrine possède une séquence PEST (^{391}KSQQSSDPDPNCVD^{404}) conservée parmi les lactoferrines et présentant un score élevé avec le serveur PEST Finder. La mutation de cette séquence augmente la durée de demi-vie de la delta-lactoferrine ce qui confirme sa fonctionnalité (article 4). Les résidus de sérine ou de thréonine des séquences PEST sont fréquemment la cible d'une phosphorylation permettant leurs activations (Rechsteiner et Rogers, 1996; Rogers *et al.*, 1986). La mutation ponctuelle de chacune des trois sérines de la séquence PEST de la delta-lactoferrine (sérines 392, 395 et 396) stabilise indifféremment la protéine, ce qui suggère que ces trois résidus pourraient être des cibles équivalentes de kinases. Les séquences PEST permettent l'interaction des protéines avec le complexe SCF qui dirige leur polyubiquitinylation en phase G1 et S. L'activité de la delta-lactoferrine, conduisant à l'arrêt du cycle en phase S (Breton *et al.*, 2004), pourrait être régulée par l'activation de sa séquence PEST conduisant à la dégradation de la protéine. Etrangement, le mutant PEST conserve un taux résiduel de polyubiquitinylation suggérant que d'autres domaines de la delta-lactoferrine pourraient contrôler l'ajout de cette modification post-traductionnelle et la dégradation de la protéine. L'analyse de la séquence primaire de la delta-lactoferrine montre la présence de plusieurs autres séquences de dégradation rapide. En effet, deux D-Box putatives de séquence RPFL (acides aminés 108 à 111) et RSNL (acides aminés 475 à 478) sont retrouvées conservées phylogénétiquement parmi les lactoferrines. Ces séquences sont connues pour interagir avec le complexe APC/C qui dirige la dégradation des protéines en phase G2 et M (Hames *et al.*, 2001; Linder *et al.*, 1989). Si la delta-lactoferrine est toujours présente pendant les phases G2 et M, ces D-Box pourraient contrôler sa dégradation. Des travaux complémentaires seront nécessaires pour confirmer la fonctionnalité des D-Box. Les résultats obtenus devraient permettre de mieux comprendre le mécanisme de dégradation de la delta-lactoferrine.

Les travaux réalisés au laboratoire ainsi que ceux de la littérature montrent que la delta-lactoferrine est effectivement présente dans le noyau (Breton *et al.*, 2004; Liu *et al.*, 2003). Mes travaux de thèse ont pu démontrer que la delta-lactoferrine possède une séquence NLS bipartite (^{417}**RR**SDTSLTWNSV**K**G**KK**^{432}) homologue à celle de la nucléoplasmine (Robbins *et al.*, 1991) et conservée parmi les lactoferrines. Cette séquence est nécessaire à

l'adressage nucléaire de la delta-lactoferrine puisque sa mutation conduit à sa rétention cytoplasmique (article 1). Néanmoins, ce mutant possède toujours une activité transcriptionnelle résiduelle suggérant qu'une faible proportion de protéine est présente dans le noyau. Une seconde séquence putative d'adressage au noyau (^{598}KRKP^{601}) pourrait contribuer à réguler son transport. Cependant, cette dernière n'est pas conservée phylogénétiquement. Si cette séquence est fonctionnelle, elle ne le sera que pour la delta-lactoferrine humaine.

Comme tous facteurs de transcription, la delta-lactoferrine se lie à l'ADN par un DBD. Deux DBD ont été proposés pour la lactoferrine : le premier localisé à l'extrémité N-terminale (Son *et al.*, 2002) est tronqué dans la delta-lactoferrine, l'autre correspondrait à la région charnière entre les deux lobes (Anderson *et al.*, 1989). Le mutant de la delta-lactoferrine délété des 14 premiers résidus d'acide aminé possède une activité transcriptionnelle supérieure à celle de l'isoforme sauvage (article 4), ce qui démontre que le DBD n'est pas localisé dans cette région. La modélisation moléculaire de la delta-lactoferrine, basée sur la structure de son homologue sécrétée, montre que cette dernière pourrait interagir avec l'ADN par la région située entre les deux lobes (communications personnelles du Dr. Christophe Mariller et du Dr. Gérard Vergoten). Des recherches complémentaires seront nécessaires afin de démontrer la fonctionnalité de cette région.

La régulation de la delta-lactoferrine par la GlcNAcylation et la phosphorylation

Les différents domaines de la delta-lactoferrine peuvent réguler son activité. Nous avons pu démontrer que cette dernière est la cible de modifications post-traductionnelles qui sont responsables de la régulation de sa dégradation et de son activité transcriptionnelle.

La technique la plus couramment utilisée pour réaliser la cartographie de sites modifiés par GlcNAcylation est le BEMAD (« β-elimination followed by Michael addition with DTT ») (Vosseller *et al.*, 2005). La position des sites modifiés est obtenue après analyse des lysats trypsiques en LC/ESI/MS-MS. Ne pouvant produire suffisamment de delta-lactoferrine purifiée pour mettre en œuvre cette méthode, nous avons réalisé la cartographie des sites de GlcNAcylation par une approche alliant immunoprécipitation de la delta-lactoferrine et de ses différentes constructions mutantes, et immunorévélation avec deux anticorps anti-motif *O*-GlcNAc (RL2 et CTD110.6). Grâce à cette stratégie, nous avons confirmé l'existence des sites de GlcNAcylation et nous les avons localisés en 10, 227 et 472 (article 4). Dans l'article 4, nous montront également que la delta-lactoferrine est phosphorylée. La recherche *in silico* des résidus modifiés par un groupement phosphate

(Figure 26, page 133) montre que la delta-lactoferrine est susceptible d'être phosphorylée sur de très nombreux sites. Les immunoprécipitations des mutants montrent que ces phosphorylations peuvent s'exercer sur d'autres sites que ceux de la GlcNAcylation. Néanmoins, nous avons pu établir que la sérine 10 est également un site de phosphorylation. Ce résidu serait alors, à la fois, la cible d'une kinase et de l'OGT. Il s'exercerait donc une relation de réciprocité entre GlcNAcylation et phosphorylation pour ce site faisant de la delta-lactoferrine le sixième exemple de protéines subissant ce type de réciprocité décrit dans la littérature (Cheng et Hart, 2001; Comer et Hart, 2001; Medina et Haltiwanger, 1998; Musicki *et al.*, 2005; Yang *et al.*, 2006).

La GlcNAcylation et la phosphorylation des séquences PEST contrôlent la durée de demi-vie des protéines sur lesquelles il peut s'exercer une compétition entre. La GlcNacylation masque l'hydroxyle des résidus de sérine ou de thréonine à la phosphorylation ce qui conduit à la stabilisation des protéines. Nous avons montré que la glycosylation de la sérine 10 contrôle la stabilité de la protéine (article 4). Cependant, ce résidu est distant de la séquence PEST de la delta-lactoferrine suggérant que la stabilisation de la protéine par la GlcNAcylation de la sérine 10 nécessite un mécanisme intermédiaire. Il a été démontré au laboratoire que les Hsp possèdent une activité lectinique leur permettant d'interagir avec les protéines GlcNAcylées (Guinez *et al.*, 2006). Le complexe protéine-*O*-GlcNAc/Hsp serait alors protégé de la dégradation protéasomale (Guinez *et al.*, 2007). Au cours de mes travaux de thèse, nous avons pu mettre en évidence la formation du complexe delta-lactoferrine/Hsp par des expériences de co-immunoprécipitation (résultats non présentés). Cependant, l'ajout de GlcNAc libre dans les échantillons ne permet pas de dissocier le complexe suggérant que les interactions, qui lient la delta-lactoferrine aux Hsp, ne sont pas dépendantes d'une activité lectinique. Le serveur Ying*O*Yang prédit que les sérines 292 et 295 pourraient être GlcNAcylées mais avec un score très bas. Bien que le mutant M4 ne soit pas glycosylé, suggérant qu'il n'existe pas d'autres sites de GlcNAcylation, nous ne pouvons pas exclure l'éventualité de cette modification sur la séquence PEST. En effet, la GlcNAcylation de ces sites pourrait dépendre de la présence de la glycosylation de la sérine 10. Sa mutation dans le mutant M4 empêcherait alors la mise en évidence d'une GlcNAcylation de la séquence PEST.

Le résidu de sérine 472 de la delta-lactoferrine modifié par GlcNAcylation se situe à proximité de la séquence NLS. Les données bibliographiques montrent que la fonctionnalité de cette séquence peut être modulée par la phosphorylation (Boulikas, 1993) et différents auteurs ont suggéré une régulation du trafic nucléocytoplasmique par la GlcNAcylation.

L'utilisation de la glucosamine (activateur de la GlcNAcylation) ou de l'acide okadaïque (inhibiteur des phosphatases) n'a, cependant, pas permis de modifier la répartition intracellulaire de la delta-lactoferrine (article 4). La répartition nucléocytoplasmique de la delta-lactoferrine ne semble donc pas être régulée par la GlcNacylation ou par la phosphorylation.

Bien que nous n'ayons pas localisé le DBD ou de domaine de transactivation (Lavery et McEwan, 2005; Mapp et Ansari, 2007; Michel *et al.*, 2008) nos résultats démontrent que la GlcNAcylation et la phosphorylation de la delta-lactoferrine modifient ses interactions avec l'élément de réponse retrouvé dans le promoteur du gène *Skp1* (article 4) ainsi que son activité transcriptionnelle envers les gènes *Skp1* (article 4) et *DcpS* (article 3). En effet, la delta-lactoferrine et son isoforme M4 non glycosylable induisent la surexpression des mêmes protéines et nos résultats montrent que ce dernier est plus efficace que l'isoforme sauvage (articles 2, 3 et 4). Cependant, tous les sites de GlcNAcylation ne semblent pas avoir le même impact. En effet, la mutation des résidus de sérine 227 et 472 n'a que peu d'effets sur l'activité transcriptionnelle de la delta-lactoferrine alors que la construction mutante pour la sérine 10 présente une activité transcriptionnelle proche de celle du mutant M4 (article 4). Par ailleurs, les résultats obtenus avec la technique de ChIP suggèrent que la delta-lactoferrine interagit préférentiellement avec ses éléments de réponse sous une forme non GlcNAcylée. En outre, en réalisant deux ChIP successives sur le même échantillon, nous avons démontré que le complexe transcriptionnel fixé à l'ADN comportant la delta-lactoferrine n'est pas GlcNAcylé. Nos résultats démontrent également que cette inhibition est due à la GlcNAcylation de la sérine 10. L'isoforme mutante ne possédant que ce site se fixe moins à l'ADN lorsqu'elle est préférentiellement glycosylée. Par ailleurs, la mutation du résidu de sérine 10 augmente l'activité transcriptionnelle de la delta-lactoferrine. En outre, l'isoforme mimant une phosphorylation sur la sérine 10 (mutant S10D) a une activité transcriptionnelle supérieure d'un facteur 4,5 par rapport à l'isofome sauvage mais ne semble pas être capable d'occuper plus d'éléments de réponse que l'isoforme sauvage (article 4). Ces résultats suggèrent que la phosphorylation de la sérine 10 ne serait pas indispensable aux interactions de la delta-lactoferrine avec son élément de réponse mais serait requise pour la transactivation de ces gènes cibles.

L'élément de réponse à la delta-lactoferrine

A travers les différentes études réalisées durant ma thèse, nous avons pu démontrer que la delta-lactoferrine interagit *in vivo* avec le promoteur des gènes *Skp1* (article 1), *Bax*

(article 2) et *DcpS* (article 3) au niveau d'un élément de réponse spécifique appelé ΔLfRE (Tableau 6). En outre, quelque soit le promoteur considéré, la mutation de cette séquence abolit la capacité de la delta-lactoferrine à transactiver un gène rapporteur. Les ΔLfRE des gènes cibles de la delta-lactoferrine sont homologues à la séquence S1 décrites par He et Furmanski (1995) pour la lactoferrine (Tableau 6). Il n'est pas possible de comparer directement les résultats de ces travaux portant sur l'étude de la séquence S1 avec nos résultats car, au cours de leurs expérimentations, ces auteurs ont utilisé un gène rapporteur CAT (He et Furmanski, 1995). Cependant, les données obtenues lors de cette étude et les nôtres nous permettent de définir une séquence optimale de fixation de la delta-lactoferrine. Les expériences de gène rapporteur réalisées par ces auteurs montrent la présence d'un « noyau » de six nucléotides (GGCACT) dont la mutation ponctuelle diminue fortement ou abolit l'activité transcriptionnelle (He et Furmanski, 1995). Bien que les mutations des

Tableau 6 : Séquences des différents ΔLfRE classées en fonction de leur efficacité. Ce tableau présente l'alignement des éléments de réponse retrouvés in vivo dans les gènes cibles de la delta-lactoferrine avec la séquence S1 décrite par He et Furmanski (1995), ainsi que les séquences mutées qui en sont issues, et leur capacité à transactiver l'expression d'un gène rapporteur. Les nucléotides différents de ceux de la séquence S1 sont en caractère gras.

	Séquence										**Efficacité (en %)**
	1	2	3	4	5	6	7	8	9	10	
ΔLfRE	G/A	G	C	A	C	T	(G)	T	A/G	C/G	
ΔLfRE ^{Skp1}	G	G	C	A	C	T	**G**	T	A	C	100
ΔLfRE ^{DcpS}	**A**	G	C	A	C	T	-	T	G	G	90
ΔLfRE ^{Bax}	G	G	C	A	C	T	-	T	A	**T**	60
ΔLfRE ^{Fas}	G	G	C	A	C	T	-	**A**	A	C	5
Elément de réponse défini par He et Furmanski (1995)											
S1	G	G	C	A	C	T	-	T	A/G	C	++
GmT	G	**T**	C	A	C	T	-	T	G	C	++
CmA	G	G	**A**	A	**A**	T	-	T	G	C	+
GmC	**C**	G	C	A	C	T	-	T	G	C	+
CmG	G	G	C	A	**G**	T	-	T	G	C	+/-
CTdmGG	G	G	C	A	**G**	**G**	-	T	G	C	+/-

cytosines de ce consensus en alanine (séquence CmA) permettent de conserver l'activité CAT, la mutation ponctuelle de la cinquième cytosine en guanine (séquence CmG) diminue fortement l'efficacité de l'élément de réponse. La conversion de la guanine en position 1 par une cytosine (séquence GmC) conduit également à la perte d'efficacité. La transactivation du ΔLfRE retrouvé dans le promoteur du gène *Skp*1 conduit à la détection de l'activité luciférase la plus importante, suggérant qu'il est le plus proche de l'élément de réponse optimal de la delta-lactoferrine. Cette séquence présente une insertion par rapport à la séquence S1. Il semble donc que l'ajout d'une guanine en position 7 n'interfère pas sur l'interaction de la delta-lactoferrine avec son élément de réponse. Comparée à l'expression optimale obtenue avec le ΔLfRE contenu dans le gène *Skp1*, celui présent dans le gène *Fas* n'induit quasiment pas de transactivation. Hormis la présence d'une adénine à la place d'une thymine en huitième position, le ΔLfRE de *Fas* est identique à la séquence S1 ce qui démontrent l'importance de cette thymine. Les ΔLfRE des promoteurs des gènes *Bax* et *DcpS* permettent respectivement de mesurer une activité luciférase de 60 % et 90 % par rapport à celui de *Skp1* en présence de delta-lactoferrine. Le ΔLfRE de *DcpS* possède une adénine et une guanine respectivement en position 1 et 10. Comme pour le mutant GmC, cette séquence est fonctionnelle mais moins efficace suggérant que cette première base module la reconnaissance du ΔLfRE par la delta-lactoferrine. Les ΔLfRE de *Bax* et *DcpS* possèdent respectivement une thymine et une guanine en dernière position à la place d'une cytosine. Leurs efficacités respectives suggèrent que la présence d'une guanine en position terminale interfère peu alors que la présence d'une thymine est plus préjudiciable. Ces différentes données nous ont permis de mieux comprendre l'importance de certains nucléotides et de définir la séquence 5'-G/A G C A C T (G) T A/G C/G-3' comme le ΔLfRE consensus.

D'autres séquences ont été décrites par He et Furmanski (1995) (Tableau 2, page 21) et nous avons retrouvé *in vivo* une séquence homologue à S2 dans le promoteur du gène *Skp1* (ΔLfRE 2). Bien que moins efficace que le ΔLfRE, nous avons pu démontrer la fonctionnalité de cette séquence. Le ΔLfRE 2 n'a pas été retrouvé *in vivo* dans le promoteur des gènes *Bax* et *DcpS* montrant que la présence des deux ΔLfREs n'est pas indispensable à l'activité transcriptionnelle de la delta-lactoferrine. La découverte d'autres gènes cibles de la delta-lactoferrine possédant le ΔLfRE 2 dans leurs promoteurs permettrait de mieux comprendre les relations entre la delta-lactoferrine et les promoteurs de ses gènes cibles. Les gènes cibles de p53 peuvent être classés en deux catégories : ceux impliqués dans l'arrêt du cycle cellulaire et la réparation de l'ADN qui possèdent dans leur promoteur un élément de réponse dit de « haute affinité », et ceux impliqués dans le déclenchement de l'apoptose qui

possèdent dans leur promoteur un élément de réponse dit de « base affinité » (Weinberg *et al.*, 2005). En fonction de ses modifications post-traductionnelles, p53 transactive une catégorie plutôt que l'autre (Das *et al.*, 2008; Knights *et al.*, 2006). En fonction des signaux cellulaires et de ses modifications post-traductionnelles, la delta-lactoferrine, de manière similaire à p53, pourrait exercer son activité transcriptionnelle sur les gènes possédant dans leurs promoteurs les deux ΔLfREs plutôt que sur ceux ne possédant que le ΔLfRE.

Le rôle de la delta-lactoferrine

Le maintien d'un phénotype normal est le résultat de l'expression tissu-spécifique de différentes protéines régulatrices et la delta-lactoferrine pourrait être l'une d'entre elles. En effet, elle est soupçonnée d'être un suppresseur de tumeur depuis qu'il a été démontré qu'elle est sous-exprimée dans les cas de cancer et que son expression est corrélée à un effet anti-prolifératif (Benaissa *et al.*, 2005; Breton *et al.*, 2004; Siebert et Huang, 1997). Cependant, le rôle de la delta-lactoferrine n'était pas connu. Les études concernant la surexpression de Skp1 (article 1) et Bax (article 2) nous ont permis de mieux appréhender le rôle de la delta-lactoferrine dans la physiologie cellulaire. En effet, dans un premier temps, la suexpression de Skp1 pourrait expliquer l'arrêt du cycle cellulaire dû à l'expression de la delta-lactoferrine (Breton *et al.*, 2004). Skp1 est un des composant du complexe SCF qui est impliqué dans la dégradation des acteurs moléculaires de la transition G1/S (Ang et Wade Harper, 2005). L'augmentation de la quantité de Skp1 pourrait avoir d'importantes conséquences sur l'activité du complexe SCF ainsi que sur la dégradation des protéines. Par exemple, Skp2 est lui-même la cible du complexe SCF (Wirbelauer *et al.*, 2000) et sa dégradation pourrait conduire à l'accumulation des cyclines et à l'arrêt du cycle. Par ailleurs, la fonctionnalité de la séquence PEST de la delta-lactoferrine (article 4) fait que cette dernière pourrait également être la cible du complexe SCF, la surexpression de Skp1 pourrait alors conduire à sa dégradation ce qui constituerait un mécanisme de rétro-contrôle de son activité. La réversion de la surexpression de Skp1, induit par la delta-lactoferrine, par l'utilisation de siRNA dirigés contre les transcrits Skp1 devrait permettre de tester ces différentes hypothèses. Dans un deuxième temps, la surexpression de Bax apporte une explication supplémentaire à l'activité anti-prolifération de la delta-lactoferrine. Bax est une protéine pro-apoptotique, dont la surexpression par rapport à certaines protéines comme Bcl-2 ou Bcl-xl, conduit au déclenchement de la mort cellulaire par la voie intrinsèque (Bleicken *et al.*, 2009). L'implication de la delta-lactoferrine dans le contrôle de la mort cellulaire renforce l'idée qu'elle pourrait être un gène suppresseur de tumeur.

Afin de confirmer définitivement son rôle de suppresseur de tumeur, des études complémentaires sur l'activité de la delta-lactoferrine *in vivo* sont en cours sur des modèles animaux. Pour ce faire, je suis parti trois mois au Japon dans le laboratoire du Pr. Tsuda pour développer, en collaboration avec le Dr. Alexander, un vecteur adénoviral permettant la réexpression de la delta-lactoferrine dans différentes lignées de cellules tumorales humaines (Long *et al*., 2009; annexe 2). Des vecteurs dérivés du génome de l'adénovirus sont de plus en plus développés pour leur utilisation en thérapie anti-tumorale dans différents laboratoires. Des tests cliniques sont actuellement en cours avec un vecteur adénoviral exprimant p53 (Peng, 2005). Les particularités de la construction développée dans le laboratoire du Pr. Tsuda sont de cibler spécifiquement les cellules cancéreuses et de ne pas induire de réponse immunitaire. Lors de mon séjour, nous avons également réalisé un vecteur adénoviral contenant l'isoforme « super delta-lactoferrine » (isoforme M4) dont l'activité n'est plus inhibée par la GlcNAcylation (articles 3 et 4) dans le but d'évaluer son intérêt en thérapie génique.

Annexes

En parallèle de mes travaux de thèse, j'ai pu participer au développement d'une méthode de quantification spécifique des transcrits de la lactoferrine et de la delta-lactoferrine par PCR en temps réel par la technologie TaqMan.

Ces travaux ont fait l'objet de la publication suivante :

Hoedt, E., Hardivillé, S., Mariller, C., Elass, E., Perraudin, J.-P., et Pierce, A., « Discrimination and evaluation of lactoferrin and delta-lactoferrin gene expression levels in cancer cells and under inflammatory stimuli using TaqMan real time PCR », *Biometals*. Sous presse (2010)

J'ai également eu l'opportunité d'effectuer un stage pré-doctoral de trois mois dans le laboratoire du Pr. Tsuda dans le cadre d'une collaboration visant à démontrer le rôle de la delta-lactoferrine *in vivo* et à évaluer le potentiel du mutant M4 en thérapie génique. Durant ce stage, j'ai pu participer au développement d'un nouveau de type de vecteur de thérapie génique basé sur le génome de l'adénovirus.

La stratégie de cette technologie a, d'ores et déjà, fait l'objet de la publication suivante :

Long, N., Shirai, T., Hardivillé, S., Pierce, A., Fukamachi, K., Futakuchi, M., Alexander, D. B., et Tsuda, H. « Construction of a Multi-Functional Helper-Dependent Adenovirus Based System », 2010, *Asian Pac J Cancer Prev.* 10(5) : 939-960

Biometals
DOI 10.1007/s10534-010-9305-5

Discrimination and evaluation of lactoferrin and delta-lactoferrin gene expression levels in cancer cells and under inflammatory stimuli using *TaqMan* real-time PCR

Esthelle Hoedt · Stephan Hardivillé · Christophe Mariller · E. Elass · Jean-Paul Perraudin · Annick Pierce

Received: 14 January 2010 / Accepted: 3 February 2010

Abstract The lactoferrin gene is known to be expressed either constitutively or under inducible conditions such as hormonal stimuli or inflammation. Its transcription from alternative promoters leads to two products, lactoferrin (Lf) and delta-lactoferrin (ΔLf) mRNAs the expressions of which are altered during oncogenesis. The comparison of the two enhancer/promoter regions revealed that the two isoforms might be differentially trans-activated. Nevertheless, concomitant expression of both transcripts has been found in some normal tissues and in a subset of breast cancer cell lines and biopsies. Moreover, we found putative inflammatory response elements in both P1 and P2 promoter regions suggesting that both Lf and ΔLf might be upregulated under inflammatory stimuli. Therefore, a duplex Taqman gene expression assay has been developed and used to profile mRNA expression of the Lf gene in the case of cancer and under inflammatory conditions. Discrimination between the two transcripts is achieved by using a primer pairs/probe set within exon 1β for ΔLf and a primer pairs/probe set within exon 1 and exon 2 for Lf. In this study, we confirmed that Lf/ΔLf Taqman gene expression assay is a powerful tool to investigate the expression of both Lf and ΔLf transcripts. We also showed that lymphocytes and leukocytes isolated from fresh human blood expressed an extremely high level of ΔLf messengers. An extensive series of cancer cell lines has been studied confirming that both P1 and P2 promoter regions of the Lf gene are downregulated or silenced in the case of cancer. Furthermore, using stimulation by bacterial lipopolysaccharides (LPS), we showed that in MDA-MB-231 and HT-29 epithelial cells, Lf expression is strongly increased with a higher expression level in MDA-MB-231 whereas ΔLf expression is not. These results suggest that the NF-κB/cRel response elements present in the P1 promoter region are functional whereas those present in the P2 promoter region are not and show that ΔLf is not regulated in inflammatory conditions.

Keywords Lactoferrin · Delta-lactoferrin · TaqMan real-time PCR · Q-PCR · Cancer · LPS

Abbreviations

Lf	Lactoferrin
ΔLf	Delta-lactoferrin

E. Hoedt · S. Hardivillé · C. Mariller · E. Elass · A. Pierce
Univ Lille Nord de France, 59000 Lille, France

E. Hoedt · S. Hardivillé · C. Mariller · E. Elass · A. Pierce
USTL, UGSF, 59655 Villeneuve d'Ascq, France

E. Hoedt · S. Hardivillé · C. Mariller · E. Elass · A. Pierce (✉)
Unité de Glycobiologie Structurale et Fonctionnelle, Université des Sciences et Technologies de Lille, UMR 8576 CNRS, IFR-147, 59655 Villeneuve d'Ascq, France
e-mail: annick.pierce@univ-lille1.fr

J.-P. Perraudin
Taradon Laboratory s.p.r.l, 1180 Bruxelles, Belgium

Published online: 14 February 2010

LPS	Lipopolysaccharides
FCS	Fetal calf serum
BSA	Bovine serum albumin
PBS	Phosphate buffer saline
NF-κB	Nuclear factor κB
STAT3	Signal transducer and activator of transcription 3

Introduction

The Lf gene appeared in mammals and is highly conserved among species with an identical organization (Teng 2002). It is mapped to human chromosome 3 at 3p21.3 (McCombs et al. 1988). Its transcription leads mainly to two products, lactoferrin (Lf) and delta-lactoferrin (ΔLf) mRNAs, which result from the use of alternative promoters, the P1 promoter for Lf and P2, located in the first intron of the Lf gene, for ΔLf (Siebert and Huang 1997; Liu et al. 2003). Therefore, alternative splicing concerns the first exon; exon 1 which characterizes the Lf mRNA is replaced by exon 1β in the ΔLf messenger (Siebert and Huang 1997). Selection of the P2 promoter and the use of the first available AUG codon in frame produce an alternative N-terminus. Thus, compared to Lf, its secretory counterpart, ΔLf is a cytoplasmic protein able to enter the nucleus (Mariller et al. 2007). The secretory Lf isoform is a powerful antimicrobial agent with immunodulatory activity which is also able to protect organisms against cancer and metastasis (Ward et al. 2005; Legrand et al. 2008; Pierce et al. 2009) whereas ΔLf, its intracellular counterpart, is a transcription factor that controls cell cycle progression and survival (Mariller et al. 2007; Pierce et al. 2009).

The two promoter regions of the Lf gene contain numerous constitutive and inducible regulatory elements that modulate both Lf and ΔLf transcriptions (Siebert and Huang 1997; Teng 2002, 2006; Liu et al. 2003). The P1 promoter is highly sensitive to nuclear receptors such as the estrogen receptor in the reproductive organs and to retinoic acid receptors. Concerning the P2 promoter, potential upstream regulatory elements different from those of P1, elevated expression in lymphoid cells and upregulation by Ets have been described (Siebert and Huang 1997; Liu et al. 2003). However, in many normal tissues Lf and ΔLf mRNAs coexist but are expressed at differential levels (Siebert and Huang 1997). In breast cancer cell lines and biopsies, both transcripts were always observed (Benaïssa et al. 2005). Taken together these data suggest that some regulatory elements might direct both transcriptions.

Lf expression is also increased during early steps of embryogenesis, following oxidative damage and in response to infection (Ward et al. 2005) and we previously demonstrated that Lf expression by human microglia was up-regulated in inflammatory conditions and under oxidative stress (Fillebeen et al. 2001). Moreover, the Lf gene is regulated by innate immune stimuli and cRel/NF-κB binding sites have been identified in the promoter of the bovine Lf gene and proved to be highly responsive to LPS induction (Zheng et al. 2005). Recently, it has been shown that Lf expression is induced in the bovine mammary gland in response to *Staphylococcus aureus* and *Escherichia coli* infection (Griesbeck-Zilch et al. 2008) and in normal murine mammary gland epithelial cells by LPS, a major component of the outer membrane of gram negative bacteria, or dsRNA which mimics viral infection (Li et al. 2009). Until now, no data are available concerning the putative presence of inflammatory response elements and their eventual functionality in the human Lf and ΔLf promoter regions.

Downregulation of the expression of Lf and ΔLf has been found in numerous tumors (Siebert and Huang 1997; Liu et al. 2003; Benaïssa et al. 2005; Ward et al. 2005). This is mainly due to structural alterations such as mutations, allelic loss of part of chromosome 3 and modification of the degree, as well as the pattern of methylation (Teng et al. 2004; Iijima et al. 2006). This genetic and epigenetic inactivation of the Lf gene in cancer may therefore provide the tumor cells with a selective growth advantage. Interestingly, we previously showed, using RT–PCR, that the expression level of either Lf or ΔLf mRNAs was of good prognosis value in human breast cancer with high concentrations being associated with longer relapse-free and overall survival, suggesting the usefulness of the detection of either Lf or ΔLf transcripts as markers for the follow-up of breast cancer patients (Benaïssa et al. 2005).

Since Lf isoforms may coexist in the same cells, be expressed under the same regulatory elements, downregulated in case of cancer and used as

prognosis tools, we have developed a very specific and highly sensitive method to discriminate and evaluate both Lf isoform transcripts in the same sample using duplex TaqMan real-time PCR.

Materials and methods

Cell culture and reagents

Human cervical carcinoma HeLa cells (ATCC CCL-2), human prostate cancer cell line DU 145 (ATCC HTB-81) and human hepatocarcinoma cell line Hep G2 (ATCC HB-8065), were a kind gift from Pr. T. Lefebvre (UGSF, UMR 8576 CNRS, Villeneuve d'Ascq, France). Human breast cancer MDA-MB-231 (ATCC HTB-26) and MCF7 (ATCC HTB-22) cell lines were kindly provided by Pr X. Lebourhis (INSERM U908, Villeneuve d'Ascq, France). Human melanoma cells SK-MEL-28 (ATCC HTB-72), human breast cancer BT-20 (ATCC HTB-19) and T-47D (ATCC HTB-133) cell lines were a generous gift of Pr. P. Delannoy (UGSF, UMR 8576 CNRS, Villeneuve d'Ascq, France). Human gastric cancer cells KATO III (ATCC HTB-103) were provided by Dr. A. Hardouin (UGSF, UMR 8576 CNRS, Villeneuve d'Ascq, France). Human colon carcinoma HT-29 cells (ATCC HTB-38), human promonocytic leukemia THP-1 cells (88081201, ECACC), normal human lung epithelial cells NL20 (ATCC CRL-2503) and lung adenocarcinoma cells Calu-3 (ATCC HTB-55) were a kind gift from Drs. C. Masselot and S. Degroux (UGSF, UMR 8576 CNRS, Villeneuve d'Ascq, France). Human leukemia acute T Jurkat cell line (ATCC TIB-152) and MCF 10A normal-like human breast cells (ATCC CRL-10318) were purchased from the American type culture collection (ATCC).

All these cell lines were routinely grown in Dulbecco's modified Eagle's medium or RPMI, containing 10% (v/v) FCS, 2 mM L-glutamine and 1% (w/v) penicillin/streptomycin. Cell lines were cultured at 37°C in a humidified atmosphere with 5% CO_2. The DAMI megakaryocytic cells were kindly provided by Dr S. Greenberg (Greenberg et al. 1990) and cultured as described (Nillesse et al. 1994).

Cell culture materials were obtained from Dutscher, and culture media and additives from Cambrex Corporation and Invitrogen. Mouse monoclonal FITC-conjugated anti-human CD14 antibody and mouse IgG isotype were purchased from Miltenyi Biotec and BSA from Sigma.

LPS induction

Cells were plated the day before induction in 6-well plates at a density of 10^6 cells/well. Prior to induction, cells were gently washed by fresh medium supplemented or not with FCS. Cells cultured without FCS were grown 24 h with FCS-free medium prior LPS treatment in FCS-free medium. Cells were then exposed or not to LPS (*E. coli* 055B5, Sigma) at a concentration of 100 ng/well during 24 h (Li et al. 2009). All experiments were conducted in triplicate. After washing with PBS, cells were harvested, centrifuged and lysed in an appropriate buffer for RNA assays.

RNA extraction and cDNA preparation

Total RNA extract from human normal mammary gland (HNMG) was purchased from Clontech. Total RNA from all cancer cell lines was isolated from cells using a NucleoSpin RNA II kit, according to the instructions of the manufacturer (Macherey–Nagel). The purity and integrity of each extract were checked using the nanodrop ND-1000 spectrophotometer (Labtech International) and the Bioanalyzer 2100 (Agilent Technologies). Reverse transcription was performed from 2 µg of total RNA with an oligo-dT primer and M-MLV reverse transcriptase (Promega). First strand cDNA from isolated human blood cells (leukocytes, lymphocytes, macrophages) were a kind gift of Dr. A. Denys (UGSF, UMR 8576 CNRS, USTL, Villeneuve d'Ascq, France).

