

N° d'ordre : 40792

Université des Sciences et Technologies de Lille

Thèse de doctorat

présentée par

Esthelle HOEDT

en vue de l'obtention du titre de

Docteur de l'Université de Lille
Spécialité Sciences de la Vie et de la Santé

**Activité antitumorale des isoformes
de la lactoferrine, une approche
comparative et quantitative**

Soutenue le 15 mars 2012 devant la commission d'examen :

Président : Pr. Tony LEFEBVRE (Université Lille1)
Rapporteurs : Pr. Jérôme LEMOINE (Université de Lyon)
Pr. Piera VALENTI (Université de Rome, Italie)
Examinateurs : Dr. Bernard MONSARRAT (Directeur de recherche CNRS, Toulouse)
Dr. Christophe MARILLER (Université Lille1)
Pr. Annick PIERCE (Université Lille1)

À mes parents,

À ma famille

Remerciements

Je tiens à remercier l'ensemble des membres du jury pour avoir accepté de juger mon travail de thèse.

Monsieur le Professeur Jérôme Lemoine

qui a accepté de juger cette thèse en qualité de rapporteur. Je vous remercie pour l'intérêt que vous portez à ce travail. Je suis particulièrement honorée de votre présence dans ce jury de soutenance de thèse et je vous exprime ici ma plus haute considération.

Madame le Professeur Piera Valenti

pour m'avoir fait l'honneur d'être rapporteur de cette thèse. Je vous remercie pour l'intérêt et la considération que vous apportez à ces travaux. Je vous adresse mes sincères remerciements ainsi que ma profonde gratitude.

Monsieur le Docteur Bernard Monsarrat

pour m'avoir acceptée au sein de son laboratoire et m'avoir permis de réaliser des analyses de spectrométrie de masse dans les meilleures conditions. Vos remarques et suggestions m'ont permis d'améliorer la qualité de ce mémoire. Merci d'avoir accepté d'examiner mon manuscrit et de faire partie de mon jury de thèse.

Monsieur le Professeur Tony Lefebvre

que je remercie chaleureusement pour avoir accepté de présider le jury de cette soutenance de thèse. Après tout le temps que tu m'as consacré, les discussions animées et ta capacité à tester ma susceptibilité, c'est pour moi un réel plaisir que de te compter parmi les membres du jury.

Monsieur le Docteur Christophe Mariller

que je remercie sincèrement et chaleureusement pour m'avoir encadrée depuis mon stage de Master 1 et tout au long de ces (nombreuses) années de thèse. Je te remercie pour tes conseils, ta patience et tes encouragements. Merci également de m'avoir fait entrer dans le monde précieux et merveilleux des Macintosh et pour les nombreuses discussions scientifiques ou non que nous avons pu avoir. Un grand merci.

Madame le Professeur Annick Pierce

à qui j'exprime ma plus profonde gratitude, pour m'avoir accueillie (et nous savons à quel point cela a été long et fastidieux), soutenue, encouragée et donné les moyens de réaliser cette thèse dans un cadre de travail idéal. Merci pour la confiance que tu m'as accordée et pour toutes ces discussions ô combien passionnées et enrichissantes que nous avons eues. Ces années passées au sein de ton équipe ont été un réel plaisir tant sur le plan professionnel que humain. Merci pour tout.

Sommaire

AVANT PROPOS	1
GENERALITES	6
I. BIOSYNTHÈSE DE LA LACTOFERRINE	6
A. GENE DE LA LACTOFERRINE	6
B. REGULATION DE L'EXPRESSION DU GENE DE LA LACTOFERRINE HUMAINE	6
1) REGULATION PAR LES FACTEURS DE TRANSCRIPTION	7
2) REGULATION PAR DES <i>STIMULI</i> INFLAMMATOIRES	9
3) REGULATION PAR METHYLATION DU PROMOTEUR	9
4) DEGRADATION DES TRANSCRITS DU GENE DE LA LACTOFERRINE HUMAINE	9
II. STRUCTURE DES LACTOFERRINES	10
A. SEQUENCE PRIMAIRE DE LA LACTOFERRINE	10
B. STRUCTURE TRIDIMENSIONNELLE DE LA LACTOFERRINE	10
C. CARACTERISTIQUES DE SURFACE DE LA LACTOFERRINE	12
D. STRUCTURE DES GLYCANNES	13
III. LES DIFFERENTES FONCTIONS DES LACTOFERRINES	14
A. ACTIVITE ANTIMICROBIENNE.	15
1) ACTIVITE ANTIBACTERIENNE.	15
2) ACTIVITE ANTIVIRALE	17
3) ACTIVITES ANTIFONGIQUE ET ANTIPARASITAIRE	18
B. MODULATION DE LA REPONSE IMMUNE ET INFLAMMATOIRE	18
1) ACTIVITE ANTI-INFLAMMATOIRE	18
2) ACTIVITE PRO-INFLAMMATOIRE	20
3) ACTIVITE ANTI-OXYDANTE	20
C. ROLE DE LA LACTOFERRINE DANS LE CANCER ET LA PROLIFERATION CELLULAIRE	21
1) ACTIVITE ANTICANCEREUSE	21
2) ACTIVITE PROLIFERATIVE DE LA LACTOFERRINE	26
D. ACTIVITE DE FACTEUR DE TRANSCRIPTION	27
1) ACTIVITE PUTATIVE DE FACTEUR DE TRANSCRIPTION DE LA LACTOFERRINE SECRETEE	27
2) LA DELTA-LACTOFERRINE EST UN FACTEUR DE TRANSCRIPTION FONCTIONNEL	29
IV. ANALYSE PROTEOMIQUE, DEFINITION ET ORIGINES	33
A. DEFINITION	33

B. ORIGINES	33
V. L'ANALYSE PROTEOMIQUE	34
A. FRACTIONNEMENT DES PROTEINES ET DES PEPTIDES	35
1) L'ELECTROPHORESE BIDIMENSIONNELLE (2-D)	35
2) LA SEPARATION DES PEPTIDES PAR CHROMATOGRAPHIE LIQUIDE HAUTE PRESSION (CLHP)	36
VI. ANALYSE PROTEOMIQUE ET SPECTROMETRIE DE MASSE	37
A. LES SOURCES D'IONISATION EN SPECTROMETRIE DE MASSE	38
1) L'IONISATION PAR UN SPECTROMETRE DE MASSE A SOURCE MALDI	38
2) IONISATION PAR UN SPECTROMETRE DE MASSE A SOURCE ESI	39
B. LES ANALYSEURS EN SPECTROMETRIE DE MASSE (MS)	40
1) L'ANALYSEUR A TEMPS DE VOL (TOF)	41
2) LE QUADRIPOLE	42
3) LA TRAPPE IONIQUE 3-D (<i>ION TRAP</i>)	43
4) ANALYSEUR A RESONANCE IONIQUE CYCLOTRONIQUE PAR TRANSFORMEE DE FOURIER (FT-ICR)	44
5) TRAPPE ORBITALE	45
C. LA SPECTROMETRIE DE MASSE EN TANDEM (MS/MS)	46
1) PRINCIPE	46
2) LES ANALYSEURS UTILISES EN SPECTROMETRIE DE MASSE MS/MS	48
VII. IDENTIFICATION DES PROTEINES	51
A. IDENTIFICATION PAR EMPREINTE PEPTIDIQUE MASSIQUE	52
B. IDENTIFICATION PAR FRAGMENTATION MS/MS	52
1) L'APPROCHE PAR RECHERCHE DANS LES BANQUES PROTEIQUES	53
2) L'APPROCHE PAR SEQUENÇAGE <i>DE NOVO</i>	54
3) L'APPROCHE HYBRIDE	54
C. INTERROGATION DES BANQUES DE DONNEES	55
VIII. L'ANALYSE PROTEOMIQUE DIFFERENTIELLE	56
A. APPROCHES GEL-BASED	56
1) QUANTIFICATION PAR ELECTROPHORESE BIDIMENSIONNELLE	56
2) QUANTIFICATION PAR L'APPROCHE DIGE	57
B. QUANTIFICATION PAR MARQUAGES ISOTOPHIQUES	58
1) MARQUAGE CHIMIQUE	60
2) MARQUAGE ENZYMATIQUE	64

3) MARQUAGE METABOLIQUE : SILAC	65
C. APPROCHES SANS MARQUAGE <i>LABEL-FREE</i>	68
1) COMPARAISON DES INTENSITES DES SIGNAUX MS	68
2) COMPARAISON DU NOMBRE DE SPECTRES (<i>SPECTRAL COUNTING</i>)	68
IX. QUANTIFICATION ABSOLUE	69
A. PEPTIDE AQUA	70
B. QCONCAT	71
C. PSAQ	71

TRAVAUX PERSONNELS **72**

PREMIERE PARTIE : DISCRIMINATION DES ISOFORMES DE LA LACTOFERRINE GRACE A LA CHIMIE TAQMAN

I. PROBLEMATIQUE ET OBJECTIFS	72
II. MOYENS MIS EN ŒUVRE	72
III. RESULTATS	73
IV. CONCLUSION	74
I. PROBLEMATIQUE ET OBJECTIFS	75

DEUXIEME PARTIE : DECOUVERTE D'UNE NOUVELLE CIBLE DE L'ACTIVITE TRANSCRIPTIONNELLE DE LA DELTA-LACTOFERRINE

II. MOYENS MIS EN ŒUVRE	75
III. RESULTATS	75
IV. CONCLUSION	77

TROISIÈME PARTIE : DECOUVERTE D'UNE NOUVELLE CIBLE DE L'ACTIVITE TRANSCRIPTIONNELLE DE LA DELTA-LACTOFERRINE

I. INTRODUCTION	79
------------------------	-----------

II. RESULTATS	80
A. IDENTIFICATION ET QUANTIFICATION DU PROTEOME DES CELLULES MDA-MB-231 EN REPONSE AUX ISOFORMES DE LA LF PAR L'APPROCHE SILAC	80
1) MARQUAGE DES CELLULES PAR SILAC ET IDENTIFICATION DES PROTEINES	80
2) QUANTIFICATION DES PROTEINES ET ANALYSE DES DONNEES	82
B. FONCTIONS CELLULAIRES DES PROTEINES SUR- OU SOUS-EXPRIMEES AN REPONSE AUX LF	88
1) STRESS OXYDANT	88
2) MODULATION DE L'ADHESION CELLULAIRE, MIGRATION ET INVASION	89
3) DEGRADATION DES PROTEINES	90
C. EXPRESSION DES ISOFORMES DE LF	91
D. ACTIVITE TRANSCRIPTIONNELLE DES ISOFORMES DE LA LF	91
E. POURSUITE DE L'ETUDE PAR SILAC	96
III. DISCUSSION	97
IV. MATERIEL ET METHODES	99
A. MISE AU POINT DE CELLULES EXPRIMANT LA ΔLF DE MANIERE STABLE ET INDUCTIBLE	99
B. PURIFICATION DE LA hLF	99
C. MARQUAGE DES CELLULES PAR SILAC (STABLE ISOTOPE LABELLING AMINO ACID IN CELL CULTURE)	
D. TRAITEMENT DES CELLULES	100
E. FRACTIONNEMENT SUBCELLULAIRE	100
F. ANALYSE DES ECHANTILLONS PAR SDS-PAGE	101
G. DIGESTION TRYPSIQUE ET EXTRACTION DES PEPTIDES	101
H. ANALYSE PAR LC-MS/MS	102
I. TRAITEMENT BIOINFORMATIQUE DES DONNEES	102
J. ANALYSE PAR WESTERN-BLOT	103
K. QUANTIFICATION DE L'EXPRESSION DES ARNm PAR QPCR	103
DISCUSSION ET CONCLUSION	105
I. DISCRIMINATION DES ISOFORMES DE LA LF	105
II. ΔLFRE, ELEMENT DE REPONSE A LA LF ET A LA ΔLF	106
III. REGULATION DE L'ACTIVITE TRANSCRIPTIONNELLE DE LA ΔLF	108
IV. LES LACTOFERRINES ET LE STRESS OXYDANT	111

V. LES LACTOFERRINES ET L'INVASION CELLULAIRE ? **111**

VI. LES LACTOFERRINES, REGULATION DU CYCLE CELLULAIRE ET APOPTOSE **112**

ANNEXES **113**

Hardivillé, S., Hoedt, E., Mariller, C., Benaïssa, M., and Pierce, A. (2010). O-GlcNAcylation/phosphorylation cycling at Ser10 controls both transcriptional activity and stability of delta-lactoferrin. *Journal of Biological Chemistry* 285, 19205–19218.

Mariller, C., Hardivillé, S., Hoedt, E., Huvent, I., Pina-Canseco, S., and Pierce, A. (2012). Delta-lactoferrin, an intracellular lactoferrin isoform that acts as a transcription factor (1) (1) This article is part of a Special Issue entitled Lactoferrin and has undergone the Journal's usual peer review process. *Biochem Cell Biol.*

ABREVIATIONS **113**

REFERENCES **114**

Avant-propos

La Lactoferrine de sécrétion (Lf) et la delta-lactoferrine intracellulaire (Δ Lf), issues de la transcription du gène de la Lf sont des suppresseurs de tumeur potentiels. Elles sont toutes deux sous-exprimées ou absentes dans le cas de cancer et leur expression est un facteur de bon pronostic dans le cas de cancer du sein, ce qui est en accord avec leurs propriétés anti-cancéreuses. En effet, elles provoquent toutes les deux un arrêt du cycle cellulaire et une induction de l'apoptose. Notre laboratoire s'intéresse depuis plusieurs années à la régulation de l'activité transcriptionnelle de la Δ Lf qui lui confère ces fonctions anti-cancéreuses. Cette thèse a été financée par des partenaires industriels producteurs de Lf bovine (bLf). Ses activités antibactérienne et immunomodulatrice font que la bLf est produite à raison de 60 tonnes par an pour être essentiellement commercialisée en Asie où elle est rajoutée à des laits diététiques pour bébés, des laits, yaourts, suppléments nutritionnels, ... Il a été démontré que l'administration orale de bLf à des enfants améliore leur flore intestinale (Ochoa *et al.*, 2008). Lorsqu'elle est délivrée de manière similaire à des femmes enceintes atteintes d'anémie, une augmentation du taux d'hématocrite a pu être observée (Paesano *et al.*, 2009). En 1994, Bezault et collaborateurs montrent que la bLf est capable de réduire la croissance tumorale et le développement des métastases grâce à sa capacité à séquestrer le fer et à activer les cellules de l'immunité. A partir de 1997, en collaboration avec le Dr. Tsuda, plusieurs compagnies se sont intéressées aux effets préventifs de la bLf contre le cancer du colon. Ainsi, la bLf est la Lf la plus utilisée dans les expériences *in vitro* avec des modèles cellulaires humains, *in vivo* avec le plus souvent avec la mise en place de modèles murins et dans les essais cliniques. La hLf (lactoferrine humaine) et la bLf sont très proches en terme de structure et de fonction. C'est dans ce cadre, qu'au cours de ma thèse, j'ai voulu connaître l'impact des isoformes humaines de la Lf, hLf et Δ Lf, dans un système cellulaire homologue en utilisant un modèle humain, les cellules de glande mammaire cancéreuses MDA-MB-231 qui présentent un caractère particulièrement invasif.

Ce travail a été rendu possible par les considérables progrès qu'a connu l'analyse protéomique par spectrométrie de masse au cours de ces dernières années, justifiant son statut de méthode de choix pour l'étude de l'expression des protéines. Les électrophorèses bidimensionnelles ont longtemps été utilisées pour la comparaison de protéomes sans *a priori* quant à la découverte de protéines d'intérêt grâce à leur *empreinte peptidique massique*. Les domaines d'application de l'analyse protéomique sont aujourd'hui devenus très vastes et, parmi eux, l'analyse protéomique quantitative apporte une avancée significative à la comparaison des protéomes. Cet aspect quantitatif est tout particulièrement mis en valeur par l'utilisation du marquage isotopique, qui sera d'autant plus profitable qu'il sera effectué en amont de l'expérience.

Dans ce contexte, le travail qui m'a été confié a consisté dans un premier à développer une approche quantitative qui permette de discriminer les deux transcrits du gène de la

lactoferrine et à analyser le protéome de cellules exprimant la Δ Lf, protéine dont nous étudions l'activité transcriptionnelle depuis plusieurs années au laboratoire. De là, mes travaux ont visé à mettre en évidence l'impact de la réintroduction de la hLf ou de la Δ Lf sur le protéome des cellules MDA-MB-231.

Après avoir présenté l'état des connaissances concernant les lactoferrines et l'analyse protéomique, j'aborderai ensuite les résultats obtenus au cours de mon année de Master II Recherche et des trois années de thèse de doctorat qui ont suivi. Le chapitre *travaux personnels* sera composé de trois parties : une partie traitant la discrimination des deux isoformes, hLf et Δ Lf, une partie mettant en évidence la comparaison de protéomes exprimant ou non la Δ Lf et une partie dans laquelle nous avons étudié les protéomes des cellules de glande mammaire cancéreuse MDA-MB-231 exprimant la Δ Lf ou soumis à la hLf.

Le premier volet de mes travaux présente la mise en place d'une méthode de quantification des transcrits par PCR quantitative Taqman qui soit capable de discriminer les transcrits hLf des transcrits Δ Lf dans deux modèles cellulaires exprimant les deux isoformes de lactoferrine.

Le second volet porte sur la comparaison des protéomes de cellules exprimant ou non la Δ Lf grâce l'électrophorèse bidimensionnelle et l'empreinte peptidique massique. Ce travail a permis de mettre en évidence une nouvelle cible de l'activité transcriptionnelle de la Δ Lf, la protéine DcpS impliquée dans la maturation des transcrits.

Le dernier volet de ma thèse aborde l'impact de la réintroduction de l'expression de protéines à activité anti-tumorale, Δ Lf et hLf, sur le protéome global des cellules de glande mammaire cancéreuse MDA-MB-231. La principale difficulté consistait à apporter un aspect quantitatif rigoureux à cette étude. Pour y parvenir, j'ai réalisé une analyse protéomique quantitative et systématique basée sur le SILAC (*Stable Isotope Labelled Amino-acids in Cell culture*) et la spectrométrie de masse en collaboration avec l'équipe du Dr. Bernard Monsarrat au sein de l'Institut de Protéomique et de Biologie Structurale de Toulouse. La comparaison des protéomes a révélé de nouvelles cibles potentielles de l'activité transcriptionnelle de la Δ Lf. La hLf présente les mêmes cibles, ce qui relance la polémique sur l'activité de facteur de transcription de cette isoforme.

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., 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) : 2038-53.

2. 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* 23 (3) : 441-52.
3. 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», 2010, *Journal of Biological Chemistry* 285 (25) : 19205-18.
4. Mariller C., Hardivillé, S., Hoedt, E., Huvent, I., Pina-Canseco, S., et Pierce, A., « Delta-lactoferrin, an intracellular lactoferrin isoform which acts as a transcription factor », 2011, revue acceptée dans *Biochemistry and Cell Biology*

Publications sans comité de lecture

1. 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
2. Pierce, A., Hardivillé, S., Hoedt, E., 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, October, 21-25th, 2007
2. 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, Beijing, China, October, 18-22th, 2009
3. 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, Beijing, October 18-22th, 2009
4. Hardivillé, S., Hoedt, E., Mariller, C., Benaïssa, M., et Pierce, A., « O-GlcNAcylation/Phosphorylation at Ser10 control the activity and stability of delta-lactoferrin, a transcription factor that regulates cell cycle progression and apoptosis »,

The annual International Conference of the RSBMB, Bucharest, September 23-25th, 2010

5. Pierce, A., Mariller, C., Hardivillé, S., **Hoedt, E.**, Pina-Canseco, S., Huvent, I., « Posttranslational modifications regulate delta-lactoferrin transcriptional activity », Xth International Conference on Lactoferrin: Structure, Function and Applications, Mazatlan, Mexico, May 8-12th, 2011
6. **Hoedt, E.**, Mariller, C., Chaoui, K., Tastet, C., Huvent, I., Burlet-Schiltz, O., Monsarrat, B., et Pierce, A., « Proteome profiling comparison of lactoferrin treated or delta-lactoferrin expressing MDAMB-231 cells using SILAC, a quantitative and systematic proteomic approach » Xth International Conference on Lactoferrin: Structure, Function and Applications, Mazatlan, Mexico, May 8-12th, 2011 (Obtention d'un financement "Student Awards")

Communications par affiche

1. 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-acetylglucosaminylation », 8^{ème} Journée André VERBERT Colloque annuel des Doctorants, Lille, Faculté de Médecine, 18 septembre 2008
2. **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, Beijing, October 18-22th, 2009
3. **Hoedt, E.**, Hardivillé, S., Mariller, C., et Annick Pierce, « DcpS, enzyme de décoiffage des ARNm, nouvelle cible de l'activité transcriptionnelle de la delta-lactoferrine », 10^{ème} Journée André VERBERT Colloque annuel des Doctorants, Lille, Faculté de Médecine 15 Septembre 2010
4. **Hoedt, E.**, Hardivillé, S., Mariller, C., et Annick Pierce « The use of Taqman real time PCR to discriminate and evaluate the variation in lactoferrin and delta lactoferrin expression levels », The annual International Conference of the RSBMB, Bucharest, September 23-25th, 2010
5. **Hoedt, E.**, Chaoui, K., Mariller, C., Tastet, C., Huvent, I., Burlet-Schiltz, O., Monsarrat, B., et Pierce, A., « Proteome profiling comparison of lactoferrin treated or delta-lactoferrin expressing MDAMB-231 cells using SILAC, a quantitative and systematic proteomic approach » HUPO, Geneva, September 4-7th, 2011

Introduction

La Lf a été mise en évidence simultanément au sein de trois laboratoires en 1960 (Montreuil *et al.*, 1960; Johansson; Grove, 1960). Purifiée à partir du lait de femme (hLf) et de vache (bLf), elle a été décrite comme une glycoprotéine non héminique capable de fixer le fer. Il s'agit d'une protéine de la famille des transferrines (Park *et al.*, 1985; Pentecost et Teng, 1987; Bowman *et al.*, 1988) qui est largement représentée dans l'organisme de nombreuses espèces de mammifères (Baldwin, 1993). Alors que la transferrine est produite par le foie et sécrétée dans le sang, la Lf est synthétisée dans les glandes exocrines et sécrétée dans de nombreux fluides biologiques. En effet, la Lf est présente dans le lait et le colostrum, les larmes (Broekhuyse, 1974; Mason et Taylor, 1978), la salive (Miyauchi, 1984), la bile (van Vugt *et al.*, 1975) (van Vugt *et al.*, 1975), les fluides nasal (Fukami *et al.*, 1993), bronchique (Biserte *et al.*, 1963; Masson *et al.*, 1965), gastro-intestinal (Inoue *et al.*, 1993), vaginal (Masson *et al.*, 1968) et le liquide séminal (Wichmann *et al.*, 1989). Elle est présente en grande quantité au sein des granules secondaires des neutrophiles (Masson *et al.*, 1969; Barton *et al.*, 1988; Baggolini *et al.*, 1970) desquels elle est libérée en réponse à l'inflammation. La Lf est également synthétisée par les cellules de la microglie (Fillebeen *et al.*, 2001). C'est une protéine dotée de nombreuses fonctions biologiques parmi lesquelles les activités antibactérienne et anti-inflammatoire sont les mieux définies (Figure 1). De par ses activités et sa localisation, elle est la première ligne de défense de l'organisme vis-à-vis de nombreux pathogènes.

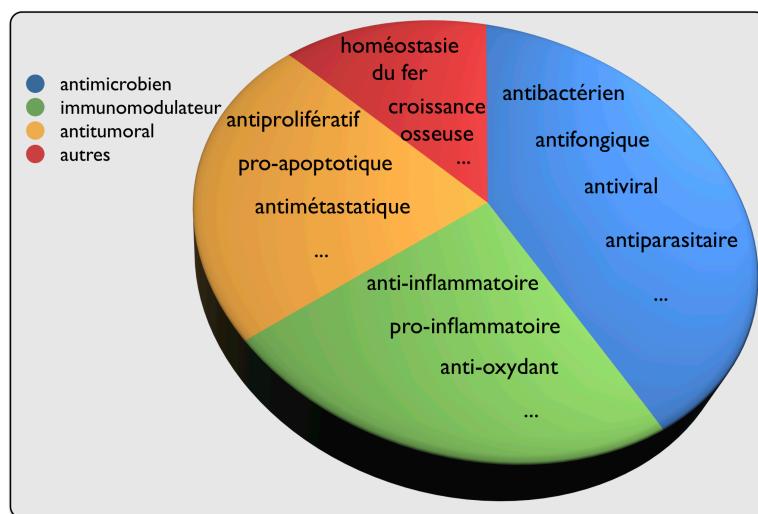


Figure 1 : Récapitulatif des différents rôles de la lactoferrine. Les rôles multiples de la lactoferrine sont regroupés par catégories : rôles antimicrobiens, immunomodulateurs, anti-cancéreux, dans l'absorption du fer et dans la croissance osseuse (d'après Pierce *et al.*, 2009).

La Δ Lf est l'isoforme intracellulaire de la hLf. Si la Lf est connue depuis plus de 50 ans, la Δ Lf a été mise en évidence pour la première fois il ya une quinzaine d'années (Siebert et Huang, 1997). Contrairement à la lactoferrine et ses multiples fonctions, la delta-lactoferrine semblerait ne posséder qu'un rôle antitumoral, via son activité de facteur de transcription (Mariller *et al.*, 2007; 2009; Hardivillé *et al.*, 2010).

I. Biosynthèse de la lactoferrine

A. Gène de la lactoferrine

Le clonage de l'ADN complémentaire de la Lf murine (Pentecost et Teng, 1987) a facilité les études visant à localiser le gène de la Lf au niveau du chromosome 9 murin (Teng *et al.*, 1987) et du chromosome 3 humain (McCombs *et al.*, 1988), en position q21.31 (Klintworth *et al.*, 1998). Le gène de la bLf a, quant à lui, été localisé au niveau du chromosome 22 bovin (Schwerin *et al.*, 1994). La hLf est encodée par un gène unique, lui même, organisé en 17 exons, entrecoupés de 16 introns (Kim *et al.*, 1998). L'organisation introns/exons du gène est très similaire à celle des gènes codant pour la sérotransferrine et l'ovotransferrine, suggérant la duplication d'un gène ancestral commun pour la famille des transferrines (Shirsat *et al.*, 1992; Baldwin, 1993).

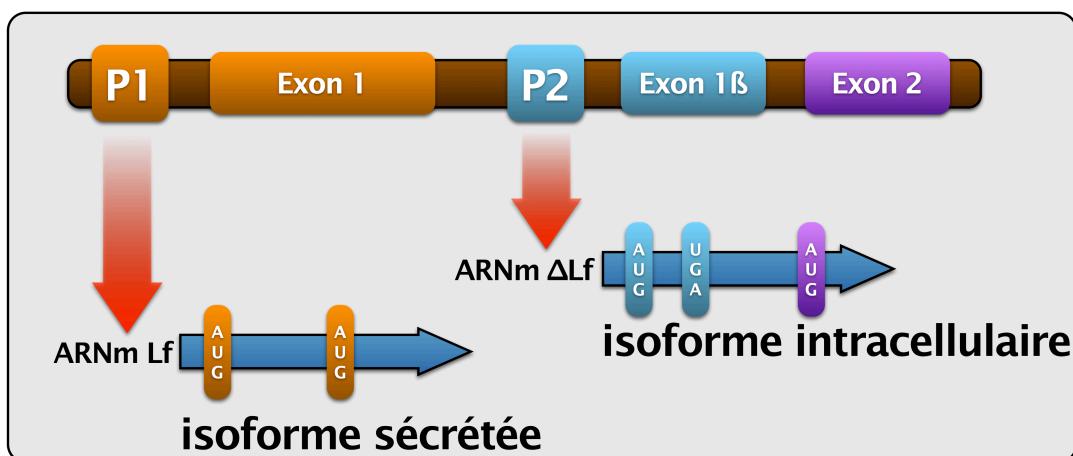


Figure 2 : Représentation schématique du gène des lactoferrines humaines et de son épissage alternatif à partir du promoteur P1 et du promoteur P2. La transcription du gène de la lactoferrine à partir du promoteur P1 permet la synthèse de l'isoforme sécrétée grâce à la traduction du peptide signal encodé 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).

B. Régulation de l'expression du gène de la lactoferrine humaine

Le gène de la hLf possède deux promoteurs, P1 et P2 (Siebert et Huang, 1997). La transcription de ce gène peut donc conduire à l'obtention de deux transcrits : l'ARNm de la hLf, issu de la transcription à partir du promoteur P1 et l'ARNm de la ΔLf, issu de la transcription à partir du promoteur P2 (Figure 2). L'ARNm de la ΔLf étant le produit de l'épissage alternatif du gène de la hLf, hLf et ΔLf 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 de la traduction est immédiatement suivi d'un codon stop. Il existe, cependant, un autre site d'initiation de la traduction en phase avec le cadre de lecture ouvert de la hLf dans l'exon 2. La traduction de la ΔLf débute donc à partir de l'exon 2 conduisant à une protéine identique à la hLf mais tronquée de ses 45

premiers résidus d'acides aminés dont la séquence peptide signal (Siebert et Huang, 1997) (Figure 2, page 6). La hLf, issue du transcrit comportant l'exon 1, est donc secrétée alors que la ΔLf est intracellulaire.

Liu *et al* (2003) ont montré que les promoteurs P1 et P2 étaient régulés différemment selon le type cellulaire. De nombreuses études ont révélé que l'expression du gène de la hLf pouvait être régulée par des récepteurs nucléaires, des hormones, des facteurs de transcription ou encore par des *stimuli* inflammatoires.

L'analyse des régions promotrices du gène de la hLf a permis de mettre en évidence la présence de séquences consensus nécessaires à l'initiation de la transcription. En effet, le promoteur P1 possède une séquence ATAAA (TATA box imparfaite) en position -32 à -27 pb, une séquence CAAT imparfaite (CAAC) en position -55 à -52 pb ainsi qu'une région très riche en GC (77 %) sur laquelle se fixe l'élément Sp1 en position -26 à +1 pb (Teng *et al.*, 1992) (Figure 3A). La ΔLf possède au sein de son promoteur, une boîte TATA inversée en position -29 pb par rapport à son site d'initiation de la transcription (Liu *et al.*, 2003) (Figure 3B).

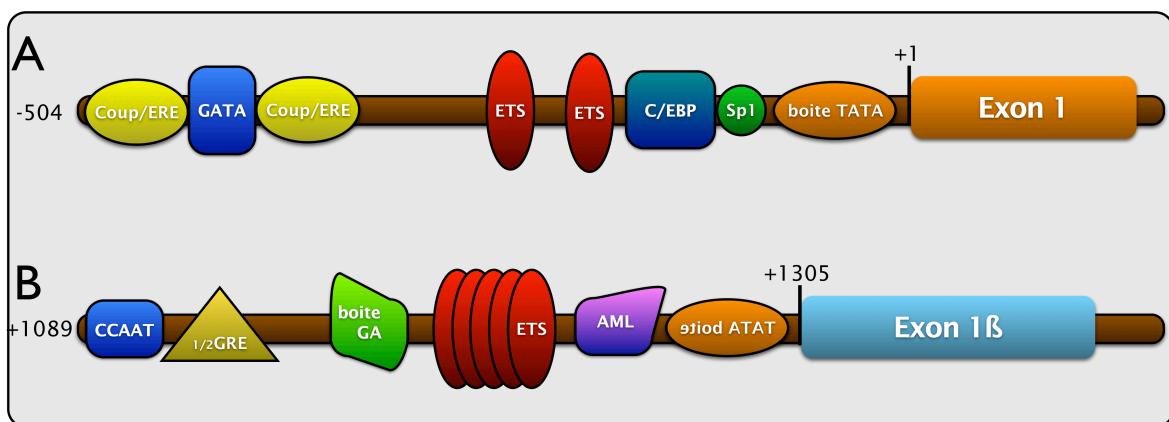


Figure 3 : Représentation schématique des différents éléments de réponse du promoteur P1 (A) (Lee *et al.*, 1995; Teng *et al.*, 1992) et du promoteur P2 (B) (Liu *et al.*, 2003; Siebert et Huang, 1997) du gène de la lactoferrine humaine.