RT–PCR conditions

RT–PCR assays were performed in triplicate as already described (Mariller et al. 2007). Negative control reactions were performed using sterile water instead of cDNA template. Contaminations of genomic DNA were excluded by performing 35 cycles of amplification without retrotranscription. GAPDH was used as internal control. Primer pairs were synthesized by Eurogentec. RT–PCR conditions specific to each primer pair are reported in Table 1. PCR products were separated onto a 1.5% agarose gel stained with

Table 1 Oligonucleotides used for TaqMan Q-PCR and RT–PCR

Method	Target DNA	Oligonucleotides (5′ → 3′)	Tm (°C)	Cycle number	Amplicon size (pb)
Q-PCR	Lf primers	Forward: cagaccgcagacatgaaacttg	60	40	117
		Reverse: gatacggcgcaccactgaa			
	Lf probe	ccctcggactgtgtct			
Q-PCR	ΔLf primers	Forward: ctgaagtttgaatcctgcagtcaat	60	40	66
		Reverse: aagggttgcaatggcacttt			
	ΔLf probe	tggtggcttgtaccc			
Q-PCR	HPRT primers	Forward: gccctggcgtcgtgatt	60	40	101
		Reverse: ctcagcataatgattaggtatgcaaaa			
	HPRT probe	tgatgatgaaccaggttat			
RT–PCR	GAPDH	Mariller et al. (2007)	69	16	240
RT–PCR	TLR4	Erridge et al. (2007)	57	25	336
RT–PCR	TNF-α	Satoh et al. (2000)	60	33	266

ethidium bromide and image acquisition was performed by UV transillumination using a Gel Doc 1000 system (Bio-Rad).

TaqMan Real-time PCR

Duplex Taqman Q-PCR and subsequent analyses were performed using the MX4000 Multiplex Quantitative System equipped with v3.0 software (Stratagene). Using Primer Express software v3.0 (Applied Biosystems) primers and probes were designed and their specificity checked using the BLAST algorithm on major nucleotide databases. Lf isoform primer pairs/probe sets are shown in Fig. 1. The transcript of HPRT (hypoxanthine guanine phosphoribosyltransferase) was used as a control to normalize the expression of genes of interest. The HPRT primer pair was designed to hybridize within exon 1 and exon 2. The nucleotide sequence of primer pairs and probes is presented in Table 1. Primer pairs were synthesized by Eurogentec and the TaqMan probes were from Applied Biosystems. The ΔLf and Lf probes were 5′-FAM-labeled and the normalizing HPRT gene probe was 5′-VIC-labeled. The 3′ non-fluorescent quencher (NFQ) (Applied Biosystems) was used for each probe. PCR reactions were performed with 12.5 µl of the QuantiTect Multiplex PCR Kit Mix (Qiagen),

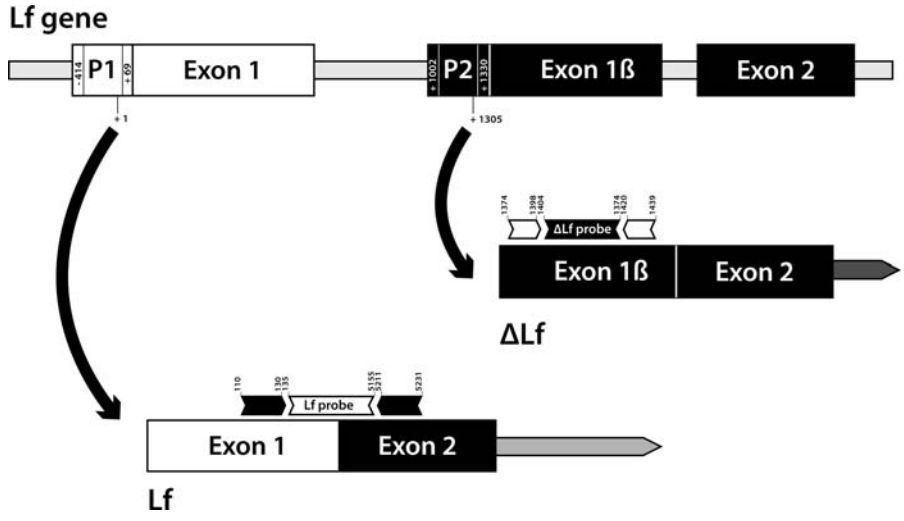


Fig. 1 Diagramatic representation of Lf and ΔLf primer pairs and TaqMan probe localization on the Lf gene. ΔLf primer pairs/probe set was designed to hybridize within exon 1β whereas for Lf, the primer pairs/probe set was designed to amplify a short region around the exon 1/exon 2 junction. The sense primer hybridizes within exon 1, the forward primer within exon 2 and the probe within the exon 1/exon 2 junction

0.6 μM of each primer, 0.2 μM of Taqman probe and 2 μl of cDNA in a final volume of 25 μl. After 95°C for 15 min, 40 PCR cycles were performed as follows: 94°C for 1 min, 60°C for 1 min (hybridization and elongation). Each TaqMan Q-PCR assay was performed in triplicate in 96-well plates. Two to three separate RNA extractions were carried out for each sample and each cDNA was analysed at least three times by Q-PCR. The amplification efficiency was generated for each primer pairs/probe set separately and in duplex with the primer pairs/probe set of the normalizing gene HPRT using 2-fold serial dilutions of cDNA. The fluorescence data were measured at the end of each cycle. The point at which the PCR product is first detected above a fixed threshold, termed cycle threshold (Ct), was determined for each sample, and the average Ct of triplicate samples was used for further calculations. In order to check amplification efficiency (E) of all three primer pairs/probe (Lf, ΔLf and HPRT), a validation experiment was performed using serial dilutions of each sample. Standard curves are obtained and the slopes of curves (*slope*) calculated. Efficiency is then established using the formula $E = 10^{[-1/slope]}$. Relative quantities of Lf and ΔLf mRNA were calculated as described (Pfaffl 2001) and expressed normalized to HPRT.

Flow cytometry analysis

THP-1, MDA-MB-231 and HT-29 cells were washed twice by centrifugation for 5 min at 1,100 rpm in cold Dulbecco's phosphate buffer saline (DPBS) (Sigma) and resuspended in DPBS supplemented with 0.5% bovine serum albumin (BSA) to obtain a final concentration of 10^6 cells/ml. Each pellet was resuspended in DPBS-BSA added with mouse monoclonal FITC-labelled anti-human CD14 (1/1,000) or mouse IgG isotype control (1/1,000) for 30 min at 4°C in the dark. Data were monitored on a Becton–Dickinson FACScan flow cytometer. The light-scatter channels were set on a linear gain, and the fluorescence channels were set on a logarithmic scale. Cells were gated for forward- and site-angle light scatters, and 10,000 fluorescent particles of the gated population were analyzed. The data collected with logarithmic amplification were analyzed by the CellQuest Pro v6.0 software.

Bioinformatics tools and promoter modelling

The P1 promoter of the Lf gene (Genbank NT_022517.18, *Homo sapiens* chromosome 3 genomic contig, 46450953-46446654) and the P2 promoter (Genbank NT_022517.18, *Homo sapiens* chromosome 3 genomic contig, 46446314-46445095) were analysed by checking matrix family assignment. All software employed were part of the Genomatix Suite/GEMS Launcher software package (http://www.genomatix.gsf.de). Potential control elements and transcription factor binding sites were analysed using MatInspector Professional software (Quandt et al. 1995). In order to ensure selection of high-quality binding sites throughout the library of binding-site matrices, MatInspector applies individually optimized matrix thresholds. The inflammatory response modules were defined as follows: NF-κB/cRel element (family V$NF-κB) with a core similarity of 1 and minimum matrix similarity of 0.87; STAT3 elements (family V$STAT) with a core similarity of 1 and minimal matrix similarity of 0.9.

Results and discussion

Design of Taqman probe and primer pairs and validation experiments

Since both human Lf and ΔLf transcripts have been often found in the same cell type at the same time, and at a low level in cancer cells, it was crucial to use a sensitive and highly specific method in order to discriminate them and assay their expression. We chose HPRT as an internal control gene since we previously checked that its expression was not altered in the cancer cell lines we used or during cell exposure to LPS. Therefore, we developed a duplex TaqMan real-time PCR assay which measures PCR-product accumulation during the exponential phase of the PCR reaction using labeled fluorogenic probes (Gibson et al. 1996). Minor groove binder (MGB) probes that incorporate a 5′ reporter dye (FAM or VIC) and a 3′ non-fluorescent quencher (NFQ) were used. The NFQ offers the advantage of a lower background signal which results in better precision in quantification (Kutyavin et al. 2000). FAM and VIC fluorescent reporter dyes were used to get a narrow spectrum and maximum of emission. Therefore, Lf

and ΔLf probes were FAM-labeled whereas HPRT was VIC-labelled at the 5′ end. Discrimination between the two transcripts was achieved as shown in Fig. 1.

Optimization of the assay was performed in order to obtain maximal efficiency of the primer/probe set. The maximum of efficiency was obtained using 600 nM primer pairs and 200 nM probe for each transcript. Efficiencies, determined for all the three primer pairs/probe sets, are 97.8% for ΔLf, 98% for Lf and 96.7% for HPRT. We confirmed that the efficiency was the same when duplex TaqMan Q-PCR assays (Lf/HPRT and ΔLf/HPRT) were carried out.

We next investigated whether the duplex TaqMan Q-PCR assay was sensitive enough to evaluate the respective expression of Lf and ΔLf transcripts in the case of downregulation as in cancerous cells and in the case of overexpression such as in inflammatory conditions.

Evaluation of Lf and ΔLf mRNA contents in cancerous and normal cells

Since both Lf and ΔLf are expressed in a wide variety of tissues, have antitumoral properties, are good prognosis markers and might act as tumor suppressors, detection of their expression levels should be extremely useful to follow tumor development. Therefore, we studied Lf and ΔLf expression levels in breast, lung, stomach, prostate, colon, skin, cervix, hepatocyte, megakaryocyte, monocyte and lymphocyte cancer cell lines, and seven normal human cells or tissues (Fig. 2) using the duplex TaqMan real-time reverse transcription PCR assay we developed.

We demonstrated that Lf mRNA was downregulated in three of the four breast cancer cell lines tested when compared with normal breast tissue (HNMG) or MCF 10A normal-like cell line. We found that only MDA-MB-231 cells exhibit a higher level of Lf mRNA compared to normal counterparts. Lf was at a very low level expressed by T-47D and BT-20 cells. Of the other cancer cell lines tested, only HT-29 and HeLa cells produced detectable Lf mRNA expression levels. Lf was expressed neither in normal (NL20) nor cancerous lung cells (Calu-3). For each sample the Lf relative expression normalized to HPRT is $\geq$0.5 confirming strong gene silencing. Lf mRNA was found in the lymphocyte population and not in the leukocyte population. The absence of Lf expression in leukocytes, which produce and store it in their secondary granules (Masson et al. 1969), is in accordance with its synthesis occurring only during granulocytic differentiation (Rado et al. 1987).

Concerning ΔLf, its mRNA was only absent from T-47D, THP-1 and Jurkat cells among the fourteen cancer cell lines tested. Its expression levels were very low with the highest expression found in DU 145 prostate cancer cell line. For each sample the ΔLf relative expression normalized to HPRT is $\geq$1.5 also confirming strong silencing of the P2 promoter. The sensitivity of our technique shows that ΔLf mRNA is detectable at very low levels in all tumour cell lines analyzed whereas it was found to be absent in all tumour cell lines examined using classical RT–PCR (Siebert and Huang 1997). Among the seven normal human tissues or isolated blood cells tested, both the leukocyte and lymphocyte populations expressed an extremely high level of ΔLf mRNA, with the hippocampus expressing a slightly lower level. The ΔLf relative expression normalized to HPRT is $\geq$60 for leukocytes and 25 for lymphocytes. The strong activity of the P2 promoter in these two cell types might be due to the presence of numerous functional Ets binding sites (Liu et al. 2003). Indeed, the proto-oncogene c-Ets, which is preferentially expressed in lymphoid cells and thymus (Chen 1985) was shown to strongly activate the P2 promoter (Liu et al. 2003). Therefore, the presence of Ets family member, which are involved in various cellular events such as cell growth, transformation, T-cell activation, hematopoietic cell differentiation, and development (Wasylyk et al. 1993), might be critical for ΔLf expression.

Presence of cRel/NF-κB/STAT3 promoter modules in the regulatory regions of the P1 and P2 promoters of the Lf gene

Bovine and murine Lf expressions are significantly up-regulated by LPS in a dose-dependent manner (Griesbeck-Zilch et al. 2008; Li et al. 2009). Recently, LPS-responsive modules were localized in the promoter region of the bovine Lf gene such as cRel/NF-κB, STAT3 and AP1 (Zheng et al. 2005). LPS through the Toll-like receptor (TLR) mediated-signaling pathway (Muzio and Mantovani 2000) triggers pro-inflammatory stimuli such as TNF-α and IL-1β (Karin 1999; Mercurio and Manning 1999) and activation of cRel/NF-κB transcription factors.

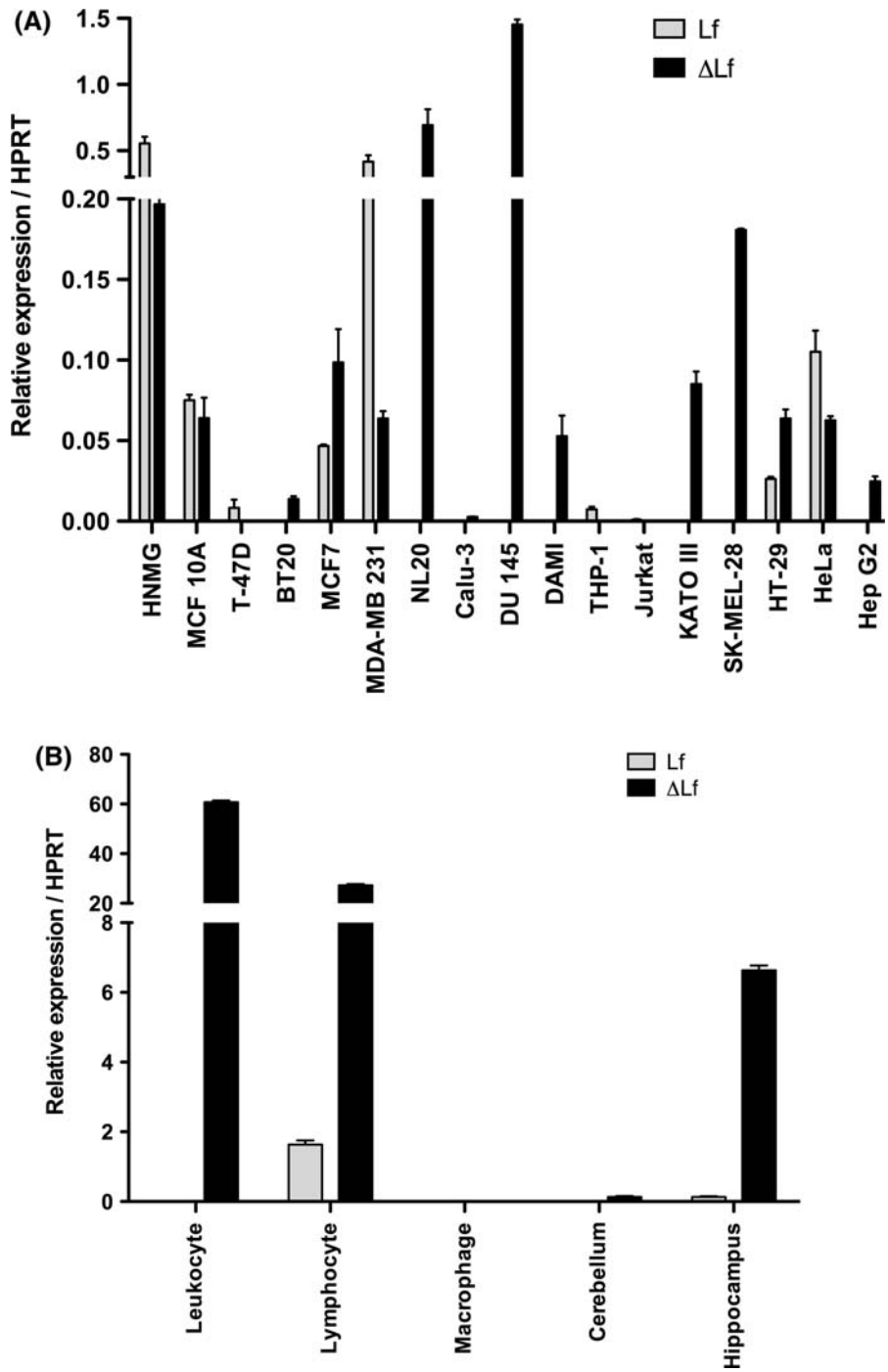


Fig. 2 Lf and ΔLf mRNA expression levels of various human cancer cell lines (**a**) and normal cells and tissues (**b**) using duplex Taqman Q-PCR. *Black* bars, ΔLf transcript; *grey* bars, Lf transcript. Values are normalized to HPRT gene expression. Data are means ± SD of n ≥ 6

The cRel/NF-κB transcription factor family regulate several physiological processes such as cellular homeostasis (Barkett and Gilmore 1999; Baldwin 2001) and the host immune response (Hayden and Ghosh 2008) by controlling the expression of an extremely high number of target genes (Pahl 1999). STAT3 is known as an acute phase response factor (Wegenka et al. 1994) and is also implicated in a variety of cellular functions, including inflammatory processes (Leonard and O'Shea 1998). AP1

regulates gene expression in response to a variety of stimuli, including cytokines, growth factors, stress, and bacterial and viral infections and in turn controls a number of cellular processes including differentiation, proliferation, and apoptosis.

Therefore, we investigated the Lf gene for the presence of such inflammatory response elements and subsequently checked whether Lf and/or ΔLf were overexpressed following LPS stimulation. The data presented in Fig. 3 show that potential NF-κB/cRel/STAT3 promoter modules are present in the regulatory regions of the P1 and P2 promoter regions of the Lf gene. Three NF-κB sites, five cRel sites and four STAT3 sites were found in the P1 promoter (Fig. 3a). The P2 promoter region also contains potential LPS-responsive elements, including two NF-κB sites, two cRel sites and two STAT3 sites (Fig. 3b). Our findings suggest that the human Lf promoters may respond to infection *via* the NF-κB pathway but further investigations will be required to ensure the biological significance of these putative responsive elements.

No AP1 site was found in either the P1 and P2 promoter regions although it was present in the bovine Lf promoter region (Zheng et al. 2005). BLAST alignment of the human and the bovine Lf promoter regions showed that neither the presence nor the localization of these elements is conserved (data not shown).

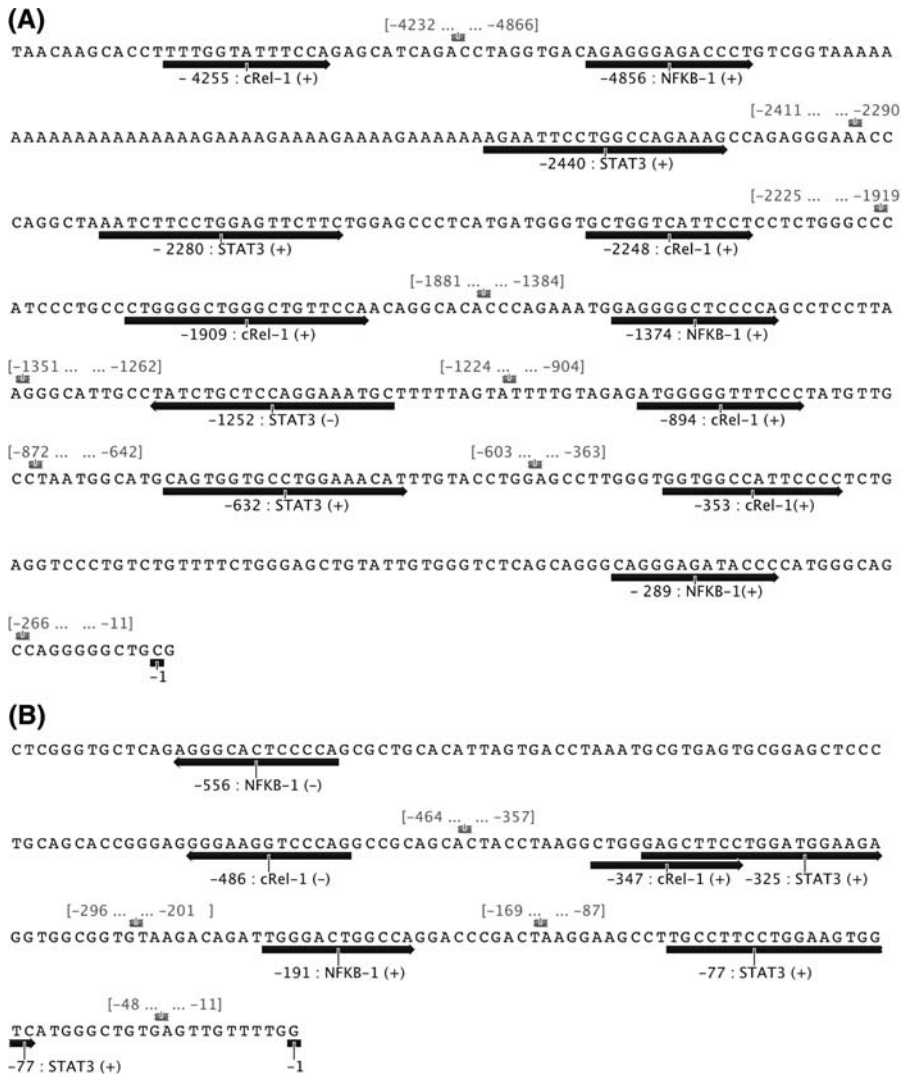


Fig. 3 Putative inflammatory response elements in the human Lf gene. The human P1 (**a**) and P2 (**b**) promoter regions were analyzed using MatInspector (Genomatix). Binding sites of infection responsive transcription factors (NF-κB, cRel1 and STAT3) are *underlined*. (+) plus strand, (−) minus strand. For easier reading, omitted sequences are presented in *grey* between brackets above sequence

LPS induces Lf but not ΔLf expression in human mammary gland epithelial MDA-MB-231 cells and in human colonic carcinoma HT-29 cells

Since inflammatory response modules are present in both the P1 and the P2 promoter regions we next investigated whether or not both human Lf isoforms are upregulated in inflammatory conditions. LPS is recognized by Toll-like receptor 4 (TLR4) (Akira and Hoshino 2003; Beutler 2005) that interacts with different proteins such as LPS binding protein (LBP), soluble CD14 (sCD14) or membrane CD14 (mCD14) to induce a signaling cascade leading to the activation of NF-κB and the production of proinflammatory cytokines such as TNF-α and IL-8 leading to a strong inflammatory response (Arditi et al. 1993; Hailman et al. 1994; Underhill and Ozinsky 2002; Pålsson-McDermott and O'Neill 2004). Therefore, human mammary gland epithelial MDA-MB-231 cells (Zaks-Zilberman et al. 2001) and human colonic epithelial HT-29 cells (Cario et al. 2000; Böcker et al. 2003; Lee et al. 2005) which are known to respond to LPS stimuli, were studied. The optimal LPS concentration (100 ng/10^6 cells) was chosen from previous experiments (data not shown) and in accordance to the literature (Li et al. 2009).

We first determined whether both cell lines expressed the CD14 epitope using flow cytometry. THP-1 were used as a known source of mCD14 (Fig. 4a, black shaded region). Unshaded and grey shaded regions overlapped the isotype control showing that MDA-MB-231 and HT-29 cell surfaces lack mCD14. Therefore, the next experiments were conducted with and without fetal calf serum (FCS) since normal plasma is a source of sCD14.

We next confirmed LPS-cell responsiveness by following TNF-α induction using RT–PCR. As shown in Fig. 4b, TNF-α is overexpressed following LPS induction compared to GAPDH. In both cell types the overexpression of TNF-α is higher in the presence of FCS confirming the importance of providing plasma sCD14 which acts as an opsonin that captures pathogenic microbes, facilitating recognition and binding of LPS to mCD14-negative intestinal and mammary gland epithelia (Pugin et al. 1993).

We then investigated whether TLR4 was overexpressed in the presence of LPS on these two cell lines.

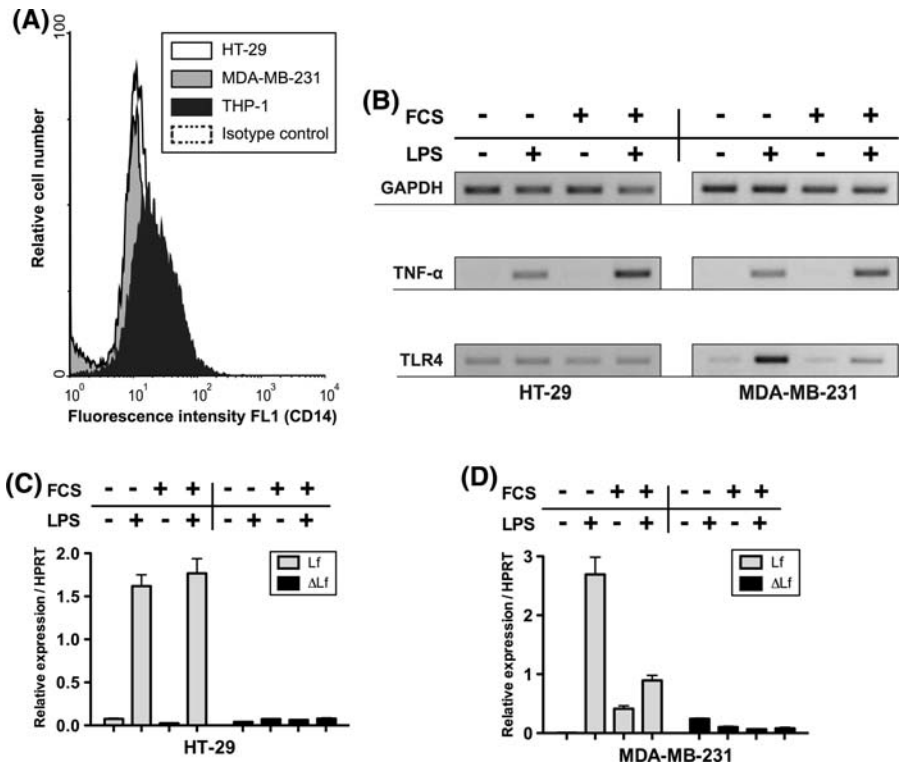


Fig. 4 Lf and not ΔLf is overexpressed under LPS stimulation. **a** Absence of mCD14 on HT-29 and MDA-MB-231 cells compared to differentiated THP-1 by flow cytometry. **b** Expression levels of TNF-α and TLR4 mRNAs compared to the expression level of the housekeeping gene GAPDH using RT–PCR. **c–d** Lf and ΔLf mRNA expression levels expressed as relative intensity compared to HPRT gene expression using duplex TaqMan Q-PCR in HT-29 (**c**) and in MDA-MB-231 (**d**) cells

Responsiveness of HT-29 cells to LPS is known to be correlated with the presence of TLR4 (Böcker et al. 2003). Here, we showed that TLR4 expression is not modulated in the presence of LPS and/or FCS suggesting that it might be constitutively expressed (Fig. 4b). Breast cancer cell responsiveness to LPS has been described previously. TLR4 has been described as being present using a DNA array assay (Merrell et al. 2006) or absent using RT–PCR (Xie et al. 2009) in MDA-MD-231 cells. Here, we showed that TLR4 is effectively expressed but at a very low level and overexpressed following LPS exposure in MDA-MB-231 cells. Expression was higher in the absence of FCS while the inflammatory response was not increased in the same conditions (Fig. 4b).

We further determined whether the human Lf gene was upregulated in inflammatory conditions using our duplex TaqMan Q-PCR assay. Figure 4c shows that Lf expression is poorly expressed without LPS stimulation but is increased 20-fold when HT-29 cells are exposed to LPS in the absence of FCS and 70-fold in the presence of FCS and LPS. Our results confirm that the human Lf gene, like its mouse and bovine counterparts, is upregulated following LPS induction and that the presence of FCS containing sCD14 favors this upregulation. ΔLf, which is also feebly expressed in HT-29 cells, was not upregulated in inflammatory conditions.

Figures 2a and 4d showed that Lf expression is poorly expressed in the absence of stimulation in MDA-MB-231 cells. Overexpression is very marked when cells are exposed to LPS in the absence of FCS (450-fold) whereas this overexpression is dramatically reduced (2-fold) in the presence of FCS. This result is mainly due to the 70-fold higher expression of the Lf gene in the MDA-MB-231 cells cultured in FCS. Lf upregulation may be correlated to TLR4 inducible expression higher in serum free media and LPS exposure. In all cases, Lf is upregulated in presence of LPS whereas under the same experimental conditions, no upregulation of ΔLf expression is detectable.

Our data show that the human MDA-MB-231 mammary gland epithelial cell line and the human colonic HT-29 cancer cell line might be used to study responses to innate immune stimuli. We also confirmed that the inflammation-responsive elements present in the P1 promoter of the Lf gene are functional. Moreover we showed for the first time that even if cRel/NF-κB/STAT3 response modules are present in the P2 promoter, they are not functional and that ΔLf expression is not regulated in response to bacterial infection. We also demonstrated that our duplex TaqMan Q-PCR assay is able to evaluate extremely variable quantities of Lf isoform transcripts.

Our results are in accordance with the function of both Lf isoforms. Lf, which is a multitasking protein, is mainly a powerful antimicrobial agent with immunodulatory activities involved in host defense (Legrand et al. 2008; Pierce et al. 2009). Its production under LPS stimulation will reduce LPS and its proinflammatory effects (Li et al. 2009). In contrast ΔLf is a transcription factor involved in the control of cell cycle progression, mRNA decay and apoptosis (Mariller et al. 2007, 2009) the expression of which is not influenced in response to inflammatory conditions.

Acknowledgments This investigation was supported in part by the CNRS UMR 8576 (Unité de Glycobiologie Structurale et Fonctionnelle), the Institut Fédératif de Recherche n° 147, the Université des Sciences et Technologies de Lille I, the Comité du Nord de la Ligue Nationale contre le Cancer and the Association pour la Recherche sur le Cancer («*Etude du rôle de la delta-lactoferrine, des gangliosides et des neurotrophines dans le développement et la progression du cancer du sein*»). We are grateful to Dr. A. Denys (UGSF, UMR 8576 CNRS, USTL, Villeneuve d'Ascq, France) for her help with the flow cytometer analyses. We would like to thank Dr. R. J. Pierce (CIIL, Institut Pasteur de Lille, France) for critical reading of this manuscript. Esthelle Hoedt is supported by Grain-Root Corporation (Taiwan), Dalian F.T.Z. New baili Int'L Industry and Trade Co Ltd (China) and Clever-Net LLC (USA).