1) Régulation par les facteurs de transcription

De nombreux facteurs de transcription régulent le gène *Lf*, dont la superfamille des récepteurs nucléaires qui influence directement la transcription à partir du promoteur P1. En effet, après avoir constaté que le gène de la hLf était capable de répondre à une stimulation par les œstrogènes (Pentecost et Teng, 1987; Yang *et al.*, 1996), Christina Teng (2002) a pu démontrer que, comme tous les gènes cibles des œstrogènes, la *hLf* possède une séquence ERE (*Estrogen Response Element*) située en -418 à -378 (Yang *et al.*, 1996) et chevauchant une séquence similaire à l'élément COUP (*Chicken Ovalbumin Upstream element*). La fixation du récepteur aux œstrogènes à la séquence COUP régule négativement la réponse aux œstrogènes sur le gène de la hLf (Teng, 2002). Le promoteur P1 possède également un hémi-site ERE appelé ERRE (*Extended Estrogen Response Element half site*), qui en plus du site ERE, augmente sensiblement la réponse aux œstrogènes dans les cellules de glande

mammaire cancéreuse MCF7 (Stokes, 2004). De nombreuses études ont démontré l'implication des éléments ER α et ERR α dans l'activation du promoteur P1 au niveau de l'épithélium endométrial (Teng *et al.*, 1992; Yang *et al.*, 1996; Teng *et al.*, 2002). Le module ERE/COUP situé en -362 à -328 (Teng *et al.*, 1992) contient également deux éléments RARE (*Retinoic Acid Response Element*) de réponse aux acides rétinoïques (Lee *et al.*, 1995). D'ailleurs, le processus de maturation myélocyttaire peut être régulé par les acides rétinoïques et leurs récepteurs qui sont apparus comme étant des inducteurs potentiels de la synthèse de la hLf dans les cellules myéloïdes (Rado *et al.*, 1987). De plus, les acides rétinoïques sont capables de stimuler directement la transcription du gène de la hLf dans les cellules souches (Geng *et al.*, 1998). Une autre hormone, la prolactine pourrait être responsable de la synthèse de bLf au sein de cellules épithéliales de glande mammaire bovine (Nakajima *et al.*, 2008).

Le promoteur P1 est essentiellement régulé par des facteurs de transcription tels que Sp1, Ets, PU1, C/EBP (α et ϵ), CDP/cut et KLF5. En effet, deux séquences consensus pour le facteur de transcription Sp1 (*Specific protein 1*) ainsi qu'une PU box ont été retrouvées en position -85 à -35 pb (Teng, 2002). Khanna-Gupta *et al.* ont montré que la fixation du facteur CDP/cut (*CCAAT displacement protein*) sur le promoteur du gène de la hLf agissait comme un répresseur transcriptionnel mais les mécanismes d'action n'ont pas été mis en évidence (Khanna-Gupta *et al.*, 1997; 2003). Au contraire, les protéines C/EBP α et ϵ (*CCAAT Enhancer Binding Protein*), régulateurs clefs de la dégranulation mais également de l'expression du gène de la hLf sont capables de se fixer au niveau du promoteur P1 et d'en induire sa surexpression (Khanna-Gupta *et al.*, 2003; 2000). Sp1 et C/EBP régulent essentiellement P1 lors de la différenciation granulocyttaire (Khanna-Gupta *et al.*, 1997). Il est à noter que l'expression de C/EBP α et ϵ est régulée positivement par les acides rétinoïques au sein des cellules myéloïdes (Gery *et al.*, 2004). A proximité de la boîte TATA, deux séquences GGAA fonctionnelles peuvent servir à la fixation des facteurs de transcription Ets (Liu *et al.*, 2003).

Pour Furmanski *et al.*, la hLf elle-même pourrait réguler sa propre transcription (Furmanski *et al.*, 1989).

Le promoteur P2, quant à lui, ne possède pas d'éléments de réponse aux œstrogènes mais un élément de réponse aux glucocorticoïdes (GRE) dont la fonctionnalité n'a pas été démontrée (Liu *et al.*, 2003). Des séquences putatives de fixation pour des facteurs de transcription tels que les XRE (*Xenobiotic Response Element*), les $1/2$ HRE (*Hormone Response Element half site*) et AML-1a (*Acute Myeloid Leukemia-1*) y ont été mises en évidence mais leur fonctionnalité n'a pas été démontrée (Liu *et al.*, 2003). Christina Teng et collaborateurs ont montré que le facteur de transcription Ets est nécessaire à l'activation de P2 qui contient onze motifs GGAA de fixation de Ets et P2 est très actif dans les cellules lymphoïdes mais pas dans les cellules utérines (Liu *et al.*, 2003).

Ainsi, tous ces facteurs de transcription sont impliqués dans l'expression différentielle du gène de la Lf, de manière cellule-spécifique.

2) Régulation par des *stimuli* inflammatoires

Plusieurs études ont révélé que des *stimuli* de la réponse innée conduisent à une surexpression du gène de la Lf. Ceci suggère 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. De plus, des éléments de réponse à l'infection ont été mis en évidence dans le promoteur du gène de la bLf (Zheng *et al.*, 2005). Leur fonctionnalité a été démontrée car l'infection de cellules de glande mammaire par *E. coli* conduit à une surexpression du gène de la bLf (Griesbeck-Zilch *et al.*, 2008). Enfin, Moriuchi et Moriuchi (2006) ont démontré que l'infection par le virus HTLV-1 (*Human T-cell leukemia virus type 1*) conduit à une transactivation du gène de la hLf dans les cellules myéloïdes HL-60 et dans les cellules de glande mammaire cancéreuse MCF7 (Moriuchi et Moriuchi, 2006). Il est à noter qu'en amont du promoteur P2, des éléments de réponses putatifs NF-κB, une boîte GA, retrouvée dans le gène c-Myc et un site de liaison de l'IRF (*Interferon Regulatory Factor*) ont également été mis en évidence mais leur fonctionnalité n'a pas encore été démontrée.

Par contre, une répression de la transcription du gène de la hLf a été observée suite à l'expression du TNF- α (*Tumor Necrosis Factor α*) dans les cellules de la moelle osseuse (Srivastava *et al.*, 1991).

3) Régulation par méthylation du promoteur

L'état de méthylation du gène de la Lf est inversement proportionnel à son expression dans de nombreux tissus (Teng *et al.*, 2004b; Shaheduzzaman *et al.*, 2007; Zhang *et al.*, 2011). De plus, il existe une expression différentielle des deux transcrits due à une méthylation différente des deux promoteurs. En effet, à la différence du promoteur P2, le promoteur P1 est sous-méthylé dans la majorité des tissus étudiés, en particulier la rate et le rein. D'ailleurs, une forte expression de transcrits hLf a été retrouvée au sein de ces tissus. Par contre, le promoteur P2 est hyperméthylé dans la plupart des tissus, ce qui explique sa faible expression, sauf dans les leucocytes où l'on trouve une faible méthylation des sites CpG, corrélée à une expression plus abondante de transcrits ΔLf (Teng *et al.*, 2004a).

4) Dégradation des transcrits du gène de la lactoferrine humaine

Les micro-ARN sont connus pour réguler la traduction et la dégradation des transcrits. Récemment, suite à la transfection dans les cellules de glande mammaire cancéreuse MCF7, du micro-ARN miR214 qui contient une région qui s'apparie parfaitement avec la région 3'UTR commune aux transcrits de Lf et ΔLf, l'équipe de Bo Lonnerdal a pu constater une diminution de l'expression de la Lf en 48 h (Liao *et al.*, 2010). Nous avons d'ailleurs établi

une collaboration avec cette équipe (Université de Californie, Davis, USA) pour savoir si les transcrits Δ Lf seraient eux aussi soumis à cette régulation négative.

II. Structure des lactoferrines

La Lf est une glycoprotéine bilobée possédant une masse moléculaire proche de 80 kDa. C'est une protéine extrêmement basique, dont la composition en acides aminés de son domaine N-terminal. En effet, son pHi est supérieur à 9, ce qui la différencie des autres transferrines. La Δ Lf possède une masse moléculaire apparente de 73 kDa. La délétion des 26 premiers résidus d'acides aminés de l'extrémité N-terminale de la hLf ne diminue que d'une unité le pHi de la Δ Lf. Les séquences polypeptidiques de la hLf et de la Δ Lf étant identiques, les deux isoformes devraient avoir la même conformation spatiale.

A. Séquence primaire de la lactoferrine

La séquence complète de la hLf isolée du lait maternel a été déterminée par séquençage peptidique (Metz-Boutigue *et al.*, 1984) et par déduction après séquençage de l'ADNc correspondant à la hLf (Powell et Ogden, 1990; Rey *et al.*, 1990). La protéine mature de hLf est composée de 692 résidus d'acides aminés comprenant des résidus de cystéine qui forment 16 ponts disulfures (Metz-Boutigue *et al.*, 1984). Aujourd'hui, les séquences primaires de la Lf de dix espèces ont pu être caractérisées : homme, souris, vache, cheval, chèvre, mouton, cochon, singe, buffle et chameau (pour revue Baker et Baker, 2009). Les Lf des différentes espèces partagent des homologies de séquences, allant de 65 à presque 100 %. Par exemple, les Lf murine et humaine partagent respectivement 65-70 % et 70-74 % d'identité avec celles des autres espèces (pour revue Baker et Baker, 2009). Il existe 37 % d'homologie de séquence entre les domaines N- et C-terminaux qui constituent les deux lobes de la hLf (Metz-Boutigue *et al.*, 1984).

La Δ Lf dont la séquence a été déterminée par séquençage de l'ADNc correspondant est composée de 666 résidus d'acides aminés (Siebert et Huang, 1997). Elle ne devrait pas posséder de ponts disulfures du à l'environnement réducteur du cytoplasme.

B. Structure tridimensionnelle de la lactoferrine

La cristallisation de la hLf saturée en fer (holo-Lf) et son analyse par diffraction aux rayons X (Anderson *et al.*, 1987; 1989) ont montré que la Lf est organisée en deux lobes globulaires N- (résidus 1 à 333) (résidus 319 à 307 pour la Δ Lf) et C-terminaux (résidus 345 à 692) (résidus 319 à 666 pour la Δ Lf). Ces deux lobes reliés entre eux par une courte hélice α formée par les acides aminés 334 à 344 (308 à 318 pour la Δ Lf) peuvent être dissociés par digestion enzymatique. L'hydrolyse enzymatique ménagée de la hLf libère deux fragments de 30 et 50 kDa correspondants respectivement aux domaines N- et C-terminaux de la protéine.

Ces deux fragments peuvent se réassocier pour former un complexe de 80 kDa dont les comportements chromatographique et spectroscopique sont identiques à ceux de la Lf native, suggérant des interactions non covalentes et réversibles entre les deux lobes (Legrand *et al.*, 1986). Chaque lobe est organisé en deux domaines d'environ 160 résidus d'acides aminés à l'interface desquels se trouve le site de fixation du fer dans le cas de la Lf. Les deux domaines possèdent la même organisation structurale basée sur l'alternance de feuillets β et d'hélices α (Anderson *et al.*, 1989; 1987) (Figure 4). Le lobe N-terminal est constitué des domaines N1 (résidus 1 à 90 et de 252 à 320 pour la Lf, résidus 1 à 64 et de 226 à 294 pour la Δ Lf) et N2 (résidus de 91 à 251 pour la lactoferrine, de 65 à 225 pour la Δ Lf), le lobe C-terminal est constitué du domaine C1 (résidus 345 à 433 et 596 à 663 pour la Lf, résidus 319 à 407 et 570 à 637 pour la Δ Lf) et du domaine C2 (résidus 434 à 595 pour la Lf, résidus 408 à 569 pour la Δ Lf) (Figure 4).

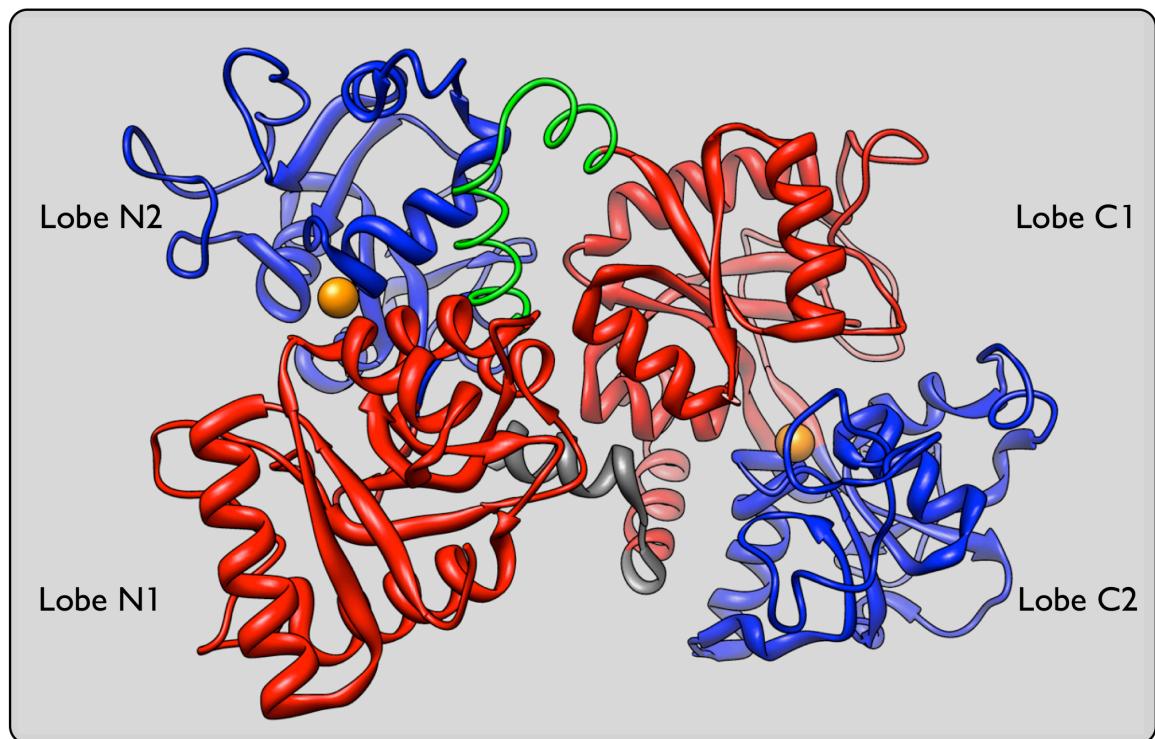


Figure 4 : Représentation schématique de la structure tridimensionnelle de la lactoferrine humaine saturée en fer (d'après Anderson *et al.*, 1989). Les lobes N1 et C1 et les lobes N2 et C2 sont respectivement représentés en rouge et en bleu. La région inter-lobe apparaît en vert. Les deux points en orange représentent les atomes de fer chélatisés par la hLf. La représentation moléculaire a été réalisée en utilisant le logiciel Chimera (<http://cgl.ucsf.edu/chimera/>).

La Lf, principalement présente dans le lait sous forme apo (dépourvue en fer, apo-Lf) est capable de fixer réversiblement un ion ferrique sur chaque lobe.

L'analyse des cristaux d'apo-hLf par diffraction des rayons X a été réalisée avec une résolution de 2,8 Å, puis affinée à 2 Å (Anderson *et al.*, 1990; Jameson *et al.*, 1998). La structure tridimensionnelle de la l'apo-hLf révèle que le site de fixation du fer du lobe Nt est ouvert, tandis que celui du lobe Ct est fermé. La fixation du fer dans le lobe Nt est accompagnée d'une faible rotation qui permet de protéger le fer de l'environnement extérieur.

Il existerait un équilibre entre les formes ouvertes et fermées de l'apo-hLf en solution (Mizutani *et al.*, 2011).

En 1997, Moore *et al.* ont déterminé la structure tridimensionnelle de la bLf avec une résolution de 2,8 Å (Moore *et al.*, 1997). Cette étude a mis en évidence que malgré des séquences primaires assez différentes (69% d'homologie) (Pierce *et al.*, 1991), les conformations des hLf et bLf sont similaires, mis à part le fait que les lobes N-terminal et C-terminal de la bLf sont plus fermés respectivement de 5,3° et 3,7° que ceux de la hLf (Moore *et al.*, 1997; Jameson *et al.*, 1998).

Cette structure dynamique de la hLf est importante pour la compréhension des activités biologiques liées au fer, bactériostatiques et antioxydante. La fixation du fer maintient les deux domaines de chaque lobe à proximité. L'holo-hLf est donc dotée d'une conformation rigide et très stable, l'apo-hLf au contraire, possède une conformation flexible et plus encline à la dénaturation thermique et à la dégradation protéolytique (Baker et Baker, 2009). Cela confère à la hLf un rôle de chélateur plutôt que de donneur de fer, par rapport aux transferrines.

C. Caractéristiques de surface de la lactoferrine

Toutes les Lf dont la structure a été caractérisée à ce jour possèdent une forte concentration de charges positives (Figure 5, page 13). Ces charges positives se situent dans le lobe N-terminal, plus particulièrement au niveau de la séquence ¹GRRRRS⁶ de la hLf, au niveau de l'extrémité C-terminale de la première hélice (résidus 12 à 30 pour la hLf) et du cluster de résidus d'acides aminés basiques associé à la neuvième hélice (résidus 263 à 279 pour la hLf). Les deux dernières hélices constituent le site de liaison proposé pour les LPS, le lipide A, les porines, l'acide lipoteïchoïque, ainsi que les adhésines bactériennes et virales (Mann *et al.*, 1994; van Berkel *et al.*, 1997). Elles sont donc le siège de l'activité antimicrobienne de la Lf. En effet, les résidus 1 à 47 forment un site de fixation pour les porines bactériennes (Sallmann *et al.*, 1999); les résidus 1 à 6 et 28 à 31 ont été caractérisés comme étant des sites de fixation aux protéglycans (Mann *et al.*, 1994; van Berkel *et al.*, 1997) ainsi qu'un site de fixation au sCD14/LPS (Elass-Rochard *et al.*, 1995); les résidus 19 à 29 forment un site de fixation au LPS (Japelj, 2005); les résidus 25 à 30 permettent la fixation au *Lipoprotein Receptor-related Protein* (LRP) (Meilinger *et al.*, 1995) ainsi qu'au récepteur lymphocytaire (Legrand *et al.*, 1992). Ainsi, les régions basiques de la Lf sont impliquées dans la fixation de nombreux récepteurs de surface (Tableau I, page 14). La région inter-lobe associée à la courte hélice α reliant les deux lobes est également composée de nombreux résidus basiques.

Cette séquence ¹GRRRRS⁶ est déletée de la Δ Lf, il n'existe pas de sites d'interaction connus pour la Δ Lf.

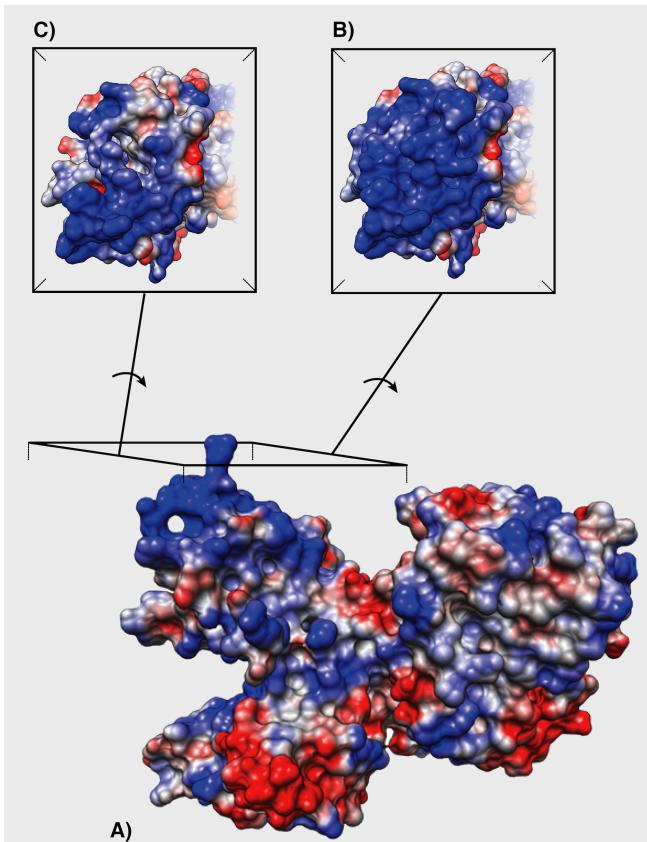


Figure 5 : Répartition des charges à la surface de la lactoferrine humaine (A, B) et de la delta-lactoferrine (C). (Mariller *et al.*, 2012). Les charges positives sont représentées en bleu et les charges négatives en rouge. La photographie présentée en A montre la surface de la molécule exposée au solvant 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 hLf (B) et de la Δ Lf (C). Le modèle moléculaire, basé sur la structure tridimensionnelle définie par Anderson *et al.* (1987), a été réalisé en utilisant le logiciel Chimera (<http://cgl.ucsf.edu/chimera/>)

D. Structure des glycannes

Les glycannes des Lf, localisés dans les domaines N2 et C2 (Figure 4, page 11), sont de type N-acétyllactosaminique biantennés, sialylés et fucosylés (Spik *et al.*, 1988). Ils peuvent être mono- ou disialylés et fucosylés en α -1,6 ou en α -1,3 sur la N-acétylglucosamine (Spik *et al.*, 1982; Coddeville *et al.*, 1992). La bLf, quant à elle, possède en plus des glycannes de type N-acétyllactosaminique, des glycannes de type *high-mannose* (Coddeville *et al.*, 1992). Il existe trois sites de glycosylation possibles pour la hLf et cinq pour la bLf. Ainsi, il existe une grande diversité au niveau de la structure primaire des glycannes entre espèces qui reste vraie au niveau du nombre et de la localisation des sites de glycosylation. Il existe de plus une hétérogénéité interindividuelle dans le nombre de sites glycosylés. Le rôle de ces glycannes semble être de protéger la protéine contre la protéolyse et de diminuer son immunogénicité (Spik *et al.*, 1988; van Veen *et al.*, 2004; Legrand *et al.*, 1990).

Tableau I : Micro-organismes contre lesquels s'exerce l'activité anti-infectieuse des lactoferrines humaine et bovine (d'après Pierce *et al.*, 2009)

Bactéries	Virus	Champignons Levures	Parasites
<i>Actinobacillus</i>	Adénovirus	<i>Aspergillus fumigates</i>	<i>Babesia caballi</i>
<i>Actinomycetemcomitans</i>	Cytomégavirus (CMV)	<i>Aspergillus niger</i>	<i>Eimeria stiedai</i>
<i>Bacillus stearothermophilus</i>	Entérovirus	<i>Candida albicans</i>	<i>Entamoeba histolytica</i>
<i>Bacillus subtilis</i>	Hantavirus	<i>Candida glabrata</i>	<i>Plasmodium berghei</i>
<i>Bacillus stearothermophilus</i>	Poliovirus	<i>Candida krusei</i>	<i>Plasmodium falciparum</i>
<i>Burkholderia cenocepacia</i>	Polyomavirus	<i>Phoma exigua</i>	<i>Toxoplasma gondii</i>
<i>Clamydophila psittaci</i>	Rotavirus	<i>Sclerotinia sclerotiorum</i>	<i>Tritrichomonas foetus</i>
<i>Clostridium spp.</i>	Virus de Friend	<i>Sclerotium rolfsii</i>	<i>Trypanosoma brucei</i>
<i>Escherichia coli</i> spp.	Virus de l'hépatite B (HBV)	<i>Rhizoctonia solani</i>	<i>Trypanosoma cruzi</i>
<i>Enterococcus spp.</i>	Virus de l'hépatite C (HCV)	<i>Trichoderma viride</i>	
<i>Haemophilus influenzae</i>	Virus de l'herpès simplex 1 (HSV-1)	<i>Trychophyton mentagrophyte</i>	
<i>Helicobacter felis</i>	Virus de l'herpès simplex 2 (HSV-2)		
<i>Helicobacter pylori</i>	Virus de l'herpès simplex félin (FSV-1)		
<i>Legionella pneumophila</i>	Virus de l'immunodéficience humaine (HIV)		
<i>Listeria monocytogenes</i>	Virus Sindbis		
<i>Micrococcus spp.</i>	Virus Semliki Forest		
<i>Mycobacterium tuberculosis</i>			
<i>Neisseriaceae</i>			
<i>Pasteurellacea</i>			
<i>Pseudomonas aeruginosa</i>			
<i>Salmonella typhimurium</i>			
<i>Shigella dysenteriae</i>			
<i>Shigella flexneri</i>			
<i>Staphylococcus aureus</i>			
<i>Staphylococcus spp.</i>			
<i>Streptococcus spp.</i>			
<i>Vibrio cholerae</i>			
<i>Vibrio parahaemolyticus</i>			

En termes de différences glycanniques, le cas de la ΔLf est le plus marquant. En effet, de part sa localisation cytoplasmique, elle ne sera pas modifiée par N-glycosylation avec des glycannes de type complexes comme la Lf de sécrétion. Cependant, l'étude *in silico* de sa structure primaire par le serveur YingOYang 1.2 (<http://www.cbs.tdu.dk/services/YinOYang>) montre qu'elle pourrait être modifiée par O-glycosylation cytosolique ou GlcNAcylation sur les résidus S¹⁰, S²²⁷, S⁴⁷² et T⁵⁵⁹. La présence d'une GlcNAcylation des sites S¹⁰ et S²²⁷ a pu être confirmée (Hardivillé *et al.*, 2010). Contrairement à la N-glycosylation, la GlcNAcylation résulte du transfert d'un monosaccharide unique, la N-acétyl- α -D-glucosamine, sur l'hydroxyle d'un résidu Ser ou Thr. Il existe très souvent une réciprocité entre la modification par O-GlcNAc et la phosphorylation (Hart *et al.*, 1989).

III. Les différentes fonctions des lactoferrines

Les différentes fonctions décrites pour la hLf sont également partagées par la bLf. D'ailleurs, la bLf est la molécule la plus utilisée dans les expériences *in vitro*, *in vivo*, ainsi que dans les essais cliniques. Ce sont des protéines multifonctionnelles qui ont été appelées protéines « couteau suisse » (Shimazaki et Kushida, 2010). Leurs propriétés antimicrobiennes sont les plus étudiées. Elles possèdent un large spectre antimicrobien qui s'étend des bactéries Gram négatives et positives, aux levures (Xu *et al.*, 1999), champignons, à certains virus et protozoaires (Tableau I, page 14). L'activité antimicrobienne de la Lf

semble s'installer dès les premiers stades de l'infection. La Lf interagit avec de nombreux microorganismes et leurs récepteurs spécifiques à la surface de la cellule hôte et inhiberait ainsi l'entrée de ces agents pathogènes au sein de la cellule hôte. Cependant, ce mode d'action de la Lf au niveau des cellules hôtes doit faire intervenir des mécanismes plus complexes qu'une simple compétition pour les sites de fixation aux microbes (Legrand *et al.*, 2008). La Lf est capable en parallèle d'activer le système immunitaire inné. Récemment, le groupe de Teng a fait une découverte primordiale en montrant que l'expression de la Lf était directement induite par un stimulus pro-inflammatoire, LPS ou ARN double brin, mimant une attaque virale (Li *et al.*, 2009). Parmi les autres fonctions attribuées à la Lf, certaines, comme l'immunomodulation, l'activité de facteur de transcription et l'activité antitumorale sont parfois sujet à controverse. S'il est indéniable aujourd'hui que les Lf de sécrétion sont des immunomodulateurs, les conclusions des premières expériences décrites dans la littérature concernant la part de la Lf dans l'immunomodulation pouvaient être erronées car les préparations de Lf isolée du lait étaient contaminées par des LPS. Seules les résultats des expérimentations réalisées avec des Lf détoxifiées sont à prendre en compte. D'autre part, l'activité de facteur de transcription attribuée à la Lf pourrait être due à la Δ Lf. Il est en effet difficile de comprendre comment une molécule sécrétée pourrait être, après internalisation et adressage aux endosomes, relocalisée au noyau. De nombreuses études suggèrent que la Lf possède également un rôle antitumorigénique qui s'exercerait par la régulation de l'activité des cellules NK, la modulation de l'expression des protéines en phase G1 et le déclenchement de l'apoptose. Récemment, la Lf a été impliquée dans un processus protumoral *via* l'augmentation de l'invasivité de certaines cellules cancéreuses (Ha *et al.*, 2011).

La Δ Lf, isoforme intracytoplasmique de la hLf, est un facteur de transcription impliqué dans la surveillance cellulaire. La délétion des 26 premiers résidus d'acides aminés l'empêche d'interagir avec les récepteurs de surface de la hLf (Siebert et Huang, 1997 ; Liu *et al.*, 2003 ; Mariller *et al.*, 2007), ce qui explique pourquoi contrairement à la hLf, elle ne présente pas de fonction antimicrobienne ou immunomodulatrice.

A. Activité antimicrobienne

1) Activité antibactérienne

La capacité à fixer deux atomes de fer ferrique a longtemps laissé penser que l'unique rôle de la Lf était son activité antibactérienne (Arnold *et al.*, 1977; Weinberg, 1993). Cette activité passe par la séquestration du fer nécessaire à la nutrition de la plupart des bactéries pathogènes, ce qui conduit à l'inhibition de leur croissance (Weinberg, 1977). En effet, plusieurs études ont démontré que seule l'apo-Lf possédait cette activité qui de plus était réduite lorsqu'elle était en compétition avec l'holo-Lf (Arnold *et al.*, 1980; Yamauchi *et al.*, 1993). En addition à cette activité bactériostatique, la Lf est dotée d'une activité bactéricide,

distincte d'une simple déprivation en fer. La présence de fortes charges positives au niveau de la structure de la Lf facilite son interaction avec le lipide A, composé du LPS des bactéries Gram négatives (Appelmelk *et al.*, 1994). Cette interaction est responsable d'une rigidification des chaînes acyls des lipopolysaccharides (Brandenburg *et al.*, 2001). La Lf peut ainsi endommager la perméabilité de la paroi de la bactérie, entraînant un relargage de LPS (Ellison *et al.*, 1988) sous la dépendance d'ions divalents tels que le Ca²⁺ et Mg²⁺ (Ellison *et al.*, 1990). En déstabilisant la paroi bactérienne, la Lf est ainsi capable d'augmenter l'effet antibactérien délivré par des antibiotiques comme la rifampicine (Ellison *et al.*, 1988). Les caractéristiques structurales de la Lf lui permettent également d'interagir avec les acides lipotéchoïques présents à la surface de bactéries Gram positives comme *Staphylococcus epidermidis*, conduisant à une diminution des charges négatives présentes au niveau de la barrière cellulaire bactérienne associée au peptidoglycane (Ellison et Giehl, 1991; Leitch et Willcox, 1999; Singh *et al.*, 2002).