References

Akira S, Hoshino K (2003) Myeloid differentiation factor 88-dependent and -independent pathways in toll-like receptor signaling. J Infect Dis 187(Suppl 2):S356–S363

Arditi M, Kabat W, Yogev R (1993) Antibiotic-induced bacterial killing stimulates tumor necrosis factor-alpha release in whole blood. J Infect Dis 167:240–244

Baldwin AS (2001) Control of oncogenesis and cancer therapy resistance by the transcription factor NF-kappaB. J Clin Investig 107:241–246

Barkett M, Gilmore TD (1999) Control of apoptosis by Rel/NF-kappaB transcription factors. Oncogene 18:6910–6924

Benaïssa M, Peyrat J-P, Hornez L, Mariller C, Mazurier J, Pierce A (2005) Expression and prognostic value of lactoferrin mRNA isoforms in human breast cancer. Int J Cancer 114:299–306

Beutler B (2005) The toll-like receptors: analysis by forward genetic methods. Immunogenetics 57:385–392

Böcker U, Yezerskyy O, Feick P, Manigold T, Panja A, Kalina U, Herweck F, Rossol S, Singer MV (2003) Responsiveness of intestinal epithelial cell lines to lipopolysaccharide is correlated with toll-like receptor 4 but not toll-like receptor 2 or CD14 expression. Int J Colorectal Dis 18: 25–32

Cario E, Rosenberg IM, Brandwein SL, Beck PL, Reinecker HC, Podolsky DK (2000) Lipopolysaccharide activates distinct signaling pathways in intestinal epithelial cell lines expressing toll-like receptors. J Immunol 164:966–972

Chen JH (1985) The proto-oncogene c-ets is preferentially expressed in lymphoid cells. Mol Cell Biol 5:2993–3000

Erridge C, Webb DJ, Spickett CM (2007) Toll-like receptor 4 signalling is neither sufficient nor required for oxidised phospholipid mediated induction of interleukin-8 expression. Atherosclerosis 193:77–85

Fillebeen C, Ruchoux MM, Mitchell V, Vincent S, Benaïssa M, Pierce A (2001) Lactoferrin is synthesized by activated microglia in the human substantia nigra and its synthesis by the human microglial CHME cell line is upregulated by tumor necrosis factor alpha or 1-methyl-4-phenylpyridinium treatment. Brain Res Mol Brain Res 96:103–113

Gibson UE, Heid CA, Williams PM (1996) A novel method for real time quantitative RT-PCR. Genome Res 6:995–1001

Greenberg SM, Chandrasekhar C, Golan DE, Handin RI (1990) Transforming growth factor beta inhibits endomitosis in the Dami human megakaryocytic cell line. Blood 76:533–537

Griesbeck-Zilch B, Meyer HHD, Kühn CH, Schwerin M, Wellnitz O (2008) *Staphylococcus aureus* and *Escherichia coli* cause deviating expression profiles of cytokines and lactoferrin messenger ribonucleic acid in mammary epithelial cells. J Dairy Sci 91:2215–2224

Hailman E, Lichenstein HS, Wurfel MM, Miller DS, Johnson DA, Kelley M, Busse LA, Zukowski MM, Wright SD (1994) Lipopolysaccharide (LPS)-binding protein accelerates the binding of LPS to CD14. J Exp Med 179:269–277

Hayden MS, Ghosh S (2008) Shared principles in NF-kappaB signaling. Cell 132:344–362

Iijima H, Tomizawa Y, Iwasaki Y, Sato K, Sunaga N, Dobashi K, Saito R, Nakajima T, Minna JD, Mori M (2006) Genetic and epigenetic inactivation of LTF gene at 3p21.3 in lung cancers. Int J Cancer 118:797–801

Karin M (1999) The beginning of the end: IkappaB kinase (IKK) and NF-kappaB activation. J Biol Chem 274:27339–27342

Kutyavin IV, Afonina IA, Mills A, Gorn VV, Lukhtanov EA, Belousov ES, Singer MJ, Walburger DK, Lokhov SG, Gall AA, Dempcy R, Reed MW, Meyer RB, Hedgpeth J (2000) 3′-Minor groove binder-DNA probes increase sequence specificity at PCR extension temperatures. Nucleic Acids Res 28:655–661

Lee SK, Il Kim T, Kim YK, Choi CH, Yang KM, Chae B, Kim WH (2005) Cellular differentiation-induced attenuation of LPS response in HT-29 cells is related to the down-regulation of TLR4 expression. Biochem Biophys Res Commun 337:457–463

Legrand D, Pierce A, Elass E, Carpentier M, Mariller C, Mazurier J (2008) Lactoferrin structure and functions. Adv Exp Med Biol 606:163–194

Leonard WJ, O'Shea JJ (1998) Jaks and STATs: biological implications. Annu Rev Immunol 16:293–322

Li Y, Limmon GV, Imani F, Teng C (2009) Induction of lactoferrin gene expression by innate immune stimuli in mouse mammary epithelial HC-11 cells. Biochimie 91:58–67

Liu D, Wang X, Zhang Z, Teng CT (2003) An intronic alternative promoter of the human lactoferrin gene is activated by Ets. Biochem Biophys Res Commun 301:472–479

Mariller C, Benaissa M, Hardiville S, Breton M, Pradelle G, Mazurier J, Pierce A (2007) Human delta-lactoferrin is a transcription factor that enhances Skp1 (S-phase kinase-associated protein) gene expression. FEBS J 274:2038–2053

Mariller C, Hardivillé S, Hoedt E, Benaïssa M, Mazurier J, Pierce A (2009) Proteomic approach to the identification of novel delta-lactoferrin target genes: characterization of DcpS, an mRNA scavenger decapping enzyme. Biochimie 91:109–122

Masson PL, Heremans JF, Schonne E (1969) Lactoferrin, an iron-binding protein in neutrophilic leukocytes. J Exp Med 130:643–658

McCombs JL, Teng CT, Pentecost BT, Magnuson VL, Moore CM, McGill JR (1988) Chromosomal localization of human lactotransferrin gene (LTF) by in situ hybridization. Cytogenet Cell Genet 47:16–17

Mercurio F, Manning AM (1999) Multiple signals converging on NF-kappaB. Curr Opin Cell Biol 11:226–232

Merrell MA, Ilvesaro JM, Lehtonen N, Sorsa T, Gehrs B, Rosenthal E, Chen D, Shackley B, Harris KW, Selander KS (2006) Toll-like receptor 9 agonists promote cellular invasion by increasing matrix metalloproteinase activity. Mol Cancer Res 4:437–447

Muzio M, Mantovani A (2000) Toll-like receptors. Microbes Infect 2:251–255

Nillesse N, Pierce A, Lecocq M, Benaissa M, Spik G (1994) Expression of the lactotransferrin receptor during the differentiation process of the megakaryocyte Dami cell line. Biol Cell 82:149–159

Pahl HL (1999) Activators and target genes of Rel/NF-kappaB transcription factors. Oncogene 18:6853–6866

Pålsson-McDermott EM, O'Neill LAJ (2004) Signal transduction by the lipopolysaccharide receptor, toll-like receptor-4. Immunology 113:153–162

Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29:e45

Pierce A, Legrand D, Mazurier J (2009) Lactoferrin: a multifunctional protein. Med Sci (Paris) 25:361–369

Pugin J, Schürer-Maly CC, Leturcq D, Moriarty A, Ulevitch RJ, Tobias PS (1993) Lipopolysaccharide activation of human endothelial and epithelial cells is mediated by lipopolysaccharide-binding protein and soluble CD14. Proc Natl Acad Sci USA 90:2744–2748

Quandt K, Frech K, Karas H, Wingender E, Werner T (1995) MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res 23:4878–4884

Rado TA, Wei XP, Benz EJ (1987) Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA during normal and leukemic myelopoiesis. Blood 70:989–993

Satoh M, Nakamura M, Satoh H, Saitoh H, Segawa I, Hiramori K (2000) Expression of tumor necrosis factor-alpha-converting enzyme and tumor necrosis factor-alpha in human myocarditis. J Am Coll Cardiol 36:1288–1294

Siebert PD, Huang BC (1997) Identification of an alternative form of human lactoferrin mRNA that is expressed differentially in normal tissues and tumor-derived cell lines. Proc Natl Acad Sci USA 94:2198–2203

Teng CT (2002) Lactoferrin gene expression and regulation: an overview. Biochem Cell Biol 80:7–16

Teng CT (2006) Factors regulating lactoferrin gene expression. Biochem Cell Biol 84:263–267

Teng C, Gladwell W, Raphiou I, Liu E (2004) Methylation and expression of the lactoferrin gene in human tissues and cancer cells. Biometals 17:317–323

Underhill DM, Ozinsky A (2002) Phagocytosis of microbes: complexity in action. Annu Rev Immunol 20:825–852

Ward PP, Paz E, Conneely OM (2005) Multifunctional roles of lactoferrin: a critical overview. Cell Mol Life Sci 62: 2540–2548

Wasylyk B, Hahn SL, Giovane A (1993) The Ets family of transcription factors. Eur J Biochem 211:7–18

Wegenka UM, Lütticken C, Buschmann J, Yuan J, Lottspeich F, Müller-Esterl W, Schindler C, Roeb E, Heinrich PC, Horn F (1994) The interleukin-6-activated acute-phase response factor is antigenically and functionally related to members of the signal transducer and activator of transcription (STAT) family. Mol Cell Biol 14:3186–3196

Xie W, Wang Y, Huang Y, Yang H, Wang J, Hu Z (2009) Toll-like receptor 2 mediates invasion via activating NF-kappaB in MDA-MB-231 breast cancer cells. Biochem Biophys Res Commun 379:1027–1032

Zaks-Zilberman M, Zaks TZ, Vogel SN (2001) Induction of proinflammatory and chemokine genes by lipopolysaccharide and paclitaxel (Taxol) in murine and human breast cancer cell lines. Cytokine 15:156–165

Zheng J, Ather JL, Sonstegard TS, Kerr DE (2005) Characterization of the infection-responsive bovine lactoferrin promoter. Gene 353:107–117

RESEARCH COMMUNICATION

Construction of a Multi-Functional Helper-Dependent Adenovirus Based System for Cancer Gene Therapy

Ne Long[1], Stephan Hardiville[2], Annick Pierce[2], Katsume Fukamachi[3], Mitsuru Futakuchi[3], David B Alexander[3]*, Tomoyuki Shirai[1], Hiroyuki Tsuda[3,4]

Abstract

Adenovirus holds great promise as a gene delivery system; it can hold large amounts of exogenous DNA and can be chemically and genetically modified to improve targeting to specific cells and tissues. A recombinant adenovirus construct expressing p53 is currently in clinical use as a cancer therapy in China. However, the use of adenovirus constructs in therapy is limited due to patients' strong immune response against these viruses and their gene products. To overcome this problem helper-dependent adenoviruses which do not express any viral gene products have been developed. Because the helper-dependent viruses do not express any viral gene products, a helper virus is required for their replication and encapsidation into infectious particles. This manuscript describes the construction of a prototype helper-dependent adenovirus system built such that it can be easily modified. The helper-dependent virus described here is built of a series of four cassettes, each with its own function. Furthermore, each individual cassette can be removed and replaced with a cassette with a different function. In this way, different helper-dependent viruses can be readily created. This type of system could be very useful in cancer therapy: For example, libraries of different cassettes could be maintained, allowing rapid assembly of constructs able to provide therapy for individual tumor types.

Key Words: Adenovirus - helper-dependent - cancer gene therapy - ADMM - EVE

Asian Pacific J Cancer Prev, **10**, 939-960

Introduction

Adenovirus based vectors are being developed in many laboratories for use in gene therapy clinical trials (Cross and Burmester, 2006; Kim et al., 2008; Shirakawa, 2008; also see http://www.wiley.co.uk/genmed/clinical/), and in China a recombinant adenoviral vector expressing p53, marketed under the name gendicine, is in clinical use as a cancer therapy (Peng, 2005). The use of adenovirus itself as the basis of a gene therapy construct is limited due to patients' strong immunogenic response against these viruses and their gene products. To decrease the patient's immunogenic response against adenovirus based constructs, constructs are being made which do not express any viral gene products. These constructs are known as helper-dependent viruses (Alba et al., 2005; Jozkowicz and Dulak, 2005; Xu et al., 2005). Because helper-dependent vectors do not encode any viral gene products, they require a helper virus, which does encode and express viral proteins, for propagation and encapsidation into infectious particles.

This manuscript describes the construction of a prototype helper-dependent adenovirus based system which can be easily modified for use as a cancer therapy. We refer to our helper-dependent constructs as adenovirus-dependent molecular medicine (ADMM), and the helper virus is referred to as EVE. Our ADMM construct consists of two adenovirus inverted terminal repeats (ITRs) which act as DNA replication origins, the adenovirus packaging site which is required for encapsidation of the ADMM DNA into infectious particles, and four cassettes. Each of these cassettes has its own function. By building a vector able to manifest multiple functions, a versatile vector with high specificity and cancer killing ability can be created. Moreover, each cassette can be removed and replaced with a cassette with a different function, allowing ready construction of ADMMs with specificity for different tumor types.

The prototype ADMM described here contains cassettes with the following functions: (1) expression of a reporter gene, LacZ, under the control of a tetracycline responsive promoter; (2) expression of a suicide gene, HSV-TK, enabling killing of the host cell, also under the control of a tetracycline responsive promoter; (3) spacer DNA; and (4) expression of a transcription factor, tTA, which binds to tetracycline responsive promoters in the absence of tetracycline (Freundlieb et al., 1999) and activates transcription.

*[1]Department of Experimental Pathology and Tumor Biology, [3]Department of Molecular Toxicology, [4]Nanotoxicology Project, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan, [2]Université des Sciences et Technologies de Lille, Villeneuve d'Ascq cedex, France *For Correspondence: dalexand@med.nagoya-cu.ac.jp*

Materials and Methods

Materials

pGEM®-T Easy (Cat. No. A3600) was purchased from Promega. The BD Adeno-X™ Expression System 1 kit (Cat. No. K1650-1) was purchased from BD Bioscience (Adeno-X™ Expression System 1 is now sold by Clontech). pIRES (Cat. No. 631605) was purchased from BD Biosciences (pIRES is now sold by Clontech). pEGFP-N1 (Cat. No. 6085-1) was purchased from BD Biosciences (pEGFP-N1 was sold by Clontech but EGFP has been replaced by more highly optimized versions of green fluorescent protein, such as ZsGreen1 and AcGFP1; pEGFP-N1 Cat. No. 6085-1 is now discontinued). The BD™ Tet-Off Gene Expression System (Cat. No. 630921) was purchased from BD Biosciences (630921 has been replaced by an improved Tet-Off Advanced Inducible Gene Expression System, 630934, and is now sold by Clontech). pxCANCre (Cat. No. 1675) was purchased from Riken BioResource Center with the permission of the depositor Izumu Saito. The Q-mate™ Inducible Expression System was purchased from Q-BIOgene (Q-mate™ Inducible Expression System is now sold by Krackeler Scientific, Inc.). pTRE-Tight (Cat. No. 631059) was purchased from Clontech. pCM-TK (Cat. No. 5953) was purchased from Riken BioResource Center with the permission of the depositor Hirofumi Hamada. FuGENE 6 was purchased from Roche Applied Science. SuperScript™ III reverse transcriptase was purchased from Invitrogen.

Molecular Biology

Unless otherwise noted, standard molecular biology techniques were used. Amplification of DNA sequences used in the construction of EVE and ADMM was initially performed using TaKaRa Ex TaqTM or TaKaRa LA TaqTM (RR006 and RR042; Takara Bio Inc.), and the amplicons were ligated into pGEM®-T Easy vectors. However, during the construction of ADMM, Takara Bio. Inc. introduced a new PCR polymerase named PrimeSTAR (Cat. No. R044). This polymerase has an extremely low mistake rate and was used in the construction of ADMM after it became available.

Amplification with PrimeSTAR results in blunt ended amplicons, consequently, these amplicons must be phosphorylated and ligated into blunt-ended vectors. Whenever PCR amplicons or DNA oligomers were inserted into blunt-ended plasmid backbones, the inserts were first phosphorylated (and the blunt-ended backbones were dephosphorylated); the DNA oligomers and amplicons were phosphorylated using polynucleotide kinase (Cat. No. 2021; Takara Bio Inc.) according to the manufacturer's instructions. DNA oligomers were purchased from Invitrogen. When constructing plasmids of 20 kb or greater, ligations were performed using TaKaRa DNA Ligation Kit LONG (Cat. No. 6024; Takara Bio Inc.) according to the manufacturer's instructions. Sequencing analysis was performed using an ABI Prism 3100 Genetic Analyzer (Applied Biosystems) according to the manufacturer's instructions.

DNA less than 10 kb was purified using DNA purification kits purchased from Qiagen and Promega. Larger plasmids were purified using DNA purification kits purchased from Qiagen. DNA fragments > 10 kb were purified as follows: DNA fragments were separated by electrophoresis through an agarose gel; the fragment of interest was cut out of the gel; the gel was dissolved in Membrane Binding Solution from Wizard®SV Gel and PCR Clean-Up System (Promega); the dissolved gel containing the DNA fragment of interest was applied to a QIAprep Spin Miniprep column (Qiagen) and centrifuged for 30 sec; the column was washed once with Buffer PB and washed twice with Buffer PE; and the DNA was eluted with 50 µl Buffer EB heated to 80˚C. Because the DNA tended to be dilute after this purification, > 10 kb restriction digest fragments were obtained from approximately 2 µg of plasmid DNA.

Plasmids were propagated in Stbl3 cells purchased from Invitrogen (Cat. No. C7373-03). These cells were grown at 37˚C. If the plasmid of interest could not be propagated at 37˚C, cells were grown at 32˚C. If the plasmid of interest could not be propagated at 32˚C, cells were grown at 28˚C: In general, larger, complicated plasmids such as pEVE, pC4-H_Tight-LacZ, and pADMM were propagated in cells grown at 28˚C.

Stuffer DNA

Genomes were scanned for sequences similar to 5'-CCANNAGNNGGC-3' using BLAST (NCBI). Intergenic regions containing positive hits were considered acceptable candidates for stuffer DNA.

Construction of the helper adenovirus (EVE)

Construction of pBS_PITR_Pack-wt:

Oligomer 5'-TCACGTGAA-3' was annealed to itself and then ligated into pGEM®-T Easy to generate pGEM_closed. Oligomers 5'-GGCCGCTGGCCATTTTCGCGGGAAAACTGAATA-3' and 5'-CTCTTATTCAGTTTTCCCGCGAAAATGGCCAGC-3' were annealed together and oligomers 5'-AGAGGAAGTGAAATCTGAATAATTTTGTGTTACTCATAGCGCGCG-3' and 5'-TCGACGCGCGCTATGAGTAACACAAAATTATTCAGATTTCACTTC-3' were annealed together. pGEM_closed was digested with NotI and SalI, and the annealed oligomers were ligated into the pGEM_closed backbone (triple ligation). Oligomers 5'-CTAGAACGCGTAATATTTGTCTAGGGCCGCGGGG-3' and 5'-AAGTCCCCGCGGCCCTAGACAAATATTACGCGTT-3' were annealed together and oligomers 5'-ACTTTGACCGTTTACGTGGAGACTCGGGATCCTTC-3' and 5'-CCGGGAAGGATCCCGAGTCTCCACGTAAACGGTCA-3' were annealed together. pBS_KS+ was digested with XbaI and XmaI, and the annealed oligomers were ligated into the pBS_KS+ backbone (triple ligation). The pGEM_oligomer construct was digested with ScaI and BssHII. The pBS_oligomer construct was digested with ScaI and MluI, and the insert from the pGEM_oligomer was ligated into the pBS_oligomer backbone to generate pBS_Pack-3'. The adenovirus inverted terminal repeat (ITR) and a 5' PacI site and the 5' portion of the adenovirus packaging site was amplified from Adeno-X (Adeno-X is part of the BD Adeno-X™ Expression System 1 kit) using primers

GGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATC
AAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAG
TATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTC
AGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTAC
GATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTC
ACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGG
TCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAG
TAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTC
ACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTAC
ATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAG
AAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTAC
TGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTG
AGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGC
GCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACT
CTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTG
ATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAA
TGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTT
TCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATG
TATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTAA
ATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTT
TTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATA
GGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAAC
GTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAA
TCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCC
CGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCG
AAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACA
CCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAAC
TGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGA
TGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAA
ACGACGGCCAGTGAGCGCGCGAATTAACCCTCACTAAAGGGAACAAAAGCTGTTACTACGT
TAATTAA**CATCATCAATAATATACCTTATTTTGGATTGAAGCCAATATGATAATGAGGGG**
GTGGAGTTTGTGACGTGGCGCGGGGCGTGGGAACGGGGCGGGTGACGTAGTAGTGTGGCG
GAAGTGTGATGTTGCAAGTGTGGCGGAACACATGTAAGCGACGGATGTGGCAAAAGTGAC
GTTTTTGGTGTGCGCCGGTGTACACAGGAAGTGACAATTTCGCGCGGTTTTAGGCGGAT
GTTGTAGTAAATTTGGGCGTAACCGAGTAAGATTTGGCCATTTTCGCGGGAAAACTGAAT
AAGAGGAAGTGAAATCTGAATAATTTTGTGTTACTCATAGCGCGTAATATTTGTCTAGGG
CCGCGGGGACTTTGACCGTTTACGTGGAGACTCGGGATCCTTCCCGGGCTGCAGGAATTC
CTGCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAATTC
GCCCTATAGTGAGTCGTATTACGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGT
GTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGGAGCATAAAGTGTAA
AGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGC
TTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAG
AGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGT
CGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGA
ATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCG
TAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAA
AAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTT
TCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCT
GTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCT
CAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCC
CGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTT
ATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGC
TACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTAT
CTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAA
ACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAA
AAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTAC

Figure 1. Putative Sequence of pBS_PITR_Pack-wt. The ampicilin resistance gene is in blue; the PacI site is in green; the inverted terminal repeat is in black; the packaging site is in yellow; and the origin of replication is in cyan. During insertion of the PacI site into the blunted KpnI site of pBS_rKS a "G" residue was lost during the blunting reaction. (Sequence data is best viewed by copying the sequence and pasting it into a program such as Gene Construction Kit or into a word processing program)

5'-CCCCAGGCTTTACACTTTATGC-3' and 5'-CTTTCGCTACCTTAGGACCGTTAT-3' and ligated into pGEM®-T Easy to generate pGEM_PITR_Pack-5' (pGEM_[PacI-Inverted Terminal Repeat]_[5' portion of the wild type packaging site]). pGEM_PITR_Pack-5' was digested with MscI and NotI. pBS-Pack-3' was digested with MscI and NotI, and the PITR-Pack-5' insert was ligated into the pBS-Pack-3' backbone to generate pBS_PITR_Pack-wt_Ex (pBS_[PITR]_[wild type packaging site]_[with extraneous sequences]). pBS_rKS was constructed by digesting pBS_KS+ with BssHII and ligating the BssHII fragment back into the parent pBS_KS+ backbone, and selecting clones with the MCS oriented in the reverse direction. pBS_rKS was digested with KpnI, blunted and dephosphorylated, and a PacI site (5'-TTACTACGTTAATTAACGATT-3') ligated into the pBS_rKS backbone to generate pBS_rKS-PacI. pBS-PITR_Pack-wt_Ex was digested with PacI and EcoRI. pBS_rKS-PacI was digested with PacI and EcoRI, and the PITR_Pack-wt insert was ligated into the pBS_rKS-PacI backbone to generate pBS_PITR_Pack-wt. The

GGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATC
AAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAG
TATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTC
AGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTAC
GATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTC
ACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGG
TCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAG
TAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTC
ACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTAC
ATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAG
AAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTAC
TGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTG
AGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGC
GCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACT
CTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTG
ATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAA
TGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTT
TCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATG
TATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTAA
ATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTT
TTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATA
GGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAAC
GTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAA
TCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCC
CGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCG
AAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACA
CCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAAC
TGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGA
TGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAA
ACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCACCGC
GCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGTTACTACG**TTAATTAA****CATCATCAA**
TAATATACCTTATTTTGGATTGAAGCCAATATGATAATGAGGGGGTGGAGTTTGTGACGT
GGCGCGGGGCGTGGGAACGGGGCGGGTGACGTAGTAGTGTGGCGGAAGTGTGATGTTGCA
AGTGTGGCGGAACACATGTAAGCGACGGATGTGGCAAAAGTGACGTTTTTGGTGTGCGCC
GGATAACTTCGTATAATGTATGCTATACGAAGTTATCCGGTGTACACAGGAAGTGACAAT
TTTCGCGCGGTTTTAGGCGGATGTTGTAGTAAATTTGGGCGTAACCGAGTAAGATTTGGC
CATTTTCGCGGGAAAACTGAATAAGAGGAAGTGAAATCTGAATAATTTTGTGTTACTCAT
AGCGCGTGGATCCACTAATAACTTCGTATAATGTATGCTATACGAAGTTATAGTTCTAGA
GCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCATCTATGTCGGGTGCG
GAGAAAGAGGTAATGAAATGGCATCGACTCGAAGATCTGAATTTAACTATAACGGTCCTA
AGGTAGCGAAAGCTCAGATCCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTT
GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA
CAACATACGAGCCGGGAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT
CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT
GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC
TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA
CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG
AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA
TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA
CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC
TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC
GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT
GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG
TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG
GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA
CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG
AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT
TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT
TTCTAC

Figure 2. Putative Sequence of pBS_PITR_LΔ5L_SC. The ampicilin resistance gene is in blue; the PacI site is in green; the inverted terminal repeat is in black; LoxP sites are in magneta; the packaging site is in yellow; the PI-SceI and I-CeuI sites are in green; and the origin of replication is in cyan. During insertion of the second LoxP site into the blunted SpeI site of pBS_PITR_LoxP-PackΔ5 a "G", a "C", and a "T" residue were lost during the blunting reaction

sequence of pBS_PITR_Pack-wt is shown in Figure 1.

Construction of pBS_PITR_LΔ5L_SC (pBS_[PITR]_[LoxP-PackΔ5-LoxP]_[PI-SceI_I-CeuI]):

pBS_PITR_Pack-wt was used to generate a sequence containing a PITR and a shortened packaging site by PCR amplification using primers 5'-GTCGAGGTGCCGTAAAGCA-3' and 5'-GGATCCACGCGCTATGAGTAACACAAAA-3'. The amplicon was ligated into pGEM®-T Easy to generate pGEM_PITR_PackΔ5. pGEM_PITR_PackΔ5 was digested with PacI and BamHI. pBS_rKS-PacI was digested with PacI and BamHI, and the PITR_PackΔ5 insert was ligated into the pBS_rKS-PacI backbone to generate pBS_PITR_PackΔ5. pBS_PITR_PackΔ5 was digested with SgrAI, blunted and dephosphorylated, and a loxP site (5'-ATAACTTCGTATAATGTATGCTATACGAAGTTAT-3') was ligated into the pBS_PITR_PackΔ5 backbone to generate pBS_PITR_LoxP-PackΔ5. pBS_PITR_LoxP-PackΔ5 was digested with SpeI, blunted and dephosphorylated, and a loxP site (5'-ATAACTTCGT ATAATGTATGCTATACGAAGTTAT-

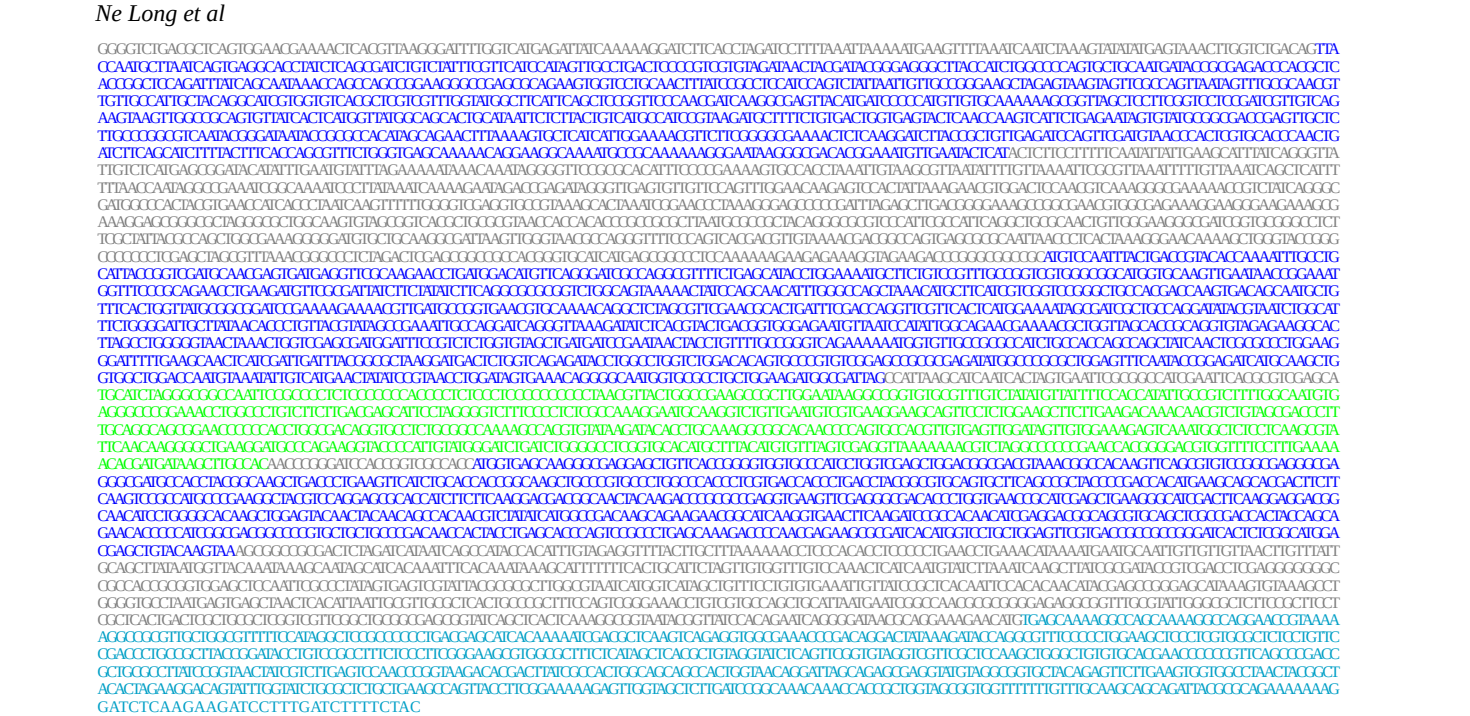


Figure 3. Putative Sequence of pBS_Cre-IRES-EGFP. The ampicilin resistance gene is in blue; the Cre and EGFP open reading frames are in blue; IRES is in green; and the origin of replication is in cyan

3') was ligated into the pBS_PITR_LoxP-PackΔ5 backbone to generate pBS_PITR_LoxP-PackΔ5-LoxP (pBS_PITR_LPΔ5L). A PI-SceI site was generated by annealing together the oligomers 5'-AATTCATCTATGTCGGGTGCGGAGAA AGAGGTAATGAAATGGCATCGACTCGAAGATCT-3' and 5'-AGATCTTCGAGTCGATGCCATTTCATTA CCTCTTTCTCCGCACCCGACATAGATG-3'. pBS_KS+ was digested with EcoRI and EcoRV, and the PI-SceI site was ligated into the pBS_KS+ backbone to

Figure 4. Putative Sequence of pBS_PITR-LΔ5L-CoE-CIE-SC. The ampicilin resistance gene is in blue; the PacI site is in green; the inverted terminal repeat is in black; LoxP sites are in magneta; the packaging site is in yellow; the CMV promoter is in black and the TATA box is in green; the CymR binding site is in red; the E2 Late Leader is in cyan; the Cre and EGFP open reading frames are in blue; IRES is in green; the PI-SceI and I-CeuI sites are in green; and the origin of replication is in cyan

Figure 5. Putative Sequence of pSH_SPC. The kanamycin resistance gene is in blue; the PacI site is in green; the PI-SceI and I-CeuI sites are in green; and the origin of replication is in cyan

generate pBS_PI-SceI. An I-CeuI site was generated by annealing together the oligomers 5'-GATCTGAATTTAA CTATAACGGTCCTAAGGTAGCGAAAGCTCAGATC-3' and 5'-GATCTGAGCTTTCGCTACCTTAGGACC GTTATAGTTAAATTCA-3'. pBS_PI-SceI was digested with KpnI, blunted, then digested with BglII, and the I-CeuI site was ligated into the pBS_PI-SceI backbone to generate pBS_SC. pBS_PITR_LΔ5L was digested with BssHII, blunted, and then digested with NotI. pBS_SC was digested with AleI and NotI, and the PITR_LΔ5L insert was ligated into the pBS_SC backbone to generate pBS_PITR_LΔ5L_SC. The sequence of pBS_PITR_ LΔ5L_SC is shown in Figure 2.

Construction of pBS_Cre-IRES-EGFP:

Previously we had constructed pBS_tTA-IRES-EGFP as follows. The NheI site containing oligomer 5'-TCGAGCTAGC-3' was annealed to itself. pBS_rKS was digested with XhoI and dephosphorylated, and the NheI containing oligomer was ligated into the pBS_rKS backbone to generate pBS_rN: When ligated into the XhoI site, the annealed NheI oligomers generate the sequence 5'-CTCGAGCTAGCTCGAG-3'. pIRES was digested with EcoRI and SmaI. pBS_rN was digested with EcoRI and SmaI, and the IRES sequence was ligated into the pBS_rN backbone to generate pBS_IRES. pEGFP-N1 was digested with SacI and AflII and blunted. pBS_KS+ was digested with EcoRV and dephosphorylated, and the EGFP sequence was ligated into the pBS_KS+ backbone to generate pBS_EGFP. pBS_EGFP was digested with ClaI, blunted, and recircularized. This construct was then digested with SmaI and ApaI and blunted. pBS_IRES was digested with SmaI and EagI, blunted and dephosphorylated, and the EGFP insert was ligated into the pBS_IRES backbone to generate pBS_IRES-EGFP. The tTA sequence of pTet-Off (pTet-Off is part of the BD™ Tet-Off Gene Expression System) was amplified using primers 5'-ATCGATCCGTCAGATCGCCTGGAGAC-3' and 5'-GATATCTTGTCCAAACTCATCAATGTATCTT ATCA-3' and ligated into pGEM®-T Easy to generate pGEM_tTA. pGEM_tTA was digested with NotI and blunted. pBS_IRES-EGFP was digested with EcoRV, and the tTA insert was ligated into the pBS_IRES-EGFP backbone to generate pBS_tTA-IRES-EGFP.

pShuttleX was previously constructed by annealing together the oligomers 5'-CTGGAATCTAGACTCGAGC GGCCGCCACGGGTAC-3' and 5'-ACACGACCTTAG ATCTGAGCTCGCCGGCGGTGCC-5' and ligating them into BstXI-KpnI digested pShuttle2 (pShuttle2 is part of the BD Adeno-X™ Expression System 1 kit).

pShuttle_Cre was previously constructed as follows. Cre recombinase was amplified from pxCANCre using primers 5'-CCCGGGTGCATCATGAGCGGCCC-3' and 5'-TCTAGAGCGCTTAATGGCTAATCGCCATCTTCC-3', and the amplicon was digested with SmaI and XbaI. pShuttleX was digested with BstXI and blunted, then digested with XbaI and dephosphorylated, and the Cre recombinase insert was ligated into the pShuttleX backbone to generate pShuttle_Cre.

pShuttle_Cre was digested with NheI and AfeI. pBS_tTA-IRES-EGFP was digested with NheI and EcoRV, and the Cre insert was ligated into the pBS_(tTA)-IRES-EGFP backbone to generate pBS_Cre recombinase-IRES-EGFP (pBS_CIE). The sequence of pBS_CIE is shown in Figure 3.

Construction of pBS_PITR_LΔ5L_CoE_CIE_SC:

The E2 late leader was cloned as follows. HEK293 cells were transfected with an Adeno-X_EGFP: the Adeno-X_EGFP was made according to the manufacturer's instructions and transfected into cells using FuGENE 6 according to the manufacturer's instructions. Total RNA was extracted using Isogen according to the manufacturer's instructions. cDNA was made using SuperScript III according to the manufacturer's instructions. The E2 Late Leader was amplified using primers 5'-GGATCCACAGCCCCGGAGTGAGTT-3' and 5'-ATCGATTTCCTTCTCCTGATATCGCCTC-3' and ligated into pGEM®-T Easy to generate pGEM_E2. The CMV-ML-CuO sequence from pCMV5CuO (pCMV5CuO is part of the Q-mate™ Inducible Expression System) was amplified using primers 5'-CCAAACTCA TCAATGTATCTTATCATG-3' and 5'-AAATTTCCTTT ATTAGCCAGAGGTC-3' and ligated into pGEM®-T Easy to generate pGEM_CMV-ML-CuO. pGEM_E2 was digested with EcoRI and blunted. pGEM_CMV5-CuO was digested with SacII, blunted and dephosphorylated, and the E2 late leader was ligated into the pGEM_CMV-(ML)-CuO backbone to generate pGEM_CoE (pGEM_[CMV promoter]-[Cym operator][with the E2 Late Leader inserted]). pGEM_CoE was digested with SpeI and blunted. pBS_CIE was digested with PmeI and dephosphorylated, and the CoE sequence was ligated into the pBS_CIE backbone to generate pBS_CoE-CIE. pBS_CoE-CIE was digested with SacI and blunted, and then digested with NheI. pBS_PITR_LΔ5L_SC was

Figure 6. Putative Sequence of pEVE

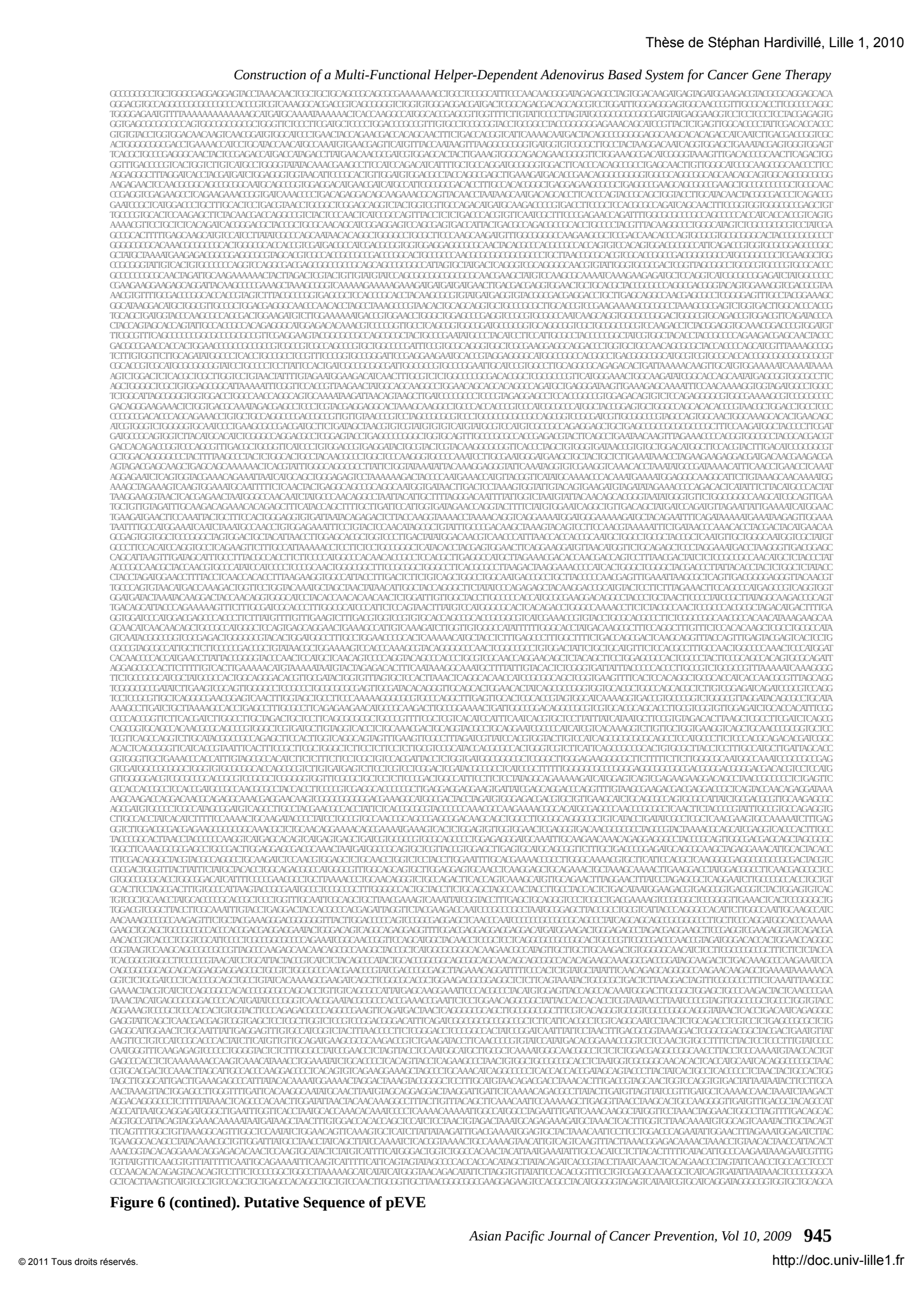


Figure 6 (contined). Putative Sequence of pEVE

Figure 6 (cont). Putative Sequence of pEVE. The ampicilin resistance gene is in blue; the PacI sites are in green; the inverted terminal repeats are in black; LoxP sites are in magneta; the packaging site is in yellow; the CMV promoter is in black and the TATA box is in green; the CymR binding site is in red; the E2 Late Leader is in cyan; the Cre and EGFP open reading frames are in blue; IRES is in green; the PI-SceI and I-CeuI sites are in green; and the origin of replication is in cyan. The sequence between the PI-SceI site and the 3' inverted terminal repeat are viral sequences from Adeno-X: Adeno-X has not been sequenced in its entirety; Adeno-X sequences were obtained from the BD Sciences web site

digested with SpeI and SmaI, and the CoE-CIE insert was ligated into the pBS_PITR_LΔ5L_SC backbone to generate pBS_PITR_LΔ5L_CoE-CIE_SC. The sequence is shown in Figure 4.

Construction of pSH_SPC:

pShuttle2 was digested with EcoRI and SapI, blunted and dephosphorylated, and an I-CeuI site (5'-TAACTATAACGGTCCTAAGGTAGCGA-3') was ligated into the pShuttle backbone. This construct was then digested with BstXI and KpnI, blunted and dephosphorylated, and a PacI site (5'-GTTACTACGTTA ATTAACGATT-3') was ligated into the backbone. This construct was then digested with DraIII and XbaI, blunted

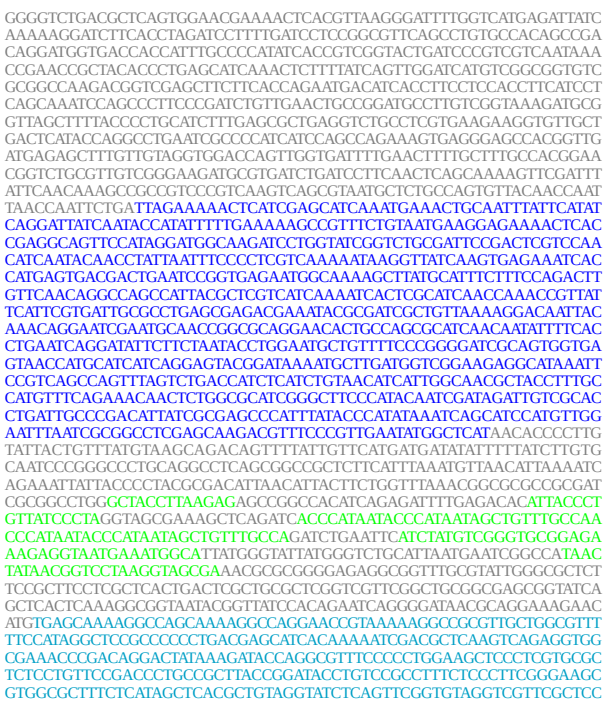


Figure 7. Putative Sequence of pSH_PSPSC. The kanamycin resistance gene is in blue; the I-PpoI, I-SceI, PI-PspI, PI-SceI and I-CeuI sites are in green (the PI-PspI site inserted in duplicate); and the origin of replication is in cyan. After digestion of pSH_CS with BsrDI, a "G" residue was lost during the blunting reaction. After digestion of pSH_PSPS with FauI, a "C" residue was lost during the blunting reaction

Figure 8. Putative Sequence of pSHCB. The kanamycin resistance gene is in blue; the PI-Scei and I-CeuI sites are in green; the sequences obtained from pSH_CST-BsrDI are in cyan; and the origin of replication is also in cyan

Figure 9. Putative Sequence of pSHPB. The kanamycin resistance gene is in blue; the PI-SceI and I-CeuI sites are in green; the sequences obtained from pSH_CST-BsrDI are in cyan; the PmeI site contained within this sequence is in black; the PmeI sites inserted into the HpaI site of pSHCB are also in black; and the origin of replication is in cyan

Figure 10. Putative Sequence of pSHcnB. The kanamycin resistance gene is in blue; the PI-SceI and I-CeuI sites are in green; the inserted oligomers are in cyan; and the origin of replication is also in cyan

and dephosphorylated, and a PI-SceI site (5'-ATCTATGT CGGGTGCGGAGAAAGAGGTAATGAAATGGCA-3') was ligated into the backbone to generate pSH_SPC. The sequence of pSH_SPC is shown in Figure 5.

Construction of pSH_AdX:

pAdeno-X_EGFP was digested with PI-SceI and PacI. pSH_SPC was digested with PI-SceI and PacI, and the AdX sequence was ligated into the pSH_SPC backbone to generate pSH_AdX.

Construction of pEVE:

pSH_AdX was digested with PI-SceI and I-CeuI. pBS_PITR_LΔ5L_CoE-CIE_SC was digested with PI-SceI and I-CeuI, and the AdX sequence was ligated into the pBS_PITR_LΔ5L_CoE-CIE_SC backbone to generate pEVE. The sequence of pEVE is shown in Figure 6.

Construction of the helper adenovirus-dependent vector (ADMM)

During construction of ADMM, newly cloned DNA was sequenced in its entirety. However, in the later stages of construction, only ligation junctions were sequenced and restriction digestion (primarily with SacI) was used assess the constructs.

Construction of pSH_PSPSC:

pShuttle2 was digested with XbaI and SpeI and recircularized to generate pSH_CS. The oligomers 5'-GG CCGGCCAGGCCGCGATCGCGGCGCGCCGTTTAA ACCAGAAGTAATGTTAATGTCGCGTAGGGGTAAT AA-3' and 5'-CTACGCGACATTAACATTACTTCTGG TTTAAACGGCGCGCCGCGATCGCGGCCTGGCCGG CC-3' were annealed together, and the oligomers 5'-TTT CTGATTTTAATGTTAACATTTAAATGAAGAGCGGC CGCTGAGGCCTGCAGGGCCCGGG-3' and 5'-CCCG GGCCCTGCAGGCCTCAGCGGCCGCTCTTCATTTAA ATGTTAACATTAAAATCAGAAATTATTACCC-3' were annealed together. pSH_CS was digested with BsrDI, blunted and dephosphorylated, and the oligomers were ligated into the pSH_CS backbone (triple ligation) to generate pSH_CST-BsrDI. pSH_CST-BsrDI was digested with BglII, blunted and dephosphorylated, and a PI-PspI site (5'-ACCCATAATACCCATAATAGCTGTTTGCCA-3') was ligated into the pSH_CST-BsrDI backbone to generate pSH_CPS. pSH_CPS was digested with FseI, blunted and dephosphorylated, and an I-PpoI site (5'-GCTACCTTAAGAGAGCCGG-3') was ligated into the pSH_CPS backbone to generate pSH_PCPS. pSH_PCPS was digested with DraIII and I-CeuI, blunted and dephosphorylated, and an I-SceI site (5'-ATTACCCTGT

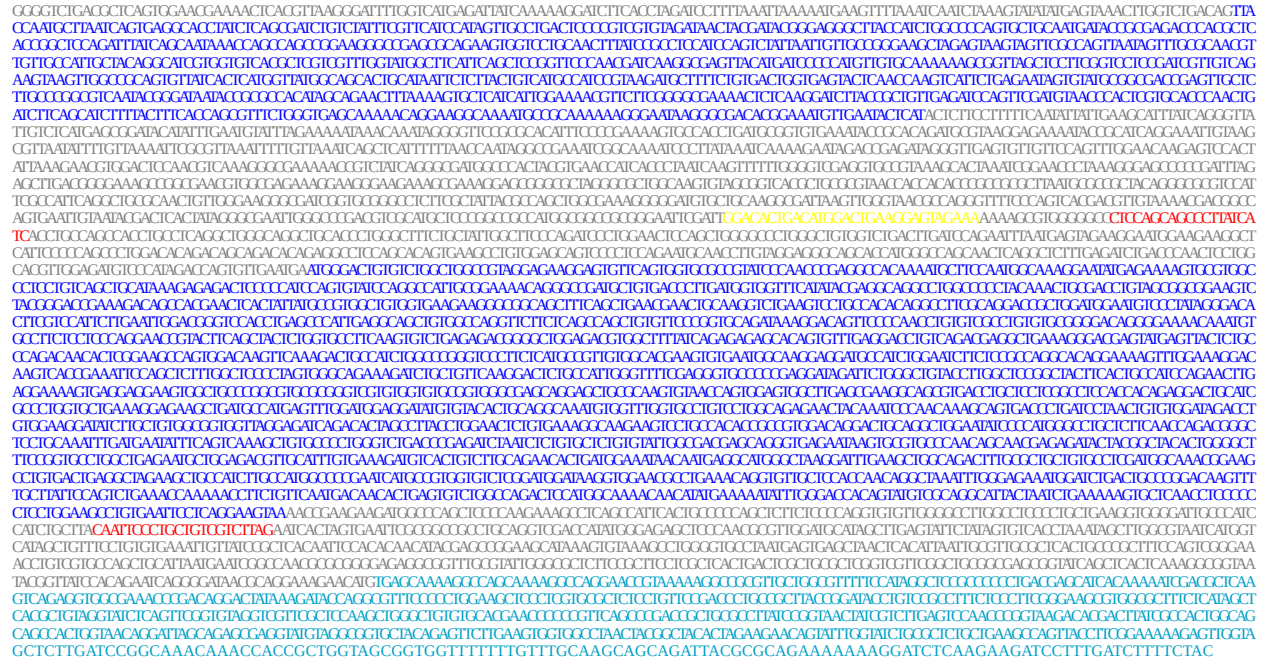


Figure 11. Putative Sequence of pGEM_gamma-lactoferrin (pGEM_gLF). The ampicillin resistance gene is in blue; the Gene Racer primer used to clone the 5' end of gLF is in yellow; the primers used to clone sub-gLF are in red; the ORF of gamma-lactoferrin is in blue; and the origin of replication is in cyan

TATCCCTA-3') was ligated into the pSH_PCPS backbone to generate pSH_PSPS. pSH_PSPS was digested with FauI, blunted and dephosphorylated, and an I-CeuI site (5'-TAACTATAACGGTCCTAAGGTAGCGA-3') was ligated into the pSH_PSPS backbone to generate pSH_PSPSC. The sequence of pSH_PSPSC is shown in Figure 7.

Construction of pSHCB:

pSH_CST-BsrDI was digested with DraIII and SmaI and blunted. pSH_CS was digested with BglII, blunted and dephosphorylated, and the insert from pSH_CST-BsrDI was ligated into the pSH_CS backbone to generate pSHCB. The sequence of pSHCB is shown in Figure 8.

Construction of pSHPB:

pSHCB was digested with HpaI and dephosphorylated, and a PmeI site (5'-AATTGGTTTAAACC-3') was inserted into the pSHCB backbone to generate pSHPB. The sequence of pSHPB is shown in Figure 9.

Construction of pSHcnB:

The oligomers 5'-GTGTCTCAAAATCTCTGATGTG GCCGGCCAGGCCGCGATCGCTGATTTGTTTAAAC TGATATTTAAATTCTGATCCACGTGTGAATTGGTT AACC-3' and 5'-CAATTCACACGTGGATCAGAATTT AAATATCAGTTTAAACAAATCAGCGATCGCGGCCT GGCCGGCCACATCAGAGATTTTGAGACAC-3' were annealed together, and the oligomers 5'-AAAAGACACG TGGATTACTATTTAAATATTAGTTTAAACTGATTAC GAAGAGCTGAGGCCTGCAGGGCCCGGGCATTGCA CAAGATAAAAATATATCATCATG-3' and 5'-CATGAT GATATATTTTTATCTTGTGCAATGCCCGGGCCCTGC AGGCCTCAGCTCTTCGTAATCAGTTTAAACTAAT ATTTAAATAGTAATCCACGTGTCTTTTGGTTAAC-3' were annealed together. pSH_CS was digested with BglII, blunted and dephosphorylated, and the annealed oligomers were ligated into the pSH_CS backbone (triple ligation) to generate pSHcnB. The sequence of pSHcnB is shown in Figure 10.

Construction of pSHCB_gSTOP, pSHPB_gSTOP, and pSHcnP_gSTOP:

We had previously cloned a putative intracellular form of lactoferrin which we named gamma-lactoferrin (gLF) (Alexander et al., 2007). Briefly, the cap-selective 5'-RACE procedure using Invitrogen's GeneRacer Kit was carried out according to the manufacturer's instructions: Total RNA was extracted from MRC5 cells using the Isogen RNA extraction procedure, dephosphorylated, treated with tobacco acid pyrophosphate, and an RNA oligonucleotide of known sequence (supplied in the Gene Racer Kit) was ligated to the 5' end of the RNA. The RNA was then reverse transcribed using a human lactoferrin specific primer (5'-ATCCTCCTTGCCATTCACAC-3'). Finally, the 5' region of the RNA was amplified using standard (Left primer, supplied by Invitrogen, 5'-CGACTGGAGCACGAGGACACTGA-3'; Right primer, anneals to lactoferrin, 5'-GCCAGATGGCAGTCTTTGA-3') and nested PCR (Nested left primer, supplied by Invitrogen, 5'-GGACACTGACATGGCTGAAGGAGT A-3'; Nested right primer, anneals to lactoferrin, 5'-TTGTCCACTGGCTTCCGAGTGTTGTC-3') and ligated into pGEM® -T Easy (pGEM_5'gLF). The 5' region was sequenced using an ABI Prism 3100 Genetic Analyzer (Applied Biosystems) according to the manufacturer's instructions. Nearly full length gLF was amplified using 5'-CTCCAGCAGCCCTTATCATC-3' (the left primer binds to the 5' UTR of gamma-lactoferrin) and 5'-CTAAG ACGACAGCAGGGAATTG-3' (the right primer binds to the 3' UTR of gamma-lactoferrin) and ligated into pGEM® -T Easy (pGEM_sub-gLF) and sequenced. Full

Figure 12. Putative Sequence of pSHCB_gSTOP. The kanamycin resistance gene is in blue; the PI-Scei and I-CeuI sites are in green; the Gene Racer primer used to clone the 5' end of gLF is in yellow; the gSTOP sequence is in red (the first 9 bases of the gLF ORF are in blue); and the origin of replication is in cyan

Figure 13. Putative Sequence of pSHPB_gSTOP. The kanamycin resistance gene is in blue; the PI-Scei and I-CeuI sites are in green; the Gene Racer primer used to clone the 5' end of gLF is in yellow; the gSTOP sequence is in red (the first 9 bases of the gLF ORF are in blue); the PmeI sites inserted into the HpaI site are in black; and the origin of replication is in cyan

length gLF was generated by digesting pGEM_5'gLF with AatII and TthIII and ligating the 5'gLF sequence into AatII and TthIII digested pGEM_sub-gLF. The sequence of pGEM_gLF is shown in Figure 11.

pGEM_gLF was digested with SpeI and SalI, blunted and recircularized. The recircularized plasmid was digested with NotI, blunted and dephosphorylated, and a PmeI site (5'-AATTGGTTTAAACC-3') was ligated into the backbone to generate pGEM_PLF. pGEM_PLF was digested with PmeI and TthIII and blunted. pSHCB was digested with PmeI and dephosphorylated, and the STOP sequence from gLF was ligated into the pSHCB backbone to generate pSHCB_gSTOP. The sequence of pSHCB_gSTOP is shown in Figure 12.

pSHCB_gSTOP was digested with HpaI and dephosphorylated, and a PmeI site was ligated into the pSHCB_gSTOP backbone to generate pSHPB_gSTOP. The sequence of pSHPB_gSTOP is shown in Figure 13.

pSHPB_gSTOP was digested with AscI and SwaI and blunted. pSHcnB was digested with PmlI and dephosphorylated, and the gSTOP sequence ligated into the pSHcnB backbone to generate pSHcnP_gSTOP. The sequence of pSHcnP_gSTOP is shown in Figure 14.

Construction of pSHcnB_Ins50:

Stuffer DNA 50 was amplified from chicken genomic DNA using 5'-AGAGAAAGGGAAGGAACTTGTCAT-3' and 5'-CATTAGCGGCAGCACCG-3'. pSHcnB was digested with SwaI and dephosphorylated, and the stuffer DNA 50 amplicon was ligated into the pSHcnB backbone to generate pSHcnB_Ins50. The sequence of pSHcnB_Ins50 is shown in Figure 15.

Construction of pSHcnB_Ins50G:

Stuffer DNA G-3016 (InsG-3016) was amplified from rat genomic DNA using 5'-GGGCACTCCATACCAACCA-3' and 5'-GGACCCAGCAGCCTTTCAT-3'. pSHcnB_Ins50 was digested with SacII, blunted and dephosphorylated, and the InsG-3016 amplicon was ligated into the pSHcnB_Ins50 backbone to generate pSHcnB_Ins50G. The sequence of pSHcnB_Ins50G is shown in Figure 16.

Construction of pRcnSS:

pBS_KS+ was digested with BamHI, blunted and

Figure 14. Putative Sequence of pSHcnP_gSTOP. The kanamycin resistance gene is in blue; the PI-SceI and I-CeuI sites are in green; the PmeI sites inserted into the HpaI site are in black; the gSTOP sequence is in red (the first 9 bases of the gLF ORF are in blue); the Gene Racer primer used to clone the 5' end of gLF is in yellow; and the origin of replication is in cyan

dephosphorylated, and an I-CeuI site (5'-TAACTATAAC GGTCCTAAGGTAGCGA-3') was ligated into the pBS_KS+ backbone to generate pBSC. mSTOP was amplified from mouse genomic DNA using 5'-TTTTTG AAATGGCTCGTTGC-3' and 5'-TGGTAACCGTATC GGTACAAGAT-3'. pBSC was digested with SmaI and dephosphorylated, and the mSTOP amplicon was ligated into the pBSC backbone to generate pBSCsma_mSTOP. pBSCsma_mSTOP was digested with PstI and NotI and blunted. pSHcnB was digested with I-CeuI, blunted and dephosphorylated, and the I-CeuI-mSTOP insert was ligated into the pSHcnB backbone to generate pRcnSx. pGEM_closed was digested with NcoI, blunted and dephosphorylated, and a PI-SceI site (5'-ATCTATGTCGGGTGCGGAGAAAGAGGTAATGAA ATGGCA-3') was ligated into the pGEM_closed backbone to generate pGPS. dSTOP was amplified from human genomic DNA using 5'-CCTTAGGTAGATGCC GTAAGACAA-3' and 5'-AGCAAGACCGAGCCTTTA-3'. pGPS was digested with NotI, blunted and dephosphorylated, and the dSTOP amplicon was ligated

Figure 15. Putative Sequence of pSHCB_Ins50. The kanamycin resistance gene is in blue; the PI-Scei and I-CeuI sites are in green; Ins50 is in magneta; and the origin of replication is in cyan

into the pGPS backbone to generate pGPS_dSTOP. pGPS_dSTOP was digested with ApaI and BstXI and blunted. pRcnSx was digested with PI-SceI, blunted and dephosphorylated, and the PI-SceI-dSTOP insert was ligated into the pRcnSx backbone to generate pRcnSS. The sequence of pRcnSS is shown in Figure 17.

Construction of pRcnSS_Ins57-111:

Stuffer DNA 57-111 was amplified from chicken genomic DNA using primers 5'-TGCCAATCCATCAAG TTCCAT-3' and 5'-TGTCCAGCCTCATATCCAGTCA-3'. RcnSS was digested with PmlI and dephos-phorylated, and the stuffer DNA 57-111 amplicon was ligated into the pRcnSS backbone. The sequence of pRcnSS_Ins57-111 is shown in Figure 18.

Construction of pSH_PPS50G-111s_PSPSC_PITR:

The 3' PacI-ITR sequence (PITR) of Adeno-X was amplified from EVE using primers 5'-TTACTCCGCCCT

Figure 16. Putative Sequence of pSHcnB_Ins50G. The kanamycin resistance gene is in blue; the PI-Scei and I-CeuI sites are in green; Ins50 is in magneta; InsG-3016 is in yellow; and the origin of replication is in cyan

AAAACCTAC-3' and 5'-GCTATTGCTTTATTTGTAAC CATT-3'. pSHcnB was digested with SwaI and dephosphorylated, and the PITR amplicon was ligated into the pSHcnB backbone to generate pSHcnB_PITR-3'. pSHcnB_PITR-3' was digested with PmeI. pSH_PSPSC was digested with PciI, blunted and dephosphorylated, and the PITR insert was ligated into the pSH_PSPSC backbone to generate pSH_PSPSC_PITR. pBS_PITR_ Pack-wt was digested with BssHII and SmaI and blunted. pSH_PSPSC_PITR was digested with SwaI and dephosphorylated, and the PITR insert was ligated into the pSH_PSPSC_PITR backbone to generate pSH_PP_ PSPSC_PITR. pSHcnP_gSTOP was digested with PmeI. pSH_PP_PSPSC_PITR was digested with BamHI, blunted and dephosphorylated, and the gSTOP insert was ligated into the pSH_PP_PSPSC_PITR backbone to generate pSH_PPS_PSPSC_PITR. pSHcnB_Ins50G was digested with PmeI. pSH_PPS_PSPSC_PITR was digested with HpaI and dephosphorylated, and the Ins50G insert was ligated into the pSH_PPS_PSPSC_PITR backbone to generate pSH_PPS50G_PSPSC_PITR. pRcnSS_Ins57-111 was digested with PI-SceI, blunted and dephosphorylated, and an I-PpoI site (5'-GCTACCTTA AGAGAGCCGG-3') was ligated into the pRcnSS_Ins57-

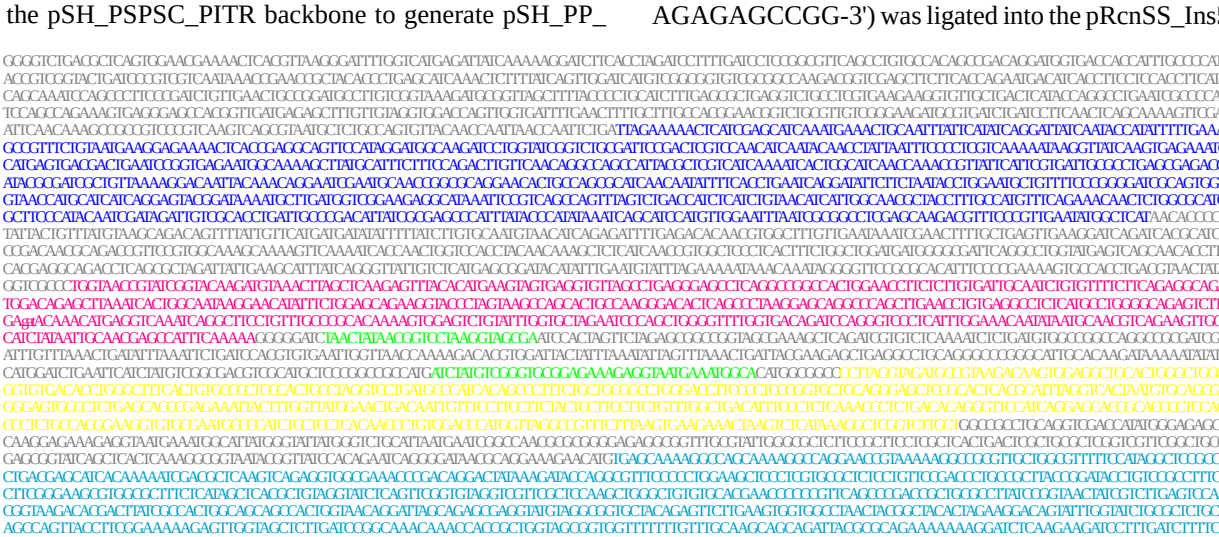


Figure 17. Putative Sequence of pRcnSS. The kanamycin resistance gene is in blue; mSTOP is in magneta; the PI-SceI and I-CeuI sites are in green; dSTOP is in yellow; and the origin of replication is in cyan. After digestion of pBS_KS+ with BamHI, a "G" residue was lost during the blunting reaction

Figure 18. Putative Sequence of pRcnSS_Ins57-111. The kanamycin resistance gene is in blue; the PI-Scei and I-CeuI sites are in green; mSTOP is in magenta; Ins57-111 is in cyan; dSTOP is in yellow; and the origin of replication is in cyan

111 backbone to generate pRcnP_Ins57-111. pRcnP_Ins57-111 was digested with I-CeuI, blunted, and then digested with I-PpoI. pSH_PPS50G_PSPSC_PITR was digested with PmeI and I-PpoI, and the Ins57-111 insert was ligated into the pSH_PPS50G_PSPSC_PITR backbone to generate pSH_PPS50G-111_PSPSC_PITR. pSH_PPS50G-111_PSPSC_PITR was digested with SpeI and recircularized to generate pSH_PPS50G-111s_PSPSC_PITR. The sequence of pSH_PPS50G-111s_PSPSC_PITR is shown in Figure 19.

Construction of pØØ_PSPSC:

pSH_PSPSC was digested with SmaI and PciI. pBS_KS+ was digested with ApoI and blunted, then digested with PciI, and the PSPSC insert was ligated into the pBS backbone to generate pØØ_PSPSC. The sequence of pØØ_PSPSC is shown in Figure 20.

Construction of the C-I cassette:

The LacZ ORF was amplified from pShuttle_LacZ (pShuttle_LacZ is part of the BD Adeno-X™ Expression System 1 kit) using primers 5'-CGAGGGGGGATCGAAAGAG-3' and 5'-ATGTAGCCAAATCGGGAAAAACG-3'. pSHPB_gSTOP was digested with SwaI and dephosphorylated, and the LacZ amplicon was ligated into the pSHPB_gSTOP backbone to generate pSHPB_gSTOP_LacZ. The poly(A) addition signal of rat glutathione S-transferase pi subclass (GST-P) was amplified from rat genomic DNA using primers 5'-CAGACTAATAAAGTTTGTAAGGCA-3' and 5'-GATCACAGCATTTGGGAGA-3'. pBS was digested with EcoRV and dephosphorylated, and the GSTP poly(A) addition signal was ligated into the pBS backbone to generate pBS_(rGSTP)pA. pBS_(rGSTP)pA was digested with SalI and SacII and blunted. pSHCB_gSTOP_LacZ was digested with NotI, blunted and dephosphorylated, and the (rGSTP)pA insert was ligated into the pSHPB_gSTOP_LacZ backbone to generate pSHPB_gSTOP_LacZ-pA. pTRE-Tight was digested with XhoI and EcoRI and blunted. pSHPB_gSTOP_LacZ-pA was digested with AscI, blunted and dephosphorylated, and the TRE-Tight promoter was ligated into the pSHPB_gSTOP_LacZ-pA backbone to generate pSHPB_Tight_gSTOP_LacZ-pA. pSHPB_Tight_gSTOP_LacZ-pA was digested with I-CeuI, blunted and dephosphorylated, and an I-PpoI site (5'-GCTACCTTAAGAGAGCCGG-3') was ligated into the pSHPB_Tight_gSTOP_LacZ-pA backbone to generate pSHP_Tight_gSTOP_LacZ-pA. pSHP_Tight_gSTOP_LacZ-pA was digested with PI-SceI, blunted and dephosphorylated, and an I-SceI site (5'-TAGGGATTACCCTGTTATCCCTACAGGGTAAT-3') was ligated into the pSHP_Tight_gSTOP_LacZ-pA backbone to generate

Figure 19. Putative Sequence of pSH_PPS50G-111s_PSPSC_PITR. The kanamycin resistance gene is in blue; the PacI sites are in green; the inverted terminal repeats are in black, the packaging site is in yellow; the Gene Racer primer used to clone the 5' end of gLF is also in yellow; the gSTOP sequence is in red (the first 9 bases of the gLF ORF are in blue); Ins50G is in magenta; Ins57-111s is in cyan; the I-PpoI, I-SceI, PI-PspI, PI-Scei and I-CeuI sites are in green; and the origin of replication is in cyan

pSHPS_Tight_gSTOP_LacZ-pA. Stuffer DNA 52 was amplified from chicken genomic DNA using primers 5'-TCATAGGGTTGTTGGGTTGTTT-3' and 5'-ATCCTCT TCTCACTGCTGCTACT-3'. pBS was digested with EcoRV and dephosphorylated, and the stuffer DNA 52 amplicon was ligated into the pBS backbone to generate pBS_Ins52. pBS_Ins52 was digested with EcoRI and HindIII and blunted. pSHPS_Tight_gSTOP_LacZ-pA was digested with PstI, blunted and dephosphorylated, and the Ins52 insert was ligated into the pSHPS_Tight_gSTOP_ LacZ-pA backbone to generate pSHPS_Tight_gSTOP_ LacZ-pA_Ins52. pSHPS_Tight_gSTOP_LacZ-pA_Ins52 was digested with PmeI (to remove the gSTOP sequence) and recircularized to generate pC-I_Tight-LacZ. The sequence of pC-I_Tight-LacZ is shown in Figure 21.