Bien que l'activité antibactérienne de la bLf ait été associée à son interaction directe avec des composés de la surface bactérienne, Fischer *et al* (2006) affirment que ces effets découlent indirectement de son activité immunomodulatrice (Fischer *et al.*, 2006). Mosquito et collaborateurs (2010) ont montré que l'ingestion de bLf par des souris infectées par *Salmonella typhimurium* mène à une réduction de l'inflammation dans l'intestin, le foie, la rate et le cerveau (Mosquito *et al.*, 2010). La Lf peut activer l'immunité innée en inhibant la formation du biofilm qui correspond à une organisation en colonie des cellules bactériennes à la surface de la cellule hôte. En effet, la bLf inhibe la formation de biofilm de *Streptococcus mutans* (Berlutti *et al.*, 2004) et de *Pseudomonas aeruginosa* (Singh *et al.*, 2002; Kamiya *et al.*, 2011) à des concentrations plus faibles que celles nécessaires pour inhiber la croissance bactérienne (Singh *et al.*, 2002). De même, Wakabayashi *et al.* (2010) ont montré que la bLf était capable de réduire le biofilm établi par les bactéries périondopathiques telles que *Porphyromonas gingivalis* et l'étude clinique effectuée a confirmé que l'ingestion de bLf par des patients réduit la plaque subgingivale et la périodontite chronique (Wakabayashi *et al.*, 2010). Certaines souches bactériennes nécessitent de fortes concentrations en fer pour former des biofilms, c'est pourquoi alors que l'apo-Lf inhibe cette formation par séquestration du fer (Weinberg, 2004), l'holo-Lf au contraire, stimule l'agrégation de biofilms chez *Pseudomonas aeruginosa* (Jesaitis *et al.*, 2003).

D'autre part, la bLf empêche l'adhésion et l'invasion par les bactéries dans les cellules hôtes grâce à une interaction de ses glycannes de type *high mannose* avec les adhésines bactériennes (Dial et Lichtenberger, 2002; Bessler *et al.*, 2006) ou une interaction entre la bLf et les récepteurs bactériens à la surface de la cellule hôte.

Enfin, une activité sérine-protéase *like* dont le site actif serait localisé dans la partie N-terminale de la Lf est responsable de la protéolyse des protéines clefs de l'adhésion

bactérienne par les cellules de mammifères (Qiu *et al.*, 1998; Plaut *et al.*, 2000; Hendrixson *et al.*, 2003). Cette activité serait à l'origine de l'inhibition de l'invasion intracellulaire de pathogènes comme *Escherichia coli*, *Listeria monocytogenes* (Lee *et al.*, 2005), *Shigella flexneri* (Gomez *et al.*, 2003; Ochoa et Clearly, 2004) ; et *Haemophilus influenzae* (Hendrixson *et al.*, 2003).

2) Activité antivirale

La Lf exerce cette activité contre des virus à ADN et à ARN, enveloppés comme ceux de l'hépatite C (Yi *et al.*, 1997), de l'herpès (Marchetti *et al.*, 1996; 1998) et du VIH (virus de l'immunodéficience humaine) (Swart *et al.*, 1996) ou nus comme les rota-, polio-, adéno-, et entérovirus (Superti *et al.*, 1997; Lin *et al.*, 2002; Arnold *et al.*, 2002; Marchetti *et al.*, 1999; Di Biase *et al.*, 2003; Pietrantonio *et al.*, 2003). Le mécanisme de l'action antivirale de la Lf n'est pas complètement élucidé. Néanmoins, dans la plupart des études réalisées *in vitro* et *in vivo*, la Lf agit aux premiers stades de l'infection virale (Puddu *et al.*, 1998) en inhibant l'attachement et l'entrée des virus *via* une interaction avec leurs récepteurs spécifiques à la surface des cellules hôtes ou en interagissant directement avec les particules virales. En effet, l'interaction de la bLf ou de la hLf avec les héparanes sulfates (Hasegawa *et al.*, 1994; Damiens *et al.*, 1998; Andersen *et al.*, 2004; Marchetti *et al.*, 2004) ou les chondroïtanes sulfates (Hasegawa *et al.*, 1994; Marchetti *et al.*, 2004) empêche l'adhésion et l'entrée du virus de l'herpès dans les cellules étudiées. De la même manière, la hLf entre en compétition avec le virus HIV pour la fixation aux récepteurs héparanes sulfates (Saidi *et al.*, 2006) aux lectines DC-SIGN (Harmsen *et al.*, 1995; Groot *et al.*, 2005), aux récepteurs CXCR4 ou CCR5 (Berkhout *et al.*, 2004; Carthagena *et al.*, 2011) ainsi qu'à la nucléoline de surface (Legrand *et al.*, 2004), inhibant ainsi l'entrée du virus au sein des cellules épithéliales. Par ailleurs, de nombreuses expériences *in vitro* ont montré une interaction directe de la hLf avec les particules virales du virus de l'herpès (Marchetti *et al.*, 1996), du virus de l'hépatite C (Yi *et al.*, 1997), des rotavirus (Superti *et al.*, 1997), de l'entérovirus 71 (Weng *et al.*, 2005), de l'adénovirus (Pietrantonio *et al.*, 2003), de l'échovirus (Ammendolia *et al.*, 2007) ou encore du VIH (Swart *et al.*, 1996). La hLf semble également pouvoir inhiber la réplication virale du virus VIH-1 (Swart *et al.*, 1996; Legrand *et al.*, 2004; Groot *et al.*, 2005), du virus de l'herpès HSV-1 (Marchetti *et al.*, 1996; Siciliano *et al.*, 1999) et celui de la grippe (Qiu *et al.*, 1998) *in vitro*.

Les études réalisées *in vivo* montrent que l'activité antivirale de la Lf peut être directement reliée à sa capacité à réguler le système immunitaire. En effet, l'administration de bLf par voie intraveineuse à des rats immunodéprimés empêche l'infection par le cytomégalovirus (CMV) (Beljaars *et al.*, 2004). D'autre part, suite à l'administration orale de bLf chez des souris BALB/c infectées par le CMV, une nette augmentation de l'activité des cellules *Natural Killer* (NK) a été détectée, conduisant à l'augmentation de la survie des

souris traitées (Shimizu *et al.*, 1996). Lors des essais cliniques il a été montré que l'administration orale de Lf diminuait la charge virale chez des patients infectés par le virus de l'herpès traités par l'acyclovir (Andersen *et al.*, 2004) ainsi que l'amélioration des réponses immunes des enfants infectés par le VIH (Zuccotti *et al.*, 2007).

3) Activités antifongique et antiparasitaire

La majorité des recherches sur l'activité antifongique de la Lf est effectuée sur *Candida albicans*. L'activité anti-*candida* est associée à l'altération de la perméabilité de la membrane cellulaire qui est directement en interaction avec la Lf (Bellamy *et al.*, 1993; Wakabayashi *et al.*, 1996). D'ailleurs, la candidose buccale est associée à une diminution de sécrétion de Lf (65%) par les glandes salivaires (Tanida *et al.*, 2003). La prise orale de bLf réduit significativement la candidose buccale, allant jusqu'à la guérison totale de certains des animaux affectés (Takakura *et al.*, 2003). Il est à noter que l'activité antimycotique de la Lf semble en relation avec l'activation du système immunitaire (Yamaguchi *et al.*, 2004). Les mécanismes moléculaires de l'activité antiparasitaire de la Lf sont souvent complexes car les niches écologiques des pathogènes diffèrent souvent d'un organisme à l'autre. Il semblerait néanmoins que les interactions entre la Lf et les parasites tels que *Toxoplasma gondii* (Dzitko *et al.*, 2007) et *Entamoeba histolytica* (León-Sicairos *et al.*, 2006) ne les empêcheraient pas d'entrer au sein de la cellule hôte mais seraient responsables de l'inhibition de la croissance du parasite ou de l'altération de sa membrane (pour revues, Jenssen et Hancock, 2009; Pierce *et al.*, 2009). Pour exemple, il a été démontré que l'apo-hLF et l'apo-bLf privent les amibes *Entamoeba histolytica* du fer nécessaire à leur croissance et ce de manière dose-dépendante (León-Sicairos *et al.*, 2006).

B. Modulation de la réponse immune et inflammatoire

La Lf est un puissant modulateur de la réponse inflammatoire protégeant l'organisme des agressions de pathogènes mais jouant également un rôle clef dans les pathologies inflammatoires comme les allergies, l'arthrose et le cancer (pour revues Legrand *et al.*, 2008; 2005; Ward *et al.*, 2005).

1) Activité anti-inflammatoire

En interagissant avec les LPS (Appelmelk *et al.*, 1994; Elass-Rochard *et al.*, 1995) et de nombreux récepteurs à la surface des cellules endothéliales et immunes (Tableau II, page 19), la Lf module la production de cytokines et le recrutement de cellules immunes au site de l'inflammation et protège ainsi du choc septique (Kruzel *et al.*, 2002; 2010; Zimecki *et al.*, 2004). En effet, en réponse à une activation par le LPS au sein des cellules monocytaires *in vitro* et *in vivo*, la Lf inhibe la production de cytokines pro-inflammatoires telles que le TNF- α (*Tumor Necrosis Factor*) et les interleukines IL-1 β et IL-6 (Håversen, 2002; Kruzel *et al.*, 2002; Crouch et Fletcher, 1992; Mattsby-Baltzer *et al.*, 1996). Cela conduit à la sécrétion

de cytokines anti-inflammatoires telles l'IL-10 et l'IL-4 et au recrutement de cellules immunes, notamment des leucocytes dans le sang (Kurose *et al.*, 1994). De nombreuses études chez l'animal montrent qu'une prise orale de bLf protège efficacement contre des doses létales de LPS (Togawa *et al.*, 2002). Les études utilisant des modèles animaux d'inflammation expérimentale touchant différents tissus, montrent que l'administration orale de bLf réduit cette inflammation (Griesbeck-Zilch *et al.*, 2008; Zheng *et al.*, 2005; Togawa *et al.*, 2002). Par ailleurs, en cas de pneumonie, le nombre de leucocytes infiltrants est diminué en présence de Lf, supprimant ainsi l'hyperréaction de l'hôte.

Le mécanisme d'action illustré par la Figure 6 (page 21), passe par une inhibition des interactions de haute affinité entre le LPS et le CD14 membranaire (mCD14), récepteur majeur du LPS sur les macrophages-monocytes, leucocytes et lymphocytes (Elass-Rochard *et al.*, 1998; Japelj, 2005). Au laboratoire, Dominique Legrand et collaborateurs ont montré qu'il existe également des interactions de haute affinité entre la lactoferrine et le CD14 soluble provenant du sérum (sCD14) et entre la hLf et le complexe sCD14/LPS (Baveye *et al.*, 2000). L'interaction entre la hLf et le sCD14 est responsable de l'inhibition de l'expression de molécules d'adhésion telles que la E-sélectine, I-CAM1, VCAM et la chimiokine IL-8 dans les cellules endothéliales ombilicales (Elass *et al.*, 2002). En outre, la fixation et l'internalisation de la Lf au sein de cellules monocytaire est capable d'inhiber la fixation du facteur de transcription NF- κ B sur le promoteur du gène TNF- α et ainsi, de diminuer la production de cytokines induites par le LPS (Håversen, 2002). Récemment, Puddu *et al.* ont démontré que l'exposition constante de cellules monocytaire à de la bLf contribue à leur différenciation en cellules dendritiques. Ces cellules potentialisent l'activité anti-inflammatoire de la bLf en sécrétant de l'IL-6, (Puddu *et al.*, 2011). Enfin, l'administration intra-nasale de Lf est également responsable d'une diminution du recrutement de cellules éosinophiles et d'une réduction de l'inflammation induite par le pollen dans des modèles d'asthme murins (Kruzel *et al.*, 2006).

Tableau II : Les différents récepteurs de la Lf présents sur les cellules de mammifères et leurs fonctions.

Tissu, Cellule	Récepteur	Fonction	Référence
Petit intestin (Caco2)	Intélectine	Absorption du fer	(Suzuki <i>et al.</i> , 2001)
	LfR	Internalisation de la Lf	(Ashida <i>et al.</i> , 2004)
Foie	LRP des asialoglycoprotéines	Régulation de l'inflammation	(Ziere <i>et al.</i> , 1993)
			(McAbee <i>et al.</i> , 1998)
Lymphocytes	CD14	Maturation des lymphocytes	(Bi <i>et al.</i> , 1994)
Monocytes	CD14	Régulation de l'inflammation	(Baveye <i>et al.</i> , 2000; 2000)
Macrophages	TLR4	Régulation de l'inflammation	(Curran <i>et al.</i> , 2006; 2006)
Fibroblastes	LRP	Synthèse du collagène	(Takayama <i>et al.</i> , 2003)
Os	LRP	Mitogénèse des ostéoblastes	(Grey <i>et al.</i> , 2004)
Cerveau	LRP	Transport de la Lf	(Fillebein <i>et al.</i> , 1999)
Glande mammaire	Nucléoline	Internalisation de la Lf	(Legrand <i>et al.</i> , 2004)
Nombreux tissus	Protéoglycannes	Fixation de la Lf	(Legrand <i>et al.</i> , 2008)

2) Activité pro-inflammatoire

Les expériences *in vitro* réalisées par Shinoda *et al.* montrent que la Lf est responsable de la production de superoxydes du relargage de molécules pro-inflammatoires telles que le NO (*nitric oxyde*) (Shinoda *et al.*, 1996), TNF- α et IL-8 (Sorimachi *et al.*, 1997). En outre, des expériences réalisées chez l'animal montrent que la bLf est responsable de l'augmentation de l'activité des cellules NK (Shimizu *et al.*, 1996), de l'activation de la phagocytose (Kai *et al.*, 2002), de la stimulation de la maturation et de la différenciation des lymphocytes T, de la présentation d'antigènes aux cellules dendritiques (la Rosa *et al.*, 2008) et de la modulation de la myélopoïèse (Broxmeyer *et al.*, 1987; Sawatzki et Rich, 1989). De plus, la prise orale de bLf augmente le nombre de lymphocytes T CD4 $^{+}$ dans les tissus lymphoïdes (Dhennin-Duthille *et al.*, 2000; Artym *et al.*, 2003), et entraîne l'induction de la production d'IL-18 et la polarisation des lymphocytes en cellules T auxiliaires de type 1 (Th-1) (Suzuki *et al.*, 2005). Ces dernières conduisent ensuite à l'élimination du virus de l'hépatite C lors d'une thérapie par l'IFN- γ (interféron), favorisent la phagocytose de *Staphylococcus aureus* et de certains champignons et augmentent l'efficacité d'une vaccination par le BCG (Hwang *et al.*, 2005). Par ailleurs, l'ingestion de hLf augmente la production de splénocytes (Zimecki *et al.*, 1995), d'IFN- γ et d'IL-12 (Wakabayashi *et al.*, 2004) en réponse à une infection par le virus de l'herpès.

3) Activité anti-oxydante

De par sa capacité à séquestrer le fer libre, la Lf n'est pas uniquement responsable de l'inhibition de la croissance des microorganismes, elle est également impliquée dans la réduction du niveaux des espèces réactives de l'oxygène (ROS) et par là même, dans la diminution du stress oxydatif (Baldwin *et al.*, 1984). En effet, les niveaux de ROS engendrés par le pollen au niveau des cellules épithéliales bronchiques sont fortement réduits en présence de hLf qui empêche le développement des cellules productrices de mucines ainsi que la production de mucines elles-mêmes (Kruzel *et al.*, 2010; 2006).

Ainsi, les voies par lesquelles la Lf transmet ses effets semblent être toutes aussi variées que ses activités biologiques et dépendantes du type cellulaire étudié (Tableau III, page 27)

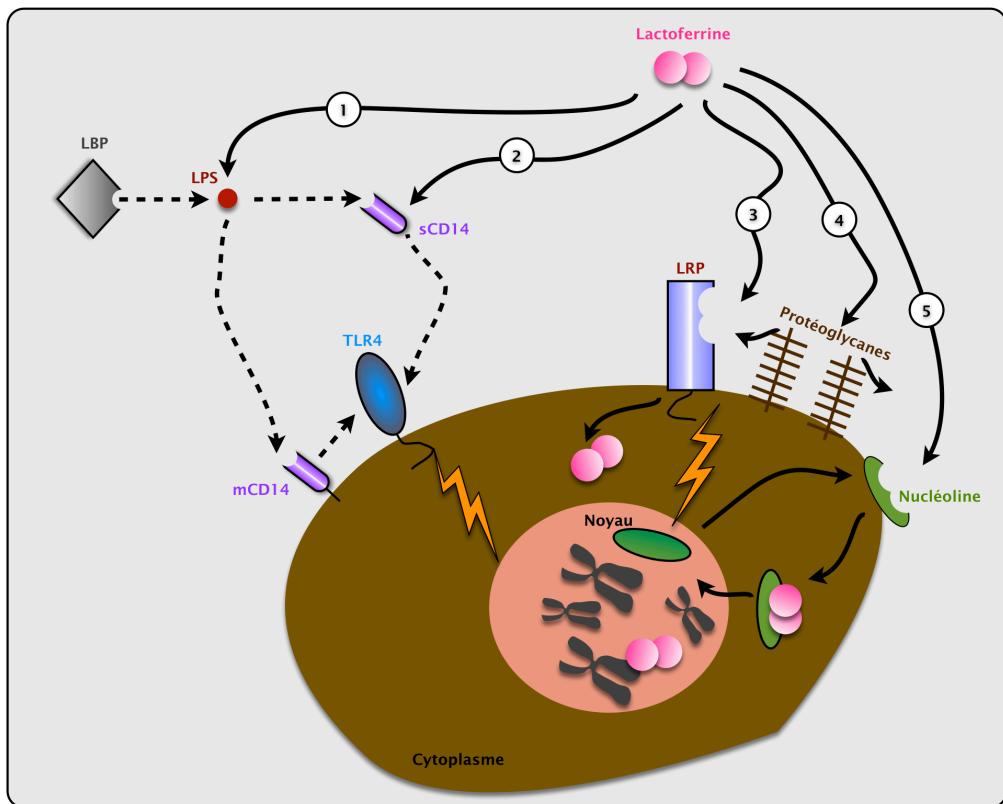


Figure 6 : Représentation schématique des interactions par lesquelles la lactoferrine pourrait exercer ses fonctions vis-à-vis des cellules immunes (schéma inspiré de (Legrand *et al.*, 2006)). Les LPS libérés dans la circulation sanguine s'associent à la LBP (*LPS-binding protein*) et au récepteur membranaire CD14 présent sur les monocytes et les macrophages. **1.** L'activation de ces cellules nécessitera l'interaction du complexe ternaire avec le récepteur TLR4 (*cell-signalling Toll-like receptor*). **2.** Les cellules endothéliales dépourvues de CD14 seront activées par la présentation au TLR4 du complexe LPS-LBP-CD14 soluble. Le CD14 soluble est produit soit par les cellules activées, soit par clivage enzymatique du CD14 membranaire. L'activation des monocytes/macrophages et des cellules endothéliales conduit à la production de cytokines proinflammatoires. La Lf, en piégeant le LPS, inhibe la formation des différents complexes et s'oppose aux effets délétères du LPS, offrant ainsi une protection contre le choc septique (Legrand *et al.*, 2005; 2006). Les flèches **3**, **4** et **5** illustrent les interactions de la Lf avec ses différents récepteurs de surface : le LRP (*Lipoprotein Receptor-related Protein*), les protéoglycanes et la nucléoline de surface respectivement. La fixation de la Lf au LRP est responsable de l'endocytose de la Lf et/ou de l'activation de voies de signalisations (éclairs oranges). La fixation de la Lf à la nucléoline conduit à une endocytose et un ciblage nucléaire des deux protéines (Legrand *et al.*, 2004).

C. Rôle de la lactoferrine dans le cancer et la prolifération cellulaire

La lactoferrine est impliquée dans la régulation de la prolifération cellulaire, de la tumorigénèse ainsi que dans la réduction de la formation de métastases.

1) Activité anticancéreuse

De nombreux travaux ont montré que les activités anticancéreuses sont partagées par la Lf et Δ Lf, bien que les deux molécules vont agir de manière différente, ceci dû à leur structure ainsi qu'à leur localisation. En effet, la Lf va utiliser différents mécanismes d'action pour exercer son rôle d'agent anticancéreux tels que des effets anti-angiogéniques, une régulation de la prolifération cellulaire ainsi qu'une immunomodulation. La Δ Lf, quant à elle,

exerce des effets antiprolifératifs et pro-apoptotiques *via* son activité de facteur de transcription.

Une des premières expériences montrant l'activité anti-tumeur de la Lf a été publiée en 1994. Dans cette étude, les auteurs ont montré que l'injection intratumorale de bLf à des souris présentant une tumeur chimiquement induite conduit à une diminution significative de sa taille ainsi qu'à une réduction du nombre de métastases (Bezault *et al.*, 1994). Par la suite, de nombreux modèles animaux de carcinogénèse chimique de différents organes ont été mis en place, et pour tous les animaux traités à la bLf ou hLf, une nette régression de la taille des tumeurs a pu être observée (Tsuda *et al.*, 1998; 2000; 2002; 2004; Ushida *et al.*, 1998; 1999; Tanaka *et al.*, 2000; Li *et al.*, 2011; Wang *et al.*, 2011a). Les hLf et bLf sont non seulement capable d'empêcher la formation de métastases *in vitro* mais également *in vivo* (Varadhachary *et al.*, 2004; Masuda *et al.*, 2000; Iigo *et al.*, 1999). Cependant, les mécanismes ne sont pas très clairs puisque des expériences *in vitro* ont montré que la hLf est responsable de la surexpression de la métalloprotéase MMP-1 (Oh, 2001) et de l'augmentation la capacité de migration et d'invasion des cellules de glande mammaire MCF7, ZR-75, MDA-MB-231, MDA-MB-468 et SK-BR-3 (Ha *et al.*, 2011).

a) Activité anti-angiogénique

L'injection intra-péritonéale ainsi que l'ingestion de bLf chez le rat, conduit à une inhibition de l'angiogenèse par l'intermédiaire de la sous-expression du VEGF (*Vascular Endothelial Growth Factor*) (Norrby *et al.*, 2001; Shimamura *et al.*, 2004). De la même manière, la bLf saturée en fer est un adjuvant naturel potentiel qui renforce les effets de la chimiothérapie en régulant négativement l'angiogénèse *in vivo* (Kanwar *et al.*, 2008). Curieusement, la bLf et la hLf semblent exercer des effets totalement opposés vis-à-vis de l'angiogénèse. Alors que l'administration orale de bLf inhibe l'angiogénèse (Norrby *et al.*, 2001), la hLf a plutôt des activités pro-angiogéniques (Norrby, 2004). L'activité anti-angiogénique de la Lf est encore sujette à controverse puisqu'elle semble dépendre de l'origine de la Lf utilisée pour le traitement ainsi que du type cellulaire étudié. D'ailleurs, Kim *et al.* ont montré que la bLf était capable de stimuler l'angiogénèse *in vivo*, *via* la surexpression du récepteur au VEGF, conduisant ainsi à la prolifération et la migration des cellules endothéliales (Kim *et al.*, 2006).

b) Régulation de la prolifération cellulaire

L'expression de la hLf et de son isoforme intracellulaire Δ Lf, est sous-régulée dans les cellules cancéreuses. La hLf est sous-exprimée ou absente de nombreux tissus cancéreux (Siebert et Huang, 1997; Benaïssa *et al.*, 2005). Cette sous-expression est due à des mutations, perte allélique du chromosome 3 ou modifications du degré de méthylation du promoteur (Iijima *et al.*, 2006; Teng *et al.*, 2004a). La Δ Lf est présente dans les tissus sains mais est sous-régulée dans les tissus cancéreux correspondants (Siebert et Huang, 1997 ; Liu

et al., 2003 ; Benaïssa *et al.*, 2005). Lf et Δ Lf semblent être des facteurs de bon pronostic dans le cas de cancer du sein puisqu'une forte expression de leurs transcrits est associée à une survie plus importante des patientes (Benaïssa *et al.*, 2005). Qui plus est, Lf et Δ Lf exercent une activité antiproliférative sur les cellules cancéreuses puisqu'elles sont capables de réduire la croissance tumorale.

En effet, La Lf limite la prolifération cellulaire des cellules MDA-MB-231 grâce à sa capacité à provoquer un arrêt du cycle cellulaire en G1/S. Cet arrêt du cycle passe par l'inhibition des cyclines cdk2 et cdk4 et l'induction de l'expression de la protéine p21 (*cdk inhibitor*) *via* la voie de signalisation des MAPK (Damiens *et al.*, 1999). Cette augmentation de p21 est également retrouvée après traitement des lignées cellulaires humaines H1299, MCF7 et HEK 293 aussi bien avec de la bLf que de la hLf (Son *et al.*, 2006). D'ailleurs, ces auteurs ont mis en évidence la capacité qu'a la hLf à surexprimer la protéine du rétinoblastome (Rb), suppresseur de tumeur clef impliqué dans la progression du cycle cellulaire. En effet, la hLf maintient la protéine Rb sous sa forme hypophosphorylée grâce à la surexpression de p21, contribuant ainsi à la formation du complexe Rb/E2F, donc à une inhibition de l'expression des gènes répondant à E2F qui sont responsables du passage de la transition G1 à S (Son *et al.*, 2006). Xiao *et al.*, (2004) ont également noté un arrêt du cycle en G1/S de cellules issues de résections de tumeurs de la tête et du cou en réponse à la Lf, *via* une diminution de la phosphorylation de la protéine Akt qui conduit à une surexpression de la protéine p27 et de la cycline E, ainsi qu'à une diminution de la phosphorylation de Rb. Varadhachari et collaborateurs ont mis en évidence un arrêt du cycle en G1/S dans des cellules squameuses cancéreuses de la tête et du cou en réponse à la hLf, avec une sous-expression de la cycline D (Wolf *et al.*, 2007). De plus, la surexpression de la hLf dans les cellules HeLa induit une surexpression du suppresseur de tumeur p53, mais également de mdm2 et de p21, due à l'activation du facteur de transcription NF- κ B (Oh *et al.*, 2004). La Lf agit donc sur de nombreuses voies pour contrôler la transition G1/S (Figure 7, page 25). Par ailleurs, des modèles animaux de carcinogénèse chimique au niveau de la glande mammaire (Wang *et al.*, 2011) et du col de l'utérus (Li *et al.*, 2011) ont été réalisés et les animaux ont subi ou non l'injection d'un vecteur adénoviral déficient en réPLICATION, contenant le gène de la hLf, directement au niveau de la tumeur. L'analyse des cellules primaires issues de ces tumeurs a aussi clairement mis en évidence un arrêt de la prolifération cellulaire des cellules traitées par rapport aux cellules contrôles, reflété par un arrêt du cycle en G1/S pour les cellules de glande mammaire (Wang *et al.*, 2011) et un arrêt en G2/M pour les cellules du col utérin (Li *et al.*, 2011).

Parallèlement, une étude réalisée au laboratoire a montré que la surexpression de la Δ Lf conduit à un arrêt du cycle cellulaire à la transition G1/S (Breton *et al.*, 2004). Cet arrêt

du cycle serait dû à la transactivation du gène *Skp1* par la Δ Lf (Mariller *et al.*, 2007). Activité pro-apoptotique

De nombreuses études ont montré que la Lf favorise l'apoptose des cellules cancéreuses. Des études *in vivo* ont montré que l'administration orale de bLf inhibe la tumorigénèse et augmente l'apoptose grâce à l'induction de l'expression des récepteurs de mort Fas, des protéines pro-apoptotiques Bax et Bid, membres de la famille Bcl-2, l'activation des caspases 8 et 3 ainsi que l'induction de la fragmentation de l'ADN caractéristique d'un état apoptotique (Fujita *et al.*, 2004a; 2004b; Chandra Mohan *et al.*, 2006). Par ailleurs, il a été démontré que le domaine $^{679}\text{Y-K}^{695}$ de la hLf est capable de se fixer à la pro-caspase 3 *in vitro* et de conduire à sa maturation. D'ailleurs, le peptide synthétique de ce domaine entre en compétition avec la hLf et bloque la maturation de la pro-caspase 3 (Katunuma *et al.*, 2006).

Des études *in vitro* ont montré que le traitement des cellules Jurkat par de la hLf conduit à une augmentation significative de l'apoptose de manière dose dépendante, avec une activation des caspases 9 et 3 ainsi qu'un clivage de la PARP (*Poly ADP Ribose Polymerase*) (Lee *et al.*, 2009). La voie de signalisation JNK1/2 est impliquée dans l'apoptose des cellules Jurkat en réponse à la hLf, ainsi qu'une phosphorylation de Bcl2 qui conduit à son inactivation. À de fortes concentrations de hLf dans ces cellules, Lee *et al.* (2010) ont pu observer une surexpression de Rb maintenue sous sa forme phosphorylée grâce à l'augmentation de l'expression de la kinase Cdk6 (Lee *et al.*, 2010b). Ceci inhibe l'interaction entre Rb et E2F, permettant ainsi l'expression du facteur pro-apoptotique Bim. Par ailleurs, l'injection de l'adénovirus contenant le gène de la hLf directement au niveau de tumeurs chimiquement produites chez la souris a montré une augmentation de l'expression de Bax, Fas et de la caspase 3 activée ainsi qu'une sous-régulation de Bcl2 (Li *et al.*, 2011 ; Wang *et al.*, 2011). D'autre part, les études réalisées sur les cellules neuronales PC12 montrent que la hLf est capable de promouvoir ou d'inhiber l'apoptose selon la dose appliquée. En effet, de fortes concentrations de hLf stimulent l'activation des caspases 8 et 3 et régulent négativement l'expression de Bcl2 et ERK1/2 sous sa forme phosphorylée. Au contraire, de faibles concentrations de hLf conduisent à une surexpression de ERK1/2 et Bcl2 et ne conduisent pas à l'apoptose des cellules traitées (Lin *et al.*, 2005; Lee *et al.*, 2009).

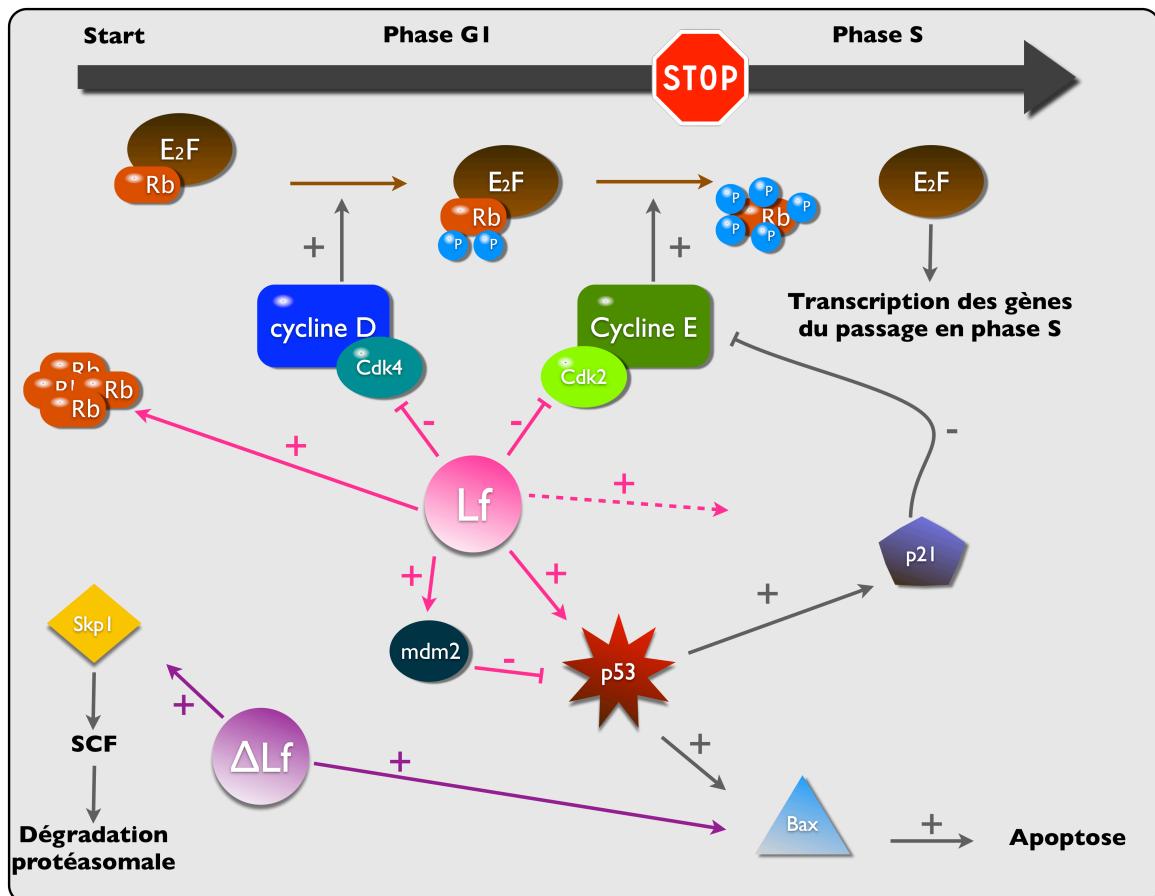


Figure 7 : Effets des lactoferrines sur la régulation de la transition G1/S. La lactoferrine intervient à différents niveaux de régulation de cette transition. La lactoferrine induit la surexpression de Rb qui s'accumule majoritairement sous une forme hypophosphorylée conduisant à la rétention de E2F. Elle induit également la surexpression de p53 qui en retour conduit à celle de p21. En activant p53, la Lf va également conduire à la transcription d'un facteur anti-angiogénique, la maspine. 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. La Δ Lf transactive les gènes *Bax* et *Skp1*. La protéine Skp1 appartient au complexe SCF à activité ubiquitine ligase responsable de la dégradation protéasomale de nombreux acteurs du cycle cellulaire de la transition G1/S (d'après Pierce *et al.*, 2009).

c) Immunomodulation

Les modèles animaux de carcinogénèse chimique traités à la bLf ont montré une augmentation du nombre de cellules NK, CD8⁺, CD4⁺ et INF- γ ⁺ au niveau des muqueuses de l'intestin ainsi que le relargage de cytokines telles que l'IL-18, l'IL-12 et l'INF- γ (Iigo *et al.*, 1999). En outre, l'administration orale de hLf recombinante à des souris atteintes de tumeur chimiquement induite conduit à une réponse immune mucosale et systémique (Varadhachari *et al.*, 2004). De plus, le relargage de la Lf par les cellules neutrophiles est considéré comme une alarmine recrutant les leucocytes et activant les cellules de l'immunité (la Rosa *et al.*, 2008; Yang *et al.*, 2009). En effet, une alarmine est définie comme étant un facteur cellulaire libéré par un stress chimique mécanique ou biologique, pouvant activer le processus inflammatoire et agissant comme un facteur de danger. Les alarmines sont reconnues par les cellules de l'immunité innée et de l'immunité acquise (Bianchi, 2006).

2) Activité proliférative de la lactoferrine

Le processus d'ossification est géré par des cellules de la moelle osseuse appelées ostéoblastes. Au contraire, il existe des cellules appelées ostéoclastes qui détruisent la substance osseuse. De manière intéressante, Cornish *et al.* ont montré que la Lf est un facteur anabolique pour l'os (Cornish *et al.*, 2004).

En effet, des études *in vivo* montrent que l'injection intraclaviculaire de bLf à des souris mâles adultes augmente la croissance osseuse de manière localisée avec une activité supérieure à la plupart des autres facteurs de croissance ou hormones ostéogéniques normalement utilisés dans les stratégies anti-ostéoporotiques ou de réparation de fracture osseuse (Cornish *et al.*, 2004). De plus, l'administration orale de bLf à des rats ayant subi une ovariectomie améliore la densité osseuse (Blais *et al.*, 2009; Guo *et al.*, 2009). Par ailleurs, de la bLf encapsulée dans des liposomes a été délivrée à des rats atteints de maladies parodontales induites par le LPS. Les auteurs ont ainsi pu constater une suppression de l'augmentation des ostéoclastes et une diminution de l'inflammation par l'intermédiaire du TNF α dans l'os (Ishikado *et al.*, 2005).

D'autre part, la bLf *in vitro*, stimule la formation de modules calcaires et augmente le nombre d'ostéoblastes matures issus de fémurs de rat (Cornish *et al.*, 2004; Takayama et Mizumachi, 2009; Blais *et al.*, 2009). En outre, la bLf est capable d'inhiber non seulement l'activité résorptive des ostéoclastes *in vitro* (Lorget *et al.*, 2002) mais également la différenciation des pré-ostéoclastes en ostéoclastes (Cornish *et al.*, 2006; Blais *et al.*, 2009). Ces données font de la bLf un agent thérapeutique prometteur contre l'ostéoporose.

Tableau III : Les lactoferrines bovines et humaines activent de nombreuses voies de signalisation dépendantes des tissus ou types cellulaires

	Cellules	Voies de signalisation	Références
<i>in vitro</i>	Cellules de la tête et du cou	PI3K/AKT	Xiao <i>et al.</i> , 2004
	MDA-MB-231	MAP kinase	Damiens <i>et al.</i> , 1999
	Cellules cancéreuses stomachales	AKT	Xu <i>et al.</i> , 2010
	HeLa	NFkB Récepteur au TGF-β et Smad	Oh <i>et al.</i> , 2004 Zemann <i>et al.</i> , 2010
	PC12	ERK, Fas P38 MAPK	Lin <i>et al.</i> , 2005 Ishii <i>et al.</i> , 2009
	Cellules primaires de glande mammaire bovine (MeBo), MCF7	Acides rétinoïques	Baumrucker <i>et al.</i> , 2006
	Jurkat	JNK-associée à Bcl-2	Lee <i>et al.</i> , 2009
	Jurkat, H1299, MCF7, HEK293	Hypophosphorylation de Rb	Son <i>et al.</i> , 2006
<i>in vivo</i>	Foie	maturation de la procaspase-3	Katunuma <i>et al.</i> , 2006
	Muqueuse du colon, poumon	Fas	Fujita <i>et al.</i> , 2004a
	Petit intestin	IFN-α/IL-7	Iigo <i>et al.</i> , 2009
	Glande mammaire	Fas, Caspase 3	Wang <i>et al.</i> , 2011
	Utérus	Caspase 3	Li <i>et al.</i> , 2011

D. Activité de facteur de transcription

1) Activité putative de facteur de transcription de la lactoferrine sécrétée

En 1989, Hutchens et collaborateurs ont mis en évidence l'existence d'une interaction entre la Lf et de l'ADN double brin et utilisé cette propriété pour de purifier et isoler la hLf issue de fluides biologiques comme le lait (Hutchens *et al.*, 1989). Garré *et al.* ont montré qu'après traitement des cellules érythrocytaires humaines K-562 par de la hLf, ils assistaient à la fixation de la hLf à ses récepteurs spécifiques de surface, une internalisation rapide puis une relocalisation nucléaire (Garré *et al.*, 1992). En outre, la partie N-terminale de la Lf expose une région fortement basique analogue au signal de localisation nucléaire de

SV-40 (Meier et Blobel, 1990). En effet, le ¹GRRRR⁵ contenant 4 résidus d'arginine consécutifs a été décrit comme étant un signal de localisation nucléaire (Penco *et al.*, 2001). He et Furmanski (1995) ont tenté de donner une signification biologique à cette présence de Lf sécrétée, interagissant avec l'ADN au sein du noyau. Pour ce faire, ils ont incubé la Lf dans des conditions stringentes en présence de séquences oligonucléotidiques aléatoires d'ADN. Les complexes Lf/ADN résultants ont été purifiés et après dissociation de l'ADN, ce dernier a été amplifié par PCR (*Polymerase Chain Reaction*). Soixante-quatre séquences d'ADN ont pu être isolées et l'alignement de ces séquences a permis de mettre en évidence trois séquences nucléotidiques consensus interagissant spécifiquement avec la Lf *in vitro* (Tableau IV). Parmi ces séquences, seules les séquences S1 et S3 permettent l'expression d'un gène rapporteur CAT (*Chloramphenicol Acetyl Transferase*) en présence de Lf ou d'une construction de Lf amputée de son peptide signal (proche de la Δ Lf). S1 (Tableau IV) est la séquence *cis* la plus efficace (He et Furmanski, 1995).

Tableau IV : 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	+

Le gène de l'IL-1 β humain possède au sein de son promoteur une séquence identique à S1 ($^{-3202}$ GGCACTTGC $^{-3193}$) et 4 séquences imparfaites, similaires à S1, situées en positions -3137 à -1043. Des expériences de retard sur gel ont permis de montrer que la hLf était capable de se fixer aux cinq séquences. La séquence la plus distale qui d'ailleurs est la seule séquence optimale, conduit à une activité transcriptionnelle *in vitro* plus importante (Son *et al.*, 2002). Par ailleurs, l'expression d'un gène rapporteur CAT en présence de fragments de différente taille de hLf a permis de montrer que ce sont les 90 premiers résidus d'acides aminés de la hLf (incluant la séquence de localisation nucléaire ¹GRRRR⁵) qui présentent la plus grande activité transcriptionnelle (Son *et al.*, 2002), suggérant que ce fragment devrait à la fois posséder un DBD (*DNA Binding Domain*) et un TAD (*TransActivation Domain*).

Récemment, le gène de l'endothéline-1 (*ET-1*) a été décrit comme étant une cible de l'activité transcriptionnelle de la hLf (Ha *et al.*, 2011). Après traitement de cellules de glande mammaire cancéreuses MDA-MB-231 par de la hLf, les extraits nucléaires ont été soumis à des expériences de retard sur gel. Les résultats obtenus mettent en évidence l'existence d'une interaction spécifique entre la hLf et la séquence d'ADN GGCACCTTGG contenue dans le promoteur du gène *ET-1* (Ha *et al.*, 2011) qui est identique à la séquence consensus S1 décrite par He et Furmanski (1995).

La localisation nucléaire de la Lf est extrêmement controversée. En effet, sa présence n'a pas été détectée au sein des noyaux des cellules Jurkat, K-562, Raji, U-937, HL-60, HeLa et Hec-1B traitées par de la Lf-GFP (Liu *et al.*, 2003). De la même manière, après traitement, il ne semble pas y avoir de relocalisation de Lf-GFP dans les noyaux des cellules embryonaires de fibroblastes de souris (Goldberg *et al.*, 2005). Par ailleurs, Mulligan *et al.*, infirment non seulement la présence de hLf dans les noyaux des cellules K-562 mais également sa fixation sur des séquences spécifiques et consensuelles d'ADN (Mulligan *et al.*, 2006). Au contraire, sept autres équipes confirment l'existence d'une endocytose de Lf suivie d'une translocation au noyau. En effet, une faible quantité de hLf a pu être détectée par immunofluorescence au sein des noyaux des cellules monocytaire Thp1 (Håversen, 2002) ainsi que par immunohistochimie dans des cellules provenant de tumeurs hépatiques (Tuccari *et al.*, 2002). Suite à l'interaction avec son récepteur spécifique à la surface des cellules épithéliales intestinales Caco-2 (Tableau II, page 19), côté apical, la hLf est internalisée puis relocalisée au noyau (Ashida *et al.*, 2004). Baumrucker *et al* (2006) ont mis en évidence la présence de bLf préalablement marquée au FITC dans les noyaux de cellules de glande mammaire MCF-7. Par ailleurs, suite à la fixation de la hLf aux protéoglycans de surface et à la nucléoline de surface, la hLf est endocytée dans les endosomes précoces, de manière clathrine dépendante et relocalisée dans le noyau de cellules MDA-MB-231 et CHO. La nucléoline est une protéine nucléolaire de 105 kDa faisant la navette entre la surface cellulaire et le noyau (Legrand *et al.*, 2004). Recemment, Puddu *et al.* (2011) ont montré par microscopie confocale à l'aide d'anticorps dirigés contre la bLf, qu'elle est internalisée dans les monocytes. En effet, dix minutes après l'endocytose, les auteurs ont pu détecter la bLf dans le cytoplasme, une heure après, elle s'accumule dans la zone périnucléaire et enfin, trois heures après elle est retrouvée dans le noyau. Cependant, les mécanismes d'internalisation restent encore à déterminer.

La localisation nucléaire de la Lf serait donc vraisemblablement cellule-dépendante mais également dépendante des récepteurs de surface présents (Tableau II, page 19). Cette activité de facteur de transcription n'a cependant jamais été démontrée *in vivo* pour la Lf de sécrétion.

2) La delta-lactoferrine est un facteur de transcription fonctionnel

Comme tout facteur de transcription, la ΔLf devrait posséder non seulement un DBD mais également un domaine de transactivation (TAD). Bien que l'extrémité N-terminale de la ΔLf soit tronquée, elle conserve l'extrémité C-terminale de la première hélice (résidus 27 à 30 de la lactoferrine, 2 à 5 pour la delta-lactoferrine) qui constitue l'un des sites 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 seconde région basique présente au niveau de la région inter-lobe a également été proposée comme un autre DBD (Figure 8, page 31). L'entrée de la Δ Lf au noyau va dépendre de la reconnaissance par des importines d'une séquence NLS (*Nuclear Localisation Sequence*) autre que le $^1\text{GRRRR}^5$ précédemment décrit comme NLS pour la hLf. L'analyse *in silico* de la séquence primaire des Lf a révélé la possible présence d'une séquence NLS bipartite conservée parmi les Lfs de différentes espèces. Il s'agit d'une séquence NLS bipartite $^{417}\text{RRSDTSLTWN}^{\text{SVKGKK}}^{432}$ localisée au niveau du lobe C de la protéine (Figure 8, page 31). Elle est proche des séquences NLS fonctionnelles retrouvées pour Rb (Efthymiadis *et al.*, 1997), IL-5 (Jans *et al.*, 1997) et la nucléoplasmine (Robbins *et al.*, 1991). Cette séquence est pleinement fonctionnelle pour la Δ Lf puisqu'au laboratoire, des expériences d'immunorévélation et de microscopie confocale ont révélé que la Δ Lf avait une localisation essentiellement cytoplasmique mais également nucléaire (Breton *et al.*, 2004). Par ailleurs, l'utilisation d'une protéine de fusion Δ Lf-GFP a clairement démontré que la localisation subcellulaire de la Δ Lf pouvait être nucléaire (Liu *et al.*, 2003 ; Goldberg *et al.*, 2005 ; Hardivillé *et al.*, 2010). Enfin, la mutation des deux résidus d'acides aminés de part et d'autre de cette séquence bipartite conduit à une inhibition du trafic nucléaire de la Δ Lf (Mariller *et al.*, 2007).

Une étude réalisée au laboratoire a permis de démontrer, pour la première fois, que la Δ Lf est un facteur de transcription responsable de la transactivation du gène *Skp1* (Mariller *et al.*, 2007). *Skp1* est une protéine qui participe à la formation de deux complexes différents agissant au niveau de la transition G1/S du cycle cellulaire et au cours de la mitose (Kitagawa et Hieter, 2001). La Δ Lf est également capable de transactiver le gène de *Bax* (Hardivillé, communication personnelle). Les promoteurs des gènes *Bax* et *Skp1* présentent des séquences homologues à S1, que nous avons pu définir comme étant des éléments de réponse fonctionnels à la Δ Lf (*Δ Lf Response Element*, *Δ LFRE*) grâce à des expériences de gènes rapporteurs et d'immunoprécipitation de la chromatine (Mariller *et al.*, 2007 ; Hardivillé *et al.*, 2010).

La majorité des facteurs de transcription sont la cible de modifications post-traductionnelles qui vont contrôler leur activité, leur localisation subcellulaire, leur durée de demi-vie ainsi que leurs interactions avec d'autres protéines. Les travaux réalisés au laboratoire ont permis de mettre en évidence les modifications qui pouvaient contrôler l'activité transcriptionnelle et la stabilité de la Δ Lf. En effet, la phosphorylation du résidu de S¹⁰ conduit à une interaction entre la Δ Lf et le promoteur du gène *Skp1*, ainsi qu'à une forte activité transcriptionnelle, alors que la modification par O-GlcNAc du même site empêche la fixation de la Δ Lf au promoteur de *Skp1* et conduit à une forte diminution de la transactivation (Hardivillé *et al.*, 2010).

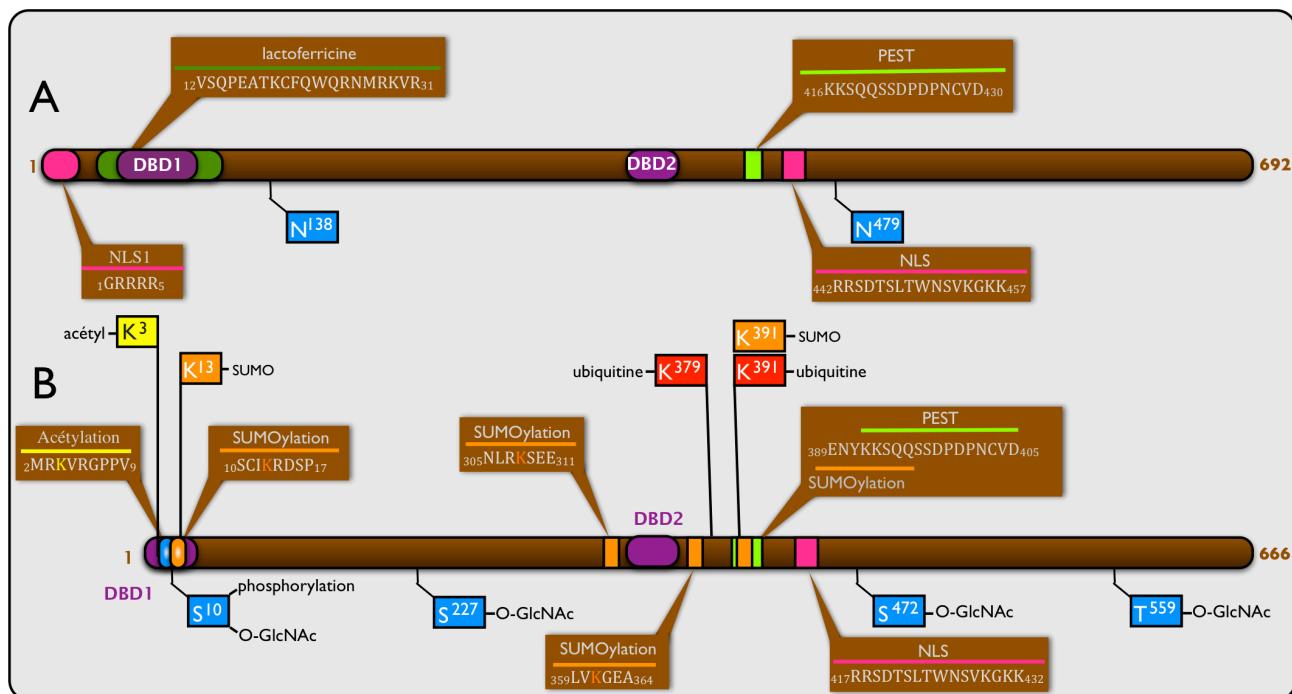


Figure 8 : Représentation schématique de la structure primaire de la lactoferrine (A) et de la delta-lactoferrine (B) présentant les différents domaines ainsi que leurs positions respectives. Les sites de N- et O-glycosylation respectivement pour la lactoferrine et la delta-lactoferrine sont représentés dans des carrés bleus. Les sites d'ubiquitination sont représentés dans les carrés rouges. L'acide aminé cible de l'acétylation est représenté en jaune, celui de la SUMOylation est représenté en orange. Les DBD1 et DBD2 sont représentés par des boîtes violettes. Les séquences PEST et NLS sont mises en évidence par une ligne horizontale verte clair et rose respectivement (D'après Mariller *et al.*, 2012).

Par ailleurs, la difficulté à isoler la Δ Lf, même lorsqu'elle est surexprimée, laisse suggérer que cette protéine est rapidement dégradée. L'analyse de la séquence primaire nous a permis de mettre en évidence un domaine de dégradation rapide (PEST) potentiel 391PVLAENYKSQQSSDPDPNCVD405 en position C-terminale. Les résidus de sérine ou de thréonine des séquences PEST sont fréquemment la cible d'une phosphorylation qui conduit généralement à la reconnaissance du substrat par le complexe SCF (Skp1-Cul1-F-box-protein) à activité ubiquitine ligase, signal de dégradation protéasomale de la protéine (Rechsteiner et Rogers, 1996). La mutation de cette séquence PEST augmente la durée de demi-vie de la Δ Lf, ce qui confirme sa fonctionnalité (Hardivillé *et al.* 2010). Cette séquence de dégradation rapide, reconnue par le système ubiquitine/protéasome, est conservée phylogénétiquement parmi les Lf. Nous avons d'autre part démontré que la GlcNAcylation du résidu de S¹⁰ protège la Δ Lf de la polyubiquitylation (Hardivillé *et al.*, 2010). Les résidus de K³⁷⁹ et K³⁹¹ proches de la séquence PEST, sont les cibles de l'ubiquitylation, leur mutation nous a permis de montrer que le résidu de K³⁷⁹ est la cible majeure de la polyubiquitylation. Deux autres séquences de dégradation rapide putatives sont retrouvées et conservées phylogénétiquement parmi les Lf, il s'agit de deux D-Box de séquence 108RPFL¹¹¹ et 475RSNL⁴⁷⁸ dont la fonctionnalité reste à confirmer. 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).

D'autre part, nous avons voulu savoir si la ΔLf pouvait être modifiée par SUMOylation. La recherche *in silico* a révélé 4 sites putatifs de SUMOylation aux positions 13, 308, 361 et 391. SUMO (*Small Ubiquitin-like Modifier*) est un membre de la famille des ubiquitines (Johnson, 2004; Kirkin et Dikic, 2007). La SUMOylation est une modification post-traductionnelle dynamique et réversible, impliquant des voies de SUMO-conjugaison et SUMO-déconjugaison qui conduisent à une modification de la fonction de la protéine cible. Les principales cibles de la SUMOylation sont les protéines impliquées dans le transport nucléaire, la répression de la transcription ou encore la réparation de l'ADN (Gareau et Lima, 2010; Hannoun *et al.*, 2010). Nous avons pu déterminer à l'aide d'expériences d'immunoprecipitation que la ΔLf est effectivement SUMOylée (Figure 9) (Hoedt *et al.*, communication personnelle). Des mutants ponctuels ont été construits afin de vérifier la fonctionnalité de chacun des 4 sites.

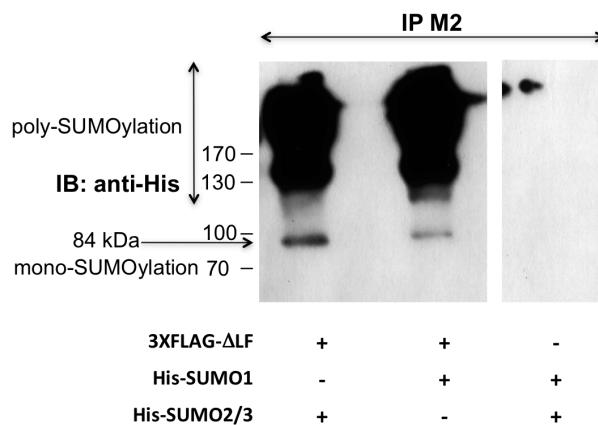


Figure 9 : La delta-lactoferrine est modifiée par SUMOylation. Les cellules HEK 293 ont été cotransférées avec les vecteurs d'expression 3XFLAG- ΔLf et His-SUMO. Les anticorps M2 dirigés contre l'étiquette 3XFLAG ont été utilisés afin d'immunoprecipiter la ΔLf 24 h après la transfection. Les fractions purifiées ont été analysées par immuno-électrotransfert avec des anticorps dirigés contre l'étiquette histidine. Il existe des motifs de polySUMOylation avec une première bande à 84 kDa qui correspond à la ΔLf monoSUMOylée et les autres bandes reflètent la polySUMOylation de la ΔLf .

La ΔLf possède également un site putatif d'acétylation en position N-terminale, au niveau du résidu de K³, cible potentielle des acétyltransférases (Figure 8, page 31). L'acétylation est une modification post-traductionnelle réversible qui neutralise la charge de la chaîne latérale du résidu de Lysine cible, entraînant ainsi un changement de conformation de la protéine cible. Cette modification modifie donc l'interaction de la protéine cible avec ses partenaires intracellulaires, en particulier l'ADN qui est chargé négativement (pour revue : (Arif *et al.*, 2010). De plus, il a été démontré que l'acétylation d'un résidu de lysine spécifique du facteur de transcription p53 est responsable de la modification conformationnelle de p53 qui est à l'origine du recrutement de co-activateurs transcriptionnels (Mujtaba *et al.*, 2004; Bode et Dong, 2004). L'acétylation d'un facteur de transcription peut entrer en compétition

avec sa phosphorylation et sa SUMOylation et ainsi induire une réactivation transcriptionnelle (pour revue : (Yang et Seto, 2008).

IV. Analyse protéomique, définition et origines

A. Définition

Le terme protéome est défini comme l'ensemble des protéines exprimées par le génome. Il a été employé pour la première fois en 1994 par Marc Wilkins lors d'un congrès et publié en 1996 (Wilkins *et al.*, 1996). L'analyse protéomique correspond à l'étude de l'ensemble des protéines exprimées par un génome pour une cellule, un tissu, un organe ou organisme à un moment donné et dans des conditions données. A la différence du génome qui reste constant (si l'on ne tient pas compte du polymorphisme génétique) dans les cellules d'un organisme, le protéome varie selon le type de cellules, l'activité des cellules ou leur environnement. En effet, en réponse à différents signaux, les protéines peuvent subir des modifications post-traductionnelles, changer de compartiment cellulaire, être synthétisées ou dégradées. En conséquence, l'étude du protéome doit s'effectuer dans un cadre bien défini, afin d'obtenir une meilleure compréhension de la complexité du fonctionnement cellulaire à partir de l'expression protéique dans un contexte exhaustif. Le volume de travaux publiés depuis l'apparition de l'analyse protéomique en 1995 ne cesse de croître et ce, de manière exponentielle (Figure 1). La première publication en analyse protéomique date de 1995 et l'on pouvait dénombrer plus de 12000 pour l'année 2010.

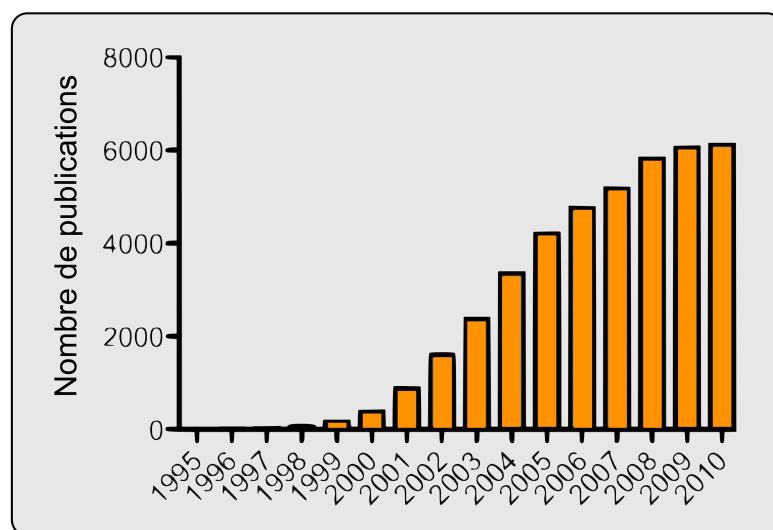


Figure 10 : Publications dans le domaine de l'analyse protéomique en recherchant les mots *proteomic* ou *proteomics* ou *proteome* depuis la base de donnée de NCBI Pubmed (<http://www.ncbi.nlm.nih.gov/pubmed>).

B. Origines

Les prémisses de l'analyse protéomique sont apparus en 1975 lorsque O'Farrell a décrit pour la première fois la séparation d'un mélange complexe de protéines d'*Escherichia coli*

grâce à l'électrophorèse bidimensionnelle (2-D) (O'Farrell, 1975). Néanmoins, les techniques de l'époque ne permettaient pas d'identifier aisément les protéines dans ce contexte. Ce n'est que dans les années 1990 qu'on a pu assister à un réel avènement de l'analyse protéomique avec l'identification de protéines dans un contexte intégré. Sans le développement simultané de techniques complémentaires telles que la séparation des protéines (électrophorèse bidimensionnelle, chromatographie liquide), la bio-informatique (bases de données en ligne, algorithmes de recherche,...) et la spectrométrie de masse (amélioration de la sensibilité, la résolution, la précision de masse,...), l'émergence de l'analyse protéomique n'aurait pu avoir lieu. De plus, l'augmentation exponentielle de génomes séquencés disponibles a largement contribué à cet essor.

Le premier séquençage complet d'un génome fut celui d'*Haemophilus influenzae*, publié en 1995 (Fleischmann *et al.*, 1995). Depuis, les outils de l'analyse génomique ont permis le séquençage de nombreux génomes dont celui de l'Homme en 2003. Le génome humain est composé de 23 000 gènes qui codent pour environ 500 000 protéines. En effet, si le génome fournit l'ensemble des gènes de chaque cellule d'un organisme, la réalité cellulaire dépendra du choix des gènes exprimés, du niveau d'expression de leurs produits ainsi que de la modulation de la fonctionnalité des protéines, par exemple par la maturation des ARN messagers et les modifications post-traductionnelles (glycosylation, phosphorylation, SUMOylation, ...). De plus, les protéines n'agissent pas dans la cellule de façon isolée mais le plus souvent sous forme de complexes. Il est donc clair que la seule séquence génique ne suffit pas à rendre compte de toutes les protéines exprimées à un moment donné dans la cellule. L'analyse génomique et l'analyse protéomique sont des domaines complémentaires qui, avec le séquençage complet du génome humain, ont ouvert un nouveau système de référence en biologie : les *omics*. Le terme *omic* réfère à la biologie qui s'intéresse aux interactions dans et entre des ensembles vivants complexes (espèces, populations, individus, cellules, protéines, ARN, ADN). C'est surtout l'introduction de nouvelles technologies qui a permis de développer les sciences *omics*. Ainsi, grâce à l'analyse génomique, l'analyse transcriptomique, l'analyse protéomique et l'analyse métabolomique, de nouveaux biomarqueurs de maladie et de nouvelles cibles thérapeutiques peuvent être identifiés (Ricquier, 2005; Adourian *et al.*, 2008; Zamboni *et al.*, 2010; Griffin *et al.*, 2011; Wu *et al.*, 2011).