Construction of the C-II cassette:

The HSVTK ORF was amplified from pCM-TK using primers 5'-TCTAGATTGGTGGCGTGAAACTCCC-3' and 5'-GTCGACGGTTCCTTCCGGTATTGTCTCC-3' and ligated into pGEM®-T Easy to generate pGEM_HSVTK. pGEM_HSVTK was digested with

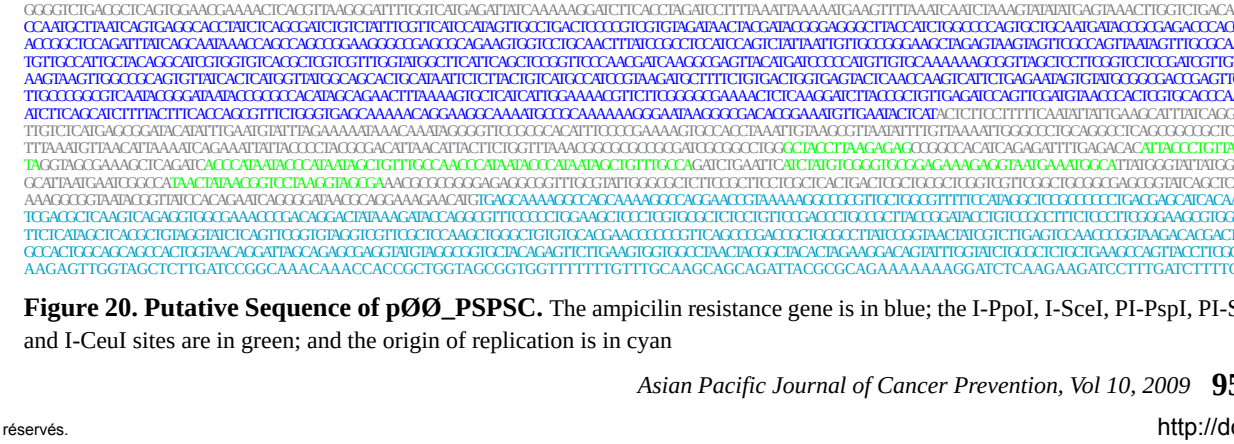


Figure 20. Putative Sequence of pØØ_PSPSC. The ampicilin resistance gene is in blue; the I-PpoI, I-SceI, PI-PspI, PI-Scei and I-CeuI sites are in green; and the origin of replication is in cyan

Figure 21. Putative Sequence of pC-I_Tight-LacZ. The kanamycin resistance gene is in blue; the I-PpoI site is in green; the TRE-Tight promoter is in black, the TATA box is in green; the LacZ ORF is in blue; the sequence containing the (rGSTP)pA is in red; Ins52 is in yellow; the I-SceI site is in gren; and the origin of replication is in cyan

EcoRI and blunted. pSHCB was digested with SwaI and dephosphorylated, and the HSVTK was ligated into the pSHCB backbone to generate pSHCB_HSVTK. The poly(A) addition signal of rat guanylate kinase 1 was amplified from rat genomic DNA using primers 5'-ACATCTATTCTCCCTGGGCTATTT-3' and 5'-TCC TTCCTGGCAAGTGGTGT-3'. pBS_KS+ was digested with EcoRV and dephosphorylated, and the guanylate kinase 1 poly(A) addition signal was ligated into the pBS backbone to generate pBS_(rGUK)pA. pBS_(rGUK)pA was digested with EcoRI and HindIII and blunted. pSHCB_HSVTK was digested with NotI, blunted and dephosphorylated, and the (rGUK)pA insert was ligated into the pSHCB_HSVTK backbone to generate pSHCB_HSPTK-pA. pTRE-Tight was digested with XhoI and EcoRI and blunted. pSHCB_stopHSVTK-pA was digested with AscI, blunted and dephosphorylated, and the TRE-Tight promoter was ligated into the pSHCB_HSVTK-pA backbone to generate pSHCB_ Tight-HSVTK-pA. Stuffer DNA 53 was amplified from chicken genomic DNA using primers 5'-AACCGTCCA CCTACCACCAAC-3' and 5'-CTGTCAGCCTCATTTC ACAAGATT-3'. pBS_KS+ was digested with EcoRV and dephosphorylated, and the stuffer DNA 53 amplicon was ligated into the pBS backbone to generate pBS_Ins53. pBS_Ins53 was digested with SmaI and SalI and blunted. pSHCB_Tight-HSVTK-pA was digested with FseI, blunted and dephosphorylated, and the Ins53 insert was ligated into the pSHCB_Tight-HSVTK-pA backbone to generate pSHCB_Ins53_Tight-HSVTK-pA. pSHCB_ Ins53_Tight-HSVTK-pA was digested with I-CeuI, blunted and dephosphorylated, and a PI-PspI site (5'-ACCCATAATACCCATAATAGCTGTTTGCCA-3') was ligated into the pSHCB_Ins53_Tight-HSVTK-pA backbone to generate pSHP_Ins53_Tight-HSVTK-pA. pSHP_Ins53_Tight-HSVTK-pA was digested with PI-SceI, blunted and dephosphorylated, and an I-SceI site (5'-ATTACCCTGTTATCCCTA-3') was ligated into the pSHP_Ins53_Tight-HSVTK-pA backbone to generate pSHPS_Ins53_Tight-HSVTK-pA. pSHPS_Ins53_Tight-HSVTK-pA was digested with HpaI and PmeI and recircularized to generate pC-II_Tight-HSVTK. The sequence of pC-II_Tight-HSVTK is shown in Figure 22.

Construction of the C-III cassette:

Stuffer DNA 56 was amplified from Chicken DNA using primers 5'-AGGCTGAATAAGACAGTAGTGG TGA-3' and 5'-GTTTCGGAAGTGCCTGCTAAG-3'. pSHcnB was digested with PmlI and dephosphorylated, and the stuffer DNA 56 amplicon was ligated into the

Figure 22. Putative Sequence of pC-II_Tight-HSVTK. The kanamycin resistance gene is in blue; the PI-PspI site is in green; Ins53 is in yellow; the TRE-Tight promoter is in black, the TATA box is in green; the HSVTK ORF is blue; the sequence containing the (rGUK)pA is in red; the I-SceI site is in green; and the origin of replication is in cyan

pShcnB backbone to generate pSHcnB_Ins56. pSHcnB_Ins56 was digested with I-CeuI, blunted and dephosphorylated, and a PI-PspI site (5'-ACCCATAATAC CCATAATAGCTGTTTGCCA-3') was ligated into the pSHcnB_Ins56 backbone to generate pSHP_Ins56. pSHP_56 was digested with PacI, blunted, and recircularized to generate pC-III_Ins56. The sequence of pC-III_Ins56 is shown in Figure 23.

Construction of the C-IV cassette:

The tTA ORF was amplified from pTet-Off using primers 5'-ATCGATCCGTCAGATCGCCTGGAGAC-3' and 5'-GATATCTTGTCCAAACTCATCAATGTATCT TATCA-3' and the amplicon was ligated into pGEM®-T Easy to generate pGEM_tTA. pGEM_tTA was digested with SpeI and EcoRV and blunted. pSHCB was digested with SwaI and dephosphorylated, and the tTA insert was ligated into the pSHCB backbone to generate pSHCB_tTA. The poly(A) addition signal of rat glucokinase was amplified from rat genomic DNA with primers 5'-GCAGGAATCATCTCCAACACTC-3' and 5'-TGTTCACCCGAAGGCATATTA-3'. pBS_KS+ was digested with EcoRV and dephosphorylated, and the glucokinase poly(A) addition signal was ligated into the pBS backbone to generate pBS_(rGck)pA. pBS_(rGck)pA was digested with EcoRI and HindIII and blunted. pSHCB_tTA was digested with NotI, blunted and dephosphorylated, and the (rGck)pA insert was ligated into the pSHCB_tTA backbone to generate pSHCB_tTA-pA. pTRE2hyg (pTRE2hyg is part of the BD™ Tet-Off Gene Expression System) was digested with XhoI and EcoRI. pBS_rN was digested with XhoI and EcoRI, and the TRE insert was ligated into the pBS_rN backbone to generate pBS_rN_TRE. pBS_rN_TRE was digested with XhoI and EcoRi and blunted. pSHCB_tTA-pA was digested with AscI, blunted and dephosphorylated, and the TRE insert was ligated into the pSHCB_tTA-pA backbone to generate

Figure 23. Putative Sequence of pC-III_Ins56. The kanamycin resistance gene is in blue; the PI-PspI site is in green; Ins56 is in yellow; the PI-SceI site is in green; and the origin of replication is in cyan

pSHCB_TRE-tTA-pA. pSHCB_ TRE-tTA-pA was digested with HpaI and PmeI and dephosphorylated, and the sequence 5'-GAAGTTCCTATACTTTCTAGA GAATAGGAACTTC-3' was ligated into the pSHCB_TRE-tTA-pA backbone to generate pSH_TRE-tTA-pA. pSH_TRE-tTA-pA was digested with I-CeuI, blunted and dephosphorylated, and an I-CeuI site was ligated into the pSH_TRE-tTA-pA backbone (I-CeuI sites in the reverse orientation were selected) to generate pSHC_TRE-tTA-pA. pSHC_TRE-tTA-pA was digested with PI-SceI, blunted and dephosphorylated, and a PI-SceI site was ligated into the pSHC_TRE-tTA-pA backbone (PI-SceI sites in the reverse orientation were selected) to generate pSHCS_TRE-tTA-pA. Part A of stuffer DNA 54 was amplified from chicken genomic DNA using primers 5'-GGCGGCTGCCTCTTCATA-3' and 5'-TCCACATTGAGAAGGGACCATAC-3'. pBS_KS+ was digested with EcoRV and dephosphorylated, and part A of stuffer DNA 54 was ligated into the pBS backbone to generate pBS_Ins54a. Part B of stuffer DNA 54 was amplified from chicken genomic DNA using primers 5'-AGGCTGTCTGTTGGCAGTCTTT-3' and 5'-CGTCTG TTTCCCAGGTGCTC-3'. pBS_KS+ was digested with EcoRV and dephosphorylated, and part B of stuffer DNA

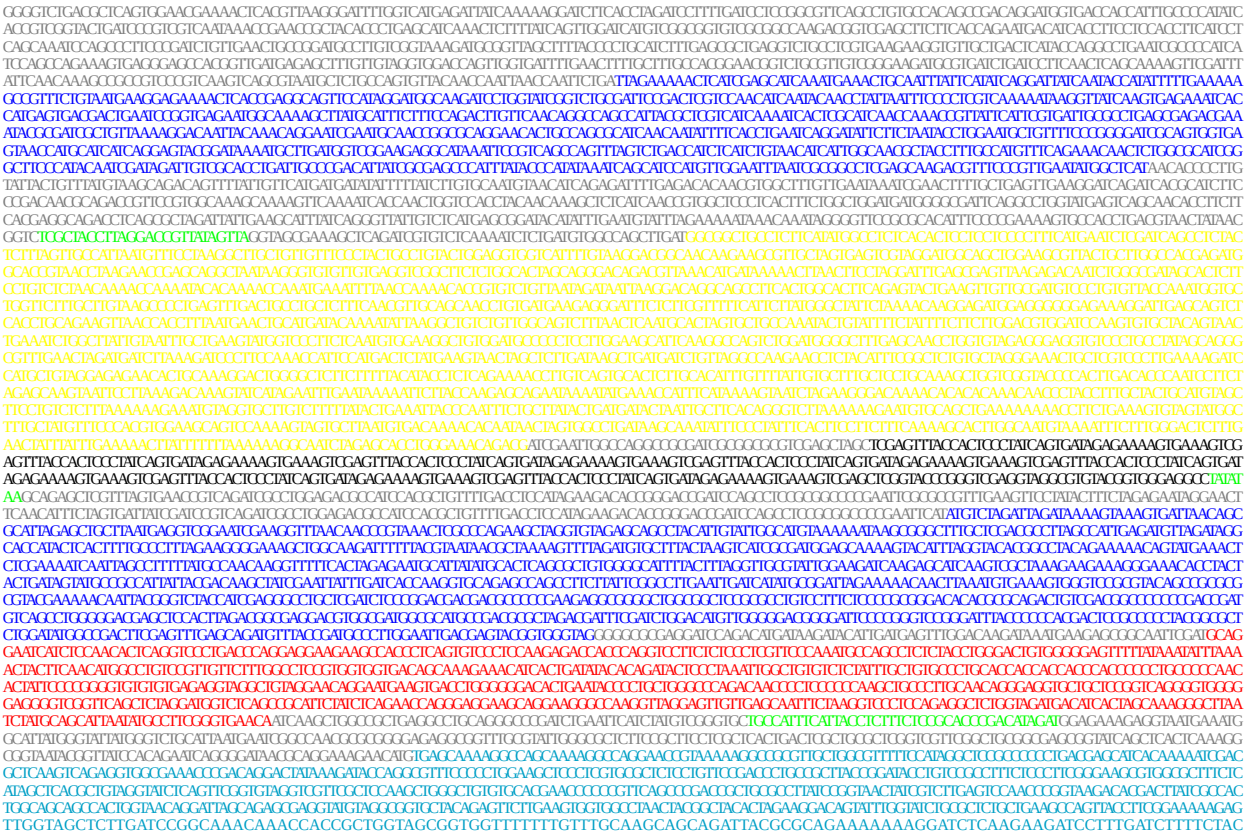


Figure 24. Putative Sequence of pC-IV_TRE-tTA. The kanamycin resistance gene is in blue; the PI-SceI site is in green; Ins54 is in yellow; the TRE promoter is in black, the TATA box is in green; the tTA ORF is in blue; the sequence containing the (rGck)pA is in red; the I-CeuI site is in green; and the origin of replication is in cyan. After digestion of pSHCB_tTA with NotI, a "C" residue was lost during the blunting reaction

54 was ligated into the pBS backbone to generate pBS_Ins54b. pBS_Ins54b was digested with BamHI. pBS_Ins54a was digested with BamHI and dephosphorylated, and the Ins54b insert was ligated into the pBS_Ins54a backbone to generate pBS_Ins54. pBS_Ins54 was digested with HindIII and EcoRI and blunted. pSHCS_TRE-tTA-pA was digested with NaeI and dephosphorylated, and the Ins54 insert was ligated into the pSHCS_TRE-tTA-pA backbone to generate pSHCS_Ins54_TRE-tTA-pA. pSHCS_Ins54_TRE-tTA-pA was digested with PacI, blunted, and recircularized to generate pC-IV_TRE-tTA. The sequence of pC-IV_TRE-tTA is shown in Figure 24.

Construction of pC4-H_Tight-LacZ:

pSH_PSPSC was digested with SmaI and PciI. pBS_KS+ was digested with ApoI and blunted, and the PSPSC insert was ligated into the pBS_KS+ backbone to generate pØØ_PSPSC. pC-IV_TRE-tTA was digested with PI-SceI and I-CeuI. pØØ_PSPSC was digested with PI-SceI and I-CeuI, and the C-IV_TRE-tTA insert was ligated into the pØØ_PSPSC backbone to generate pØØ_PSPS_tTA. pC-II_Tight-HSVTK was digested with I-SceI and PI-PspI. pØØ_PSPS_tTA was digested with I-SceI and PI-PspI, and the C-II_Tight-HSVTK insert was ligated into the pØØ_PSPS_tTA backbone to generate pØØ_PS_HSV_PS_tTA. pC-III_Ins56 was digested with PI-PspI and PI-SceI. pØØ_PS_HSV_56_tTA was digested with PI-PspI and PI-SceI, and the C-III_Ins56 insert was ligated into the pØØ_PS_HSV_56_tTA backbone to generate pØØ_PS_HSV_56_tTA. pC-I_Tight-LacZ was digested with I-PpoI and I-SceI. pØØ_PS_HSV_56_tTA was digested with I-PpoI and I-SceI, and the C-I_Tight-LacZ insert was ligated into the pØØ_PS_HSV_56_tTA backbone to generate pC4-H_Tight-LacZ (pØØ_LacZ_HSV_56_tTA).

Construction of pADMM:

pC4-H_Tight-LacZ was digested with I-PpoI and I-CeuI. pSH_PPS50G-111s_PSPSC_PITR was digested with I-PpoI and I-CeuI, and the LacZ_HSV_56_tTA insert was ligated into the pSH_PPS50G-111s_PSPSC_PITR backbone to generate pADMM-H_Tight-LacZ.

Computer analysis

Trace files generated by the ABI Prism 3100 Genetic Analyzer were viewed using 4Peaks software (A. Griekspoor and Tom Groothuis, mekentosj.com). DNA sequences were aligned using ClustalW (Larkin et al., 2007) (http://www.ebi.ac.uk/clustalw). Primers were designed using Primer Premier (Premier Biosoft International: www.premierbiosoft.com/index.html).

Results and Discussion

Vector Construction

During construction of pBS_PITR_Pack-wt: After pBS_rKS was digested with KpnI, a "G" residue was lost during the blunting reaction; this loss had no effect on the function of the construct. During construction of

pBS_PITR_LΔ5L_SC: After digestion of pBS_PITR_LoxP-PackΔ5 with SpeI, a "G" residue, a "C" residue, and a "T" residue were lost during the blunting reaction; these losses had no effect on the function of the construct. During construction of pSH_PSPSC: After pSH_CS was digested with BsrDI, a "G" residue was lost during the blunting reaction; this loss had no effect on the function of the construct. During construction of pSH_PSPSC: After digestion of pSH_PSPS with FauI, a "C" residue was lost during the blunting reaction; this loss had no effect on the function of the construct. During construction of pRcnSS: After digestion of pBS_KS+ with BamHI, a "G" residue was lost during blunting reaction; this loss had no effect on the function of the construct. During construction of the C-IV cassette (pC-IV_TRE-tTA): After digestion of pSHCB_tTA with NotI, a "C" residue was lost during the blunting reaction; this loss had no effect on the function of the construct.

During construction of ADMM, the most stable constructs often had the PI-PspI site inserted in duplicate; the reason for this is unknown. During digestion of 5'-CTCGAGCTAGCTCGAG-3' containing sequences (pBS_rN and pBS_rN_TRE) with XhoI, only one of the XhoI sites was digested; the reason for this is unknown. During construction of the C-IV cassette (pC-IV_TRE-tTA), insertion of GAAGTTCCTATACTTTCTAGAGAA TAGGAACTTC is not necessary and can be omitted.

EVE

EVE is based on Adeno-X which is a replication incompetent, ΔE1/ΔE3, human adenoviral type 5 genome (Mizuguchi and Kay, 1998). Successful generation of helper adenovirus-dependent vectors (ADMM) requires removal of the helper virus (EVE) from the ADMM preparation. To achieve this goal, we modified the packaging site of EVE, inserted loxP sites on the 5' and 3' sides of the packaging site, and inserted a regulatable expression cassette based on the Q-mateTM Inducible Expression System which expresses Cre recombinase and enhanced green fluorescent protein (EGFP).

Due to patients' strong immune response against adenovirus and its gene products and because helper adenoviruses do express adenovirus gene products, contamination of helper-dependent virus preparations by the helper virus is one of the limiting factors in the use of these preparations in gene therapy. The most common method of removing the helper virus is by flanking the packaging site of the helper virus with LoxP sites and propagating the helper-dependent virus in Cre recombinase expressing packaging cells (such as HEK293 cells transfected with a Cre recombinase expression cassette) infected with helper virus: The Cre recombinase excises the floxed packaging site from the helper virus and removal of the packaging site severely limits encapsidation of the helper virus into infectious particles, but the helper virus retains the ability to express all the viral proteins required for propagation of the helper-dependent virus and its encapsidation into infectious particles (since the packaging site of the helper-dependent virus is not flanked by LoxP sites, its packing site is not removed by Cre recombinase). The expression of high levels of Cre recombinase, however, is detrimental to most cells, resulting in chromosomal aberrations (Loonstra et al., 2001). Therefore, we constructed EVE to express Cre recombinase. This allows high levels of Cre recombinase to be expressed during propagation of ADMM, resulting in efficient removal of EVE from ADMM preparations, but the packaging cell is not damaged prior to propagation of ADMM.

During the later stages of infection adenoviruses express proteins which inhibit Cap-dependent translation (Cuesta et al., 2000; Cuesta et al., 2004), but translation of adenoviral gene products continues via ribosomal binding to sequences in adenoviral mRNA known as leader sequences (Xi et al., 2004). Therefore, we constructed EVE such that the mRNA for Cre recombinase contained the adenoviral E2 late leader sequence.

Since EVE expresses Cre recombinase and also has its packaging site flanked by LoxP sites, Cre recombinase expression by EVE must be off when EVE is being propagated. We chose the Q-mateTM Inducible Expression System (Krackeler Scientific, Inc.) to regulate Cre recombinase expression by EVE. Therefore, we constructed EVE with a CymR binding site between the CMV promoter and the Cre recombinase sequence. When EVE is propagated in CymR expressing cells, CymR binds to the CymR binding site contained in EVE and blocks CMV driven transcription of Cre recombinase, keeping the packaging site of EVE intact and allowing EVE to become encapsidated into infectious particles.

EVE was constructed such that EVE contamination of ADMM preparations could be monitored using EGFP. This was done by inserting an internal ribosomal entry site (IRES) followed by an ORF coding for EGPF downstream of the Cre recombinase ORF. Infection of cells which do not express CymR by EVE results in expression of EGFP.

Recombination between EVE and ADMM can remove LoxP sites from EVE resulting in encapsidation of EVE. To reduce recombination between EVE and ADMM, we attempted to create a packaging site in EVE that contained less DNA than the packaging site in ADMM. Initially we generated a wild type packaging site and 3 different variant packaging sites with portions of the packaging sequence deleted. The variants were generated by amplification of the wild type packaging site sequence with the following primer pairs: Pack-Δ6, 5'-GTCCCATTCGCCATTCAG G-3' and 5'-GGATCCGCGGCCCTAGACAAATA-3'; Pack-Δ5, 5'-GTCGAGGTGCCGTAAAGCA-3' and 5'-GGATCCACGCGCTATGAGTAACACAAAA-3'; and Pack-Δ3, 5'-GTCCCATTCGCCATTCAGG-3' and 5'-GGATCCGGCCAAATCTTACTCGGTTAC-3' (constructs not shown). The EVEs containing Pack-wt, Pack-Δ6, and PackΔ5 were able to propagate in the HEK293_CymR packaging cell line, but the EVE containing Pack-Δ3 did not propagate in the HEK293_CymR packaging cell line (data not shown). Therefore, the EVE described in this manuscript contains PackΔ5.

When propagating plasmids containing both a floxed packaging site and the Cre Recombinase coding region in *E. coli*, it is imperative that the *E. coli* be grown at

28˚C or less. Also, the cultures should be harvested while they are still in the log phase of growth. Finally, after harvesting, the plasmids must be screened to confirme that the packaging site has not been lost.

ADMM

We have found that plasmids containing certain sequences are extremely difficult to propagate in E. coli. One possible explanation is that these sequences code for toxic peptides. We, therefore, inserted sequences which we call STOP sequences into four plasmids (generating pSHCB_STOP, pSHPB_STOP, pSHcnP_STOP, and pRcnSS: all based on pShuttle2) and used these plasmids as recipients of hard-to-propagate DNA sequences. The STOP sequences were obtained from human and mouse DNA; they contain a number of ATG codons followed by STOP codons, making translation of downstream ORFs a rare event.

A STOP sequence is incorporated into the ADMM described here (ADMM-H_Tight-LacZ) just downstream of the packaging site. This is because the adenovirus E1 promoter overlaps the packaging site sequence (Kovesdi et al., 1987), indicating that the packaging site itself can have gene promoter activity.

Enhancer-blocking insulators block interaction between a promoter and distal enhancer elements, and CCCTC-binding factor (CTCF) is the major protein implicated in establishment of insulators in vertebrates (Gaszner and Felsenfeld, 2006; Wallace and Felsenfeld, 2007). We scanned the chicken genome and the rat GSTP locus for putative CTCF binding sites (Kim et al., 2007; Xie et al., 2007). Intergenetic sequences with putative CTCF binding sites were used as stuffer DNA in ADMM-H_Tight-LacZ. Only Ins57-111s has no putative CTCF binding sites: The putative CTCF binding sites in Ins57-111 were lost after shortening Ins57-111 to generate Ins57-111s.

Stuffer DNA is also used to constrain the size of ADMM constructs. Type 5 human adenoviruses are not efficiently encapsidated if they are less than approximately 75% or more than approximately 105% of the wild type genome length (Bett et al., 1993; Parks and Graham, 1997): the genome length of type 5 human adenoviruses is approximately 36 kb. The length of ADMM-H_Tight-LacZ is approximately 33,945 bp; note that this ADMM construct itself has not been sequenced in its entirety: during construction of ADMM-H_Tight-LacZ, all newly cloned DNA was completely sequenced, however, in the later stages of construction, only ligation junctions were sequenced and restriction digestion was used assess the constructs.

ADMM-H_Tight-LacZ consists of 2 inverted terminal repeats (ITRs), a packaging site, and four cassettes: a LacZ expressing cassette, an HSV-TK expressing cassette, a large stuffer cassette, and a tTA expressing cassette. Digestion with PacI removes the entire ADMM-H_Tight-LacZ sequence from the carrier plasmid, resulting in a sequence which can propagate and be encapsidated into infectious adenovirus type 5 particles when transfected into packaging HEK293 cells which have been infected with EVE.

The tTA expressing cassette is modeled on the tetracycline responsive system originally developed by Hermann Bujard (Freundlieb et al., 1999). In ADMM-H_Tight-LacZ, the transcriptional activator tTA is itself under the control of the tetracycline responsive element (TRE), therefore, once tTA begins to be expressed in the absence of tetracycline, a positive feedback loop is established resulting in high expression levels of tTA. These high levels in turn result in robust expression of LacZ and HSV-TK, which are under the control of another tTA/tetracycline responsive element TRE-Tight. Because of these high expression levels, ADMM-H_Tight-LacZ is propagated in the presence of tetracycline (data not shown).

Importantly, ADMMs are built of a series of cassettes, and these cassettes can be changed. Therefore, different pathways can be targeted. For example, by utilizing different promoters and different positive and negative feedback loops, ADMM activity can be targeted to cells which do not express p53 (most human tumors) or cells which overexpress glutathione S-transferases (in the case of tumors resistant to chemotherapy), thereby specifically targeting ADMM activity to different kinds of tumors. One possibility is that libraries of different cassettes could be maintained. In this case, ADMM constructs able to respond to individual tumor types could be assembled in as little as a few weeks, making ADMM a practical cancer therapy given today's technology.

Propagation of EVE and ADMM in vitro

EVE is propagated in HEK293 cells expressing the cumate repressor protein CymR: CymR represses expression of Cre recombinase by the helper virus, allowing the helper virus to propagate normally.

ADMMs are propagated in HEK293 cells (not expressing CymR) which are also infected with helper virus. In the absence of CymR expression, Cre recombinase is expressed by EVE and catalyses excision of the packaging site out of the EVE genome. Consequently, the helper virus is not efficiently encapsidated and, therefore, not purified along with ADMM. However, EVE without its packaging site is still able to express all of the viral proteins required for propagation of ADMM and its encapsidation into infectous particles (data not shown). EGFP expressed by EVE is used to help monitor helper virus contamination of ADMM preparations.

Naming Conventions

In addition to ADMM-H_Tight-LacZ, we have generated ADMMs which express human delta-lactoferrin (ΔLF) (Mariller et al., 2007; Siebert and Huang, 1997) under the control of the human actin promoter. In this construct, the TRE-Tight-HSVTK sequence was removed from cassette pC-II by digestion of pC-II_Tight-HSVTK with SalI followed by recircularization, and the TRE-tTA was removed from pC-IV by digestion of pSHCS_Ins54_TRE-tTA-pA with NheI and PacI followed by blunting and recircularization. (Also, in the ΔLF construct, the TRE-tight promoter was replaced with the human actin promoter.) This construct is named pADMM-

54_AP-ΔLF. We also have a series of constructs in which the TRE-Tight-HSVTK sequence was removed from cassette pC-II_HSVTK, but cassette pC-IV_TRE-tTA was not altered. These ADMMs are named pADMM-53_XXX: for example, pADMM-53_LacZ. Finally, we constructed an ADMM in which the TRE-Tight-[downstream ORF] sequence has been removed from cassette pC-I, but the C-II, C-III, and C-IV cassettes were not altered. This construct is named pADMM-52_Tight-HSVTK. In sum, we have four series of pADMMs which we named the pADMM-52, pADMM-53, pADMM-54, and pADMM-H series.

Acknowledgements

The authors would like to acknowledge Emily Van Tassel who inspired the work described in this manuscript: the helper virus EVE is named after her (Emily Van Tassel's Virus with the E2 late leader inserted). The authors would also like to thank Dr. Masashi Mizokami of Nagoya City University Graduate School of Medical Sciences (Nagoya, Japan) for his advice and assistance. The authors would also like to acknowledge the students of Nagoya City University who participated in this work: Kan Omi, Sae Aratani, Sunsuke Nakamura, Naomi Niwa, Yukihide Numata, ,Sunao Ito, Yuki Kamishima, Yusuke Kato, Shohei Noguchi , Sae Saigo, Yasuhiro Wakano, Katsuyoshi Shikimori, Kotaro Hayashi, and Rui Terada. We would also like to thank the Région Nord-Pas de Calais, the Institut Fédératif de Recherche 147 and the Université des Sciences et Technologies de Lille for support for Stephan Hardivillé's stay at Nagoya University.

References

Alba, R, Bosch, A, Chillon, M (2005). Gutless adenoenovirus: last-generation adenovirus for gene therapy. *Gene Ther*, **12** Suppl 1, S18-27.

Alexander DB, Iigo M, Morita Y, Takase M, Tsuda H (2007). Identification of a second intracellular isoform of lactoferrin. In "Lactoferrin 2007" (H. Tsuda, K. Shimazaki, and K. Tanaka, Eds.), pp. 73-83. Nihon Igakukan, Tokyo, Japan.

Bett, A J, Prevec, L, Graham, F L (1993). Packaging capacity and stability of human adenovirus type 5 vectors. *J Virol*, **67**, 5911-21.

Cross D, Burmester JK (2006). Gene therapy for cancer treatment: past, present and future. *Clin Med Res*, **4**, 218-27.

Cuesta R, Xi Q, Schneider RJ (2000). Adenovirus-specific translation by displacement of kinase Mnk1 from cap-initiation complex eIF4F. *EMBO J*, **19**, 3465-74.

Cuesta R, Xi Q, Schneider RJ (2004). Structural basis for competitive inhibition of eIF4G-Mnk1 interaction by the adenovirus 100-kilodalton protein. *J Virol*, **78**, 7707-16.

Freundlieb S, Schirra-Muller C, Bujard H (1999). A tetracycline controlled activation/repression system with increased potential for gene transfer into mammalian cells. *J Gene Med*, **1**, 4-12.

Gaszner M, Felsenfeld G (2006). Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat Rev Genet*, **7**, 703-13.

Jozkowicz A, Dulak J (2005). Helper-dependent adenoviral vectors in experimental gene therapy. *Acta BiochimPol*, 52, 589-99.

Kim S, Peng Z, Kaneda Y (2008). Current status of gene therapy in Asia. *Mol Ther*, **16**, 237-43.

Kim TH, Abdullaev ZK, Smith AD, et al (2007). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell*, **128**, 1231-45.

Kovesdi, I, Reichel, R, and Nevins, J R (1987). Role of an adenovirus E2 promoter binding factor in E1A-mediated coordinate gene control. *Proc Natl Acad Sci USA*, **84**, 2180-4.

Larkin MA, Blackshields G, Brown NP, et al (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947-8.

Loonstra A, Vooijs M, Beverloo HB, et al (2001). Growth inhibition and DNA damage induced by Cre recombinase in mammalian cells. *Proc Natl Acad Sci USA*, **98**, 9209-14.

Mariller C, Benaissa M, Hardiville S, et al (2007). Human delta-lactoferrin is a transcription factor that enhances Skp1 (S-phase kinase-associated protein) gene expression. *FEBS J*, **274**, 2038-53.

Mizuguchi H, Kay MA (1998). Efficient construction of a recombinant adenovirus vector by an improved in vitro ligation method. *Hum Gene Ther*, **9**, 2577-83.

Parks RJ, Graham FL (1997). A helper-dependent system for adenovirus vector production helps define a lower limit for efficient DNA packaging. *J Virol*, **71**, 3293-8.

Peng, Z (2005). Current status of gendicine in China: recombinant human Ad-p53 agent for treatment of cancers. *Hum Gene Ther*, **16**, 1016-27.

Shirakawa T (2008). The current status of adenovirus- based cancer gene therapy. *Mol Cells*, **25**, 462-6.

Siebert PD, Huang BC (1997). Identification of an alternative form of human lactoferrin mRNA that is expressed differentially in normal tissues and tumor-derived cell lines. *Proc Natl Acad Sci USA*, **94**, 2198-203.

Wallace JA, Felsenfeld G (2007). We gather together: insulators and genome organization. *Curr Opin Genet Dev*, **17**, 400-7.

Xi Q, Cuesta R, Schneider RJ (2004). Tethering of eIF4G to adenoviral mRNAs by viral 100k protein drives ribosome shunting. *Genes Dev*, **18**, 1997-2009.

Xie X, Mikkelsen TS, Gnirke A, et al (2007). Systematic discovery of regulatory motifs in conserved regions of the human genome, including thousands of CTCF insulator sites. *Proc Natl Acad Sci USA*, **104**, 7145-50.

Xu ZL, Mizuguchi H, Sakurai F, et al (2005).Approaches to improving the kinetics of adenovirus-delivered genes and gene products. *Adv Drug Deliv Rev*, **57**, 781-802.

Liste des abréviations

5'RACE	Rapid Amplification of 5' Complementary DNA Ends
A	Adénine
ADN	Acide désoxyribonucléique
ADNc	ADN complémentaire
AIF	Apoptosis-inducing factor
AML-1	Acute myeloid leukemia 1
AMPc	Adénosine monophosphate cyclique
AOA	Acide okadaïque
Apaf-1	Apoptotic peptidase activating factor 1
APC/C	Anaphase-promoting complex/cyclosome
Ara	Arabinose
ARN	Acide ribonucléique
ARNm :	ARN messager
ATP	Adénosine triphosphate
Bad	Bcl-2-associated agonist of cell death
Bak	Bcl-2-antagonist/killer
Bax	Bcl-2-associated X protein
Bcl-2	B-cell lymphoma 2
BEMAD	β-Elimination followed by Michael Addition with DTT
BH	Bcl-2 Homology domain
Bid	BH3 interacting domain death agonist
Bik	Bcl-2-interacting killer
Bim	Bcl-2 interacting mediator of cell death
Blk	B lymphoid tyrosine kinase
BNIP3	Bcl-2/adenovirus E1B 19kDa interacting protein 3
Bok	Bcl-2-related ovarian killer
C	Cytosine
CAD	Caspase Activated DNase
CASH	Carbohydrate-binding proteins and Sugar Hydrolases
CAT	Chloramphénicol acétyl transférase
CBC	Cap-binding complex
CBF3	Centromere-binding factor 3

CCT2	T-complex protein 1, beta subunit
Cdc	Cell division cycle
Cdk	Cyclin dependent kinase
CENP-E	Centromere-associated protein E
CH	Calponin homology
ChIP	Chromatin immunoprecipitation
CRE	cAMP response element
CREB	cAMP responsive element-binding protein
CTD	Carboxy-terminal domain
Cul	Culline
DBD	DNA binding domain
D-Box	Destruction box
DcpS	Scavenger decapping enzyme
DON	6-diazo-5-oxo-norleucine
DR	Death receptor
DTT	Dithiothréitol
DUB	Desubiquitinase
EGF	Epidermal growth factor
eIF4E	Eukaryotic translation initiation factor 4E
eNOS	Endothelial nitrique oxyde synthase
ERE	Estrogen response element
FADD	Fas associated death domain
FBP	F-box protein
FoxO1	Forkhead box O1
Fuc	Fucose
G	Guanine
Gal	Galactose
GalNAc	N-acétyl galactosamine
GFAT	Glutamine:fructose-6-phosphate aminotransférase
Glc	Glucose
GlcNAc	N-acétylglucosamine
GlcNAcylation	*O*-N-acétylglucosaminylation
GlcNH2	Glucosamine

GLUT	Glucose transporter
GRIF-1	GABAA receptor-interacting factor 1
HAT	Histone acétyltransférase
HECT	Homologous to the E6-AP Carboxyl Terminus
HIC-1	Hypermethylated In Cancer 1
HNRPL	Heterogeneous Nuclear RibonucleoProtein L
HRE	Hormone Response Element
Hrk	Harakiri, Bcl-2 interacting protein
Hsc	Heat shock cognate protein
Hsp	Heat shock proteins
ICAD	Inhibitor of caspase activated DNase
IRES	Internal ribosomal entry sequence
IRS-1	Insulin receptor substrate 1
kDa	Kilo Dalton
LC-MS	Liquid chromatography-mass spectrometry
LPS	Lipopolysaccharide
MALDI-TOF	Matrix assisted laser desorption ionization-time of flight
Man	Mannose
Mcl 1	Myeloid cell leukemia sequence 1
mER-β	Murine estrogen receptor-β
mOGT	Mitochondrial OGT
MTS	Mitochondrial targeting sequence
ncOGT	Nuclear and cytoplasmic OGT
NeuroD1	Neurogenic differentiation 1
NF-κB	Nuclear factor-kappa B
NK	Natural killer
NLS	Nuclear localization signal
NSF	N-ethylmaleimide sensitive fusion protein
OGA	N-acétyl β-D-glucosaminidase
OGT	Uridine diphospho-*N*-acétylglucosamine : polypeptide β-*N*-acétylglucosaminyltransférase
OIP 106	OGT-interacting protein 106
PARP	Poly (ADP-ribose) polymérase

Pax-6	Paired box gene 6
pb	Paire de base
PCBP2	Poly(rC)-binding protein 2
PCR	Polymerase chain reaction
PDIA3	Protein disulfide isomerase A
PDX-1	Pancreatic and duodenal homeobox 1
PHD	Plant homeo domain
PIP3	Phosphatidylinositol 3,4,5 triphosphate
PP1	Ppyrophosphatase
PRP19	Pre-mRNA processing factor 19
PSB3	Proteasome subunit beta type 3
PUGNAc	*O*-(2-Acetamido-2-deoxy-D-glucopyranosylidenamino)N-phenylcarbamate
qPCR	Quantitative polymerase chain reaction
Rb	Retinoblastoma susceptibility protein
Rbx	Ring box protein
RE	Response element
RING	Really interesting new gene
SCF	Skp1-Cul1-F-box-protein
siRNA	Short interfering RNA
Skp1	S-phase-kinase-associated protein-1
SNARE	Soluble N-ethylmaleimide-sensitive fusion attachment protein receptor
sOGT	small OGT
Stat5A	Signal transducer and activator of transcription 5A
SUMO	Small ubiquitin related modifier
SV40	Simian virus 40
T	Thymine
TAF110	TATA- binding- protein associated factor 110
Tau	Tubulin associated unit
TDL	Traf domain-like
TGF	Transforming growth factor
TNF	Tumor necrosis factor
TPR	Tetratricopeptide repeat
Ub	Ubiquitine

Uba	Ubiquitin-activating enzymes
Ubc	Ubiquitin-conjugating enzyme
Ubl	Ubiquitin like modifier
UDP-GlcNAc	Uridine diPhospho- N-acétylglucosamine
WGA	Wheat germ agglutinin
XRE	Xenobiotic response element
Xrn1	5'-3' exoribonuclease 1
Xyl	Xylose
YY1	YinYang-1
ΔLfRE	Delta-lactoferrin response element
µM	Micromolaire

Liste des acides aminés

Nom	Code à une lettre	Code à trois lettres
Alanine	A	Ala
Arginine	R	Arg
Asparagine	N	Asn
Aspartate	D	Asp
Cystéine	C	Cys
Glutamate	E	Glu
Glutamine	Q	Gln
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Leucine	L	Leu
Lysine	K	Lys
Méthionine	M	Met
Phénylalanine	F	Phe
Proline	P	Pro
Sérine	S	Ser
Thréonine	T	Thr
Tryptophane	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val

Bibliographie

~ A ~

Abe, Y., Yoon, S.O., Kubota, K., Mendoza, M.C., Gygi, S.P., et Blenis, J. 2009. p90 ribosomal S6 kinase and p70 ribosomal S6 kinase link phosphorylation of the eukaryotic chaperonin containing TCP-1 to growth factor, insulin, and nutrient signaling. *J Biol Chem*. 284:14939-48.

Acehan, D., Jiang, X., Morgan, D.G., Heuser, J.E., Wang, X., et Akey, C.W. 2002. Three-dimensional structure of the apoptosome: implications for assembly, procaspase-9 binding, and activation. *Mol Cell*. 9:423-32.

Akimoto, Y., Comer, F.I., Cole, R.N., Kudo, A., Kawakami, H., Hirano, H., et Hart, G.W. 2003. Localization of the O-GlcNAc transferase and O-GlcNAc-modified proteins in rat cerebellar cortex. *Brain Res*. 966:194-205.

Alexander, G.E., Chen, K., Pietrini, P., Rapoport, S.I., et Reiman, E.M. 2002. Longitudinal PET Evaluation of Cerebral Metabolic Decline in Dementia: A Potential Outcome Measure in Alzheimer's Disease Treatment Studies. *Am J Psychiatry*. 159:738-45.

Alexander, L.J., Levine, W.B., Teng, C.T., et Beattie, C.W. 1992. Cloning and sequencing of the porcine lactoferrin cDNA. *Anim Genet*. 23:251-6.

Algeciras-Schimnich, A., Barnhart, B.C., et Peter, M.E. 2002. Apoptosis-independent functions of killer caspases. *Curr Opin Cell Biol*. 14:721-6.

Anderson, B.F., Baker, H.M., Dodson, E.J., Norris, G.E., Rumball, S.V., Waters, J.M., et Baker, E.N. 1987. Structure of human lactoferrin at 3.2-A resolution. *Proc Natl Acad Sci U S A*. 84:1769-73.

Anderson, B.F., Baker, H.M., Norris, G.E., Rice, D.W., et Baker, E.N. 1989. Structure of human lactoferrin: crystallographic structure analysis and refinement at 2.8 A resolution. *J Mol Biol*. 209:711-34.

Anderson, B.F., Baker, H.M., Norris, G.E., Rumball, S.V., et Baker, E.N. 1990. Apolactoferrin structure demonstrates ligand-induced conformational change in transferrins. *Nature*. 344:784-7.

Andrali, S.S., Marz, P., et Ozcan, S. 2005. Ataxin-10 interacts with O-GlcNAc transferase OGT in pancreatic beta cells. *Biochem Biophys Res Commun*. 337:149-53.

Andrali, S.S., Qian, Q., et Ozcan, S. 2007. Glucose mediates the translocation of NeuroD1 by O-linked glycosylation. *J Biol Chem*. 282:15589-96.

Ang, X.L., et Wade Harper, J. 2005. SCF-mediated protein degradation and cell cycle control. *Oncogene*. 24:2860-70.

Arnold, C.S., Johnson, G.V., Cole, R.N., Dong, D.L., Lee, M., et Hart, G.W. 1996. The microtubule-associated protein tau is extensively modified with O-linked N-acetylglucosamine. *J Biol Chem*. 271:28741-4.

Artym, J., Zimecki, M., et Kruzel, M.L. 2003. Reconstitution of the cellular immune response by lactoferrin in cyclophosphamide-treated mice is correlated with renewal of T cell compartment. *Immunobiology*. 207:197-205.

~ B ~

Bail, S., et Kiledjian, M. 2008. DcpS, a general modulator of cap-binding protein-dependent processes? *RNA Biol*. 5:216-9.

Bailly, V., Lamb, J., Sung, P., Prakash, S., et Prakash, L. 1994. Specific complex formation between yeast RAD6 and RAD18 proteins: a potential mechanism for targeting RAD6 ubiquitin-conjugating activity to DNA damage sites. *Genes Dev*. 8:811-20.

Baker, E.N., et Baker, H.M. 2005. Molecular structure, binding properties and dynamics of

lactoferrin. *Cell Mol Life Sci*. 62:2531-9.

Baker, E.N., Baker, H.M., et Kidd, R.D. 2002. Lactoferrin and transferrin: functional variations on a common structural framework. *Biochem Cell Biol*. 80:27-34.

Baldwin, A.S., Jr. 2001. Series introduction: the transcription factor NF-kappaB and human disease. *J Clin Invest*. 107:3-6.

Baldwin, G.S. 1993. Comparison of transferrin sequences from different species. *Comp Biochem Physiol B*. 106:203-18.

Barton, J.C., Parmley, R.T., Butler, T.W., Williamson, S., MacKenzie, S., Chandler, D.B., Blackburn, W., et Heck, L.W., Jr. 1988. Neutrophil lactoferrin content: variation among mammals. *Anat Rec*. 221:567-75.

Baynes, R., Bezwoda, W., Bothwell, T., Khan, Q., et Mansoor, N. 1986. The non-immune inflammatory response: serial changes in plasma iron, iron-binding capacity, lactoferrin, ferritin and C-reactive protein. *Scand J Clin Lab Invest*. 46:695-704.

Beato, M., et Klug, J. 2000. Steroid hormone receptors: an update. *Hum Reprod Update*. 6:225-36.

Bellamy, W., Takase, M., Wakabayashi, H., Kawase, K., et Tomita, M. 1992a. Antibacterial spectrum of lactoferricin B, a potent bactericidal peptide derived from the N-terminal region of bovine lactoferrin. *J Appl Bacteriol*. 73:472-9.

Bellamy, W., Takase, M., Yamauchi, K., Wakabayashi, H., Kawase, K., et Tomita, M. 1992b. Identification of the bactericidal domain of lactoferrin. *Biochim Biophys Acta*. 1121:130-6.

Benaissa, M., Peyrat, J.P., Hornez, L., Mariller, C., Mazurier, J., et Pierce, A. 2005. Expression and prognostic value of lactoferrin mRNA isoforms in human breast cancer. *Int J Cancer*. 114:299-306.

Bennett, R.M., Eddie-Quartey, A.C., et Holt, P.J. 1973. Lactoferrin--an iron binding protein in synovial fluid. *Arthritis Rheum*. 16:186-90.

Bezault, J., Bhimani, R., Wiprovnick, J., et Furmanski, P. 1994. Human lactoferrin inhibits growth of solid tumors and development of experimental metastases in mice. *Cancer Res*. 54:2310-2.

Blagosklonny, M.V., et Pardee, A.B. 2002. The restriction point of the cell cycle. *Cell Cycle*. 1:103-10.

Bleicken, S., Classen, M., Padmavathi, P.V., Ishikawa, T., Zeth, K., Steinhoff, H.J., et Bordignon, E. 2009. Molecular details of Bax activation, oligomerization and membrane insertion. *J Biol Chem*.

Blumenfeld, M.L., et Krisman, C.R. 1986. The initiation of glycogen biosynthesis in rat heart alpha-1,4 glucans tightly associated with glycogen synthase. *Eur J Biochem*. 156:163-9.

Boulikas, T. 1993. Nuclear localization signals (NLS). *Crit Rev Eukaryot Gene Expr*. 3:193-227.

Braidman, I., Carroll, M., Dance, N., Robinson, D., Poenaru, L., Weber, A., Dreyfus, J.C., Overdijk, B., et Hooghwinkel, G.J. 1974. Characterisation of human N-acetyl-beta-hexosaminidase C. *FEBS Lett*. 41:181-4.

Brenner, C., Marzo, I., et Kroemer, G. 1998. A revolution in apoptosis: from a nucleocentric to a mitochondriocentric perspective. *Exp Gerontol*. 33:543-53.

Breton, M., Mariller, C., Benaissa, M., Caillaux, K., Browaeys, E., Masson, M., Vilain, J.P., Mazurier, J., et Pierce, A. 2004. Expression of delta-lactoferrin induces cell cycle arrest. *Biometals*. 17:325-9.

Broschat, K.O., Gorka, C., Page, J.D., Martin-Berger, C.L., Davies, M.S., Huang Hc, H.C., Gulve, E.A., Salsgiver, W.J., et Kasten, T.P. 2002. Kinetic characterization of human glutamine-fructose-6-phosphate amidotransferase I: potent feedback inhibition by

glucosamine 6-phosphate. *J Biol Chem*. 277:14764-70.

Butkinaree, C., Cheung, W.D., Park, S., Park, K., Barber, M., et Hart, G.W. 2008. Characterization of beta-N-acetylglucosaminidase cleavage by caspase-3 during apoptosis. *J Biol Chem*. 283:23557-66.

- C -

Campbell, T., Skilton, R.A., Coombes, R.C., Shousha, S., Graham, M.D., et Luqmani, Y.A. 1992. Isolation of a lactoferrin cDNA clone and its expression in human breast cancer. *Br J Cancer*. 65:19-26.

Cardozo, T., et Pagano, M. 2004. The SCF ubiquitin ligase: insights into a molecular machine. *Nat Rev Mol Cell Biol*. 5:739-51.

Castro, A., Bernis, C., Vigneron, S., Labbe, J.C., et Lorca, T. 2005. The anaphase-promoting complex: a key factor in the regulation of cell cycle. *Oncogene*. 24:314-25.

Cetinbas, N., Macauley, M.S., Stubbs, K.A., Drapala, R., et Vocadlo, D.J. 2006. Identification of Asp174 and Asp175 as the key catalytic residues of human O-GlcNAcase by functional analysis of site-directed mutants. *Biochemistry*. 45:3835-44.

Champattanachai, V., Marchase, R.B., et Chatham, J.C. 2007. Glucosamine protects neonatal cardiomyocytes from ischemia-reperfusion injury via increased protein-associated O-GlcNAc. *Am J Physiol Cell Physiol*. 292:C178-87.

Chaudhary, P.M., Eby, M., Jasmin, A., Bookwalter, A., Murray, J., et Hood, L. 1997. Death receptor 5, a new member of the TNFR family, and DR4 induce FADD-dependent apoptosis and activate the NF-kappaB pathway. *Immunity*. 7:821-30.

Cheng, X., et Hart, G.W. 2001. Alternative O-glycosylation/O-phosphorylation of serine-16 in murine estrogen receptor beta: post-translational regulation of turnover and transactivation activity. *J Biol Chem*. 276:10570-5.

Chinnaiyan, A.M., O'Rourke, K., Tewari, M., et Dixit, V.M. 1995. FADD, a novel death domain-containing protein, interacts with the death domain of Fas and initiates apoptosis. *Cell*. 81:505-12.

Chinnaiyan, A.M., O'Rourke, K., Yu, G.L., Lyons, R.H., Garg, M., Duan, D.R., Xing, L., Gentz, R., Ni, J., et Dixit, V.M. 1996. Signal transduction by DR3, a death domain-containing receptor related to TNFR-1 and CD95. *Science*. 274:990-2.

Chiu, Y.H., Sun, Q., et Chen, Z.J. 2007. E1-L2 activates both ubiquitin and FAT10. *Mol Cell*. 27:1014-23.

Chou, C.F., Smith, A.J., et Omary, M.B. 1992. Characterization and dynamics of O-linked glycosylation of human cytokeratin 8 and 18. *J Biol Chem*. 267:3901-6.

Chou, T.Y., Dang, C.V., et Hart, G.W. 1995a. Glycosylation of the c-Myc transactivation domain. *Proc Natl Acad Sci U S A*. 92:4417-21.

Chou, T.Y., Hart, G.W., et Dang, C.V. 1995b. c-Myc is glycosylated at threonine 58, a known phosphorylation site and a mutational hot spot in lymphomas. *J Biol Chem*. 270:18961-5.

Ciechanover, A., Elias, S., Heller, H., et Hershko, A. 1982. "Covalent affinity" purification of ubiquitin-activating enzyme. *J Biol Chem*. 257:2537-42.

Coller, J., et Parker, R. 2004. Eukaryotic mRNA decapping. *Annu Rev Biochem*. 73:861-90.

Comer, F.I., et Hart, G.W. 1999. O-GlcNAc and the control of gene expression. *Biochim Biophys Acta*. 1473:161-71.

Comer, F.I., et Hart, G.W. 2001. Reciprocity between O-GlcNAc and O-phosphate on the carboxyl terminal domain of RNA polymerase II. *Biochemistry*. 40:7845-52.

Comtesse, N., Maldener, E., et Meese, E. 2001. Identification of a nuclear variant of MGEA5,

a cytoplasmic hyaluronidase and a beta-N-acetylglucosaminidase. *Biochem Biophys Res Commun*. 283:634-40.

Conus, S., Kaufmann, T., Fellay, I., Otter, I., Rosse, T., et Borner, C. 2000. Bcl-2 is a monomeric protein: prevention of homodimerization by structural constraints. *EMBO J*. 19:1534-44.

Cook, J.C., et Chock, P.B. 1992. Isoforms of mammalian ubiquitin-activating enzyme. *J Biol Chem*. 267:24315-21.

Cook, J.C., et Chock, P.B. 1995. Phosphorylation of ubiquitin-activating enzyme in cultured cells. *Proc Natl Acad Sci U S A*. 92:3454-7.

Cullen, S.P., et Martin, S.J. 2009. Caspase activation pathways: some recent progress. *Cell Death Differ*. 16:935-8.

~ D ~

Damiens, E., El Yazidi, I., Mazurier, J., Duthille, I., Spik, G., et Boilly-Marer, Y. 1999. Lactoferrin inhibits G1 cyclin-dependent kinases during growth arrest of human breast carcinoma cells. *J Cell Biochem*. 74:486-98.

Damiens, E., Mazurier, J., el Yazidi, I., Masson, M., Duthille, I., Spik, G., et Boilly-Marer, Y. 1998. Effects of human lactoferrin on NK cell cytotoxicity against haematopoietic and epithelial tumour cells. *Biochim Biophys Acta*. 1402:277-87.

Das, A.K., Cohen, P.W., et Barford, D. 1998. The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions. *EMBO J*. 17:1192-9.

Das, S., Boswell, S.A., Aaronson, S.A., et Lee, S.W. 2008. P53 promoter selection: choosing between life and death. *Cell Cycle*. 7:154-7.

Dash, B.C., et El-Deiry, W.S. 2004. Cell cycle checkpoint control mechanisms that can be disrupted in cancer. *Methods Mol Biol*. 280:99-161.

Dauphinee, S.M., Ma, M., et Too, C.K. 2005. Role of O-linked beta-N-acetylglucosamine modification in the subcellular distribution of alpha4 phosphoprotein and Sp1 in rat lymphoma cells. *J Cell Biochem*. 96:579-88.

Dehennaut, V., Lefebvre, T., Sellier, C., Leroy, Y., Gross, B., Walker, S., Cacan, R., Michalski, J.C., Vilain, J.P., et Bodart, J.F. 2007. O-linked N-acetylglucosaminyltransferase inhibition prevents G2/M transition in Xenopus laevis oocytes. *J Biol Chem*. 282:12527-36.

Dennis, R.J., Taylor, E.J., Macauley, M.S., Stubbs, K.A., Turkenburg, J.P., Hart, S.J., Black, G.N., Vocadlo, D.J., et Davies, G.J. 2006. Structure and mechanism of a bacterial beta-glucosaminidase having O-GlcNAcase activity. *Nat Struct Mol Biol*. 13:365-71.

Derisbourg, P., Wieruszeski, J.M., Montreuil, J., et Spik, G. 1990. Primary structure of glycans isolated from human leucocyte lactotransferrin. Absence of fucose residues questions the proposed mechanism of hyposideraemia. *Biochem J*. 269:821-5.

Dieckmann-Schuppert, A., Bause, E., et Schwarz, R.T. 1993. Studies on O-glycans of Plasmodium-falciparum-infected human erythrocytes. Evidence for O-GlcNAc and O-GlcNAc-transferase in malaria parasites. *Eur J Biochem*. 216:779-88.

Donadio, A.C., Lobo, C., Tosina, M., de la Rosa, V., Martin-Rufian, M., Campos-Sandoval, J.A., Mates, J.M., Marquez, J., Alonso, F.J., et Segura, J.A. 2008. Antisense glutaminase inhibition modifies the O-GlcNAc pattern and flux through the hexosamine pathway in breast cancer cells. *J Cell Biochem*. 103:800-11.

Dong, D.L., et Hart, G.W. 1994. Purification and characterization of an O-GlcNAc selective N-acetyl-beta-D-glucosaminidase from rat spleen cytosol. *J Biol Chem*. 269:19321-30.

Dong, D.L., Xu, Z.S., Hart, G.W., et Cleveland, D.W. 1996. Cytoplasmic O-GlcNAc modification of the head domain and the KSP repeat motif of the neurofilament protein neurofilament-H. *J Biol Chem*. 271:20845-52.

~ E ~

Efthymiadis, A., Shao, H., Hubner, S., et Jans, D.A. 1997. Kinetic characterization of the human retinoblastoma protein bipartite nuclear localization sequence (NLS) in vivo and in vitro. A comparison with the SV40 large T-antigen NLS. *J Biol Chem*. 272:22134-9.

Emerald, B.S., Chen, Y., Zhu, T., Zhu, Z., Lee, K.O., Gluckman, P.D., et Lobie, P.E. 2007. AlphaCP1 mediates stabilization of hTERT mRNA by autocrine human growth hormone. *J Biol Chem*. 282:680-90.

~ F ~

Fang, B., et Miller, M.W. 2001. Use of galactosyltransferase to assess the biological function of O-linked N-acetyl-d-glucosamine: a potential role for O-GlcNAc during cell division. *Exp Cell Res*. 263:243-53.

Farley, J., Loup, D., Nelson, M., Mitchell, A., Esplund, G., Macri, C., Harrison, C., et Gray, K. 1997. Neoplastic transformation of the endocervix associated with downregulation of lactoferrin expression. *Mol Carcinog*. 20:240-50.

Fillebeen, C., Ruchoux, M.M., Mitchell, V., Vincent, S., Benaissa, M., et Pierce, A. 2001. Lactoferrin is synthesized by activated microglia in the human substantia nigra and its synthesis by the human microglial CHME cell line is upregulated by tumor necrosis factor alpha or 1-methyl-4-phenylpyridinium treatment. *Brain Res Mol Brain Res*. 96:103-13.

Finlay, D.R., Newmeyer, D.D., Price, T.M., et Forbes, D.J. 1987. Inhibition of in vitro nuclear transport by a lectin that binds to nuclear pores. *J Cell Biol*. 104:189-200.

~ G ~

Gao, Y., Miyazaki, J., et Hart, G.W. 2003. The transcription factor PDX-1 is post-translationally modified by O-linked N-acetylglucosamine and this modification is correlated with its DNA binding activity and insulin secretion in min6 beta-cells. *Arch Biochem Biophys*. 415:155-63.

Gao, Y., Wells, L., Comer, F.I., Parker, G.J., et Hart, G.W. 2001. Dynamic O-glycosylation of nuclear and cytosolic proteins: cloning and characterization of a neutral, cytosolic beta-N-acetylglucosaminidase from human brain. *J Biol Chem*. 276:9838-45.

Garneau, N.L., Wilusz, J., et Wilusz, C.J. 2007. The highways and byways of mRNA decay. *Nat Rev Mol Cell Biol*. 8:113-26.

Garre, C., Bianchi-Scarra, G., Sirito, M., Musso, M., et Ravazzolo, R. 1992. Lactoferrin binding sites and nuclear localization in K562(S) cells. *J Cell Physiol*. 153:477-82.

Garvie, C.W., et Wolberger, C. 2001. Recognition of specific DNA sequences. *Mol Cell*. 8:937-46.

George, N.M., Targy, N., Evans, J.J., Zhang, L., et Luo, X. Bax contains two functional mitochondrial targeting sequences and translocates to mitochondria in a conformational change- and homo-oligomerization-driven process. *J Biol Chem*.

285:1384-92.

Gewinner, C., Hart, G., Zachara, N., Cole, R., Beisenherz-Huss, C., et Groner, B. 2004. The coactivator of transcription CREB-binding protein interacts preferentially with the glycosylated form of Stat5. *J Biol Chem*. 279:3563-72.

Giancotti, V. 2006. Breast cancer markers. *Cancer Lett*. 243:145-59.

Gifford, J.L., Hunter, H.N., et Vogel, H.J. 2005. Lactoferricin: a lactoferrin-derived peptide with antimicrobial, antiviral, antitumor and immunological properties. *Cell Mol Life Sci*. 62:2588-98.

Gingras, A.C., Raught, B., et Sonenberg, N. 1999. eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu Rev Biochem*. 68:913-63.

Glotzer, M., Murray, A.W., et Kirschner, M.W. 1991. Cyclin is degraded by the ubiquitin pathway. *Nature*. 349:132-8.

Goodman, R.E., et Schanbacher, F.L. 1991. Bovine lactoferrin mRNA: sequence, analysis, and expression in the mammary gland. *Biochem Biophys Res Commun*. 180:75-84.

Greis, K.D., Gibson, W., et Hart, G.W. 1994. Site-specific glycosylation of the human cytomegalovirus tegument basic phosphoprotein (UL32) at serine 921 and serine 952. *J Virol*. 68:8339-49.

Griffith, L.S., et Schmitz, B. 1999. O-linked N-acetylglucosamine levels in cerebellar neurons respond reciprocally to pertubations of phosphorylation. *Eur J Biochem*. 262:824-31.

Groettrup, M., Pelzer, C., Schmidtke, G., et Hofmann, K. 2008. Activating the ubiquitin family: UBA6 challenges the field. *Trends Biochem Sci*. 33:230-7.

Guinez, C., Losfeld, M.E., Cacan, R., Michalski, J.C., et Lefebvre, T. 2006. Modulation of HSP70 GlcNAc-directed lectin activity by glucose availability and utilization. *Glycobiology*. 16:22-8.

Guinez, C., Mir, A.M., Dehennaut, V., Cacan, R., Harduin-Lepers, A., Michalski, J.C., et Lefebvre, T. 2008. Protein ubiquitination is modulated by O-GlcNAc glycosylation. *FASEB J*. 22:2901-11.

Guinez, C., Mir, A.M., Leroy, Y., Cacan, R., Michalski, J.C., et Lefebvre, T. 2007. Hsp70-GlcNAc-binding activity is released by stress, proteasome inhibition, and protein misfolding. *Biochem Biophys Res Commun*. 361:414-20.

~ H ~

Haas, A.L., Bright, P.M., et Jackson, V.E. 1988. Functional diversity among putative E2 isozymes in the mechanism of ubiquitin-histone ligation. *J Biol Chem*. 263:13268-75.

Haltiwanger, R.S., Blomberg, M.A., et Hart, G.W. 1992. Glycosylation of nuclear and cytoplasmic proteins. Purification and characterization of a uridine diphospho-N-acetylglucosamine:polypeptide beta-N-acetylglucosaminyltransferase. *J Biol Chem*. 267:9005-13.

Haltiwanger, R.S., Holt, G.D., et Hart, G.W. 1990. Enzymatic addition of O-GlcNAc to nuclear and cytoplasmic proteins. Identification of a uridine diphospho-N-acetylglucosamine:peptide beta-N-acetylglucosaminyltransferase. *J Biol Chem*. 265:2563-8.

Hames, R.S., Wattam, S.L., Yamano, H., Bacchieri, R., et Fry, A.M. 2001. APC/C-mediated destruction of the centrosomal kinase Nek2A occurs in early mitosis and depends upon a cyclin A-type D-box. *EMBO J*. 20:7117-27.

Han, I., et Kudlow, J.E. 1997. Reduced O glycosylation of Sp1 is associated with increased proteasome susceptibility. *Mol Cell Biol*. 17:2550-8.

Hanover, J.A., Cohen, C.K., Willingham, M.C., et Park, M.K. 1987. O-linked N-

acetylglucosamine is attached to proteins of the nuclear pore. Evidence for cytoplasmic and nucleoplasmic glycoproteins. *J Biol Chem*. 262:9887-94.

Hanover, J.A., Forsythe, M.E., Hennessey, P.T., Brodigan, T.M., Love, D.C., Ashwell, G., et Krause, M. 2005. A Caenorhabditis elegans model of insulin resistance: altered macronutrient storage and dauer formation in an OGT-1 knockout. *Proc Natl Acad Sci U S A*. 102:11266-71.

Hanover, J.A., Yu, S., Lubas, W.B., Shin, S.H., Ragano-Caracciola, M., Kochran, J., et Love, D.C. 2003. Mitochondrial and nucleocytoplasmic isoforms of O-linked GlcNAc transferase encoded by a single mammalian gene. *Arch Biochem Biophys*. 409:287-97.

Hansen, N.E., Malmquist, J., et Thorell, J. 1975. Plasma myeloperoxidase and lactoferrin measured by radioimmunoassay: relations to neutrophil kinetics. *Acta Med Scand*. 198:437-43.

Harper, J.W., Burton, J.L., et Solomon, M.J. 2002. The anaphase-promoting complex: it's not just for mitosis any more. *Genes Dev*. 16:2179-206.

Hartweck, L.M., Scott, C.L., et Olszewski, N.E. 2002. Two O-linked N-acetylglucosamine transferase genes of Arabidopsis thaliana L. Heynh. have overlapping functions necessary for gamete and seed development. *Genetics*. 161:1279-91.

He, J., et Furmanski, P. 1995. Sequence specificity and transcriptional activation in the binding of lactoferrin to DNA. *Nature*. 373:721-4.