V. L'analyse protéomique

La démarche classique de l'analyse protéomique consiste, dans un premier temps, à séparer les protéines présentes dans une cellule ou dans un compartiment cellulaire. Après extraction des protéines, ces dernières sont prétraitées : la solubilisation, la dénaturation et

la réduction permettent de supprimer toute interaction potentielle entre les protéines et d'éliminer le contenu non protéique. L'étape suivante consiste à séparer avec la meilleure résolution possible les protéines (ou peptides si la digestion a déjà été réalisée) présentes grâce à des méthodes séparatives (électrophorèse 1-D/2-D, chromatographie liquide), afin de les identifier et de les quantifier. Pour cela, chaque protéine d'intérêt est donc hydrolysée en fragments (c.à.d. peptides) par une protéase dont la spécificité est connue. La spectrométrie de masse permet de mesurer précisément la masse des différents peptides obtenus, lesquels constituent dans leur ensemble, la signature de la protéine (approche dite *d'empreinte peptidique massique*). Finalement, l'identification des protéines par interprétation des résultats des analyses de spectrométrie de masse est réalisée grâce à des outils bioinformatiques.

A. Fractionnement des protéines et des peptides

Les techniques électrophorétiques et chromatographiques sont les deux principales approches utilisées en analyse protéomique pour séparer les protéines.

1) L'électrophorèse bidimensionnelle (2-D)

Depuis la description originale de la 2-D par les travaux de O'Farrell en 1975, de nombreuses études ont contribué à l'amélioration de cette technique de séparation des protéines. Dès 1977, Anderson et Anderson ont utilisé l'électrophorèse 2-D pour séparer environ 300 protéines du plasma humain mais avec une résolution encore insuffisante (Anderson et Anderson, 1977). L'électrophorèse 2-D repose sur une séparation en deux étapes basées sur les deux principales caractéristiques physico-chimiques intrinsèques des protéines : leur point isoélectrique et leur masse moléculaire. La première dimension consiste en une isoélectrofocalisation (IEF). Les protéines sont séparées en fonction de leur point isoélectrique (pI) grâce à un gradient de pH utilisant des ampholytes, libres ou liées, dans un gel d'acrylamide. La fixation du gradient de pH au gel, qui passe par un greffage d'ampholytes directement sur l'acrylamide par des liaisons covalentes, a été mise au point afin de palier aux problèmes récurrents de reproductibilité. Ce système d'*immobilines* utilise des bandelettes à gradient de pH immobilisé, commercialisé sous le nom d'*Immobilized pH Gradient* (IPG), a aussi pour intérêt de proposer des gradients de pH étroits ou non linéaires, adaptés à la résolution d'un mélange à l'échelle du protéome (Görg *et al.*, 1985; 1988). Sous l'effet du champ électrique, les protéines vont donc migrer dans le gel selon leur charge nette. On peut ainsi focaliser les protéines sur une gamme allant de 3 à 10 unités de pH ou sur des gammes plus ciblées. Vient ensuite la seconde dimension qui consiste en une séparation en électrophorèse sur gel de polyacrylamide, en présence de dodécylsulfate de sodium (*Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis*, SDS-PAGE). Dans cette seconde étape, l'ordre de migration n'est pas déterminé par les charges électriques des protéines mais par

leur taille. Le SDS, qui est un détergent anionique, va se fixer sur les protéines, masquer leur charge propre et uniformiser leur forme. Ainsi, la séparation des protéines se fera principalement en fonction de la longueur de la chaîne polypeptidique.

A l'issue des deux étapes, les protéines sont séparées en fonction de leurs caractéristiques (pI et masse moléculaire) sur la surface du gel. Il existe différentes techniques de mise en évidence des protéines sur le gel. Chacune d'elle a des caractéristiques propres, en terme de sensibilité de détection et de linéarité de quantification. Le choix va donc dépendre de l'application ultérieure. La coloration à l'argent est considérée comme la référence en ce qui concerne la sensibilité (détection de l'ordre du ng de protéines) afin de visualiser les protéines minoritaires, mais elle présente une faible reproductibilité et elle est peu compatible avec l'analyse par spectrométrie de masse sans un protocole adapté (Rabilloud *et al.*, 2008). D'autres colorants comme le bleu de Coomassie sont utilisés pour la bonne linéarité de leur réponse, leur reproductibilité et leur bonne compatibilité avec l'analyse en spectrométrie de masse mais disposent d'une sensibilité moindre (détection de l'ordre de 10 ng de protéines) (Lin *et al.*, 2008). Les fluorochromes tels le *SYPRO Ruby* et le *Deep Purple* représentent actuellement le meilleur compromis car ils permettent une coloration avec une sensibilité équivalente à celle du nitrate d'argent, linéaire, homogène et compatible avec l'analyse par spectrométrie de masse. Néanmoins, ces réactifs sont couteux et nécessitent un appareillage spécifique pour la visualisation (Chakravarti *et al.*, 2010).

Par conséquent, la combinaison de ces deux séparations permet d'obtenir une carte 2-D où chaque spot représente une protéine de l'échantillon initial. Elle permet aussi de différencier les isoformes d'une même protéine en utilisant simultanément les deux principaux paramètres des protéines, leur masse moléculaire et leur point isoélectrique.

L'électrophorèse 2-D est une technique très utile en analyse protéomique car elle permet de séparer les protéines avec une très bonne résolution, jusqu'à 2000 protéines pouvant être analysées en routine. Une électrophorèse à haute résolution est capable de séparer 5000 protéines en une seule expérience (Wildgruber *et al.*, 2000). Cependant, il s'agit d'une approche longue à mettre en œuvre et difficile à établir en terme de reproductibilité. De plus, elle est peu adaptée à l'analyse des protéines membranaires (Rabilloud, 2002) et ne permet pas d'analyser les protéines de très haut ou très bas poids moléculaire (<10kDa ou >100kDa), ainsi que les protéines de haut ou bas pH (pI<4 ou >9) (Corthals *et al.*, 2000).

2) La séparation des peptides par chromatographie liquide haute pression (CLHP)

La chromatographie liquide haute pression (CLHP) est définie en fonction de la nature de sa phase stationnaire. Il existe des chromatographies de partage, d'exclusion, d'adsorption, d'affinité et d'échange d'ions. En analyse protéomique, les molécules à séparer sont

classiquement des peptides issus de la digestion trypsique d'un mélange plus ou moins complexe de protéines (ces peptides proviennent de la digestion d'un spot de gel 2-D, d'une bande de gel 1-D voire de la digestion d'un mélange plus complexe). Des colonnes chromatographiques en phase inverse (*Reverse Phase*, RP) sont communément utilisées. Ces colonnes sont composées de billes de silice greffée à des chaînes linéaires de 8 à 18 atomes de carbone qui leur confèrent un caractère très hydrophobe (phase stationnaire). Elles séparent les peptides trypsiques (ou peptides prototypiques) en fonction de leur hydrophobie grâce à un gradient d'un solvant organique (phase mobile) qui est généralement de l'acetonitrile. Ces colonnes ont été miniaturisées afin de travailler avec des débits plus faibles, de l'ordre de quelques nl/min, de manière à augmenter l'efficacité de l'analyse en spectrométrie de masse et ce, grâce à un couplage de la chromatographie liquide (LC) avec la spectrométrie de masse (LC-MS), on parle alors de nano-LC (Calvio *et al.*, 1995; Davis *et al.*, 1995). Les peptides élués en sortie de colonne sont directement envoyés à la source du spectromètre de masse pour être ionisés. Les faibles débits d'injection, sont préférables pour une ionisation de type électro-nébulisation (Shen *et al.*, 2002). Ainsi, à volume d'échantillon constant, cette diminution des débits d'entrée permet d'avoir un temps d'analyse plus long et donc d'améliorer la précision de mesure du spectromètre de masse. Cette miniaturisation présente également d'autres avantages tels qu'une amélioration de la tolérance aux sels et une augmentation de la sensibilité (Wilm et Mann, 1996).

La complexité des échantillons nécessite parfois plusieurs étapes de séparations en chromatographie avant l'analyse par MS. Il existe une approche basée sur une double séparation en chromatographie appelée MudPIT pour *Multidimensional Protein Identification Technology*. Elle consiste à séparer les peptides dans un premier temps sur une colonne échangeuse de cations qui est séquentiellement éluée sur une colonne en phase inverse, cette fois-ci, directement couplée au spectromètre de masse (McCormack *et al.*, 1997; Wolters *et al.*, 2001).

VI. Analyse protéomique et spectrométrie de masse

Afin d'identifier et d'étudier l'expression des protéines dans différentes conditions physiologiques ou pathologiques, l'analyse par spectrométrie de masse (MS) est devenue incontournable en protéomique. Un spectromètre de masse permet de séparer, de détecter et d'identifier des protéines (ou autres analytes) grâce à la mesure de leur masse monoisotopique. De plus, le spectromètre de masse va permettre de caractériser la structure chimique de ces molécules en les fragmentant. Après ionisation des molécules en phase gazeuse, les ions sont envoyés, grâce à un champ électrique, vers l'analyseur qui les sépare en fonction de leur rapport masse/charge (m/z). Le détecteur envoie alors l'information vers

l'ordinateur pour analyse (Figure 11). Les méthodes d'ionisation fréquemment utilisées pour les biomolécules sont la désorption-ionisation laser assistée par matrice (*Matrix-Assisted Laser Desorption/Ionisation*, MALDI) et l'électro-nébulisation (ou *electrospray*, ESI) car elles engendrent peu ou pas de fragmentation de la molécule durant le processus d'ionisation et/ou de désorption. Elles sont couplées à des analyseurs de plus en plus résolutifs et sensibles, rendant ainsi possible l'analyse de très faibles quantités d'échantillon (inférieures à la femtomole).

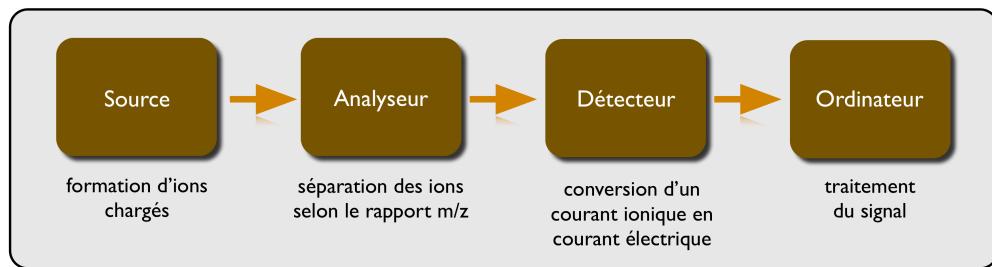


Figure 11 : Principe général de fonctionnement du spectromètre de masse.

A. Les sources d'ionisation en spectrométrie de masse

La molécule doit être ionisée car il est bien plus simple de manipuler de manière expérimentale le mouvement et la direction d'un ion par rapport à une molécule neutre. En appliquant des forces électriques et/ou magnétiques, l'énergie et la vitesse des espèces ioniques peuvent être contrôlées, ce qui va aider à leur séparation et leur détection.

1) L'ionisation par un spectromètre de masse à source MALDI

La spectrométrie de masse par désorption et ionisation laser assistée par matrice (MALDI) a été développée en 1988, simultanément par deux équipes de recherche en Allemagne (Karas et Hillenkamp, 1988) et au Japon (Tanaka *et al.*, 1988). Toutes deux ont montré que la MS peut analyser les protéines et d'autres molécules ayant une masse supérieure à 200 kDa avec une bonne sensibilité.

La première étape en MALDI consiste à co-cristalliser sur une cible l'analyte avec un large excès de matrice qui absorbe l'énergie à la longueur d'onde du laser employé. Elle va pouvoir ensuite transmettre de manière douce et contrôlée l'énergie absorbée à l'analyte. En analyse protéomique, trois matrices sont principalement utilisées : l'*alpha-cyano* (acide α -cyano-4-hydroxycinnamique), l'acide sinapinique (acide 3,5-dimethoxy-4-hydroxy cinnamique) et le DHB (acide 2,5-dihydroxybenzoïque).

L'ionisation MALDI (Figure 12, page 39) génère généralement des ions monochargés, ce qui facilite l'analyse des spectres de masse (Karas *et al.*, 2000). Les peptides ainsi ionisés seront ensuite accélérés par l'application d'une différence de potentiel en vue d'être séparés par l'analyseur à temps de vol (Figure 15, page 41).

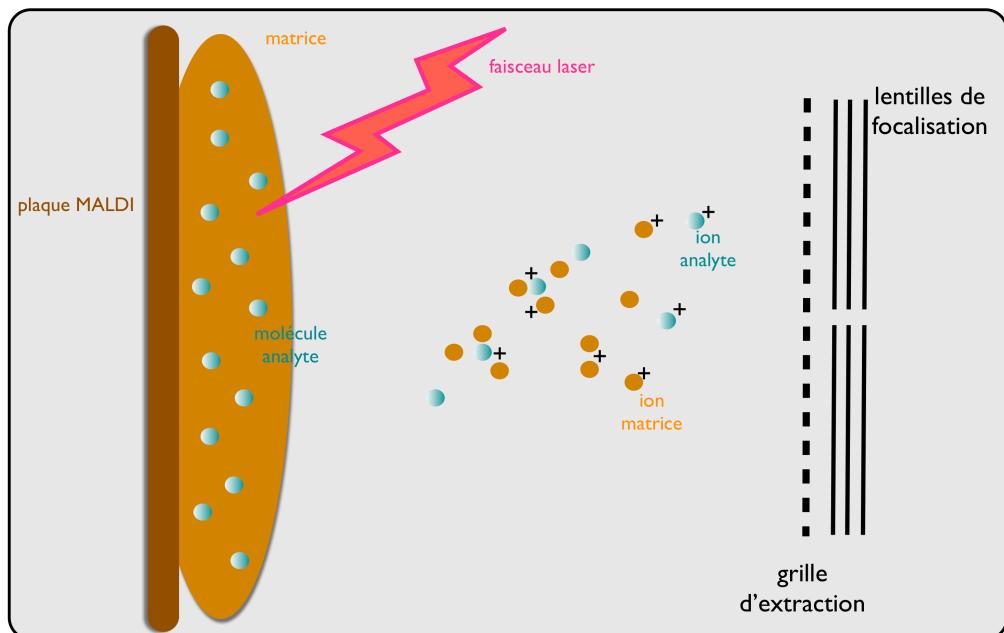


Figure 12 : Principe de l'ionisation MALDI. L'ionisation MALDI s'effectue sous vide ($\approx 10^{-7}$ mbar). Les ions sont produits en phase gazeuse sous l'effet d'une impulsion laser qui irradie la surface cristalline de la matrice entourant l'analyte. L'ionisation s'effectue en deux phases. La première consiste en l'ionisation de la matrice résultant de l'absorption de photons émis par le laser. L'ionisation de l'échantillon se réalise dans un second temps grâce au transfert d'un proton de la matrice à l'analyte.

2) Ionisation par un spectromètre de masse à source ESI

L'ionisation par *electrospray* (ESI) a été découverte par Dole *et al.* en 1968 (Dole, 1968). Cette technique fut développée par les travaux de Fenn à la fin années 80 (Fenn *et al.*, 1989; Meng et Fenn, 1990). Ils ont démontré que dans ce type d'expérience, les molécules non volatiles peuvent être ionisées en phase gazeuse sans fragmentation ni destruction et ce, à pression atmosphérique. Cette technique peut donc être directement couplée (*on line*) à une chromatographie liquide. Ces travaux lui ont d'ailleurs permis de recevoir le prix Nobel de Chimie en 2002. Le principe de l'ESI consiste à produire un nuage de gouttelettes chargées sous l'influence d'un champ électrique intense. Ces gouttelettes vont être dirigées vers l'analyseur grâce au champ électrique et au gradient de pression appliqué. Pendant ce trajet, la taille des gouttelettes diminue par évaporation du solvant qui est au contact de l'air. L'électricité fournie conduit à des explosions coulombiennes successives provoquant la libération des ions en phase gazeuse. (Figure 13, page 40) L'ionisation par electrospray crée souvent des ions multi-chargés (Nguyen et Fenn, 2007).

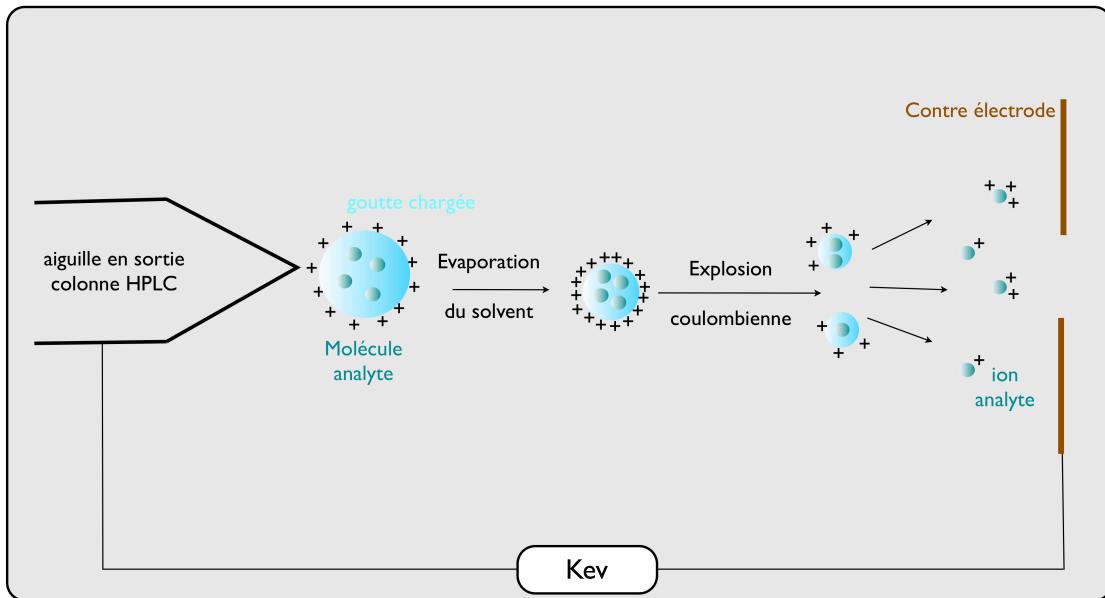


Figure 13 : Principe de l'ionisation ESI. L'ESI consiste en une cascade de fission de gouttes entraînant une diminution de la taille de la goutte donc une augmentation de la charge à sa surface jusqu'à atteindre la limite d'instabilité de Railey. A ce point, les forces répulsives de Coulomb excèdent la tension de surface de la goutte, causant une fission de la goutte en une plus petite et hautement chargée.

B. Les analyseurs en spectrométrie de masse (MS)

Les analyseurs sont caractérisés par cinq paramètres : la précision de mesure, la résolution, la vitesse de balayage, la gamme de rapports m/z balayée et la transmission (Stroobant, 2007). La précision sur la mesure des rapports m/z des ions correspond à la différence observée entre la masse théorique et la masse mesurée. Souvent exprimée en parts par million (ppm), elle est liée à la stabilité et à la résolution de l'analyseur. La résolution ou pouvoir de résolution de l'analyseur correspond à sa capacité à distinguer des ions de m/z proches (Figure 14A, page 41). Plus la résolution est élevée, plus la différence entre deux signaux distincts est faible sur le paramètre mesuré. En spectrométrie, la résolution correspond à la finesse des pics. En considérant le pouvoir séparatif, il est possible de définir la résolution comme le rapport $m/\Delta m$ où Δm est la largeur du pic à mi-hauteur (Figure 14B, page 41). Dans ce cas, une résolution de 1000 est atteinte pour un ion de masse 1000 est séparé de l'ion de masse 1001. Dans le cas d'un analyseur quadripolaire, la résolution est liée à la vitesse de balayage des m/z (en m/z par seconde) : plus le balayage est rapide et moins bonne est la résolution. La gamme de masse correspond à la masse (m/z) maximale que l'analyseur peut mesurer. La transmission caractérise la capacité de l'analyseur à conserver les ions entre leur formation et leur détection. Elle est difficile à évaluer car on ne connaît pas précisément la quantité d'ions formés en source (Stroobant, 2007).

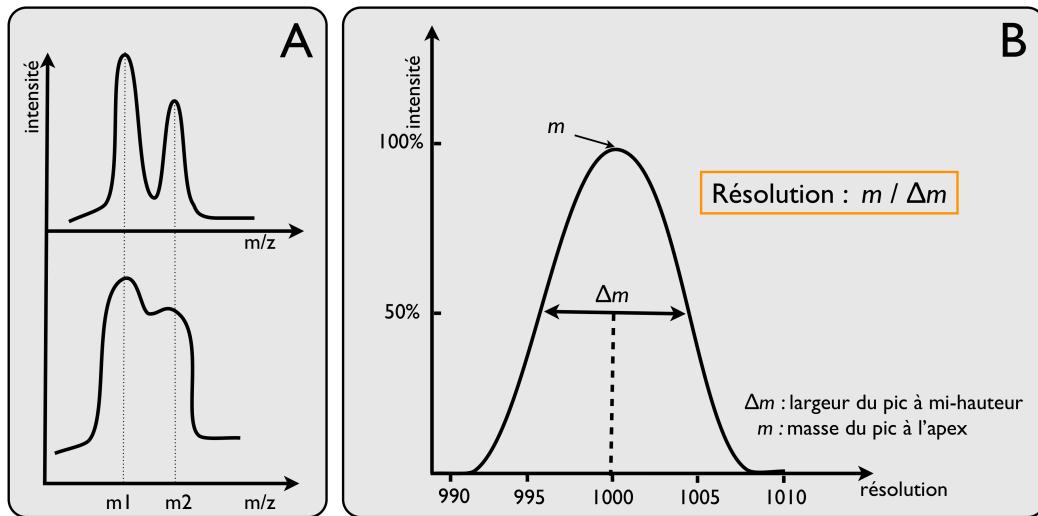


Figure 14 : Pouvoir de résolution. Le pouvoir de résolution est une caractéristique importante du spectromètre de masse, il définit sa qualité. Il représente l'aptitude à séparer deux pics adjacents (A). La résolution est définie comme le rapport entre la masse m et la largeur Δm du pic de masse mesurée à une certaine fraction de sa hauteur (souvent à 50 % : largeur à mi-hauteur) : (B) $R = m / \Delta m$.

1) L'analyseur à temps de vol (TOF)

L'analyseur à temps de vol (*Time Of Flight*, TOF) a été introduit par Dempster en 1918. Ce type d'analyseur est souvent couplé à une source d'ionisation de type MALDI, ce qui en fait la principale technique pour l'analyse de biomolécules (Cornish et Cotter, 1993). Il s'agit d'un long cylindre vide dans lequel règne une basse pression (Figure 15). Son principe est d'accélérer les ions issus de la source grâce à l'application d'une différence de potentiel qui va les séparer en fonction de leur énergie cinétique qui est inversement proportionnelle à leur rapport m/z . En conséquence, des ions de rapports m/z différents auront des temps de vol différents. Un spectre de masse sera généré après mesure du temps de vol de chaque ion.

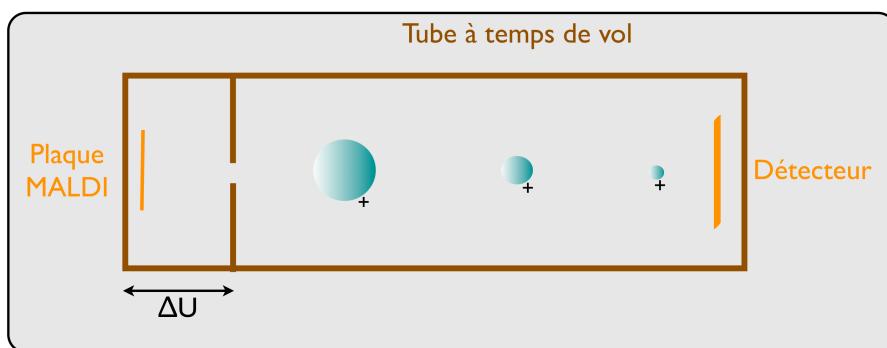


Figure 15 : L'analyseur en temps de vol (TOF). Le TOF sépare les ions selon leur différence de vitesse. Comme tous les ions possèdent la même énergie cinétique à la sortie de la source : $E_c = 1/2 mv^2$, les ions les plus légers ont une vitesse plus grande que les ions plus lourds. En réalisant, non pas un faisceau ionique continu, mais un faisceau "pulsé", il sera possible de séparer les ions dans le temps, à condition qu'ils partent tous en même temps de la fente d'entrée du système dispersif.

L'utilisation du TOF en mode linéaire est la plus simple. Elle repose sur la génération d'ions par impulsion du laser, ces ions sont alors accélérés et expulsés de la source. Ce mode n'a en théorie pas de limitation sur la masse des ions mais la résolution peut être altérée. En

effet, l'ionisation peut être à l'origine d'une dispersion angulaire du faisceau ionique qui fait que deux ions identiques avec des vitesses initiales égales, s'ils sont initialement à deux endroits différents, peuvent avoir des temps de vol différents. Leur entrée dans le tube de vol se fera alors avec des vitesses et des énergies cinétiques différentes. Or la résolution du TOF est définie par la distribution du temps de vol d'une population d'ions de même masse. Pour palier au problème de basse résolution en mode linéaire, le TOF peut être réglé en mode *rélectron*. Cette méthode consiste en l'interposition de miroirs électrostatiques appelés *rélectrons* sur le trajet ionique. Ces derniers permettent de compenser les différences de vitesses pour des valeurs de m/z identiques en imposant des trajets à parcourir d'autant plus longs que les espèces ioniques sont rapides (Figure 16). Il est donc possible d'avoir recours à la mise en place d'une extraction retardée qui permet de réaligner les ions de vitesse différente pour qu'ils pénètrent simultanément dans le tube de temps de vol, ce qui augmente ainsi sensiblement la résolution des pics. Dans ce cas, le détecteur est placé au point focal du *rélectron* (Figure 16). Ainsi, alors qu'en mode linéaire on atteint une résolution comprise entre 1000 et 5000, la résolution d'un TOF moderne en mode *rélectron* peut atteindre une résolution de 20 000 en routine. De même, cet analyseur est doté d'une précision de mesure comprise entre 1 et 5 ppm (Cole, 2010).

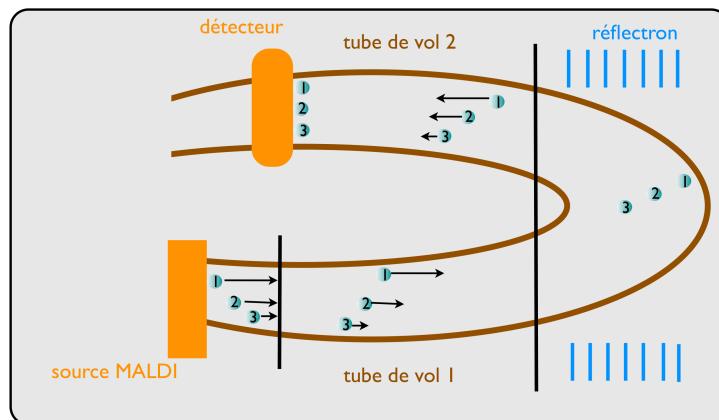


Figure 16 : Extraction retardée. La formation d'ions peut générer une certaine dispersion en énergie cinétique qui peut être corrigée avec un *rélectron*. Le *rélectron* met en œuvre un champ électrique pour réfléchir le faisceau d'ions vers le détecteur. Les ions dotés de la plus grande énergie pénètrent plus profondément dans le *rélectron*, ce qui augmente la longueur effective de l'analyseur, alors que les ions de même m/z mais moins énergétiques pénètrent moins profondément dans leur *rélectron*.

2) Le quadripôle

Le quadripôle est constitué de quatre électrodes cylindriques soumises deux à deux à une différence de potentiel qui crée un champ quadripolaire. Il est constitué d'une tension alternative (radiofréquence) et d'une tension continue (Figure 17, page 43).

Son principe de fonctionnement a été décrit en 1953 par Paul et Steinwedel et utilisé quelques années plus tard en spectrométrie de masse (Prestage *et al.*, 1991). Il consiste à utiliser la stabilité des trajectoires des ions dans le quadripôle en fonction des tensions qui y sont appliquées pour séparer ces ions selon leur rapport m/z . En fonction des valeurs de

tension alternative et de tension continue appliquées, les ions ayant un rapport m/z donné sont sélectionnés et adoptent des trajectoires stables en direction du détecteur. Ces trajectoires obéissent aux équations de Mathieu (équations qui ont été décrites en 1866 pour décrire la propagation d'ondes dans les membranes). Le spectre de masse est alors obtenu en faisant varier progressivement les valeurs de tensions pour stabiliser de manière séquentielle un ensemble d'ions compris dans une certaine gamme de masse (Wells *et al.*, 1998).

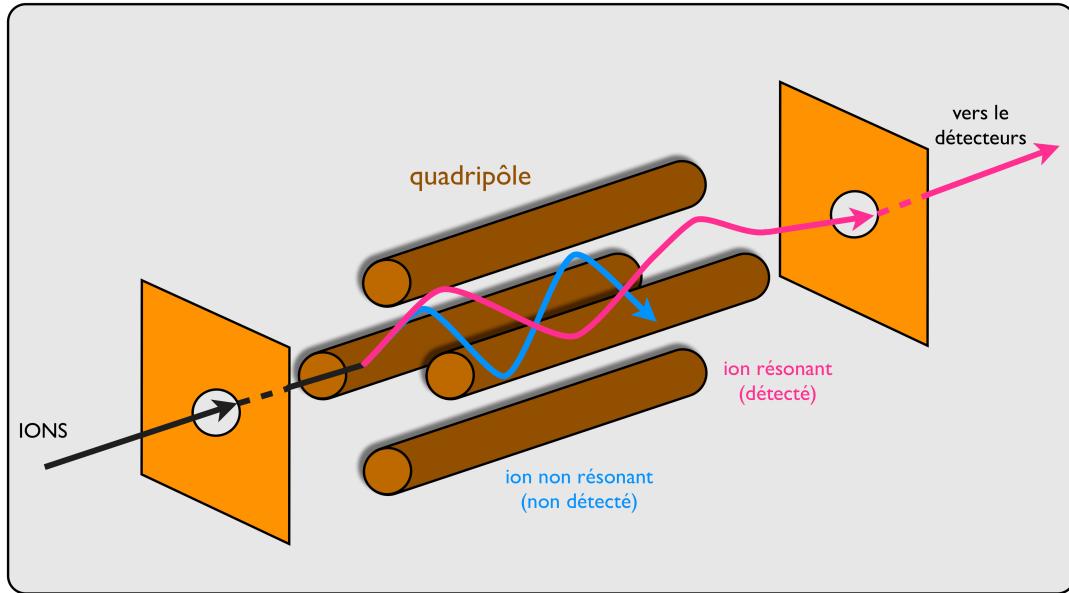


Figure 17 : Représentation schématique de l'analyseur quadripolaire. Le quadripôle est constitué de quatre barres à section cylindrique. Les ions sont faiblement accélérés dans la source ionique, de l'ordre de la centaine de volts seulement. Un champ électrique continu et un champ variable (radiofréquence) sont appliqués sur les barres diamétralement opposées. Les trajectoires des ions obéissent à des équations mathématiques de Mathieu, qui permettent de définir une probabilité de trouver l'ion recherché dans une certaine région de l'espace. Seuls les ions résonants seront détectés.

3) La trappe ionique 3-D (*Ion trap*)

La trappe ionique est constituée de trois électrodes hyperboliques (Stafford, 2002). Une des électrodes a la forme d'un anneau, elle est appelée électrode annulaire. L'électrode annulaire est encadrée par les deux autres électrodes, dites électrodes chapeau, une d'entrée et une de sortie (Figure 18, page 44).

Contrairement au quadripôle qui sélectionne uniquement les ions ayant un rapport m/z donné et confère à eux seuls une trajectoire stable, la trappe ionique piège tous les ions et mesure leurs masses. L'électrode annulaire génère un champ quadripolaire (une tension continue plus une tension en radiofréquence) qui permet de piéger les ions et de les retenir. De plus, un gaz neutre, l'hélium, est utilisé afin de stabiliser les ions dans la trappe. Il va freiner les ions, réduire leur trajectoire pour éviter les fragmentations induites par collisions, qui pourraient avoir lieu de façon non contrôlée. Comme pour le quadripôle, la trajectoire des ions à l'intérieur du piège est imposée par un champ quadripolaire et peut être définie par l'équation de Mathieu. En augmentant de façon progressive la tension alternative, les ions sont successivement déstabilisés et expulsés de façon séquentielle de la trappe, ce qui permet

l'enregistrement des spectres. La gamme de masse est traditionnellement de 50 à 2000/3000 Da. Les meilleures performances d'un piège ionique sont obtenues quand le nombre d'ions stockés est réduit (Schwartz *et al.*, 2002). Le piège à ions peut être aisément couplé à une chromatographie et il possède des avantages indéniables en terme de coût et d'encombrement.

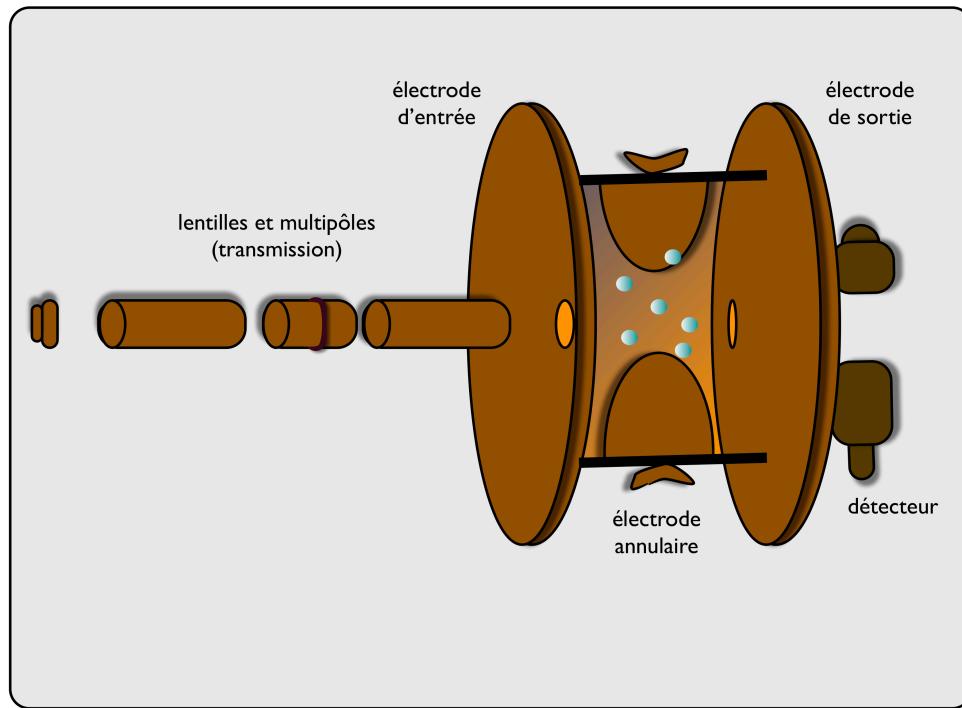


Figure 18 : Représentation schématique de la trappe ionique. La trappe ionique est un quadripôle à 3 dimensions, constitué d'une électrode annulaire et de deux électrodes chapeaux (d'entrée et de sortie) de sections hyperboliques. Une pression résiduelle d'hélium (10^{-3} Torr) freine les ions et les maintient tous dans la trappe. Les ions seront ensuite éjectés sélectivement en direction du détecteur en augmentant progressivement la tension pour déstabiliser les trajectoires des ions en fonction de leurs masses croissantes.

4) Analyseur à résonance ionique cyclotronique par transformée de Fourier (FT-ICR)

Dans cet analyseur, les ions sont piégés dans une cellule ICR (*Ion Cyclotron Resonance*) où un champ magnétique uniforme force les ions à suivre une trajectoire. L'application d'un champ électrique balayant une gamme de fréquence identique à celle de la rotation des ions rend leur mouvement cohérent. Le mouvement cyclotronique cohérent génère alors un signal électrique entre deux électrodes opposées qui sera mesuré et converti en fréquence. L'échantillon global conduit à un signal complexe, nommé interférogramme, composé de l'ensemble des fréquences induites par l'ensemble des ions de rapport m/z différents. De cet interférogramme sera déterminée par transformée de Fourier, la totalité des rapports m/z de la population d'ions. La détermination des valeurs des rapports m/z des ions est reliée à la fréquence de rotation des ions.

Cet analyseur a l'une des meilleures résolutions qui soient ($>100\ 000$), ainsi qu'une très bonne précision de mesure de masse (~ 1 ppm). Il est également doté d'une sensibilité de

l'ordre de l'atomole. Par contre, il nécessite d'avoir à disposition un aimant supraconducteur permettant la génération d'un champ magnétique élevé, ce qui engendre un coût non négligeable tant en terme d'utilisation que d'entretien (Feng et Siegel, 2007).

5) Trappe orbitale

C'est au sein de *General Electric* en 1923 qu'est né le piégeage orbital. Il fut mis en place par Langmuir Irving et Lingdon Kenneth Hay dans le but de capturer les ions en les faisant tourner dans l'orbite d'une électrode centrale. En 2000, Makarov développe un analyseur de masse électrostatique dérivé du piégeage orbital et appelé Orbitrap® (Makarov, 2000).

La géométrie de l'Orbitrap® consiste en une électrode externe en forme de tonneau partagée par une électrode centrale fuselée le long de l'axe (Figure 19, page 46). L'électrode axiale de l'Orbitrap génère un potentiel électrostatique qui est mesuré par l'électrode en forme de tonneau.

L'ion piégé analysé suit un mouvement oscillatoire dont la fréquence est inversement proportionnelle à la racine carrée de son rapport m/z. Ce rapport m/z peut ainsi être calculé par la transformée de Fourier inverse à partir de l'enregistrement du signal. L'Orbitrap® est un système moins onéreux que le FT-ICR et il présente des spécifications de résolution et de mesure de masse similaires. En effet, il a une précision de masse élevée (1 à 2 ppm), une puissance de résolution élevée (jusqu'à 200 000) et une haute gamme dynamique (autour de 5000 Da) (Makarov *et al.*, 2006). L'avantage est qu'il permet de capturer les ions dans un champ électrique généré par deux électrodes et non un champ magnétique.

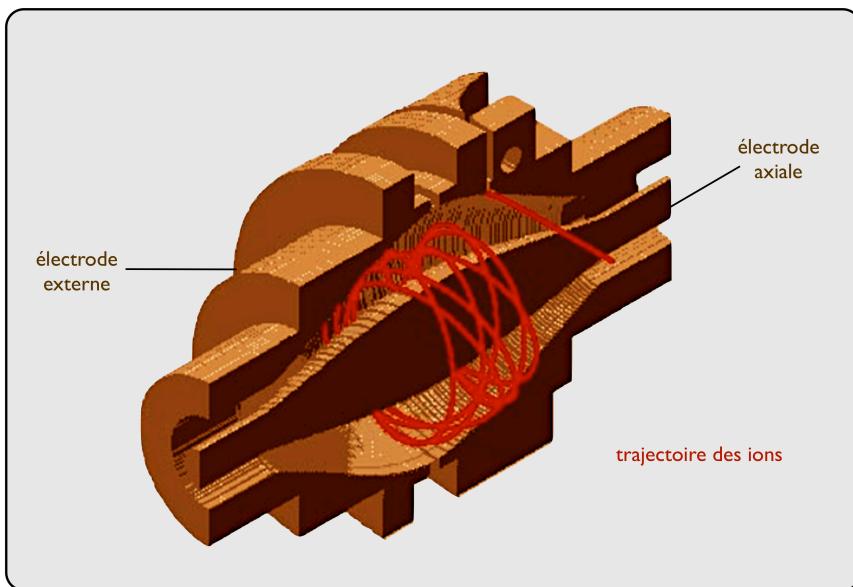


Figure 19 : Représentation schématique de la trappe orbitale (Orbitrap®). L'Orbitrap® est constitué d'un piège d'ions traversé par une électrode axiale chargée. Les ions ont à la fois un mouvement de rotation autour de l'électrode, ainsi qu'un mouvement axial le long de cette électrode. Ce mouvement axial génère un courant induit qui est enregistré par les deux électrodes. La transformée de Fourier permet de calculer la fréquence axiale des ions qui est proportionnelle à leur rapport m/z . (d'après <http://aston.chem.purdue.edu>)

C. La spectrométrie de masse en tandem (MS/MS)

1) Principe

La spectrométrie de masse en tandem (MS/MS) fait intervenir au moins deux étapes d'analyse de masse, donc au moins deux analyseurs. Son but est de fournir des informations structurales sur les molécules à analyser. En analyse protéomique, elle va donc aider à l'identification des protéines. Les ions issus de la source entrent dans le premier analyseur qui permet de sélectionner l'ion à fragmenter. Cet ion est appelé ion parent ou ion précurseur et il sera le seul à sortir de l'analyseur pour entrer en collision avec des molécules de gaz neutres (argon, azote ou hélium) dans la chambre de collision. La collision provoquera sa fragmentation en peptides plus petits, appelés ions fils, et qui permettront de déduire la séquence en acides aminés. Les ions fils seront ensuite triés dans un second analyseur selon leur rapport m/z . La formation de ces ions suit des règles précises. L'étude du spectre MS/MS obtenu donne des informations fiables sur la séquence de l'ion précurseur. En fonction de la localisation de la fragmentation et de la partie du peptide portant la charge de l'ion, différents types d'ions peuvent être générés. Il existe une nomenclature explicite qui différencie les fragments selon la rupture de la liaison peptidique et la terminaison du peptide qui retient la charge après fragmentation. Dans le cas de coupures au niveau de la chaîne peptidique, les séries a, b et c correspondent aux ions dont la charge est portée par la partie N terminale ; tandis que les séries x, y et z correspondent aux ions dont la charge est portée par la partie C terminale. (Figure 20, page 47). La nomenclature associe à ces fragments un indice qui indique le nombre de résidus portés par l'ion fragment (Biemann,

1990). La nature des fragments générés va dépendre du mode de fragmentation utilisé. Les principaux modes de fragmentation des peptides utilisés en spectrométrie de masse sont la fragmentation (ou dissociation) induite par collision (CID) et la dissociation par transfert d'électron (ETD) ou par capture d'électron (ECD).

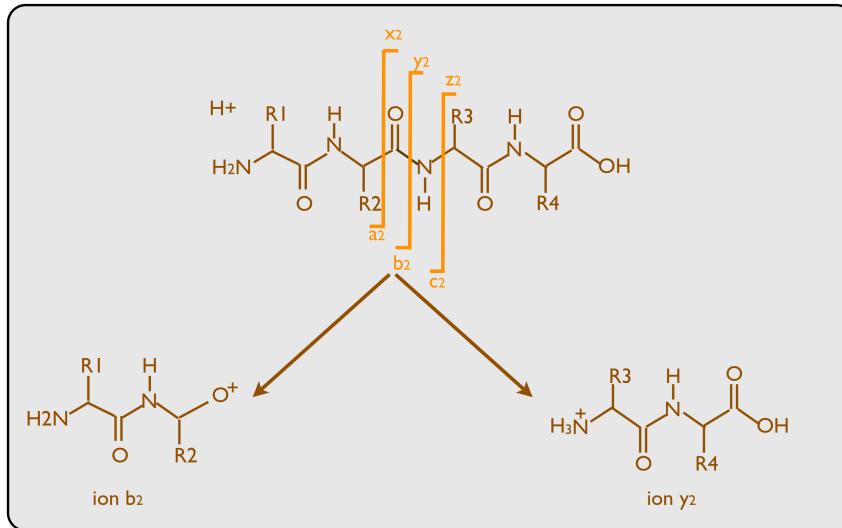


Figure 20 : Nomenclature des fragments ioniques en spectrométrie de masse inspiré de Biemann, 1990. Les ions (sous-peptides ou acides aminés) qui résultent de la fragmentation des peptides par ionisation par collision ont une nomenclature particulière. Lorsque la fragmentation a lieu au niveau d'une liaison peptidique, des ions b (contenant la partie N-terminale du peptide fragmenté) et des ions y (contenant la partie C-terminale du peptide fragmenté) sont formés. Lorsque la fragmentation a lieu au niveau d'une autre liaison des ions a ou c (contenant la partie N-terminale) et des ions x ou z (contenant la partie C-terminale) sont formés. Le chiffre qui accompagne la lettre indique le nombre d'acides aminés dans le fragment.

Dans le cas de la CID, la fragmentation est réalisée en augmentant l'énergie interne des ions par collision de ceux-ci dans un gaz inerte. Ce mode de fragmentation implique une délocalisation du proton à différents endroits du squelette peptidique. En fonction du type d'analyseur, on observe deux types de fragmentation en CID, les fragmentations basse énergie (quelques eV, IT et Q-TOF) et les fragmentations haute énergie (quelques keV, TOF-TOF). Pour les fragmentations basse énergie, les liaisons peptidiques sont les principaux sites de fragmentation et les ions des séries b et y majoritaires. Ainsi, les spectres MS/MS obtenus sont relativement simples à analyser. En effet, la différence de masse entre deux ions consécutifs d'une même série issus d'un même parent permet l'identification du résidu d'acide aminé ((Palumbo et al., 2011; Wiesner et al., 2008; Darula et al., 2010; Nagaraj et al., 2010) (Figure 21, page 48)).

Néanmoins, pour les acides aminés isobares (leucine et isoleucine) cette distinction sera impossible.

La fragmentation par ECD, spécifique aux spectromètres de masse FT-ICR et Orbitrap® est basée sur la capture exothermique d'un électron par les peptides chargés. Dans le cas de l'ETD, l'électron nécessaire à la fragmentation d'un peptide correspond à un radical anionique fixé à des molécules comme l'anthracène ou le fluoranthène (Hartmer *et al.*,

2008). Dans les deux cas, le site de fragmentation est aléatoire et toutes les séries peuvent être générées. Les spectres MS/MS obtenus sont beaucoup plus complexes à analyser mais ce type de fragmentation apporte des informations supplémentaires sur la structure des peptides, notamment au niveau des modifications post-traductionnelles (phosphorylation, glycosylation,...). La fragmentation en MS/MS est donc utilisée pour déterminer la séquence des peptides, mais aussi pour mettre en évidence et caractériser des modifications post-traductionnelles (Palumbo *et al.*, 2011; Wiesner *et al.*, 2008; Darula *et al.*, 2010; Nagaraj *et al.*, 2010).

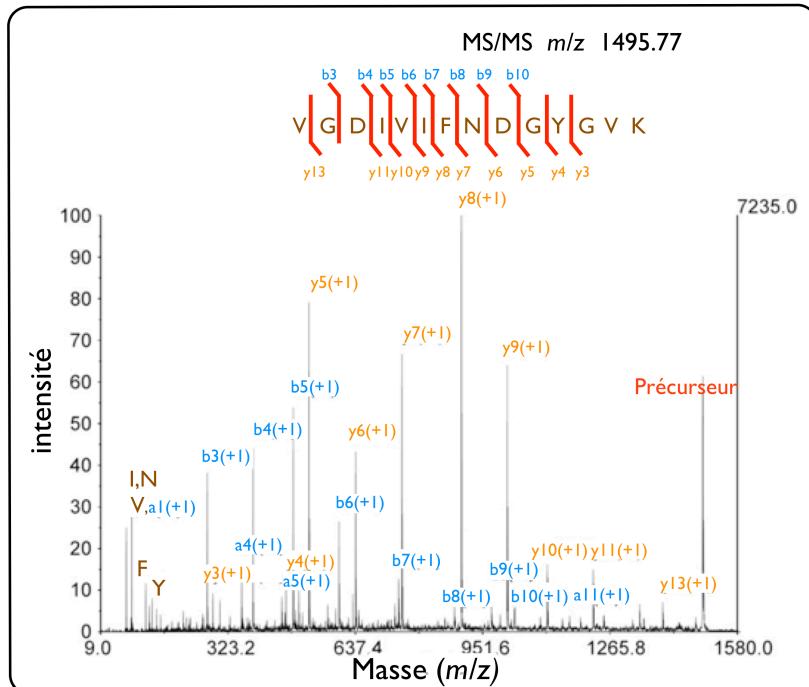


Figure 21 : Exemple de spectre de masse et d'ions générés par fragmentation de type CID. L'ion précurseur de masse 1495.77 a été fragmenté par collision de type CID. Les différentes masses des ions fils ainsi formés permettent de remonter à la séquence du peptide précurseur.

2) Les analyseurs utilisés en spectrométrie de masse MS/MS

a) La trappe ionique 3-D

Contrairement au simple quadripôle, la trappe ionique permet intrinsèquement d'effectuer également de la MS/MS. En mode MS/MS, des valeurs précises de potentiel sont appliquées aux électrodes pour ne conserver dans la trappe que des ions de la valeur de m/z choisie. Puis en appliquant au niveau de l'électrode annulaire une radiofréquence correspondant à la fréquence de résonance les ions m/z spécifiquement choisis, ces ions vont alors être fragmentés dans l'analyseur et être éjectés vers le détecteur. Ainsi, chaque tension de l'électrode annulaire correspond à l'éjection d'un m/z particulier. Ce type d'appareil permet même d'obtenir des ions fils de génération supérieure à deux et ce, par simple renouvellement du processus de fragmentation. Ce type d'expérience est appelé MSⁿ, n étant

le nombre de génération d'ions (Viseux *et al.*, 1998). Ainsi la MS² est la MS/MS, la MS³ est la MS/MS/MS et ainsi de suite.

Son couplage ais     un chromatographe et ses facilit  s d'utilisation en mode spectrom  trie de masse en tandem font de cet appareil un dispositif tr  s fr  quemment utilis   dans les laboratoires d'analyse.

b) Le TOF-TOF

D  crit pour la premi  re fois en 2000, cet appareil compos   de deux analyseurs en temps de vol, s  par  s par une cellule de collision, est principalement utilis   pour les analyses de s  quence de peptides et la caract  risation de modifications post-traductionnelles (Redeker, 2010; Medzihradszky *et al.*, 2000).

c) Le triple quadripôle

Un triple quadripôle correspond   l'association de deux analyseurs quadripolaires en s  rie (Q1 et Q3), s  par  s par une cellule de collision constitu  e d'un quadripôle plus court (Q2) (Figure 22, page 50). Cette combinaison permet de travailler en MS ou MS/MS. Lors d'une acquisition en MS/MS, la cellule de collision est remplie d'un gaz inerte   pression lev  e. Diff  rentes analyses peuvent  tre entreprises selon l'information recherch  e. Le mode descendant permet d'obtenir des informations structurales. Les modes ascendant et perte de neutre sont utiles pour rechercher des ions poss  dant des caract  ristiques communes (s  quence, modifications post-traductionnelles,...). Le quatri  me mode (*Multiple Reaction Monitoring* ou MRM), d  riv   du mode descendant, est d  di   quant   lui   la quantification.

En mode descendant, l'ion pr  curseur est s  lectionn   en Q1, fragment   en Q2 et ses ions fragments analys  s en Q3.

En mode ascendant, le Q1 balaie une gamme de masse tandis que Q2 est focalis   sur une valeur de *m/z* particuli  re. Tous les ions pr  curseurs capables de donner un ion fragment de *m/z* similaire sont d  tect  s en Q3.

En mode perte de neutre, Q1 et Q3 balaient simultan  m  t une gamme de masse avec un d  calage de masse constant. Le spectre tabli pr  sente alors tous les ions pr  curseurs capables de se fragmenter en g  nerant une masse neutre gale au d  calage impos  .

En mode MRM, l'ion pr  curseur d'int  r  t est s  lectionn   en Q1 et fragment   en Q2, comme pour le mode descendant. En revanche, Q3 est focalis   sur l'ion fragment d'int  r  t. Ce mode pr  sente une double s  lectivit   : au niveau de l'ion pr  curseur et de l'ion fragment. En outre, Q1 et Q3 sont   tensions constantes, ce qui augmente la sensibilit   de d  tection par rapport aux autres modes de balayage. En cons  quence, le mode MRM est tr  s utile pour la quantification et la d  tection de mol  cules en faible quantit   (Gstaiger et Aebersold, 2009; Yao *et al.*, 2008; Lange *et al.*, 2008).

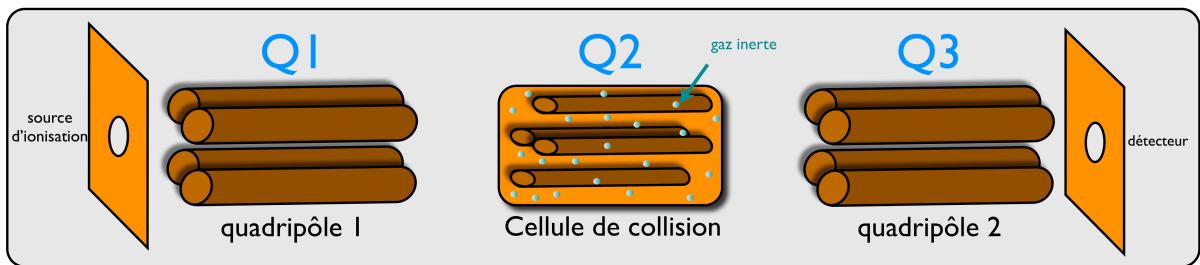


Figure 22 : Concept de spectrométrie de masse en tandem dans un analyseur triple quadripôle : Le premier analyseur (Q1) sélectionne l'ion précurseur d'intérêt. Cet ion précurseur est ensuite fragmenté par des collisions énergétiques dues à la pression d'un gaz inerte dans le quadripôle Q2 encore appelé cellule de collision. Enfin, le quadripôle Q3 permet d'effectuer l'analyse des ions ainsi formés.

Depuis quelques années, les spectromètres de masse qui apparaissent sur le marché sont de plus en plus performants. Pour augmenter la performance d'analyse de masse, sur le marché, sont apparus des appareils de type hybrides qui combinent les forces des différents analyseurs qui les composent en évitant bien sûr d'additionner leurs faiblesses. Tous les appareils hybrides permettent de faire de la MS/MS, ils sont choisis selon le type d'analyse à effectuer (Walther et Mann, 2010).

Aujourd'hui, ils sont de plus en plus nombreux, je ne décrirai que le quadripôle-TOF, et la LTQ-Orbitrap® car ce sont les appareils les plus utilisés en analyse protéomique.

d) **Le quadripôle-TOF (Q-TOF)**

Comme son nom l'indique, cet appareil de masse est composé d'un premier analyseur, le quadripôle et d'un analyseur à temps de vol TOF perpendiculaire. Entre les deux analyseurs se trouve la cellule de collision. Le Q-TOF combine la sensibilité, la résolution et la précision de mesure du TOF avec l'efficacité du quadripôle en MS/MS (Domon et Aebersold, 2006b).

e) **La LTQ-Orbitrap®**

La LTQ-Orbitrap® a été commercialisé en 2005 par la société ThermoFisher (Scigelova et Makarov, 2006). Elle consiste en deux analyseurs de masse complets qui peuvent être utilisés indépendamment ou en association. Il s'agit d'un orbitrap® et d'une trappe ionique linéaire : LIT® (*Linear Ion Trap*) ou LTQ® (*Linear Quadrupole ion Trap*). La trappe linéaire permet de combiner les avantages du quadripôle et de la trappe ionique 3-D en ajoutant de nouveaux types de balayage (Riter *et al.*, 2006). En effet, la LTQ® montre une vitesse de balayage cinq fois supérieure à la trappe ionique. De plus, elle acquiert 2,5 fois plus de spectres MS/MS qu'une trappe classique, ce qui augmente l'identification de protéines issues d'un mélange complexe (Yates *et al.*, 2006; Scigelova et Makarov, 2006). Le couplage entre ces deux analyseurs se fait via une *C-trap* qui focalise et concentre les ions. Dans le cas du couplage, tandis qu'un spectre de haute résolution est acquis dans l'Orbitrap®, la trappe ionique, plus rapide, réalise la fragmentation et la détection des spectres MS/MS des peptides les plus intenses sélectionnés lors d'un *prescan*. En une seconde, il est possible d'acquérir un spectre MS de haute résolution dans l'Orbitrap® et 3 à 5 spectres de

fragmentation MS/MS dans la trappe ionique, ce qui permet d'augmenter considérablement le nombre de peptides identifiés (Scigelova et Marakov 2006).

En 2009, ThermoFisher crée le LTQ-Orbitrap Velos® qui a une précision de masse plus élevée et augmente la vitesse et la confiance de l'identification des protéines dans des échantillons complexes grâce à une couverture de séquence plus importante et une sensibilité accrue.

VII. Identification des protéines

La première information requise pour l'étude des protéines est la connaissance de leur séquence. Aujourd'hui, la spectrométrie de masse présente une très bonne sensibilité et permet de fragmenter les peptides en quelques secondes au lieu de quelques heures voire quelques jours pour le séquençage dégradation d'Edman (Wilm *et al.*, 1996). Elle est sans conteste la méthode de choix pour la caractérisation des biomolécules (Cravatt *et al.*, 2007; Mann *et al.*, 2001; Domon et Aebersold, 2006a). La spectrométrie de masse nécessite une mesure précise de la masse des analytes. Lors d'une expérience d'analyse protéomique classique, les protéines sont lysées au préalable en peptides de façon contrôlée grâce à l'utilisation d'endoprotéases. L'endoprotéase la plus couramment utilisée est la trypsine. Elle hydrolyse les liaisons peptidiques en position carboxy-terminale d'un acide aminé basique (lysine ou arginine). Ces deux acides aminés présentent une abondance et une répartition homogène dans les différentes protéines. L'occurrence de ces résidus basiques permet de générer des peptides dont la taille moyenne est d'environ 10 résidus d'acides aminés, ce qui correspond à un poids moléculaire moyen de 1100 Da, taille idéale pour une analyse par spectrométrie de masse.

Il existe deux grandes approches pour l'identification de protéines : l'empreinte peptidique massique (*Peptide Mass Fingerprinting*, PMF) et les éléments de séquence issus de la fragmentation spécifique des peptides. Les deux reposent sur un même principe général : les données expérimentales obtenues en spectrométrie de masse sont comparées à des données théoriques stockées dans les bases de données protéiques grâce à l'utilisation d'un moteur de recherche. Si les deux types de données coïncident, la protéine est identifiée avec un score de confiance.

A. Identification par empreinte peptidique massique

Cette approche, décrite en 1993, est généralement utilisée lorsque la méthode de séparation des protéines est l'électrophorèse bidimensionnelle (Henzel *et al.*, 1993; Mann *et al.*, 1993; Yates *et al.*, 1993). Lorsque la résolution de l'électrophorèse bidimensionnelle est suffisante, chaque spot observé correspond à une seule forme protéique. Le spot protéique d'intérêt est alors excisé, la protéine contenue dans le morceau de gel est digérée par la trypsine et les différents peptides qui en sont issus sont extraits du gel. L'analyse en mode MS qui passe généralement par l'utilisation du MALDI TOF (Domon et Aebersold, 2006b), génère une liste de masses expérimentales correspondant aux peptides trypsiques observés pour la protéine analysée. Cette liste constitue la signature de la protéine, on parle alors d'empreinte peptidique massique ou PMF. La liste de masses est ensuite comparée aux listes de masses théoriques issues de la digestion *in silico* de protéines répertoriées dans des bases de données (Figure 23). Les résultats obtenus sont ensuite confrontés à l'aide d'outils statistiques afin de déterminer la meilleure correspondance (Baldwin, 2003).

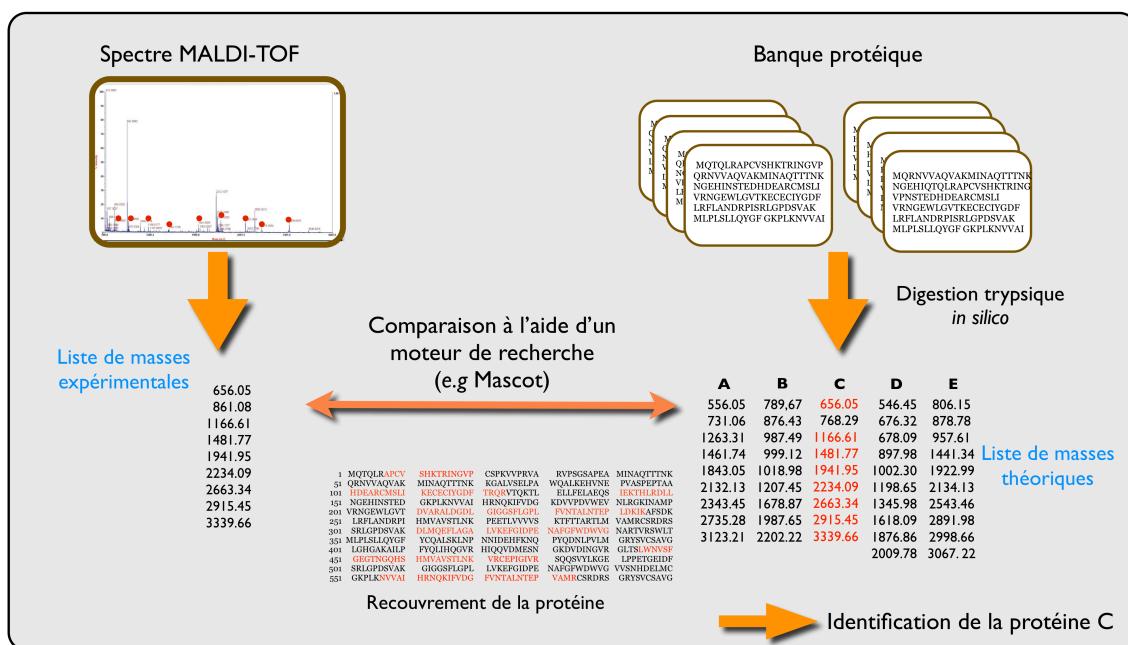


Figure 23 : Stratégie d'identification d'une protéine par empreinte peptidique massique. Après séparation des protéines par gel 2-D, digestion et analyse par MALDI-TOF-MS, les listes de masses expérimentales sont comparées aux listes de masses théoriques issues de la digestion trypsique *in silico* de toutes les protéines contenues dans une banque protéique

B. Identification par fragmentation MS/MS

L'identification de la protéine se fait via l'analyse du spectre de masse d'un peptide *séquence spécifique* (Aebersold et Goodlett, 2001). Le principe consiste à digérer un mélange complexe de protéines par la trypsine, afin de générer les peptides qui seront ensuite analysés par LC-MS/MS. L'identification des protéines repose alors sur la fragmentation spécifique de chaque peptide précurseur ainsi que sur l'analyse en spectrométrie de masse des différents

fragments générés. Lors des analyses LC- MS/MS, la sélection des ions parents qui seront fragmentés au cours de l'analyse chromatographique est réalisée de manière automatique sur la base de leur intensité, leur rapport m/z ou encore leur état de charge. Etant donné que chaque spectre MS représente l'élution chromatographique de peptides appartenant à plusieurs protéines, cette fois-ci, l'approche PMF ne peut pas être utilisée. À partir de l'ensemble des peptides dont le rapport *m/z* est mesuré par le détecteur du spectromètre de masse à l'instant t, un certain nombre sera fragmenté. Les spectres MS/MS ainsi générés, correspondent à la fragmentation des peptides dont on connaît la masse. L'interprétation de ces spectres donnera des informations de séquence sur ces peptides, permettant de remonter à l'identité des protéines dont ils sont issus. Il existe trois approches utilisant des outils informatiques et statistiques qui permettent d'attribuer automatiquement une séquence peptidique à un spectre d'ion fragment (Nesvizhskii *et al.*, 2007).