Hebert, L.F., Jr., Daniels, M.C., Zhou, J., Crook, E.D., Turner, R.L., Simmons, S.T., Neidigh, J.L., Zhu, J.S., Baron, A.D., et McClain, D.A. 1996. Overexpression of glutamine:fructose-6-phosphate amidotransferase in transgenic mice leads to insulin resistance. *J Clin Invest*. 98:930-6.

Heckel, D., Comtesse, N., Brass, N., Blin, N., Zang, K.D., et Meese, E. 1998. Novel immunogenic antigen homologous to hyaluronidase in meningioma. *Hum Mol Genet*. 7:1859-72.

Herman, J.G., et Baylin, S.B. 2003. Gene silencing in cancer in association with promoter hypermethylation. *N Engl J Med*. 349:2042-54.

Hershko, A., et Ciechanover, A. 1998. The ubiquitin system. *Annu Rev Biochem*. 67:425-79.

Hershko, A., Heller, H., Elias, S., et Ciechanover, A. 1983. Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. *J Biol Chem*. 258:8206-14.

Hiller, A.M., Koro, L.A., et Marchase, R.B. 1987. Glucose-1-phosphotransferase and N-acetylglucosamine-1-phosphotransferase have distinct acceptor specificities. *J Biol Chem*. 262:4377-81.

Hirai, Y., Kawakata, N., Satoh, K., Ikeda, Y., Hisayasu, S., Orimo, H., et Yoshino, Y. 1990. Concentrations of lactoferrin and iron in human milk at different stages of lactation. *J Nutr Sci Vitaminol (Tokyo)*. 36:531-44.

Hirschberg, C.B., et Snider, M.D. 1987. Topography of glycosylation in the rough endoplasmic reticulum and Golgi apparatus. *Annu Rev Biochem*. 56:63-87.

Hochstrasser, M. 1996. Ubiquitin-dependent protein degradation. *Annu Rev Genet*. 30:405-39.

Hochstrasser, M. 2006. Lingering mysteries of ubiquitin-chain assembly. *Cell*. 124:27-34.

Hoedt, E., Hardiville, S., Mariller, C., Elass, E., Perraudin, J.P., et Pierce, A. 2010. Discrimination and evaluation of lactoferrin and delta-lactoferrin gene expression levels in cancer cells and under inflammatory stimuli using TaqMan real time PCR. *Biometals*.

Holt, G.D., et Hart, G.W. 1986. The subcellular distribution of terminal N-acetylglucosamine moieties. Localization of a novel protein-saccharide linkage, O-linked GlcNAc. *J Biol Chem*. 261:8049-57.

Holt, G.D., Snow, C.M., Senior, A., Haltiwanger, R.S., Gerace, L., et Hart, G.W. 1987.

Nuclear pore complex glycoproteins contain cytoplasmically disposed O-linked N-acetylglucosamine. *J Cell Biol*. 104:1157-64.

Horie-Inoue, K., Takayama, K., Bono, H.U., Ouchi, Y., Okazaki, Y., et Inoue, S. 2006. Identification of novel steroid target genes through the combination of bioinformatics and functional analysis of hormone response elements. *Biochem Biophys Res Commun*. 339:99-106.

Hotchkiss, R.S., Strasser, A., McDunn, J.E., et Swanson, P.E. 2009. Cell death. *N Engl J Med*. 361:1570-83.

Hsu, S.Y., Kaipia, A., Zhu, L., et Hsueh, A.J. 1997. Interference of BAD (Bcl-xL/Bcl-2-associated death promoter)-induced apoptosis in mammalian cells by 14-3-3 isoforms and P11. *Mol Endocrinol*. 11:1858-67.

Hurley, W.L., Hegarty, H.M., et Metzler, J.T. 1994. In vitro inhibition of mammary cell growth by lactoferrin: a comparative study. *Life Sci*. 55:1955-63.

~ I ~

Iigo, M., Alexander, D.B., Long, N., Xu, J., Fukamachi, K., Futakuchi, M., Takase, M., et Tsuda, H. 2009. Anticarcinogenesis pathways activated by bovine lactoferrin in the murine small intestine. *Biochimie*. 91:86-101.

Iigo, M., Kuhara, T., Ushida, Y., Sekine, K., Moore, M.A., et Tsuda, H. 1999. Inhibitory effects of bovine lactoferrin on colon carcinoma 26 lung metastasis in mice. *Clin Exp Metastasis*. 17:35-40.

Issad, T., et Kuo, M. 2008. O-GlcNAc modification of transcription factors, glucose sensing and glucotoxicity. *Trends Endocrinol Metab*. 19:380-9.

Itoh, N., Yonehara, S., Ishii, A., Yonehara, M., Mizushima, S., Sameshima, M., Hase, A., Seto, Y., et Nagata, S. 1991. The polypeptide encoded by the cDNA for human cell surface antigen Fas can mediate apoptosis. *Cell*. 66:233-43.

Iyer, S.P., Akimoto, Y., et Hart, G.W. 2003. Identification and cloning of a novel family of coiled-coil domain proteins that interact with O-GlcNAc transferase. *J Biol Chem*. 278:5399-409.

~ J ~

Jackson, S.P., et Tjian, R. 1988. O-glycosylation of eukaryotic transcription factors: implications for mechanisms of transcriptional regulation. *Cell*. 55:125-33.

Jacobs, H.W., Keidel, E., et Lehner, C.F. 2001. A complex degradation signal in Cyclin A required for G1 arrest, and a C-terminal region for mitosis. *EMBO J*. 20:2376-86.

Jans, D.A., Briggs, L.J., Gustin, S.E., Jans, P., Ford, S., et Young, I.G. 1997. A functional bipartite nuclear localisation signal in the cytokine interleukin-5. *FEBS Lett*. 406:315-20.

Jensen, P.E. 2007. Recent advances in antigen processing and presentation. *Nat Immunol*. 8:1041-8.

Jin, J., Cardozo, T., Lovering, R.C., Elledge, S.J., Pagano, M., et Harper, J.W. 2004. Systematic analysis and nomenclature of mammalian F-box proteins. *Genes Dev*. 18:2573-80.

Jin, J., Li, X., Gygi, S.P., et Harper, J.W. 2007. Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging. *Nature*. 447:1135-8.

Jurgensmeier, J.M., Xie, Z., Deveraux, Q., Ellerby, L., Bredesen, D., et Reed, J.C. 1998. Bax

directly induces release of cytochrome c from isolated mitochondria. *Proc Natl Acad Sci U S A*. 95:4997-5002.

- K -

Kamemura, K., et Hart, G.W. 2003. Dynamic interplay between O-glycosylation and O-phosphorylation of nucleocytoplasmic proteins: a new paradigm for metabolic control of signal transduction and transcription. *Prog Nucleic Acid Res Mol Biol*. 73:107-36.

Kang, E.S., Han, D., Park, J., Kwak, T.K., Oh, M.A., Lee, S.A., Choi, S., Park, Z.Y., Kim, Y., et Lee, J.W. 2008. O-GlcNAc modulation at Akt1 Ser473 correlates with apoptosis of murine pancreatic beta cells. *Exp Cell Res*. 314:2238-48.

Karthikeyan, S., Yadav, S., Paramasivam, M., Srinivasan, A., et Singh, T.P. 2000. Structure of buffalo lactoferrin at 3.3 A resolution at 277 K. *Acta Crystallogr D Biol Crystallogr*. 56:684-9.

Kawamata, T., Tooyama, I., Yamada, T., Walker, D.G., et McGeer, P.L. 1993. Lactotransferrin immunocytochemistry in Alzheimer and normal human brain. *Am J Pathol*. 142:1574-85.

Kearse, K.P., et Hart, G.W. 1991a. Lymphocyte activation induces rapid changes in nuclear and cytoplasmic glycoproteins. *Proc Natl Acad Sci U S A*. 88:1701-5.

Kearse, K.P., et Hart, G.W. 1991b. Topology of O-linked N-acetylglucosamine in murine lymphocytes. *Arch Biochem Biophys*. 290:543-8.

Kelekar, A., et Thompson, C.B. 1998. Bcl-2-family proteins: the role of the BH3 domain in apoptosis. *Trends Cell Biol*. 8:324-30.

Kelly, W.G., et Hart, G.W. 1989. Glycosylation of chromosomal proteins: localization of O-linked N-acetylglucosamine in Drosophila chromatin. *Cell*. 57:243-51.

Kemp, Z., Carvajal-Carmona, L., Spain, S., Barclay, E., Gorman, M., Martin, L., Jaeger, E., Brooks, N., Bishop, D.T., Thomas, H., Tomlinson, I., Papaemmanuil, E., Webb, E., Sellick, G.S., Wood, W., Evans, G., Lucassen, A., Maher, E.R., et Houlston, R.S. 2006. Evidence for a colorectal cancer susceptibility locus on chromosome 3q21-q24 from a high-density SNP genome-wide linkage scan. *Hum Mol Genet*. 15:2903-10.

Khan, J.A., Kumar, P., Paramasivam, M., Yadav, R.S., Sahani, M.S., Sharma, S., Srinivasan, A., et Singh, T.P. 2001. Camel lactoferrin, a transferrin-cum-lactoferrin: crystal structure of camel apolactoferrin at 2.6 A resolution and structural basis of its dual role. *J Mol Biol*. 309:751-61.

Khanna-Gupta, A., Zibello, T., Kolla, S., Neufeld, E.J., et Berliner, N. 1997. CCAAT displacement protein (CDP/cut) recognizes a silencer element within the lactoferrin gene promoter. *Blood*. 90:2784-95.

Kim, E.J., Kang, D.O., Love, D.C., et Hanover, J.A. 2006. Enzymatic characterization of O-GlcNAcase isoforms using a fluorogenic GlcNAc substrate. *Carbohydr Res*. 341:971-82.

Kim, S.J., Yu, D.Y., Pak, K.W., Jeong, S., Kim, S.W., et Lee, K.K. 1998. Structure of the human lactoferrin gene and its chromosomal localization. *Mol Cells*. 8:663-8.

Kitagawa, K., et Hieter, P. 2001. Evolutionary conservation between budding yeast and human kinetochores. *Nat Rev Mol Cell Biol*. 2:678-87.

Klionsky, D.J. 2007. Autophagy: from phenomenology to molecular understanding in less than a decade. *Nat Rev Mol Cell Biol*. 8:931-7.

Kluck, R.M., Bossy-Wetzel, E., Green, D.R., et Newmeyer, D.D. 1997. The release of cytochrome c from mitochondria: a primary site for Bcl-2 regulation of apoptosis. *Science*. 275:1132-6.

Knights, C.D., Catania, J., Di Giovanni, S., Muratoglu, S., Perez, R., Swartzbeck, A., Quong, A.A., Zhang, X., Beerman, T., Pestell, R.G., et Avantaggiati, M.L. 2006. Distinct p53 acetylation cassettes differentially influence gene-expression patterns and cell fate. *J Cell Biol*. 173:533-44.

Kobe, B., et Kajava, A.V. 2001. The leucine-rich repeat as a protein recognition motif. *Curr Opin Struct Biol*. 11:725-32.

Kondrakhin, Y.V., Kel, A.E., Kolchanov, N.A., Romashchenko, A.G., et Milanesi, L. 1995. Eukaryotic promoter recognition by binding sites for transcription factors. *Comput Appl Biosci*. 11:477-88.

Kornfeld, R. 1967. Studies on L-glutamine D-fructose 6-phosphate amidotransferase. I. Feedback inhibition by uridine diphosphate-N-acetylglucosamine. *J Biol Chem*. 242:3135-41.

Kreppel, L.K., Blomberg, M.A., et Hart, G.W. 1997. Dynamic glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a unique O-GlcNAc transferase with multiple tetratricopeptide repeats. *J Biol Chem*. 272:9308-15.

Kreppel, L.K., et Hart, G.W. 1999. Regulation of a cytosolic and nuclear O-GlcNAc transferase. Role of the tetratricopeptide repeats. *J Biol Chem*. 274:32015-22.

Kroemer, G., et Jaattela, M. 2005. Lysosomes and autophagy in cell death control. *Nat Rev Cancer*. 5:886-97.

Kroemer, G., Zamzami, N., et Susin, S.A. 1997. Mitochondrial control of apoptosis. *Immunol Today*. 18:44-51.

Kuo, M., Zilberfarb, V., Gangneux, N., Christeff, N., et Issad, T. 2008. O-GlcNAc modification of FoxO1 increases its transcriptional activity: a role in the glucotoxicity phenomenon? *Biochimie*. 90:679-85.

Kuwata, H., Yip, T.T., Tomita, M., et Hutchens, T.W. 1998. Direct evidence of the generation in human stomach of an antimicrobial peptide domain (lactoferricin) from ingested lactoferrin. *Biochim Biophys Acta*. 1429:129-41.

~ L ~

Lahortiga, I., Vazquez, I., Agirre, X., Larrayoz, M.J., Vizmanos, J.L., Gozzetti, A., Calasanz, M.J., et Odero, M.D. 2004. Molecular heterogeneity in AML/MDS patients with 3q21q26 rearrangements. *Genes Chromosomes Cancer*. 40:179-89.

Lamb, J.R., Tugendreich, S., et Hieter, P. 1995. Tetratrico peptide repeat interactions: to TPR or not to TPR? *Trends Biochem Sci*. 20:257-9.

Lambert, L.A., Perri, H., Halbrooks, P.J., et Mason, A.B. 2005. Evolution of the transferrin family: conservation of residues associated with iron and anion binding. *Comp Biochem Physiol B Biochem Mol Biol*. 142:129-41.

Lavery, D.N., et McEwan, I.J. 2005. Structure and function of steroid receptor AF1 transactivation domains: induction of active conformations. *Biochem J*. 391:449-64.

Lavin, M.F., et Gueven, N. 2006. The complexity of p53 stabilization and activation. *Cell Death Differ*. 13:941-50.

Lawen, A. 2003. Apoptosis-an introduction. *Bioessays*. 25:888-96.

Lazarus, B.D., Love, D.C., et Hanover, J.A. 2006. Recombinant O-GlcNAc transferase isoforms: identification of O-GlcNAcase, yes tyrosine kinase, and tau as isoform-specific substrates. *Glycobiology*. 16:415-21.

Lee, D.H., Lim, M.H., Youn, D.Y., Jung, S.E., Ahn, Y.S., Tsujimoto, Y., et Lee, J.H. 2009. hnRNP L binds to CA repeats in the 3'UTR of bcl-2 mRNA. *Biochem Biophys Res Commun*. 382:583-7.

Lee, M.O., Liu, Y., et Zhang, X.K. 1995. A retinoic acid response element that overlaps an estrogen response element mediates multihormonal sensitivity in transcriptional activation of the lactoferrin gene. *Mol Cell Biol*. 15:4194-207.

Lefebvre, T., Alonso, C., Mahboub, S., Dupire, M.J., Zanetta, J.P., Caillet-Boudin, M.L., et Michalski, J.C. 1999. Effect of okadaic acid on O-linked N-acetylglucosamine levels in a neuroblastoma cell line. *Biochim Biophys Acta*. 1472:71-81.

Lefebvre, T., Baert, F., Bodart, J.F., Flament, S., Michalski, J.C., et Vilain, J.P. 2004a. Modulation of O-GlcNAc glycosylation during Xenopus oocyte maturation. *J Cell Biochem*. 93:999-1010.

Lefebvre, T., Ferreira, S., Dupont-Wallois, L., Bussiere, T., Dupire, M.J., Delacourte, A., Michalski, J.C., et Caillet-Boudin, M.L. 2003. Evidence of a balance between phosphorylation and O-GlcNAc glycosylation of Tau proteins--a role in nuclear localization. *Biochim Biophys Acta*. 1619:167-76.

Lefebvre, T., Pinte, S., Guerardel, C., Deltour, S., Martin-Soudant, N., Slomianny, M.C., Michalski, J.C., et Leprince, D. 2004b. The tumor suppressor HIC1 (hypermethylated in cancer 1) is O-GlcNAc glycosylated. *Eur J Biochem*. 271:3843-54.

Lefebvre, T., Planque, N., Leleu, D., Bailly, M., Caillet-Boudin, M.L., Saule, S., et Michalski, J.C. 2002. O-glycosylation of the nuclear forms of Pax-6 products in quail neuroretina cells. *J Cell Biochem*. 85:208-18.

Legrand, D., et Mazurier, J. A critical review of the roles of host lactoferrin in immunity. *Biometals*.

Legrand, D., Mazurier, J., Aubert, J.P., Loucheux-Lefebvre, M.H., Montreuil, J., et Spik, G. 1986. Evidence for interactions between the 30 kDa N- and 50 kDa C-terminal tryptic fragments of human lactotransferrin. *Biochem J*. 236:839-44.

Legrand, D., Pierce, A., Elass, E., Carpentier, M., Mariller, C., et Mazurier, J. 2008. Lactoferrin structure and functions. *Adv Exp Med Biol*. 606:163-94.

Lehmann, A., et Toda, T. 2004. Fission yeast Skp1 is required for spindle morphology and nuclear membrane segregation at anaphase. *FEBS Lett*. 566:77-82.

Leist, M., Single, B., Castoldi, A.F., Kuhnle, S., et Nicotera, P. 1997. Intracellular adenosine triphosphate (ATP) concentration: a switch in the decision between apoptosis and necrosis. *J Exp Med*. 185:1481-6.

Leonard, W.J., et O'Shea, J.J. 1998. Jaks and STATs: biological implications. *Annu Rev Immunol*. 16:293-322.

Leveugle, B., Spik, G., Perl, D.P., Bouras, C., Fillit, H.M., et Hof, P.R. 1994. The iron-binding protein lactotransferrin is present in pathologic lesions in a variety of neurodegenerative disorders: a comparative immunohistochemical analysis. *Brain Res*. 650:20-31.

Lewis, J.D., et Izaurralde, E. 1997. The role of the cap structure in RNA processing and nuclear export. *Eur J Biochem*. 247:461-9.

Li, H., Zhu, H., Xu, C.J., et Yuan, J. 1998. Cleavage of BID by caspase 8 mediates the mitochondrial damage in the Fas pathway of apoptosis. *Cell*. 94:491-501.

Li, L.Y., Luo, X., et Wang, X. 2001. Endonuclease G is an apoptotic DNase when released from mitochondria. *Nature*. 412:95-9.

Li, X., Lu, F., Wang, J.Z., et Gong, C.X. 2006. Concurrent alterations of O-GlcNAcylation and phosphorylation of tau in mouse brains during fasting. *Eur J Neurosci*. 23:2078-86.

Lindberg, M.K., Moverare, S., Skrtic, S., Gao, H., Dahlman-Wright, K., Gustafsson, J.A., et Ohlsson, C. 2003. Estrogen receptor (ER)-beta reduces ERalpha-regulated gene transcription, supporting a "ying yang" relationship between ERalpha and ERbeta in mice. *Mol Endocrinol*. 17:203-8.

Linder, P., Lasko, P.F., Ashburner, M., Leroy, P., Nielsen, P.J., Nishi, K., Schnier, J., et Slonimski, P.P. 1989. Birth of the D-E-A-D box. *Nature*. 337:121-2.

Littleton, J.T., Barnard, R.J., Titus, S.A., Slind, J., Chapman, E.R., et Ganetzky, B. 2001. SNARE-complex disassembly by NSF follows synaptic-vesicle fusion. *Proc Natl Acad Sci U S A*. 98:12233-8.

Liu, D., Wang, X., Zhang, Z., et Teng, C.T. 2003. An intronic alternative promoter of the human lactoferrin gene is activated by Ets. *Biochem Biophys Res Commun*. 301:472-9.

Liu, D., Zhang, N., Du, J., Cai, X., Zhu, M., Jin, C., Dou, Z., Feng, C., Yang, Y., Liu, L., Takeyasu, K., Xie, W., et Yao, X. 2006. Interaction of Skp1 with CENP-E at the midbody is essential for cytokinesis. *Biochem Biophys Res Commun*. 345:394-402.

Liu, F., Iqbal, K., Grundke-Iqbal, I., Hart, G.W., et Gong, C.X. 2004a. O-GlcNAcylation regulates phosphorylation of tau: a mechanism involved in Alzheimer's disease. *Proc Natl Acad Sci U S A*. 101:10804-9.

Liu, H., et Kiledjian, M. 2006. Decapping the message: a beginning or an end. *Biochem Soc Trans*. 34:35-8.

Liu, K., Paterson, A.J., Zhang, F., McAndrew, J., Fukuchi, K., Wyss, J.M., Peng, L., Hu, Y., et Kudlow, J.E. 2004b. Accumulation of protein O-GlcNAc modification inhibits proteasomes in the brain and coincides with neuronal apoptosis in brain areas with high O-GlcNAc metabolism. *J Neurochem*. 89:1044-55.

Liu, Y., Liu, F., Grundke-Iqbal, I., Iqbal, K., et Gong, C.X. 2009. Brain glucose transporters, O-GlcNAcylation and phosphorylation of tau in diabetes and Alzheimer's disease. *J Neurochem*. 111:242-9.

Long, N., Hardiville, S., Pierce, A., Fukamachi, K., Futakuchi, M., Alexander, D.B., Shirai, T., et Tsuda, H. 2009. Construction of a Multi-Functional Helper-Dependent Adenovirus Based System for Cancer Gene Therapy. *Asian Pac J Cancer Prev*. 10:939-960.

Love, D.C., Kochan, J., Cathey, R.L., Shin, S.H., et Hanover, J.A. 2003. Mitochondrial and nucleocytoplasmic targeting of O-linked GlcNAc transferase. *J Cell Sci*. 116:647-54.

Lubas, W.A., Frank, D.W., Krause, M., et Hanover, J.A. 1997. O-Linked GlcNAc transferase is a conserved nucleocytoplasmic protein containing tetratricopeptide repeats. *J Biol Chem*. 272:9316-24.

Lubas, W.A., et Hanover, J.A. 2000. Functional expression of O-linked GlcNAc transferase. Domain structure and substrate specificity. *J Biol Chem*. 275:10983-8.

Luo, X., Budihardjo, I., Zou, H., Slaughter, C., et Wang, X. 1998. Bid, a Bcl2 interacting protein, mediates cytochrome c release from mitochondria in response to activation of cell surface death receptors. *Cell*. 94:481-90.

Luqmani, Y.A., Campbell, T.A., Bennett, C., Coombes, R.C., et Paterson, I.M. 1991. Expression of lactoferrin in human stomach. *Int J Cancer*. 49:684-7.

Lydon, J.P., O'Malley, B.R., Saucedo, O., Lee, T., Headon, D.R., et Conneely, O.M. 1992. Nucleotide and primary amino acid sequence of porcine lactoferrin. *Biochim Biophys Acta*. 1132:97-9.

~ M ~

Maacks, S., Tegtmeyer, K.F., et Wood, W.G. 1990. The influence of blood sampling techniques on plasma concentrations of lactoferrin and elastase-alpha 1-PI. *Ann Clin Biochem*. 27 (Pt 5):512-3.

Makarova, O.V., Makarov, E.M., Urlaub, H., Will, C.L., Gentzel, M., Wilm, M., et Luhrmann, R. 2004. A subset of human 35S U5 proteins, including Prp19, function prior to

catalytic step 1 of splicing. *EMBO J*. 23:2381-91.

Malhi, H., Gores, G.J., et Lemasters, J.J. 2006. Apoptosis and necrosis in the liver: a tale of two deaths? *Hepatology*. 43:S31-44.

Mann, D.M., Romm, E., et Migliorini, M. 1994. Delineation of the glycosaminoglycan-binding site in the human inflammatory response protein lactoferrin. *J Biol Chem*. 269:23661-7.

Mapp, A.K., et Ansari, A.Z. 2007. A TAD further: exogenous control of gene activation. *ACS Chem Biol*. 2:62-75.

Margolis, R.K., Crockett, C.P., Kiang, W.L., et Margolis, R.U. 1976. Glycosaminoglycans and glycoproteins associated with rat brain nuclei. *Biochim Biophys Acta*. 451:465-9.

Margolis, R.K., Thomas, M.D., Crockett, C.P., et Margolis, R.U. 1979. Presence of chondroitin sulfate in the neuronal cytoplasm. *Proc Natl Acad Sci U S A*. 76:1711-5.

Marmorstein, R., et Fitzgerald, M.X. 2003. Modulation of DNA-binding domains for sequence-specific DNA recognition. *Gene*. 304:1-12.

Marshall, S., Bacote, V., et Traxinger, R.R. 1991. Discovery of a metabolic pathway mediating glucose-induced desensitization of the glucose transport system. Role of hexosamine biosynthesis in the induction of insulin resistance. *J Biol Chem*. 266:4706-12.

Martinez-Fleites, C., Macauley, M.S., He, Y., Shen, D.L., Vocadlo, D.J., et Davies, G.J. 2008. Structure of an O-GlcNAc transferase homolog provides insight into intracellular glycosylation. *Nat Struct Mol Biol*. 15:764-5.

Masson, P., Heremans, J.F., et Prignot, J. 1965. Immunohistochemical localization of the iron-binding protein lactoferrin in human bronchial glands. *Experientia*. 21:604-5.

Masson, P.L., et Heremans, J.F. 1971. Lactoferrin in milk from different species. *Comp Biochem Physiol B*. 39:119-29.

Masson, P.L., Heremans, J.F., et Ferin, J. 1968. Presence of an Iron-binding protein (lactoferrin) in the genital tract of the human female. I. Its immunohistochemical localization in the endometrium. *Fertil Steril*. 19:679-89.

Masson, P.L., Heremans, J.F., et Schonne, E. 1969. Lactoferrin, an iron-binding protein in neutrophilic leukocytes. *J Exp Med*. 130:643-58.

Mazurier, J., Metz-Boutigue, M.H., Jolles, J., Spik, G., Montreuil, J., et Jolles, P. 1983. Human lactotransferrin: molecular, functional and evolutionary comparisons with human serum transferrin and hen ovotransferrin. *Experientia*. 39:135-41.

McClain, D.A., Lubas, W.A., Cooksey, R.C., Hazel, M., Parker, G.J., Love, D.C., et Hanover, J.A. 2002. Altered glycan-dependent signaling induces insulin resistance and hyperleptinemia. *Proc Natl Acad Sci U S A*. 99:10695-9.

McCombs, J.L., Teng, C.T., Pentecost, B.T., Magnuson, V.L., Moore, C.M., et McGill, J.R. 1988. Chromosomal localization of human lactotransferrin gene (LTF) by in situ hybridization. *Cytogenet Cell Genet*. 47:16-7.

McGrattan, P., Logan, A., Humphreys, M., et Bowers, M. 2009. Jumping translocation in acute monocytic leukemia (M5b) with alternative breakpoint sites in the long arm of donor chromosome 3. *Med Oncol*.

Mead, P.E., et Tweedie, J.W. 1990. cDNA and protein sequence of bovine lactoferrin. *Nucleic Acids Res*. 18:7167.

Medina, L., Grove, K., et Haltiwanger, R.S. 1998. SV40 large T antigen is modified with O-linked N-acetylglucosamine but not with other forms of glycosylation. *Glycobiology*. 8:383-91.

Medina, L., et Haltiwanger, R.S. 1998. Calf thymus high mobility group proteins are nonenzymatically glycated but not significantly glycosylated. *Glycobiology*. 8:191-8.

Metz-Boutigue, M.H., Jolles, J., Mazurier, J., Schoentgen, F., Legrand, D., Spik, G.,

Montreuil, J., et Jolles, P. 1984. Human lactotransferrin: amino acid sequence and structural comparisons with other transferrins. *Eur J Biochem*. 145:659-76.

Michel, I., Borck, H., McElligott, S., Krieg, C., et Diel, F. 2008. Histamine receptor H4R-selective ligands influence the STAT6 Transcription Activation Domain (TAD) and the DNA-binding. *Inflamm Res*. 57 Suppl 1:S47-8.

Miller, M.W., et Hanover, J.A. 1994. Functional nuclear pores reconstituted with beta 1-4 galactose-modified O-linked N-acetylglucosamine glycoproteins. *J Biol Chem*. 269:9289-97.

Miura, T., Klaus, W., Gsell, B., Miyamoto, C., et Senn, H. 1999. Characterization of the binding interface between ubiquitin and class I human ubiquitin-conjugating enzyme 2b by multidimensional heteronuclear NMR spectroscopy in solution. *J Mol Biol*. 290:213-28.

Mizanur, R.M., Jaipuri, F.A., et Pohl, N.L. 2005. One-step synthesis of labeled sugar nucleotides for protein O-GlcNAc modification studies by chemical function analysis of an archaeal protein. *J Am Chem Soc*. 127:836-7.

Moore, S.A., Anderson, B.F., Groom, C.R., Haridas, M., et Baker, E.N. 1997. Three-dimensional structure of diferric bovine lactoferrin at 2.8 A resolution. *J Mol Biol*. 274:222-36.

Musicki, B., Kramer, M.F., Becker, R.E., et Burnett, A.L. 2005. Inactivation of phosphorylated endothelial nitric oxide synthase (Ser-1177) by O-GlcNAc in diabetes-associated erectile dysfunction. *Proc Natl Acad Sci U S A*. 102:11870-5.

Myers, A., Holmans, P., Marshall, H., Kwon, J., Meyer, D., Ramic, D., Shears, S., Booth, J., DeVrieze, F.W., Crook, R., Hamshere, M., Abraham, R., Tunstall, N., Rice, F., Carty, S., Lillystone, S., Kehoe, P., Rudrasingham, V., Jones, L., Lovestone, S., Perez-Tur, J., Williams, J., Owen, M.J., Hardy, J., et Goate, A.M. 2000. Susceptibility locus for Alzheimer's disease on chromosome 10. *Science*. 290:2304-5.

~ N ~

Nagata, S. 1997. Apoptosis by death factor. *Cell*. 88:355-65.

Nakayama, K.I., et Nakayama, K. 2005. Regulation of the cell cycle by SCF-type ubiquitin ligases. *Semin Cell Dev Biol*. 16:323-33.

Nanashima, N., Asano, J., Hayakari, M., Nakamura, T., Nakano, H., Yamada, T., Shimizu, T., Akita, M., Fan, Y., et Tsuchida, S. 2005. Nuclear localization of STAT5A modified with O-linked N-acetylglucosamine and early involution in the mammary gland of Hirosaki hairless rat. *J Biol Chem*. 280:43010-6.

Nguyen, M., Millar, D.G., Yong, V.W., Korsmeyer, S.J., et Shore, G.C. 1993. Targeting of Bcl-2 to the mitochondrial outer membrane by a COOH-terminal signal anchor sequence. *J Biol Chem*. 268:25265-8.

Noach, N., Segev, Y., Levi, I., Segal, S., et Priel, E. 2007. Modification of topoisomerase I activity by glucose and by O-GlcNAcylation of the enzyme protein. *Glycobiology*. 17:1357-64.

Nolte, D., et Muller, U. 2002. Human O-GlcNAc transferase (OGT): genomic structure, analysis of splice variants, fine mapping in Xq13.1. *Mamm Genome*. 13:62-4.

Nolte, D., Niemann, S., et Muller, U. 2003. Specific sequence changes in multiple transcript system DYT3 are associated with X-linked dystonia parkinsonism. *Proc Natl Acad Sci U S A*. 100:10347-52.

- O -

Oehm, A., Behrmann, I., Falk, W., Pawlita, M., Maier, G., Klas, C., Li-Weber, M., Richards, S., Dhein, J., Trauth, B.C., et et al. 1992. Purification and molecular cloning of the APO-1 cell surface antigen, a member of the tumor necrosis factor/nerve growth factor receptor superfamily. Sequence identity with the Fas antigen. *J Biol Chem*. 267:10709-15.

Oh, S.M., Pyo, C.W., Kim, Y., et Choi, S.Y. 2004. Neutrophil lactoferrin upregulates the human p53 gene through induction of NF-kappaB activation cascade. *Oncogene*. 23:8282-91.

Oltvai, Z.N., et Korsmeyer, S.J. 1994. Checkpoints of dueling dimers foil death wishes. *Cell*. 79:189-92.

Oltvai, Z.N., Milliman, C.L., et Korsmeyer, S.J. 1993. Bcl-2 heterodimerizes in vivo with a conserved homolog, Bax, that accelerates programmed cell death. *Cell*. 74:609-19.

Ozoren, N., et El-Deiry, W.S. 2003. Cell surface Death Receptor signaling in normal and cancer cells. *Semin Cancer Biol*. 13:135-47.

- P -

Pan, G., Bauer, J.H., Haridas, V., Wang, S., Liu, D., Yu, G., Vincenz, C., Aggarwal, B.B., Ni, J., et Dixit, V.M. 1998. Identification and functional characterization of DR6, a novel death domain-containing TNF receptor. *FEBS Lett*. 431:351-6.

Panella, T.J., Liu, Y.H., Huang, A.T., et Teng, C.T. 1991. Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes, leukemic cells, and breast cancer. *Cancer Res*. 51:3037-43.

Park, M.K., D'Onofrio, M., Willingham, M.C., et Hanover, J.A. 1987. A monoclonal antibody against a family of nuclear pore proteins (nucleoporins): O-linked N-acetylglucosamine is part of the immunodeterminant. *Proc Natl Acad Sci U S A*. 84:6462-6.

Park, S.Y., Ryu, J., et Lee, W. 2005. O-GlcNAc modification on IRS-1 and Akt2 by PUGNAc inhibits their phosphorylation and induces insulin resistance in rat primary adipocytes. *Exp Mol Med*. 37:220-9.