1) L'approche par recherche dans les banques protéiques :

Il s'agit d'une approche parfois appelée *Peptide Fragment Fingerprinting* (PFF) par analogie au PMF (Blueggel *et al.*, 2004). L'utilisation de moteurs de recherche permet de comparer les profils de spectres MS/MS expérimentaux obtenus aux profils de fragmentations théoriques des peptides obtenus après digestion *in silico* des protéines de la base de données ou confrontation aux spectres d'ions fragments d'une grande collection d'expériences ultérieures où l'on peut observer des spectres de masse de peptides correctement identifiés (*spectral libraries searching*). L'ensemble des peptides identifiés dans cette approche, qui est la plus fréquemment utilisée dans les études protéomiques à grande échelle, permet de remonter à l'identité des protéines présentes dans l'échantillon (Figure 24, page 54).

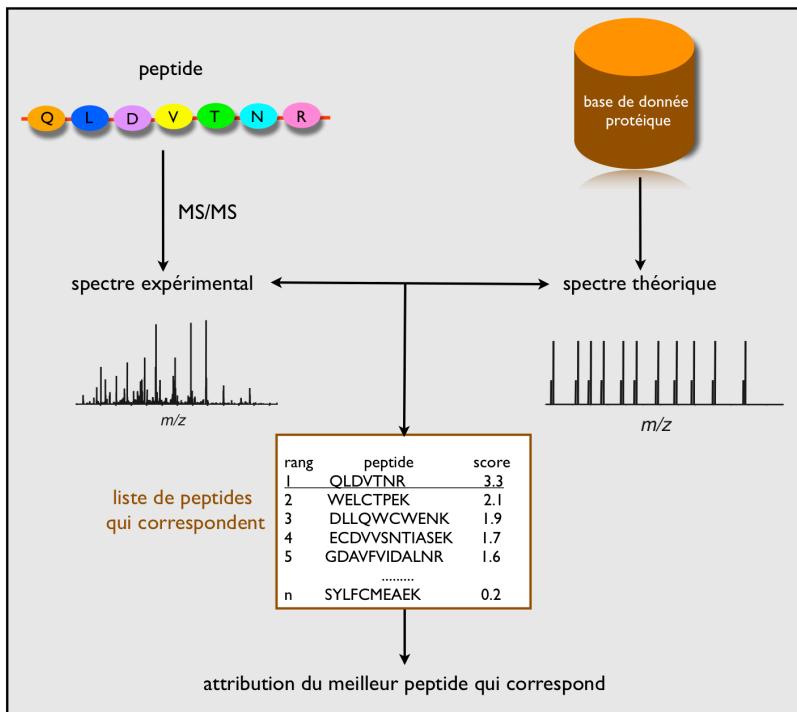


Figure 24 : Identification peptidique par recherche dans les banques protéiques. Confrontation du spectre MS/MS acquis au spectre théorique construit pour chaque banque de données peptidique ciblée par l'expérimentateur. C'est sur la base du score qu'est mesuré le degré de similarité entre les deux spectres. Le peptide ayant le meilleur score sera ensuite sélectionné pour les analyses suivantes schéma inspiré de Nesvizhskii *et al.*, 2007).

2) L'approche par séquençage *de novo*

Le séquençage *de novo* permet de déduire directement la séquence peptidique du spectre de l'ion fragment. Dans cette approche, la séquence en acides aminés des peptides est directement déterminée à partir des spectres MS/MS. A l'origine, cette approche était réalisée manuellement en tenant compte des règles de fragmentation des peptides (vu précédemment). Depuis quelques années, une série d'outils informatiques a été développée pour assister le séquençage *de novo*. Parmi les plus connus on trouve Lutefisk (Johnson et Taylor, 2002), Peaks (Ma *et al.*, 2003) ou encore Pepnovo (Frank *et al.*, 2005). Le séquençage *de novo*, qu'il soit manuel ou assisté, permet d'obtenir des fragments étiquetés de séquences en acides aminés qui sont généralement soumis à un algorithme de recherche de similarité de séquence comme MS-BLAST (Shevchenko *et al.*, 2001).

3) L'approche hybride

Cette approche combine le séquençage *de novo* et la recherche dans les banques de données protéiques. L'identification commence par la caractérisation de petites étiquettes de séquence à partir des spectres MS/MS, puis une recherche tolérante aux erreurs dans les banques protéiques est réalisée grâce aux outils informatiques Peaks (Ma *et al.*, 2003), GutenTag (Tabb *et al.*, 2003) ou encore Inspect (Tanner *et al.*, 2005).

Cette approche permet une réduction du temps de recherche pour l'identification et elle est essentiellement utilisée dans le cadre d'une analyse systématique des peptides portant des modifications post-traductionnelles.

C. Interrogation des banques de données

Les banques de données destinées à l'analyse protéomique regroupent les séquences des protéines connues avec des annotations sur leurs fonctions, leurs modifications post-traductionnelles connues ou encore les variants connus des différentes protéines. Elles offrent la possibilité d'échanger, de comparer, d'analyser à nouveau les résultats, de développer des moteurs de recherche et des méthodologies d'analyses statistiques plus performants afin d'obtenir des conclusions plus fiables (Mead *et al.*, 2007).

La banque protéique du NCBI rassemble les séquences des banques *GenPept* (*GenBank gene Products Data Bank*) (<http://www.ncbi.nlm.nih.gov/protein>) créée à partir de la traduction des régions codantes des gènes entreposés dans la *GenBank* (<http://www.ncbi.nlm.nih.gov/genbank>), EMBL (*European Molecular Biology Laboratory*) (<http://www.ebi.ac.uk/embl/>) (Stoesser, 2003), DDBJ (*DNA Data Bank of Japan*) (<http://www.ddbj.jp>) (Miyazaki, 2003) ainsi que de multiples sources incluant SwissProt (<http://expasy.org/sprot/>) (Boeckmann, 2003), PIR (*Protein Information Ressource*) (<http://pir.georgetown.edu/>) (Wu *et al.*, 2003), PRF (*Protein Research Fundation*) (<http://www.prf.or.jp/index-e.html>) (Belew *et al.*, 2008) et PDB (*Protein Data Bank*) (<http://www.pdb.org/>) (Westbrook *et al.*, 2003).

Les banques contenant des informations non redondantes les plus complètes sont *SwissProt* (Boeckmann *et al.*, 2003) et PIR (Wu *et al.*, 2003). Parmi les autres banques de données, on peut citer TrEMBL, Refseq et PDB. La TrEMBL (http://www.matrixscience.com/help/seq_db_setup_Trembl.html) correspond à la traduction des séquences nucléotidiques de la banque génomique EMBL, déduction faite des entrées de *SwissProt* et *GenPept*. La PDB, quant à elle, est une banque de structure tridimensionnelle de protéines contenant des données issues de cristallographie et/ou de RMN pour chaque protéine de la banque. Enfin, la banque Refseq (<http://www.ncbi.nlm.nih.gov/RefSeq/>) a été créée afin de fournir un ensemble non redondant de séquences protéiques de référence d'un nombre limité d'espèces (Pruitt *et al.*, 2009).

De nombreux outils bio-informatiques permettent l'identification des protéines à partir de résultats obtenus par spectrométrie de masse en utilisant l'interrogation de ces banques de données. Les sites de *Mascot* (<http://matrixscience.com>) (Perkins *et al.*, 1999), *Protein Prospector* (<http://prospector.ucsf.edu/prospector/mshome.htm>) (Clauser *et al.*, 1999), *ProFound* (<http://visites.unb.br/cbsp/paginiciais/profound.htm>) (Zhang et Chait, 2000),

Phenix (<http://phenix-ms.com> et *Sequest* (<http://fields.scripps.edu/sequest/>) (Klimek *et al.*, 2008) proposent les outils pour réaliser l'analyse de cartographies peptidiques massiques.

Les résultats de l'identification des protéines à l'aide des moteurs de recherche sont associés à un score qui représente la probabilité que la correspondance entre la liste de valeurs expérimentales et celle obtenue théoriquement avec la base de données soit aléatoire. Si le score excède le seuil de confiance défini au préalable (au moins 95 %), l'identification de la protéine est significative. En analyse MS, la valeur du score dépend de la taille de la base de données interrogée, du pourcentage de recouvrement de séquence de la protéine et du nombre de peptides qui ont permis son identification. Outre la masse expérimentale des peptides trypsiques, d'autres paramètres tels que l'origine de la protéine, ses propriétés physiques ainsi que les modifications apportées par la technique utilisée sont pris en compte par les algorithmes de recherche. Bien que sensiblement différente, l'interrogation des banques de données à partir de données MS/MS fait intervenir les mêmes éléments. Mais pour l'analyse MS/MS, la valeur du score prend également en compte la masse, la séquence des peptides ainsi que le bruit de fond. Enfin, des algorithmes de séquençage *de novo* de plus en plus performants sont disponibles et permettent d'identifier une séquence à partir de spectres MS/MS qui n'ont pas été interprétés. Par exemple, les logiciels PEAKS (Anderson *et al.*, 2003), MassLynx de Micromass et Biotools de Bruker sont à citer dans ce cadre.

VIII. L'analyse protéomique différentielle

L'analyse protéomique différentielle consiste à comparer et quantifier des protéomes distincts (par exemple les cellules normales par rapport aux cellules malades, malades traitées par rapport aux malades non traitées, etc ...). Il existe différentes approches et chacune d'elle présente des avantages et des inconvénients à la fois du point de vue conceptuel et méthodologique. Les approches utilisées impliquent généralement l'électrophorèse et/ou la chromatographie combinée à spectrométrie de masse. Elles visent à identifier et quantifier des cibles moléculaires, les protéines, impliquées dans différents états physiopathologiques.

A. Approches *gel-based*

1) Quantification par électrophorèse bidimensionnelle

L'un des objectifs principal de l'électrophorèse 2-D est d'identifier l'expression différentielle des protéines entre deux échantillons ayant migrés sur des séries distinctes de gels 2-D. L'analyse de cette expression différentielle se fait par la quantification et l'identification des spots qui changent de niveau d'expression, voire disparaissent ou apparaissent, d'un type d'échantillon à l'autre, et ce, de façon significative. La quantification se fait par le biais de l'analyse de l'intensité de coloration des spots de protéines. Il existe de

nombreux logiciels tels que *Progenesis (Nonlinear Dynamics)*, *Image Master 2D Platinum* et *Melanie Software (Amersham Biosciences)* ou encore *PDQuest (Bio-Rad)* qui permettent d'aligner automatiquement les gels à comparer spot par spot pour mettre en évidence des modifications de quantité. Ces logiciels permettent d'effectuer un pré-traitement automatisé des spots et de les quantifier, mais cette étape doit être complétée manuellement. L'analyse d'image est cruciale pour la quantification sur gel mais reste un processus long et fastidieux (Monteoliva et Albar, 2004).

Après sélection des spots d'intérêt, il faut identifier les protéines qu'ils contiennent par spectrométrie de masse. Le travail d'identification des spots est un point crucial de l'interprétation des gels. Il est nécessaire de prendre en considération les caractéristiques particulières de l'électrophorèse 2-D, ainsi que les nombreux biais liés à son caractère expérimental. Le MALDI-TOF est l'appareil de masse le plus adapté pour l'identification de protéines issues d'un gel 2-D. En effet, ces analyses sont rapides, facilement automatisables, relativement tolérantes aux sels et aux détergents, et présentent une bonne résolution et une bonne précision de masse (Thiede *et al.*, 2005). Cette identification est principalement réalisée par l'approche de type PMF mais dans certains cas, le recours à des approches par LC-MS/MS, offrant une plus grande spécificité grâce à l'obtention d'informations de séquence, est nécessaire (Horn *et al.*, 2004).

2) Quantification par l'approche DIGE

L'approche DIGE (*Differential In Gel Electrophoresis*) est basée sur l'emploi de différents esters de succinimidyle de colorants cyanines pour effectuer un marquage différentiel covalent des protéines de chaque condition par des cyanines fluorescentes spécifiques (Cy2, Cy3 et Cy5) afin de les quantifier. Une normalisation est effectuée à l'aide d'un standard interne qui correspond souvent au mélange de chaque échantillon à comparer selon un ratio 1:1, marqué par l'un des trois fluorochrome. Ensuite, il est d'usage de mélanger les protéines des différentes conditions, marquées par des fluorochromes différents et de les faire migrer au sein du même gel 2-D, de manière à éliminer les erreurs dues aux variations inter-gels (Figure 25, page 58). L'analyse différentielle est réalisée en comparant la fluorescence des différentes cyanines. Des logiciels tels que *Decyder (GE Healthcare)* et *SameSpots (Nonlinear Dynamics)* permettent d'analyser les images acquises aux longueurs d'onde d'émission de chaque cyanine à partir du même gel. La superposition des images et l'intensité des spots protéiques à chaque longueur d'onde renseigne sur l'abondance relative de chaque protéine (Alban *et al.*, 2003).

Les méthodes classiques de quantification utilisant les intensités de colorations ou les fluorophores disposent d'une très bonne sensibilité, linéarité et gamme dynamique, cependant, elles présentent deux principales limitations. D'abord, elles nécessitent une séparation à haute résolution des protéines en mélange par le biais de l'électrophorèse 2-D,

ce qui limite donc leur application aux protéines abondantes et solubles. Ensuite, elles ne permettent pas toujours de révéler l'identité de la protéine directement et avec exactitude car il peut y avoir plusieurs protéines dans un même spot. Ces deux problèmes peuvent être résolus par les applications récentes de la LC-MS/MS.

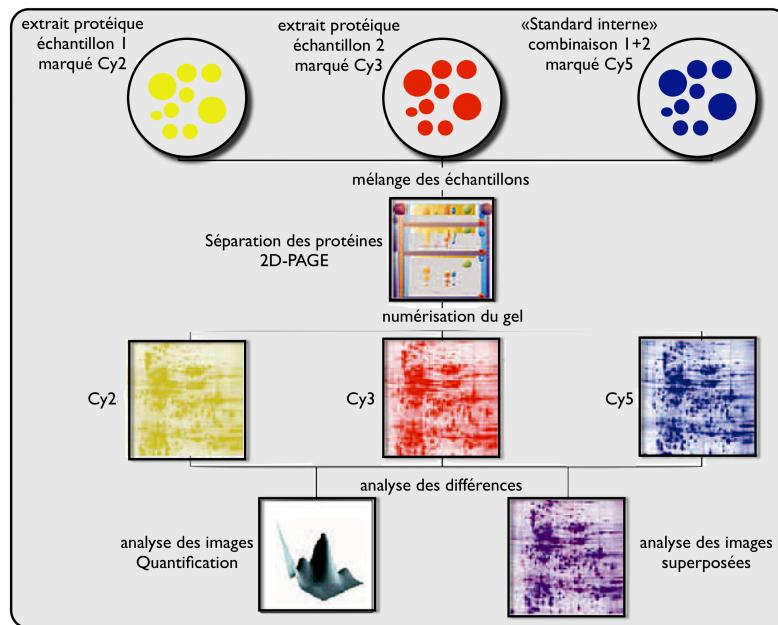


Figure 25 : Principe du marquage DIGE : la 2-D-DIGE est une électrophorèse bidimensionnelle réalisée sur des extraits préalablement marqués par des fluorochromes (Cy2, Cy3 et Cy5). Dans le but de comparer deux échantillons, leurs extraits protéiques sont marqués à l'aide de deux cyanines différentes, ici, Cy2 et Cy3. La cyanine Cy5 permet de marquer les extraits d'un mélange à parts égales des deux extraits à analyser, elle sert ainsi de standard pour tous les gels. Les trois extraits peuvent alors co-migrer sur le même gel 2-D. La fluorescence permet une quantification sur une large gamme dynamique et avec une bonne sensibilité. L'acquisition des images en fluorescence est assurée par un scanner spécifique. Deux logiciels d'analyse d'images sont disponibles pour comparer les gels : *Decyder* (GE Healthcare) et *SameSpots* (Nonlinear Dynamics). (Schéma inspiré de *Medical Proteomics* : <http://www.medicalproteomics.com/instrumentation.html>).

B. Quantification par marquages isotopiques

L'approche majeure d'analyse protéomique basée sur la spectrométrie de masse repose sur le principe selon lequel un peptide marqué par un isotope stable a les mêmes propriétés physicochimiques que son *homologue* non marqué, à l'exception de sa masse. Les deux peptides ont donc des comportements identiques en chromatographie liquide. Le spectromètre de masse peut ainsi repérer précisément la différence de masse entre la forme marquée et non marquée du même peptide. Le marquage par isotopes stables a été introduit en analyse protéomique en 1999 par trois laboratoires indépendants (Paša-Tolić *et al.*, 1999; Gygi *et al.*, 1999; Oda *et al.*, 1999) et a été grandement développé ces dernières années (Bantscheff *et al.*, 2007). Il existe différentes stratégies de marquage isotopiques qui peuvent être réalisées par des approches chimiques (ICAT, *Isotope Coded Affinity Tag* et iTRAQ™, *isobaric Tags for Relative and Absolute Quantification*), enzymatiques ($H_2^{16}O/H_2^{18}O$) ou par marquage métabolique (SILAC, *Stable Isotope Labelling by Amino acids in Cell culture*) (Figure 26, page 59).

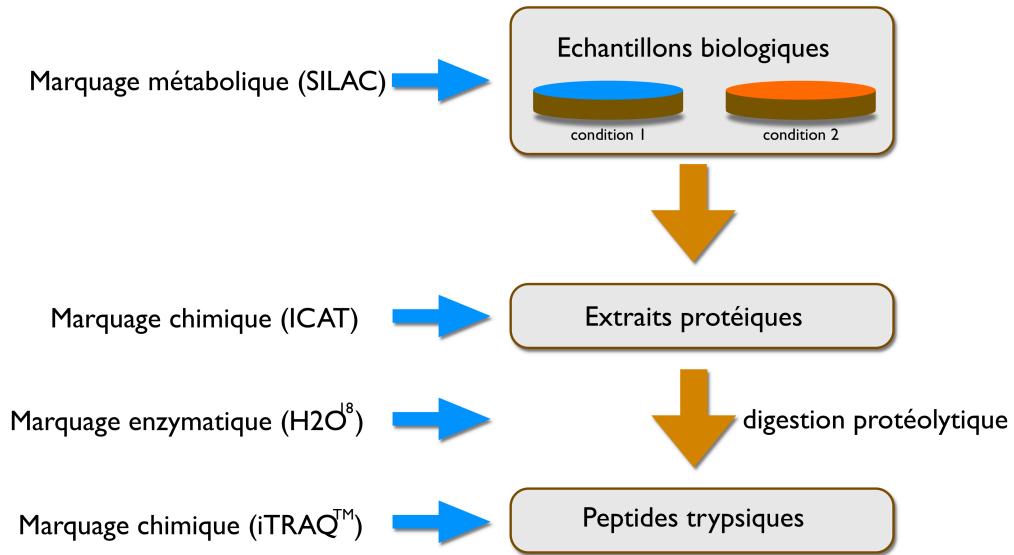


Figure 26 : Niveau d'introduction du marquage dans l'échantillon. Le marquage peut être introduit dans l'échantillon à différentes étapes du processus analytique qui mène à la quantification des protéines. Le marquage métabolique (SILAC) s'effectue directement au niveau des cellules en culture (Ong *et al.*, 2002). Après lyse des cellules, l'étiquette ICAT permet de marquer les protéines à quantifier (Gygi *et al.*, 1999) alors que le second marquage chimique, iTRAQ™, est introduit au niveau des peptides après digestion enzymatique des protéines (Ross, 2004). Enfin, le marquage enzymatique ($H_2^{16}O/H_2^{18}O$) est effectué lors de la lyse des protéines à quantifier (Stewart *et al.*, 2001).

Le marquage peut également provenir de l'ajout d'un standard externe constitué de peptides synthétiques marqués (Ong et Mann, 2005; Eyers *et al.*, 2008). La quantification s'effectue par comparaison des intensités des signaux MS. Lorsque le peptide est élué de la colonne chromatographique dans une analyse LC-MS, il peut être suivi au cours du temps. L'aire sous le pic chromatographique correspond au courant ionique extrait (XIC, *eXtracted Ion Current* ou *eXtracted Ion Chromatogram*). Chelius et Bondarenko ont été les premiers à montrer en 2002 que le ratio de l'aire sous le pic dans les XIC peut être utilisé pour réaliser une quantification de l'abondance relative des protéines d'un complexe en mélange (Chelius et Bondarenko, 2002). La quantification s'effectue par l'intégration et la comparaison des intensités de signal pour chaque forme isotopique (Figure 27, page 60). Cette technique est donc basée sur l'intégration du XIC sur l'ensemble de l'analyse chromatographique et sur l'intensité du signal de l'ion précurseur (MS) qui provient des données acquises sur des spectromètres de masse de haute précision comme les nouvelles générations de TOF, le FTICR ou encore l'Orbitrap® (Silva, 2005).

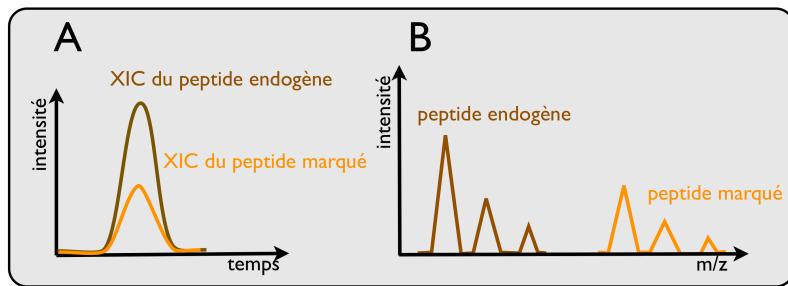


Figure 27 : Définition du XIC. Lors de l'étape de chromatographie liquide, les versions endogènes et marquées du peptide protéotypique sont élues simultanément (A). La quantification est effectuée en comparant les chromatogrammes des ions extraits (XIC) grâce à l'aire sous le pic. Par contre, les contributions isotopiques des formes endogènes et marquées du peptide protéotypique sont différenciables au niveau du spectre de masse grâce au marquage différentiel (B).

Des logiciels tels que MsQuant (Mortensen *et al.*, 2010), MaxQuant (Cox *et al.*, 2009) et MFPaQ (Bouyssié *et al.*, 2007) ont été développés afin d'extraire les intensités observées pour chaque forme isotopique dans le but de comparer leur abondance relative.

Selon la méthode utilisée, le marquage s'effectue à différentes étapes du processus d'analyse protéomique (Brewis et Brennan, 2010) (Figure 26, page 59).

Depuis quelques années, on assiste aussi à l'émergence de techniques alternatives appelées *Label-free*. Ces méthodes permettent de comparer deux ou plusieurs expériences grâce à la quantification d'un peptide donné en spectrométrie de masse.

1) Marquage chimique

L'analyse différentielle par marquage isotopique est réalisée par l'ajout d'une étiquette non protéique particulière soit au niveau protéique, soit au niveau peptidique après hydrolyse. Différentes approches ont été développées et nous présenterons ici les deux plus utilisées : l'ICAT et l'iTRAQ.

a) ICAT

L'ICAT a été le premier marquage chimique utilisé pour la quantification relative des protéines. Développé en 1999 par l'équipe de Ruedi Aebersold, cette approche repose sur l'utilisation de deux étiquettes portant des isotopes distincts pour marquer spécifiquement les résidus de cystéine des protéines par liaison covalente (Gygi *et al.*, 1999). Ces étiquettes sont composées de trois parties distinctes (Gartner *et al.*, 2007) (Figure 28, page 61) :

- un groupement réactif iodoacétamide qui réagit par alkylation avec les groupements thiols des cystéines de la protéine marquée.
- une étiquette de structure isotopique caractéristique (*Isotope coded tag*) qui induit à la différence de masse entre les deux formes de l'étiquette selon sa composition isotopique, neuf atomes de ^{12}C ou neuf atomes de ^{13}C , introduisant une différence de masse de 9 Da entre les deux étiquettes correspondant aux peptides lourd et léger.

- une étiquette biotine va permettre la purification des peptides marqués grâce à une chromatographie d'affinité.

Un bras espaceur clivable en milieu acide permet d'éliminer l'étiquette biotine pour l'analyse par spectrométrie de masse.

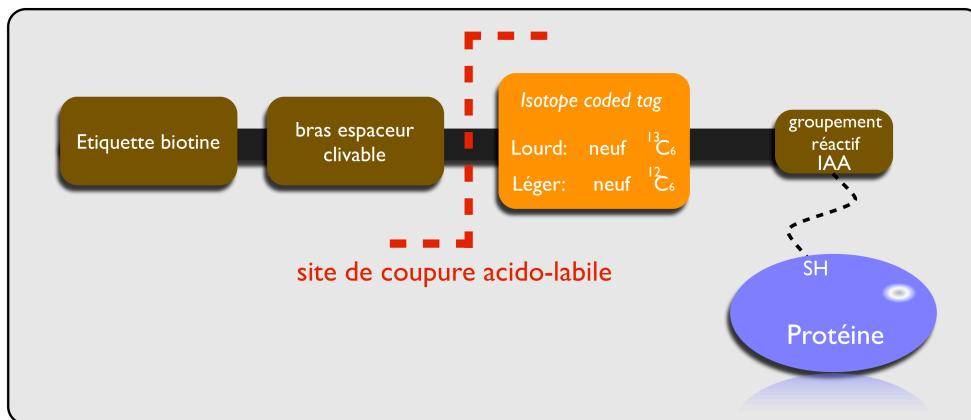


Figure 28 : Composition de l'étiquette ICAT. Elle est composée d'un groupement réactif iodoacétamide (IAA) qui réagit avec le groupement thiol des cystéines des protéines, d'une étiquette biotine qui permet de purifier les peptides marqués sur colonne d'avidine, d'un bras clivable en milieu acide et d'une étiquette de structure isotopique caractéristique (*Isotope coded tag*) qui induit à la différence de masse entre les deux formes de l'étiquette selon sa composition isotopique, neuf atomes de ^{12}C ou neuf atomes de ^{13}C .

Les protéines de chaque échantillon à comparer par une étiquette ICAT spécifique. Les échantillons sont ensuite mélangés en quantité équivalente et digérés par la trypsine. Après enrichissement des peptides marqués par purification sur colonne d'avidine, la quantification relative des peptides est réalisée par intégration du signal MS au niveau des massifs isotopiques des peptides lourds et légers (Figure 29, page 62).

Cependant, seules deux conditions expérimentales sont comparables par cette approche. De plus, cette technique est basée sur le marquage des cystéines des protéines, or la répartition des cystéines est hétérogène selon les protéines et il s'agit d'un acide aminé rare présent dans seulement 80% des protéines. Cette approche ne permet donc pas de quantifier de nombreuses protéines qui ne contiennent pas de cystéines (Schmidt *et al.*, 2005). Ce type de quantification n'est donc pas adapté à une analyse exhaustive de protéomes complexes.

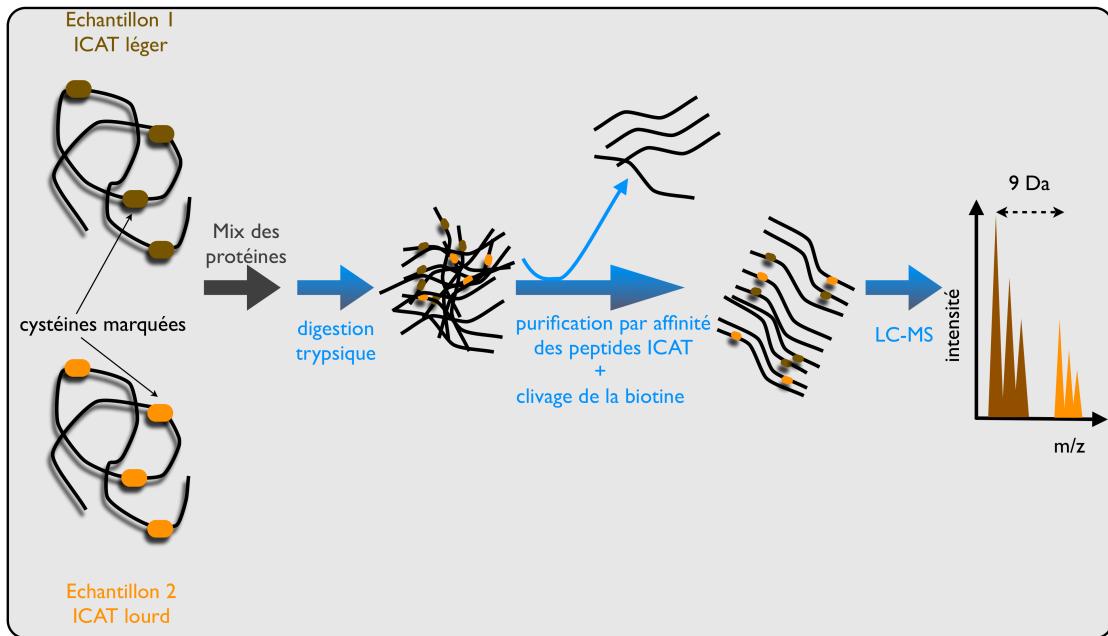


Figure 29 : Principe de la stratégie ICAT. Les protéines de deux échantillons biologiques à comparer sont d'abord marquées avec un réactif *léger* (neuf ^{12}C) ou *lourd* (neuf ^{13}C). Puis, les deux extraits sont mélangés et soumis à une hydrolyse enzymatique. Le marqueur utilisé possède à son extrémité un groupement biotine permettant ensuite la sélection spécifique des peptides contenant au moins une cystéine après le passage sur une colonne d'avidine immobilisée. La dernière étape consiste à analyser le mélange peptidique par LC-MS/MS. Deux peptides homologues issus chacun d'un des deux extraits cellulaires analysés se distinguent par un écart de 9 Da. Ainsi, le calcul du ratio de l'intensité entre les deux formes permet la quantification relative des protéines présentes dans le mélange.

b) iTRAQTM

La technologie iTRAQ a été introduite en 2003 par l'équipe de Thompson (Thompson *et al.*, 2003). Elle vise à palier les limitations de l'ICAT, à savoir qu'il devient possible de comparer jusqu'à huit conditions simultanément par cette approche (Choe *et al.*, 2007). Son principe repose sur le marquage des peptides trypsiques sur leurs groupements $\alpha\text{-NH}_2$ terminal et $\varepsilon\text{-NH}_2$ des chaînes latérales de lysines grâce à des étiquettes dites isobariques car la masse ajoutée au peptide est la même quelque soit l'étiquette utilisée. Par conséquent, les spectres MS sont relativement simples puisqu'un peptide marqué par les différentes formes du réactif présente une masse unique avec la co-élution parfaite des différentes formes du peptide marqué. Seul son spectre MS/MS atteste de ce marquage par les différentes formes du réactif. En effet, les peptides marqués génèrent des ions rapporteurs de masses spécifiques lors de la fragmentation des ions en MS/MS (Figure 30, page 63)

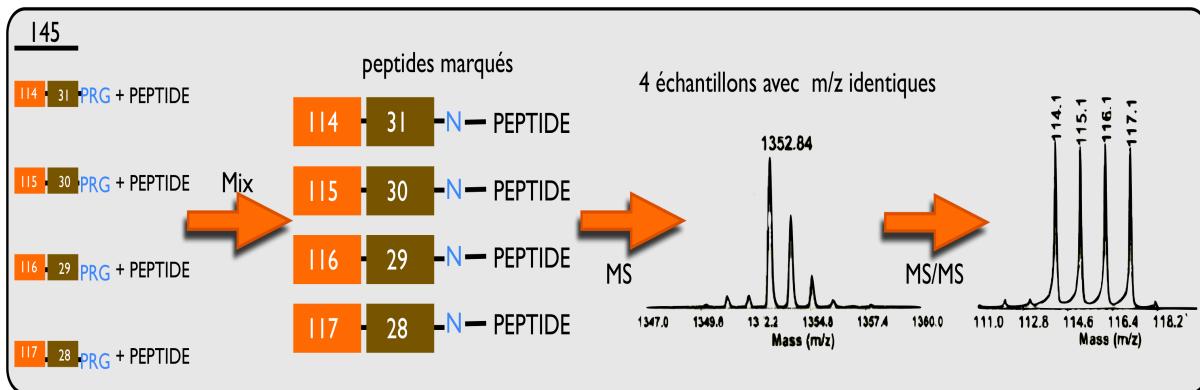


Figure 30 : Principe de l'analyse iTRAQ 4-plex. Les échantillons de protéines à comparer sont d'abord soumis à l'action enzymatique de la trypsine et les peptides résultants, marqués en N-terminal et sur la chaîne latérale des lysines avec une étiquette iTRAQ™ particulière. Après mélange et séparation des peptides par chromatographie, ils sont analysés en spectrométrie de masse. Lors de l'analyse en MS, les différents peptides marqués apparaissent sous la forme d'un seul et même pic correspondant à la masse du peptide + les 145 Da ajoutés par l'étiquette. Lors du fractionnement en tandem (MS/MS) chaque peptide ayant une étiquette différente génère une fragmentation peptidique caractéristique de l'étiquette greffée. La coupure préférentielle de part et d'autre du groupement perte de neutre provoque la libération des groupements rapporteurs de m/z 114 à 117 qui est donc proportionnelle à la quantité relative du peptide dans chacune des conditions auxquelles ils se rapportent. La quantification protéique est alors réalisée en comparant les intensités des 4 ions rapporteurs dans le spectre de masse MS/MS (inspiré de *Applied Biosystems* : <http://www.appliedbiosystems.com>).