Parker, G., Taylor, R., Jones, D., et McClain, D. 2004. Hyperglycemia and inhibition of glycogen synthase in streptozotocin-treated mice: role of O-linked N-acetylglucosamine. *J Biol Chem*. 279:20636-42.

Parker, G.J., Lund, K.C., Taylor, R.P., et McClain, D.A. 2003. Insulin resistance of glycogen synthase mediated by o-linked N-acetylglucosamine. *J Biol Chem*. 278:10022-7.

Pelzer, C., Kassner, I., Matentzoglu, K., Singh, R.K., Wollscheid, H.P., Scheffner, M., Schmidtke, G., et Groettrup, M. 2007. UBE1L2, a novel E1 enzyme specific for ubiquitin. *J Biol Chem*. 282:23010-4.

Penco, S., Caligo, M.A., Cipollini, G., Bevilacqua, G., et Garre, C. 1999. Lactoferrin expression in human breast cancer. *Cancer Biochem Biophys*. 17:163-78.

Penco, S., Pastorino, S., Bianchi-Scarra, G., et Garre, C. 1995. Lactoferrin down-modulates the activity of the granulocyte macrophage colony-stimulating factor promoter in interleukin-1 beta-stimulated cells. *J Biol Chem*. 270:12263-8.

Penco, S., Scarfi, S., Giovine, M., Damonte, G., Millo, E., Villaggio, B., Passalacqua, M., Pozzolini, M., Garre, C., et Benatti, U. 2001. Identification of an import signal for, and the nuclear localization of, human lactoferrin. *Biotechnol Appl Biochem*. 34:151-9.

Peng, Z. 2005. Current status of gendicine in China: recombinant human Ad-p53 agent for treatment of cancers. *Hum Gene Ther*. 16:1016-27.

Pentecost, B.T., et Teng, C.T. 1987. Lactotransferrin is the major estrogen inducible protein of mouse uterine secretions. *J Biol Chem*. 262:10134-9.

Petersen, B.O., Wagener, C., Marinoni, F., Kramer, E.R., Melixetian, M., Lazzerini Denchi, E., Gieffers, C., Matteucci, C., Peters, J.M., et Helin, K. 2000. Cell cycle- and cell growth-regulated proteolysis of mammalian CDC6 is dependent on APC-CDH1. *Genes Dev*. 14:2330-43.

Pfleger, C.M., et Kirschner, M.W. 2000. The KEN box: an APC recognition signal distinct from the D box targeted by Cdh1. *Genes Dev*. 14:655-65.

Pickart, C.M. 2001. Mechanisms underlying ubiquitination. *Annu Rev Biochem*. 70:503-33.

Pierce, A., Colavizza, D., Benaissa, M., Maes, P., Tartar, A., Montreuil, J., et Spik, G. 1991. Molecular cloning and sequence analysis of bovine lactotransferrin. *Eur J Biochem*. 196:177-84.

Pierce, A., Legrand, D., et Mazurier, J. 2009. [Lactoferrin: a multifunctional protein]. *Med Sci (Paris)*. 25:361-9.

Piva, R., Liu, J., Chiarle, R., Podda, A., Pagano, M., et Inghirami, G. 2002. In vivo interference with Skp1 function leads to genetic instability and neoplastic transformation. *Mol Cell Biol*. 22:8375-87.

Powell, M.J., et Ogden, J.E. 1990. Nucleotide sequence of human lactoferrin cDNA. *Nucleic Acids Res*. 18:4013.

~ R ~

Rado, T.A., Bollekens, J., St Laurent, G., Parker, L., et Benz, E.J., Jr. 1984. Lactoferrin biosynthesis during granulocytopoiesis. *Blood*. 64:1103-9.

Rado, T.A., Wei, X.P., et Benz, E.J., Jr. 1987. Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA during normal and leukemic myelopoiesis. *Blood*. 70:989-93.

Rastinejad, F. 2001. Retinoid X receptor and its partners in the nuclear receptor family. *Curr Opin Struct Biol*. 11:33-8.

Rechsteiner, M., et Rogers, S.W. 1996. PEST sequences and regulation by proteolysis. *Trends Biochem Sci*. 21:267-71.

Rey, M.W., Woloshuk, S.L., deBoer, H.A., et Pieper, F.R. 1990. Complete nucleotide sequence of human mammary gland lactoferrin. *Nucleic Acids Res*. 18:5288.

Robbins, J., Dilworth, S.M., Laskey, R.A., et Dingwall, C. 1991. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell*. 64:615-23.

Robertson, L.A., Moya, K.L., et Breen, K.C. 2004. The potential role of tau protein O-glycosylation in Alzheimer's disease. *J Alzheimers Dis*. 6:489-95.

Robertson, M., Swain, S.M., Chandler, P.M., et Olszewski, N.E. 1998. Identification of a negative regulator of gibberellin action, HvSPY, in barley. *Plant Cell*. 10:995-1007.

Rodrigo-Brenni, M.C., Thomas, S., Bouck, D.C., et Kaplan, K.B. 2004. Sgt1p and Skp1p modulate the assembly and turnover of CBF3 complexes required for proper kinetochore function. *Mol Biol Cell*. 15:3366-78.

Rodriguez, I.R., et Whelan, W.J. 1985. A novel glycosyl-amino acid linkage: rabbit-muscle glycogen is covalently linked to a protein via tyrosine. *Biochem Biophys Res Commun*. 132:829-36.

Rogers, S., Wells, R., et Rechsteiner, M. 1986. Amino acid sequences common to rapidly

degraded proteins: the PEST hypothesis. *Science*. 234:364-8.

Roos, M.D., et Hanover, J.A. 2000. Structure of O-linked GlcNAc transferase: mediator of glycan-dependent signaling. *Biochem Biophys Res Commun*. 271:275-80.

Roquemore, E.P., Chou, T.Y., et Hart, G.W. 1994. Detection of O-linked N-acetylglucosamine (O-GlcNAc) on cytoplasmic and nuclear proteins. *Methods Enzymol*. 230:443-60.

~ S ~

Saito, K., et Nakanuma, Y. 1992. Lactoferrin and lysozyme in the intrahepatic bile duct of normal livers and hepatolithiasis. An immunohistochemical study. *J Hepatol*. 15:147-53.

Saito, N., Takemori, N., Hirai, K., Onodera, R., Watanabe, S., et Namiki, M. 1993. Ultrastructural localization of lactoferrin in the granules other than typical secondary granules of human neutrophils. *Hum Cell*. 6:42-8.

Schall, T.J., Lewis, M., Koller, K.J., Lee, A., Rice, G.C., Wong, G.H., Gatanaga, T., Granger, G.A., Lentz, R., Raab, H., et et al. 1990. Molecular cloning and expression of a receptor for human tumor necrosis factor. *Cell*. 61:361-70.

Schirm, M., Kalmokoff, M., Aubry, A., Thibault, P., Sandoz, M., et Logan, S.M. 2004. Flagellin from Listeria monocytogenes is glycosylated with beta-O-linked N-acetylglucosamine. *J Bacteriol*. 186:6721-7.

Schnell, J.D., et Hicke, L. 2003. Non-traditional functions of ubiquitin and ubiquitin-binding proteins. *J Biol Chem*. 278:35857-60.

Schultz, J., et Pils, B. 2002. Prediction of structure and functional residues for O-GlcNAcase, a divergent homologue of acetyltransferases. *FEBS Lett*. 529:179-82.

Sean, P., Nguyen, J.H., et Semler, B.L. 2009. Altered interactions between stem-loop IV within the 5' noncoding region of coxsackievirus RNA and poly(rC) binding protein 2: effects on IRES-mediated translation and viral infectivity. *Virology*. 389:45-58.

Sedlak, T.W., Oltvai, Z.N., Yang, E., Wang, K., Boise, L.H., Thompson, C.B., et Korsmeyer, S.J. 1995. Multiple Bcl-2 family members demonstrate selective dimerizations with Bax. *Proc Natl Acad Sci U S A*. 92:7834-8.

Segurado, M., et Tercero, J.A. 2009. The S-phase checkpoint: targeting the replication fork. *Biol Cell*. 101:617-27.

Sekine, K., Ushida, Y., Kuhara, T., Iigo, M., Baba-Toriyama, H., Moore, M.A., Murakoshi, M., Satomi, Y., Nishino, H., Kakizoe, T., et Tsuda, H. 1997a. Inhibition of initiation and early stage development of aberrant crypt foci and enhanced natural killer activity in male rats administered bovine lactoferrin concomitantly with azoxymethane. *Cancer Lett*. 121:211-6.

Sekine, K., Watanabe, E., Nakamura, J., Takasuka, N., Kim, D.J., Asamoto, M., Krutovskikh, V., Baba-Toriyama, H., Ota, T., Moore, M.A., Masuda, M., Sugimoto, H., Nishino, H., Kakizoe, T., et Tsuda, H. 1997b. Inhibition of azoxymethane-initiated colon tumor by bovine lactoferrin administration in F344 rats. *Jpn J Cancer Res*. 88:523-6.

Semenov, D.V., Kanyshkova, T.G., Buneva, V.N., et Nevinsky, G.A. 1999. Human milk lactoferrin binds ATP and dissociates into monomers. *Biochem Mol Biol Int*. 47:177-84.

Seo, J., et Lee, K.J. 2004. Post-translational modifications and their biological functions: proteomic analysis and systematic approaches. *J Biochem Mol Biol*. 37:35-44.

Shafi, R., Iyer, S.P., Ellies, L.G., O'Donnell, N., Marek, K.W., Chui, D., Hart, G.W., et Marth, J.D. 2000. The O-GlcNAc transferase gene resides on the X chromosome and is

essential for embryonic stem cell viability and mouse ontogeny. *Proc Natl Acad Sci U S A*. 97:5735-9.

Siebert, P.D., et Huang, B.C. 1997. Identification of an alternative form of human lactoferrin mRNA that is expressed differentially in normal tissues and tumor-derived cell lines. *Proc Natl Acad Sci U S A*. 94:2198-203.

Slawson, C., Pidala, J., et Potter, R. 2001. Increased N-acetyl-beta-glucosaminidase activity in primary breast carcinomas corresponds to a decrease in N-acetylglucosamine containing proteins. *Biochim Biophys Acta*. 1537:147-57.

Slawson, C., Shafii, S., Amburgey, J., et Potter, R. 2002. Characterization of the O-GlcNAc protein modification in Xenopus laevis oocyte during oogenesis and progesterone-stimulated maturation. *Biochim Biophys Acta*. 1573:121-9.

Slawson, C., Zachara, N.E., Vosseller, K., Cheung, W.D., Lane, M.D., et Hart, G.W. 2005. Perturbations in O-linked beta-N-acetylglucosamine protein modification cause severe defects in mitotic progression and cytokinesis. *J Biol Chem*. 280:32944-56.

Smith, T.F., Gaitatzes, C., Saxena, K., et Neer, E.J. 1999. The WD repeat: a common architecture for diverse functions. *Trends Biochem Sci*. 24:181-5.

Soderberg, M., Raffalli-Mathieu, F., et Lang, M.A. 2007. Regulation of the murine inducible nitric oxide synthase gene by dexamethasone involves a heterogeneous nuclear ribonucleoprotein I (hnRNPI) dependent pathway. *Mol Immunol*. 44:3204-10.

Soesanto, Y.A., Luo, B., Jones, D., Taylor, R., Gabrielsen, J.S., Parker, G., et McClain, D.A. 2008. Regulation of Akt signaling by O-GlcNAc in euglycemia. *Am J Physiol Endocrinol Metab*. 295:E974-80.

Son, H.J., Lee, S.H., et Choi, S.Y. 2006. Human lactoferrin controls the level of retinoblastoma protein and its activity. *Biochem Cell Biol*. 84:345-50.

Son, K.N., Park, J., Chung, C.K., Chung, D.K., Yu, D.Y., Lee, K.K., et Kim, J. 2002. Human lactoferrin activates transcription of IL-1beta gene in mammalian cells. *Biochem Biophys Res Commun*. 290:236-41.

Song, M., Kim, H.S., Park, J.M., Kim, S.H., Kim, I.H., Ryu, S.H., et Suh, P.G. 2008. o-GlcNAc transferase is activated by CaMKIV-dependent phosphorylation under potassium chloride-induced depolarization in NG-108-15 cells. *Cell Signal*. 20:94-104.

Spik, G., Coddeville, B., Mazurier, J., Bourne, Y., Cambillaut, C., et Montreuil, J. 1994. Primary and three-dimensional structure of lactotransferrin (lactoferrin) glycans. *Adv Exp Med Biol*. 357:21-32.

Spik, G., Coddeville, B., et Montreuil, J. 1988. Comparative study of the primary structures of sero-, lacto- and ovotransferrin glycans from different species. *Biochimie*. 70:1459-69.

Spik, G., Strecker, G., Fournet, B., Bouquelet, S., Montreuil, J., Dorland, L., van Halbeek, H., et Vliegenthart, J.F. 1982. Primary structure of the glycans from human lactotransferrin. *Eur J Biochem*. 121:413-9.

Stephen, A.G., Trausch-Azar, J.S., Ciechanover, A., et Schwartz, A.L. 1996. The ubiquitin-activating enzyme E1 is phosphorylated and localized to the nucleus in a cell cycle-dependent manner. *J Biol Chem*. 271:15608-14.

Strasser, A., O'Connor, L., et Dixit, V.M. 2000. Apoptosis signaling. *Annu Rev Biochem*. 69:217-45.

Swain, S.M., Tseng, T.S., et Olszewski, N.E. 2001. Altered expression of SPINDLY affects gibberellin response and plant development. *Plant Physiol*. 126:1174-85.

Sykes, J.A., Thomas, M.J., Goldie, D.J., et Turner, G.M. 1982. Plasma lactoferrin levels in pregnancy and cystic fibrosis. *Clin Chim Acta*. 122:385-93.

- T -

Tanahashi, N., Tsurumi, C., Tamura, T., et Tanaka, K. 1993. Molecular structure of 20S and 26S proteasomes. *Enzyme Protein*. 47:241-51.

Tartaglia, L.A., Ayres, T.M., Wong, G.H., et Goeddel, D.V. 1993. A novel domain within the 55 kd TNF receptor signals cell death. *Cell*. 74:845-53.

Teng, C., Gladwell, W., Raphiou, I., et Liu, E. 2004. Methylation and expression of the lactoferrin gene in human tissues and cancer cells. *Biometals*. 17:317-23.

Teng, C.T. 2002. Lactoferrin gene expression and regulation: an overview. *Biochem Cell Biol*. 80:7-16.

Teng, C.T., Liu, Y., Yang, N., Walmer, D., et Panella, T. 1992. Differential molecular mechanism of the estrogen action that regulates lactoferrin gene in human and mouse. *Mol Endocrinol*. 6:1969-81.

Testoni, N., Borsaru, G., Martinelli, G., Carboni, C., Ruggeri, D., Ottaviani, E., Pelliconi, S., Ricci, P., Pastano, R., Visani, G., Zaccaria, A., et Tura, S. 1999. 3q21 and 3q26 cytogenetic abnormalities in acute myeloblastic leukemia: biological and clinical features. *Haematologica*. 84:690-4.

Thorburn, A., Thorburn, J., et Frankel, A.E. 2004. Induction of apoptosis by tumor cell-targeted toxins. *Apoptosis*. 9:19-25.

Toleman, C., Paterson, A.J., et Kudlow, J.E. 2006. Location and characterization of the O-GlcNAcase active site. *Biochim Biophys Acta*. 1760:829-39.

Toleman, C., Paterson, A.J., Whisenhunt, T.R., et Kudlow, J.E. 2004. Characterization of the histone acetyltransferase (HAT) domain of a bifunctional protein with activable O-GlcNAcase and HAT activities. *J Biol Chem*. 279:53665-73.

Torres, C.R., et Hart, G.W. 1984. Topography and polypeptide distribution of terminal N-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for O-linked GlcNAc. *J Biol Chem*. 259:3308-17.

Tsuda, H., Sekine, K., Ushida, Y., Kuhara, T., Takasuka, N., Iigo, M., Han, B.S., et Moore, M.A. 2000. Milk and dairy products in cancer prevention: focus on bovine lactoferrin. *Mutat Res*. 462:227-33.

- V -

van Berkel, P.H., Geerts, M.E., van Veen, H.A., Mericskay, M., de Boer, H.A., et Nuijens, J.H. 1997. N-terminal stretch Arg2, Arg3, Arg4 and Arg5 of human lactoferrin is essential for binding to heparin, bacterial lipopolysaccharide, human lysozyme and DNA. *Biochem J*. 328 (Pt 1):145-51.

van Berkel, P.H., van Veen, H.A., Geerts, M.E., de Boer, H.A., et Nuijens, J.H. 1996. Heterogeneity in utilization of N-glycosylation sites Asn624 and Asn138 in human lactoferrin: a study with glycosylation-site mutants. *Biochem J*. 319 (Pt 1):117-22.

van Dijk, E., Le Hir, H., et Seraphin, B. 2003. DcpS can act in the 5'-3' mRNA decay pathway in addition to the 3'-5' pathway. *Proc Natl Acad Sci U S A*. 100:12081-6.

van Leuken, R., Clijsters, L., et Wolthuis, R. 2008. To cell cycle, swing the APC/C. *Biochim Biophys Acta*. 1786:49-59.

Varecha, M., Zimmermann, M., Amrichova, J., Ulman, V., Matula, P., et Kozubek, M. 2009. Prediction of localization and interactions of apoptotic proteins. *J Biomed Sci*. 16:59.

Vaux, D.L., et Korsmeyer, S.J. 1999. Cell death in development. *Cell*. 96:245-54.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O.,

Yandell, M., Evans, C.A., Holt, R.A., Gocayne, J.D., Amanatides, P., Ballew, R.M., Huson, D.H., Wortman, J.R., Zhang, Q., Kodira, C.D., Zheng, X.H., Chen, L., Skupski, M., Subramanian, G., Thomas, P.D., Zhang, J., Gabor Miklos, G.L., Nelson, C., Broder, S., Clark, A.G., Nadeau, J., McKusick, V.A., Zinder, N., Levine, A.J., Roberts, R.J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A.E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T.J., Higgins, M.E., Ji, R.R., Ke, Z., Ketchum, K.A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G.V., Milshina, N., Moore, H.M., Naik, A.K., Narayan, V.A., Neelam, B., Nusskern, D., Rusch, D.B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., *et al.* 2001. The sequence of the human genome. *Science*. 291:1304-51.

Vijay-Kumar, S., Bugg, C.E., et Cook, W.J. 1987. Structure of ubiquitin refined at 1.8 A resolution. *J Mol Biol*. 194:531-44.

Vosseller, K., Hansen, K.C., Chalkley, R.J., Trinidad, J.C., Wells, L., Hart, G.W., et Burlingame, A.L. 2005. Quantitative analysis of both protein expression and serine / threonine post-translational modifications through stable isotope labeling with dithiothreitol. *Proteomics*. 5:388-98.

Vosseller, K., Wells, L., Lane, M.D., et Hart, G.W. 2002. Elevated nucleocytoplasmic glycosylation by O-GlcNAc results in insulin resistance associated with defects in Akt activation in 3T3-L1 adipocytes. *Proc Natl Acad Sci U S A*. 99:5313-8.

Vulliet, P.R., Hall, F.L., Mitchell, J.P., et Hardie, D.G. 1989. Identification of a novel proline-directed serine/threonine protein kinase in rat pheochromocytoma. *J Biol Chem*. 264:16292-8.

- W -

Walsh, C.T., Garneau-Tsodikova, S., et Gatto, G.J., Jr. 2005. Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew Chem Int Ed Engl*. 44:7342-72.

Webster, D.M., Teo, C.F., Sun, Y., Wloga, D., Gay, S., Klonowski, K.D., Wells, L., et Dougan, S.T. 2009. O-GlcNAc modifications regulate cell survival and epiboly during zebrafish development. *BMC Dev Biol*. 9:28.

Weinberg, R.L., Veprintsev, D.B., Bycroft, M., et Fersht, A.R. 2005. Comparative binding of p53 to its promoter and DNA recognition elements. *J Mol Biol*. 348:589-96.

Welchman, R.L., Gordon, C., et Mayer, R.J. 2005. Ubiquitin and ubiquitin-like proteins as multifunctional signals. *Nat Rev Mol Cell Biol*. 6:599-609.

Wells, L., Gao, Y., Mahoney, J.A., Vosseller, K., Chen, C., Rosen, A., et Hart, G.W. 2002. Dynamic O-glycosylation of nuclear and cytosolic proteins: further characterization of the nucleocytoplasmic beta-N-acetylglucosaminidase, O-GlcNAcase. *J Biol Chem*. 277:1755-61.

Wells, L., Kreppel, L.K., Comer, F.I., Wadzinski, B.E., et Hart, G.W. 2004. O-GlcNAc transferase is in a functional complex with protein phosphatase 1 catalytic subunits. *J Biol Chem*. 279:38466-70.

Whelan, S.A., Dias, W.B., Lakshmanan, T., Lane, M.D., et Hart, G.W. 2009. Regulation of Insulin Receptor 1 (IRS-1)/AKT Kinase Mediated Insulin Signaling by O-linked

{beta}-N-acetylglucosamine (O-GlcNAc) in 3T3-L1 Adipocytes. *J Biol Chem*.

Wirbelauer, C., Sutterluty, H., Blondel, M., Gstaiger, M., Peter, M., Reymond, F., et Krek, W. 2000. The F-box protein Skp2 is a ubiquitylation target of a Cul1-based core ubiquitin ligase complex: evidence for a role of Cul1 in the suppression of Skp2 expression in quiescent fibroblasts. *EMBO J*. 19:5362-75.

Worsham, M.J., Pals, G., Schouten, J.P., Miller, F., Tiwari, N., van Spaendonk, R., et Wolman, S.R. 2006. High-resolution mapping of molecular events associated with immortalization, transformation, and progression to breast cancer in the MCF10 model. *Breast Cancer Res Treat*. 96:177-86.

Wrabl, J.O., et Grishin, N.V. 2001. Homology between O-linked GlcNAc transferases and proteins of the glycogen phosphorylase superfamily. *J Mol Biol*. 314:365-74.

~ X ~

Xiao, Y., Monitto, C.L., Minhas, K.M., et Sidransky, D. 2004. Lactoferrin down-regulates G1 cyclin-dependent kinases during growth arrest of head and neck cancer cells. *Clin Cancer Res*. 10:8683-6.

~ Y ~

Yan, Y., et Mumby, M.C. 1999. Distinct roles for PP1 and PP2A in phosphorylation of the retinoblastoma protein. PP2a regulates the activities of G(1) cyclin-dependent kinases. *J Biol Chem*. 274:31917-24.

Yang, J., Liu, X., Bhalla, K., Kim, C.N., Ibrado, A.M., Cai, J., Peng, T.I., Jones, D.P., et Wang, X. 1997. Prevention of apoptosis by Bcl-2: release of cytochrome c from mitochondria blocked. *Science*. 275:1129-32.

Yang, W.H., Kim, J.E., Nam, H.W., Ju, J.W., Kim, H.S., Kim, Y.S., et Cho, J.W. 2006. Modification of p53 with O-linked N-acetylglucosamine regulates p53 activity and stability. *Nat Cell Biol*. 8:1074-83.

Yang, X., Ongusaha, P.P., Miles, P.D., Havstad, J.C., Zhang, F., So, W.V., Kudlow, J.E., Michell, R.H., Olefsky, J.M., Field, S.J., et Evans, R.M. 2008. Phosphoinositide signalling links O-GlcNAc transferase to insulin resistance. *Nature*. 451:964-9.

Yang, X., Zhang, F., et Kudlow, J.E. 2002. Recruitment of O-GlcNAc transferase to promoters by corepressor mSin3A: coupling protein O-GlcNAcylation to transcriptional repression. *Cell*. 110:69-80.

Yang, Y., Li, C.C., et Weissman, A.M. 2004. Regulating the p53 system through ubiquitination. *Oncogene*. 23:2096-106.

Yao, P.J., et Coleman, P.D. 1998a. Reduced O-glycosylated clathrin assembly protein AP180: implication for synaptic vesicle recycling dysfunction in Alzheimer's disease. *Neurosci Lett*. 252:33-6.

Yao, P.J., et Coleman, P.D. 1998b. Reduction of O-linked N-acetylglucosamine-modified assembly protein-3 in Alzheimer's disease. *J Neurosci*. 18:2399-411.

Yew, P.R. 2001. Ubiquitin-mediated proteolysis of vertebrate G1- and S-phase regulators. *J Cell Physiol*. 187:1-10.

Yoo, Y.C., Watanabe, S., Watanabe, R., Hata, K., Shimazaki, K., et Azuma, I. 1997. Bovine lactoferrin and lactoferricin, a peptide derived from bovine lactoferrin, inhibit tumor metastasis in mice. *Jpn J Cancer Res*. 88:184-90.

Yoo, Y.C., Watanabe, S., Watanabe, R., Hata, K., Shimazaki, K., et Azuma, I. 1998. Bovine

lactoferrin and Lactoferricin inhibit tumor metastasis in mice. *Adv Exp Med Biol*. 443:285-91.

Yu, C.H., Si, T., Wu, W.H., Hu, J., Du, J.T., Zhao, Y.F., et Li, Y.M. 2008. O-GlcNAcylation modulates the self-aggregation ability of the fourth microtubule-binding repeat of tau. *Biochem Biophys Res Commun*. 375:59-62.

~ Z ~

Zachariae, W., et Nasmyth, K. 1999. Whose end is destruction: cell division and the anaphase-promoting complex. *Genes Dev*. 13:2039-58.

Zamzami, N., Susin, S.A., Marchetti, P., Hirsch, T., Gomez-Monterrey, I., Castedo, M., et Kroemer, G. 1996. Mitochondrial control of nuclear apoptosis. *J Exp Med*. 183:1533-44.

Zebol, J.R., Hewitt, N.M., Moretti, P.A., Lynn, H.E., Lake, J.A., Li, P., Vadas, M.A., Wattenberg, B.W., et Pitson, S.M. 2009. The CCT/TRiC chaperonin is required for maturation of sphingosine kinase 1. *Int J Biochem Cell Biol*. 41:822-7.

Zhang, F., Su, K., Yang, X., Bowe, D.B., Paterson, A.J., et Kudlow, J.E. 2003. O-GlcNAc modification is an endogenous inhibitor of the proteasome. *Cell*. 115:715-25.

Zhu, H., Zhou, Z.M., Huo, R., Huang, X.Y., Lu, L., Lin, M., Wang, L.R., Zhou, Y.D., Li, J.M., et Sha, J.H. 2004. Identification and characteristics of a novel E1 like gene nUBE1L in human testis. *Acta Biochim Biophys Sin (Shanghai)*. 36:227-34.

Résumé

La GlcNAcylation est un glycosylation atypique qui met en jeu un seul résidu de N-acétylglucosamine. Elle est retrouvée dans les compartiments cytoplasmiques et nucléaires où elle cible des protéines tels que les facteurs de transcription, les oncoprotéines et les protéines du cytosquelette. C'est une modification post-traductionnelle dynamique contrôlée par deux enzymes : l'OGT transfert le monosaccharide de l'UDP-GlcNAc sur un résidu de Ser ou de Thr alors que l'OGA l'enlève. Les taux de GlcNAcylation des protéines sont corrélés aux concentrations en UDP-GlcNAc, produit final de la voie de biosynthèse des hexosamines, et varient en fonction de facteurs environnementaux et nutritionnels. Des modifications des profils de GlcNAcylation sont rapportées dans les maladies neurodégénératives, le diabète et le cancer.

La delta-lactoferrine est un suppresseur de tumeur potentiel. Elle est sous-exprimée dans le cancer du sein et son expression conduit à l'arrêt du cycle cellulaire. Nous avons pu montrer que la delta-lactoferrine induit une surexpression de différents gènes impliqués dans la régulation des acteurs moléculaires du cycle cellulaire comme Skp1. Nous avons démontré que la delta-lactoferrine est un facteur de transcription reconnaissant un ΔLfRE retrouvé fonctionnel dans le promoteur de ses gènes cibles. Nous avons ensuite réalisé une étude protéomique et identifié huit protéines qui sont surexprimées par la delta-lactoferrine. Les protéines identifiées sont essentiellement impliquées dans les processus de dégradation et le contrôle de la qualité des ARNm. L'analyse de leur promoteur respectif a permis de mettre en évidence un ΔLfRE dans le promoteur de *DcpS*. Puisque la delta-lactoferrine conduit à la mort cellulaire, nous avons vérifié qu'elle induit l'apoptose par l'observation de la fragmentation de l'ADN. Nous avons ensuite recherché quelles molécules, impliquées dans ce processus, sont transactivées par la delta-lactoferrine et nous avons identifié un ΔLfRE fonctionnel dans le promoteur du gène *Bax*. Nous avons également démontré que la delta-lactoferrine surexprime la protéine Bax et déclenche l'apoptose par la voie mitochondriale.

En tant que facteur de transcription et suppresseur de tumeur, la delta-lactoferrine doit être hautement régulée. Nous avons pu montrer que son activité transcriptionnelle est régulée par la GlcNAcylation et la phosphorylation. Nous montrons grâce à l'utilisation de mutants de glycosylation, d'inhibiteurs et d'activateurs de la voie des hexosamines que la GlcNAcylation régule négativement l'activité transcriptionnelle de la delta-lactoferrine alors que la phosphorylation a l'effet inverse. La mutation des quatre sites de GlcNAcylation conduit à une isoforme de la delta-lactoferrine constitutivement activée possédant une activité pro-apoptotique accrue comparée à l'isoforme sauvage. L'utilisation de mutants de glycosylation, pour lesquels un site unique est préservé, nous a permis de démontrer que la GlcNAcylation du site Ser 10 est crucial dans la régulation de l'activité de la delta-lactoferrine. Nous avons également démontré que la GlcNAcylation de ce résidu inhibe la polyubiquitinylation et augmente la demi-vie de la delta-lactoferrine. Les modifications post-traductionnelles du site Ser 10 contrôlent donc la stabilité de la delta-lactoferrine. Nous avons également défini une séquence PEST fonctionnelle en C-terminal ainsi que le résidu de lysine cible de l'ubiquitinylation. De plus, par des expériences de Re-ChIP, nous montrons que le complexe transcriptionnel de la delta-lactoferrine interagit avec le ΔLfRE sous une forme phosphorylée et ubiquitinylée suggérant que l'activité transcriptionnelle de la delta-lactoferrine et sa dégradation sont concomitantes.

Abstract

The GlcNAcylation is an unconventional glycosylation since it involves a single N-acetylglucosamine residue. It is mainly present into the cytoplasmic and the nuclear compartments and proteins such as targets transcription factors, oncoproteins and cytoskeleton proteins. It is a dynamic posttranslational modification controlled by two enzymes: OGT which transfers the single monosaccharide from UDP-GlcNAc to Ser or Thr residues and OGA which removes it. The GlcNAcylation rate of proteins is tightly correlated to the concentration of UDP-GlcNAc, the end product of the hexosamine biosynthetic pathway, which vary according to environmental and nutritional factors. Alterations of the GlcNAcylation profile appear in neurodegenerative disorders, diabetes and cancer.

Delta-lactoferrin is a potential tumor suppressor gene. It is down-regulated in breast cancer and its expression leads to cell cycle arrest. We first showed that delta-lactoferrin leads to the up-regulation of different genes involved in the control of cell cycle actors such as Skp1. We demonstrated that delta-lactoferrin is a transcription factor which binds a functional ΔLfRE found in its target gene promoter region. We next performed proteomic studies and eight proteins were identified by mass spectrometry the transcription of which was up-regulated by delta-lactoferrin. The identified proteins are mainly involved in processes such as mRNA turnover and quality control. Analyses of their respective promoter led to the detection of a functional ΔLfRE in the promoter of DcpS. Since cell death followed delta-lactoferrin overexpression, we verified that apoptosis was involved by observing DNA fragmentation in delta-lactoferrin-expressing cells. We next investigated whether delta-lactoferrin may transactivate key genes implicated in this programmed cell death and identified a ΔLfRE in the *Bax* promoter region which was found functional. Since the transactivation of the *Bax* promoter is extremely efficient, we followed Bax expression and further demonstrated that the intrinsic mitochondrial pathway was involved.

As a transcription factor and a tumor suppressor, delta-lactoferrin should have its activity strongly controlled. Our investigations have shown that delta-lactoferrin transcriptional activity is controlled by GlcNAcylation and phosphorylation. We showed using glycosylation mutants, inhibitors and/or activators of the hexosamine biosynthesis pathway that GlcNAcylation negatively regulates delta-lactoferrin transcriptional activity while phosphorylation activates it. The directed mutagenesis of the four GlcNAcylation sites leads to a constitutively active delta-lactoferrin isoform with increased pro-apoptotic effects compared to wild type. Using a series of different glycosylation mutants in which only one glycosylation site was preserved we showed that Ser 10 is crucial and regulates delta-lactoferrin activity. We also demonstrated that GlcNAcylation inhibits delta-lactoferrin polyubiquitination and increases its half-live. Thus, these posttranslational modifications on Ser 10 residue controlled delta-lactoferrin turnover. We also mapped a functional PEST sequence at the C-terminus and the lysine which is the ubiquitin ligase target. Moreover, using Re-ChIP assays, we showed that delta-lactoferrin transcriptional complex binds to ΔLfRE as a phosphorylated and ubiquitinated isoform suggesting that delta-lactoferrin transcriptional activity and degradation are concomitant events.