L'étiquette iTRAQ™ est composée de trois parties (**Figure 31**, page 64) :

- un groupement réactif spécifique qui permet la fixation covalente de l'étiquette aux fonctions amines N-terminales des peptides ainsi qu'aux chaînes latérales des lysines.
- un groupement *rapporateur* chargé de m/z 113, 114, 115, 116, 117, 118 ou 121 Da selon l'étiquette.
- un groupement perte de neutre, non chargé, de 36 à 28 Da permet de contrebalancer les masses du groupement rapporteur de façon à avoir, quelque soit l'étiquette, une masse finale de 145 Da. Un site de fragmentation préférentielle en MS/MS est présent de chaque côté de ce groupement. Ainsi, la coupure préférentielle de part et d'autre du groupement perte de neutre provoque la libération des groupements rapporteurs de m/z 114 à 117 qui est donc proportionnelle à la quantité relative du peptide dans chacune des conditions auxquelles ils se rapportent. La quantification protéique est alors réalisée en comparant les intensités des 4 ions rapporteurs dans le spectre de masse MS/MS (Figure 30).

L'intégration des signaux obtenus en MS/MS et leur comparaison permet d'effectuer une quantification relative de chaque peptide marqué. Cette approche, contrairement à l'ICAT, ne nécessite donc pas la présence de paires de peptides pour la quantification relative. Ce qui offre l'avantage de décomplexifier le spectre car moins de chevauchements des signaux peptidiques sont à craindre (Pflieger *et al.*, 2008). Il a été démontré que le marquage iTRAQ est plus sensible que l'ICAT, ce dernier étant aussi sensible que la 2D-DIGE (Wu *et al.*, 2006)

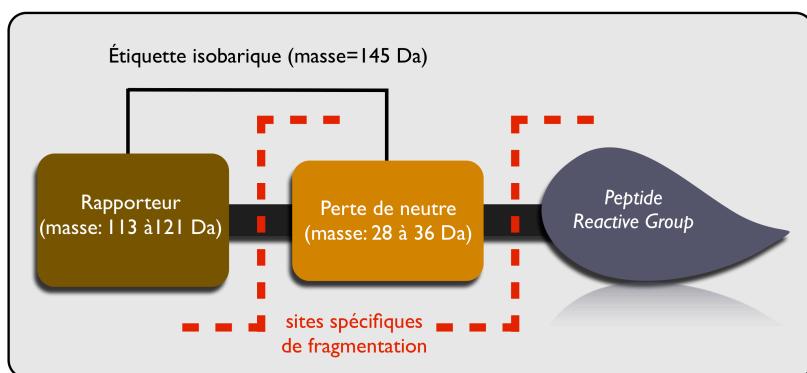


Figure 31 : Structure de l'étiquette iTRAQ™. Elle est composée d'un groupement réactif d'ester de N-hydroxysuccinimide (*Peptide Reactive Group*) qui se fixe de manière covalente aux fonctions amines des peptides et d'une étiquette isobarique qui permet de différencier les différentes formes du même peptide uniquement en analyse MS/MS grâce à la présence de sites spécifiques de fragmentation.

2) Marquage enzymatique

La technique de marquage enzymatique repose sur l'incorporation d'atomes d'oxygène à l'extrémité C-terminale des peptides générés au cours de la digestion protéolytique (Havlis *et al.*, 2003; Stewart *et al.*, 2001). Un des échantillons est soumis à l'action protéolytique de la trypsine dans de l'eau normale ($H_2^{16}O$) alors que le second est hydrolysé dans de l'eau lourde ($H_2^{18}O$). La trypsine ajoute deux atomes d'oxygène ^{18}O ou ^{16}O au niveau du groupement carboxyl-terminal (COOH) généré pour chaque peptide. L'analyse en spectrométrie de masse révèle pour chaque peptide des paires de pics, séparés de 4 Da (incorporation de deux ^{18}O). Ces paires sont utilisées pour la quantification relative du peptide dans chacun des échantillons (Gevaert *et al.*, 2008). Cette méthode, qui permet en théorie de marquer tous les peptides générés par la protéolyse, peut être appliquée à de très faibles quantités de matériel et est compatible avec pratiquement tous les types de préparation d'échantillon. Par contre, le marquage est réalisé au niveau du peptide et toutes les étapes précédentes dans la préparation des échantillons sont réalisées en parallèle et peuvent conduire à une variabilité expérimentale entre les deux échantillons à comparer. Enfin, il a également été démontré que l'incorporation des isotopes ^{18}O est très lente et incomplète pour les peptides très acides (Gevaert *et al.*, 2008).

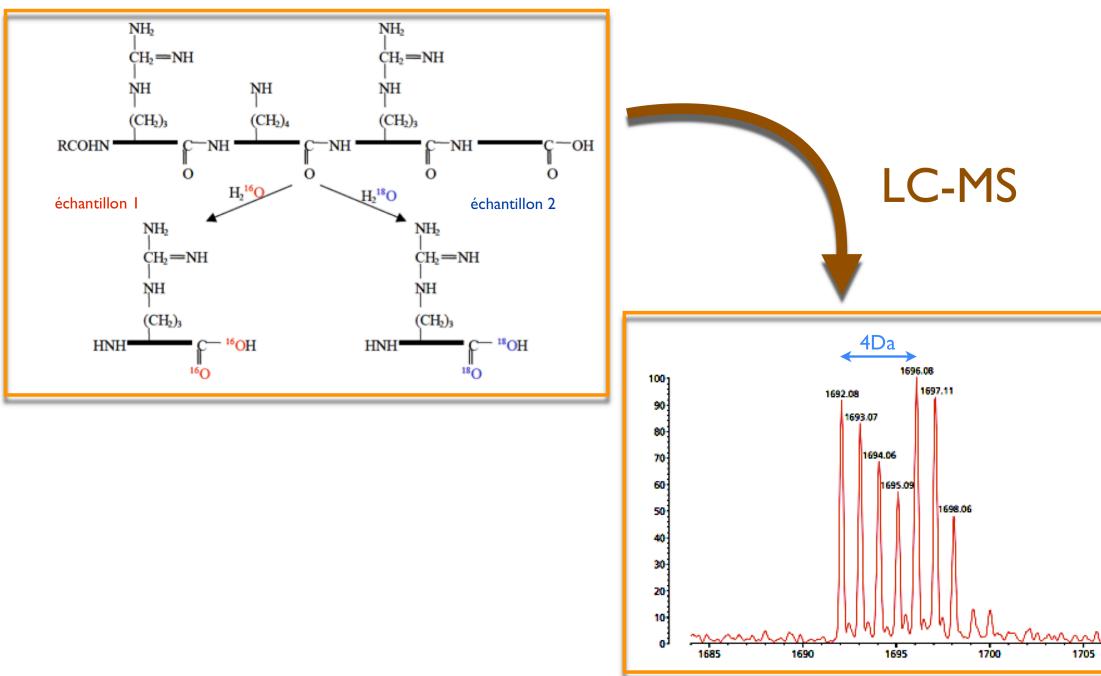


Figure 32 : Marquage enzymatique. Les deux échantillons sont préparés parallèlement jusqu'à la protéolyse incorporant une étiquette ^{16}O ou ^{18}O puis mélangés et analysés par chromatographie liquide couplée à la spectrométrie de masse. L'hydrolyse par la trypsine dans de l'eau lourde permet l'inclusion d' ^{18}O aux deux atomes d'oxygène des groupements carboxyliques en C-terminal des peptides.

3) Marquage métabolique : SILAC

Le SILAC est une stratégie qui permet de comparer les protéomes de types cellulaires maintenus en culture.

Initialement appelée AACT (*Amino Acid Coded mass Tagging*), elle a d'abord été mise au point par Oda *et al.* pour l'étude de protéines bactériennes grâce à un enrichissement à l'azote isotopiquement stable ^{15}N (Oda *et al.*, 1999). Cette approche a ensuite été adaptée à l'étude de cellules eucaryotes par l'équipe de Mathias Mann par l'ajout de leucine deutérée au milieu de culture de cellules d'une condition pour comparaison avec des cellules non marquées d'un état physiologique différent (Ong *et al.*, 2002; Gygi *et al.*, 1999). Elle a été mise en place dans le but de détecter de faibles variations d'expression spécifiques lors d'une étude protéomique fonctionnelle (Blagoev *et al.*, 2003).

Actuellement, le principe de l'approche SILAC est basé sur le remplacement d'un ou plusieurs acides aminés naturels essentiels, dans le milieu de culture de l'un des deux types cellulaires, par leur(s) équivalent(s) isotopique(s) lourd(s). La croissance cellulaire et la demi-vie des protéines conduisent les cellules à incorporer les acides aminés marqués afin de renouveler leur stock protéique et synthétiser des protéines soit *légères* soit *lourdes* dans deux cultures réalisées en parallèle. Après un nombre de doubléments cellulaires suffisant (au moins cinq) afin d'assurer un renouvellement complet du stock protéique, une quantité égale des deux populations de cellules (ou une quantité équivalente de protéines des deux conditions) est mélangée et traitée comme un seul échantillon. Après séparation par

SDS-PAGE, les protéines sont alors extraites du gel puis digérées par la trypsine. Les peptides résultants sont analysés et quantifiés par spectrométrie de masse. Sur les spectres MS obtenus, chaque peptide apparaît alors sous la forme d'une paire de massifs isotopiques utilisée pour le calcul de la quantité relative du peptide dans chaque condition. Les intensités de tous les ions-peptides contenant le marquage peuvent être comparées avec celles de leurs *homologues* non marqués. La méthode a été améliorée avec l'introduction dans le milieu de culture des résidus d'acides aminés [¹³C₆]Lysine et [¹³C₆]Arginine qui confèrent une différence de masse fixe de 6 Da entre les peptides lourds et légers, ce qui conduit au marquage de l'extrémité C-terminale de tous les peptides trypsiques et assure la co-élation des peptides lourds et légers en chromatographie liquide en phase inverse (Cicciaro et Blair, 2010) (Figure 33, page 67). C'est un avantage du SILAC par rapport aux premières approches de marquage métabolique au ¹⁵N modifiant l'extrémité N-terminale des liaisons peptidiques ainsi que les chaînes latérales des acides aminés contenant de l'azote. En effet, lors d'un marquage métabolique au ¹⁵N, l'écart de masse entre le peptide lourd et léger n'est alors pas fixe puisqu'il dépend de la séquence en acides aminés de chaque peptide et la quantification est donc plus difficile. Il est cependant nécessaire d'effectuer un double marquage Lys et Arg afin d'obtenir un maximum de peptides marqués pour une quantification précise et exhaustive (Graumann *et al.*, 2007). Enfin, il faut normaliser les ratios pour limiter tout biais dû à un renouvellement partiel du stock protéique. Il est possible de comparer jusque cinq conditions en jouant sur les combinaisons d'acides aminés marqués ([¹³C₆]Lys (+6 Da); [¹³C₆]Arg (+6 Da); [¹⁵N₄]Arg (+4 Da); [D₄]Lys (+4 Da); [¹³C₆,¹⁵N₂]Lys (+8 Da); [¹³C₆,¹⁵N₄]Arg (+10 Da); [¹³C₆,¹⁵N₂,D₉]Lys (+17 Da); [¹³C₆,¹⁵N₄,D₇]Arg (+17 Da)) (Molina *et al.*, 2009; Pan *et al.*, 2008; Seyfried *et al.*, 2010).

Le principal avantage de la méthode SILAC réside dans l'incorporation précoce de la forme isotopique sur les cellules en culture. Les échantillons sont mélangés au stade cellulaire, limitant ainsi les biais dus à des pertes inégales de protéines durant la lyse des cellules, l'extraction des protéines, la digestion ou les étapes de séparation. Par contre, étant limité à l'étude de cellules amplifiables en culture, le SILAC n'est donc pas applicable aux tissus ou échantillons cliniques (plasma,...). Toutefois, une méthode récente a été mise au point pour remédier à ce problème. Elle permet de quantifier qualitativement le protéome d'une tumeur humaine en combinant des lignées cellulaires marquées en SILAC directement avec des tissus de carcinome humain (Geiger *et al.*, 2010b).

À l'heure actuelle, la stratégie SILAC est l'une des méthodes de protéomique quantitative les plus fiables et les plus performantes. Ces atouts sont d'ailleurs reflétés par la bibliographie qui révèle que, comparativement aux autres techniques de l'analyse protéomique quantitative avec marquage, les dix dernières années ont montré un véritable essor du nombre publications reposant sur le SILAC. Cette approche est parfaitement

adaptée à l'étude quantitative des profils d'expression protéiques à grande échelle et elle est désormais une méthode de choix pour l'étude des interactomes (Hubner et Mann, 2011; Agrawal *et al.*, 2010; Hanke et Mann, 2009).

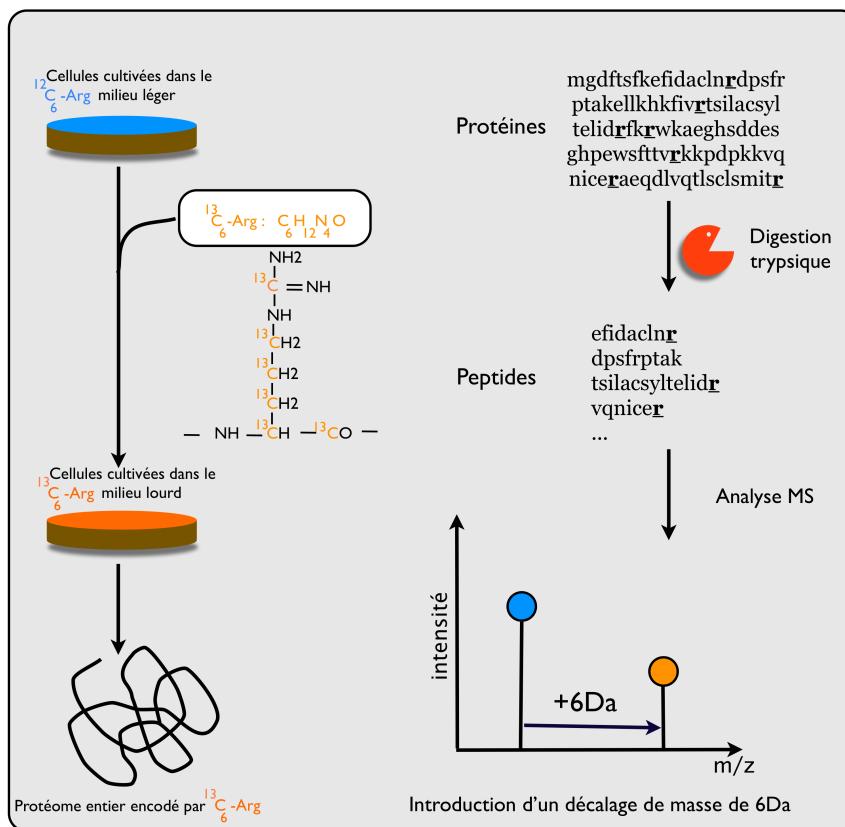


Figure 33 : Principe du SILAC. Les cellules sont cultivées dans un milieu de culture normal, milieu léger puis une partie est repiquée et cultivée dans un milieu lourd, contenant les acides aminés marqués par des isotopes stables comme l'arginine avec six ^{13}C . Pendant la division cellulaire, la synthèse des protéines et leur demi-vie conduisent à une incorporation métabolique de ces acides aminés SILAC dans tout le protéome. Les protéines sont ensuite hydrolysées par la trypsine en fragments peptidiques contenant des arginine $^{13}\text{C}_6$ (et lysines $^{13}\text{C}_6$) et un décalage de masse résidu-spécifique de 6 Da entre le peptide léger et le peptide lourd. Les deux formes utiles pour la quantification sont distinguables en MS.

Il existe de nombreux critères qui conduisent au choix de la stratégie de marquage. Le prix et la facilité ainsi que la vitesse d'introduction du marquage, la qualité de la séparation chromatographique, le nombre de peptides disponibles pour chaque protéine et la différence de masse entre les peptides marqués et non marqués peuvent être importants. Dans la plupart des cas, il est préférable d'introduire un marquage qui incrémentale la masse du peptide de manière homogène en évitant d'introduire une modification au niveau des caractéristiques d'hydrophobie pour une bonne séparation en chromatographie liquide. La stratégie doit être exhaustive et applicable à de petites quantités d'échantillons protéiques en mélange, ayant une large gamme dynamique. Il est préférable de mélanger les échantillons à comparer le plus précocement possible. Ce dernier point est l'argument majeur pour une incorporation métabolique des marquages. Notons aussi que la précision de mesure pour une quantification relative est aussi largement déterminée par le choix du spectromètre de masse.

C. Approches sans marquage *label-free*

Les techniques de marquage protéiques décrites précédemment permettent la quantification relative de protéines et peptides d'un échantillon. Leur inconvénient est qu'il s'agit de techniques onéreuses qui ne fournissent parfois pas suffisamment de données pour permettre une analyse statistique fiable et donner une signification biologique pertinente aux résultats obtenus. Cela est très sensible pour les études cliniques où un grand nombre d'échantillons doit être analysé pour expliquer la variabilité biologique. De plus, les techniques de marquage métabolique ne peuvent pas être utilisées pour tous les types d'analyse et les méthodes de dérivation de protéines nécessitent une optimisation adaptée à chaque type d'échantillon. C'est la raison pour laquelle un grand intérêt est porté à l'utilisation d'une LC-MS *label-free* quantitative en tant que technique alternative. L'approche *label-free*, peu onéreuse, vise également à déterminer l'abondance relative des protéines issues d'une ou plusieurs conditions. Deux approches sont principalement utilisées. La première consiste à comparer directement les intensités des signaux MS d'identification pour un même peptide dans différentes analyses chromatographiques (*ion current*). La seconde est basée sur le comptage et la comparaison du nombre de spectres de fragmentation (MS/MS) des peptides d'une protéine donnée (*spectral counting*) entre plusieurs analyses.

1) Comparaison des intensités des signaux MS

Dans le cas du *label free*, les échantillons à comparer sont analysés individuellement. Pour un même peptide dans les mêmes conditions expérimentales, la quantité est directement proportionnelle au XIC (Schulze et Mann, 2004; Bouyssié *et al.*, 2007; Cox *et al.*, 2009). Les logiciels MsQuant (Schulze et Mann, 2004), MaxQuant (Cox *et al.*, 2009; Geiger *et al.*, 2010a) et MFPaQ (Bouyssié *et al.*, 2007) ont également été développés afin d'extraire les intensités observées pour un même peptide dans différentes expériences.

2) Comparaison du nombre de spectres (*spectral counting*)

La stratégie de comptage et comparaison du nombre de spectres est empirique et basée sur l'observation selon laquelle le nombre de fragmentations (de MS/MS) obtenues pour une protéine est proportionnel à son abondance dans le mélange protéique (Lu *et al.*, 2007; Blondeau *et al.*, 2004; Ishihama *et al.*, 2005). Cependant, une normalisation est essentielle car une protéine de grande taille va générer plus de peptides qu'une protéine de petite taille. C'est la base du PAI (*Protein Abundance Indices*) qui est calculé en normalisant le nombre de peptides observés par le nombre de peptides théoriquement observables grâce à une hydrolyse *in silico* pour une protéine donnée (Rappsilber *et al.*, 2002; Liu *et al.*, 2004; Old *et al.*, 2005). La quantification relative entre deux peptides est donc révélée en comparant le nombre de spectres entre deux séries d'expériences. Une estimation plus juste de l'abondance relative des protéines dans un échantillon passe donc par la normalisation du nombre de

spectres MS/MS collectés pour les peptides en tenant compte des particularités des différentes protéines. Cette normalisation peut être réalisée pour chaque protéine en divisant le nombre de spectres MS/MS observés pour la protéine par la taille ou la masse de la protéine (Bernay *et al.*, 2009; Blondeau *et al.*, 2004; Bergeron et Hallett, 2007). Ce type de normalisation permet de corriger les estimations des abondances relatives des protéines de l'échantillon de manière satisfaisante. Lu et collaborateurs ont proposé dans leur approche APEX (*Absolute Protein EXpression measurement*) une méthode de normalisation plus poussée qui tient compte des propensions prédictives des différents peptides à être identifiés dans les analyses MS/MS. Les résultats obtenus avec ce mode de traitement ont montré une bonne corrélation avec ceux obtenus par d'autres méthodes. L'approche par *spectral counting* permet donc une estimation de l'abondance relative des protéines présentes dans un échantillon de manière simple. La connaissance de la quantité d'une ou plusieurs protéines présentes dans les mélanges (ou ajoutées artificiellement) permet d'estimer l'abondance absolue de l'ensemble des protéines du mélange.

Les principaux avantages du *label-free* sont l'absence de marquage, l'application à tous types d'échantillons et la réduction du coût de l'analyse. Cette technique est fréquemment utilisée et s'applique à de nombreux domaines en biologie. Cependant, le gros inconvénient est qu'elle nécessite une reproductibilité rigoureuse de la séparation en LC et une calibration précise des spectromètres de masse pour palier au risque d'erreurs multiples durant les procédés d'analyse individuelle des échantillons à comparer.

IX. Quantification absolue

Les stratégies de quantification décrites précédemment ont été développées dans le but d'étudier le dynamisme de protéomes entiers et de fournir une comparaison relative des quantités de protéines entre différentes conditions. Par contre, les stratégies de quantification absolue ont pour but de déterminer les quantités de protéines cibles dans des échantillons biologiques, permettant ainsi de comparer les données entre différents laboratoires (Viswanathan *et al.*, 2007; Pan *et al.*, 2009). Si le dosage ELISA (*Enzyme-Linked Immunosorbent Assay*), particulièrement la méthode *sandwich* reste la technique la plus utilisée de nos jours, récemment le dosage quantitatif de protéines, *via* leurs peptides prototypiques, par des techniques de spectrométrie de masse telles que le SRM (*Selected Reaction Monitoring*) ou le MRM (*Multiple Reaction Monitoring*) lorsque plusieurs dosages SRM sont réalisés simultanément, a été validé dans des échantillons biologiques complexes au sein de plusieurs équipes (Anderson et Hunter, 2006; Lee *et al.*, 2010a; Wei *et al.*, 2011). En amont du dosage par spectrométrie de masse, les protéines à doser sont hydrolysées par une endoprotéase. Les peptides spécifiques de la protéine sont alors dosés par spectrométrie

de masse. Les avantages des dosages par SRM ou MRM par rapport aux dosages ELISA, sont une réduction considérable du coût à long terme et de la durée de développement du dosage, notamment si des anticorps nécessaires au dosage ELISA doivent être développés (Zhi *et al.*, 2011). Il a été démontré qu'un grand nombre de peptides peuvent être quantifiés en une seule expérience de LC-MS grâce au mode MRM (Stahl-Zeng *et al.*, 2007; Lange *et al.*, 2008). De telles approches par spectrométrie de masse apparaissent donc comme des méthodes de choix pour le dosage de nombreuses protéines marqueurs potentiels, par les recherches en analyse protéomique différentielle (Carr et Anderson, 2008; Zhi *et al.*, 2011).

La quantification des peptides prototypiques est souvent réalisée dans un spectromètre de masse de type triple quadripôle, particulièrement adapté au mode MRM (Figure 22, page 50). Cette quantification est basée sur le principe de dilution isotopique (Bowers *et al.*, 1993) grâce à l'utilisation d'un standard interne : AQUA, QconCAT ou PSAQ.

A. Peptide AQUA

L'approche AQUA (*Absolute QUAntification*) développée en 2003 consiste à réaliser une quantification absolue de protéines choisies, contenues dans un échantillon biologique (Gerber *et al.*, 2003). Cette quantification est réalisée suite à l'incorporation, dans l'échantillon, d'une quantité connue d'un peptide synthétique nommé peptide AQUA et utilisé comme standard interne. Ce peptide possède une séquence identique au peptide natif mais contient des formes isotopiques stables de certains acides aminés (^{13}C , ^{15}N). Après protéolyse, l'échantillon est analysé en LC-MS. Les peptides natifs et AQUA ont le même temps de rétention et les mêmes propriétés d'ionisation mais ils sont distinguables en MS grâce à la différence de masse due au marquage isotopique. Le spectre MS met en évidence une paire de pics au niveau du peptide d'intérêt. Connaissant la quantité ajoutée de peptide synthétique, la quantité de peptide natif, donc de la protéine peut être déduite grâce à une comparaison des intensités de signal en MS. Les limites de cette approche sont que la quantification est faite à partir d'un seul peptide et que ce peptide AQUA est ajouté tardivement à l'échantillon (Kirkpatrick *et al.*, 2005). Cet ajout tardif est peu compatible avec un préfractionnement des échantillons. En effet, Havlis *et al.* ont montré qu'une simple étape de décomplexification de l'échantillon par électrophorèse monodimensionnelle pouvait dramatiquement affecter la précision de la quantification par l'approche peptide AQUA (Havlis et Shevchenko, 2004). Deux améliorations méthodologiques ont été mises en place. La première, appelée SISCAPA pour *Stable Isotope Standards and Capture by Anti-Peptide Antibodies* consiste à ajouter une étape d'enrichissement par immunoprécipitation du peptide prototypique et de son analogue AQUA entre la digestion trypsique et l'analyse par LC-MS (Anderson *et al.*, 2004). Comme l'approche SISCAPA augmente la sensibilité de détection, elle est particulièrement utilisée pour la quantification de protéines très peu

abondantes dans les échantillons complexes comme le plasma (Whiteaker *et al.*, 2007; Keshishian *et al.*, 2007; Brun *et al.*, 2009). La seconde permet d'incorporer des modifications covalentes comme la phosphorylation au niveau des standards AQUA, ce qui ouvre la voie à l'étude des enzymes de régulation et des voies de signalisation (Pan *et al.*, 2009; Brun *et al.*, 2009; Gerber *et al.*, 2003; Mayya *et al.*, 2006).

B. QconCAT

La technique QconCAT (*quantification concatamer*) utilise un principe similaire à l'AQUA mais a l'avantage de permettre la quantification absolue de plusieurs protéines simultanément (Beynon et Pratt, 2005). Elle repose sur la synthèse et le marquage métabolique d'un concatémère de peptides prototypiques caractéristiques des protéines d'intérêt à quantifier. Un gène synthétique codant pour ce concatémère des différents peptides doit donc être construit. Son expression permet alors la production d'une protéine chimérique constituée de l'accolement bout à bout des peptides de chaque protéine à doser (Rivers *et al.*, 2007; Pratt *et al.*, 2006). Comme pour la technique AQUA, cette protéine artificielle est synthétisée en y incorporant des formes isotopiques stables de certains acides aminés, permettant ainsi la quantification à partir du spectre MS.

C. PSAQ

La méthode de quantification PSAQ (*Protein Standard Absolute Quantification*) utilise comme étalons internes des protéines recombinantes entières de séquences identiques aux protéines naturelles à doser. Ces étalons PSAQ sont alourdis par incorporation de résidus de [¹³C₆, ¹⁵N₂]Lysine et [¹³C₆, ¹⁵N₄]Arginine lors de leur biosynthèse *in vitro*. Les étalons internes PSAQ, se comportant comme leurs *contre-parties* naturelles, peuvent être ajoutés préalablement à tout traitement de l'échantillon, ce qui permet d'éviter tout biais dû au processus analytique. L'échantillon est ensuite fractionné, digéré à la trypsine puis analysé en LC-MS. Chaque peptide issu d'un étalon PSAQ produit un massif isotopique caractéristique décalé du peptide naturel à doser. L'intégration des signaux *légers* et *lourds* résultants permet de déterminer la concentration exacte du peptide cible. La technique PSAQ est de fait, la seule qui prenne en compte les rendements des étapes de pré-fractionnement, de décomplexification et de digestion trypsique. En effet, des études comparatives ont montré que seule la méthode PSAQ permettait une quantification exacte des protéines (Lebert *et al.*, 2011; Huillet *et al.*, 2011; Brun *et al.*, 2007).

Travaux personnels

Première partie

**Discrimination des isoformes de la lactoferrine
grâce à la chimie Taqman**

I. Problématique et objectifs

L'expression de la Lf a souvent été étudiée dans de nombreux tissus et cellules mais les techniques utilisées ne permettent pas toujours de discerner les deux isoformes. De plus, leur expression simultanée dans de nombreux tissus sains complique la mise en évidence de l'une ou l'autre des isoformes. En effet, la différence entre les séquences primaires de ces protéines réside uniquement dans l'absence des 26 résidus d'acides aminés N-terminaux de la Δ Lf. C'est pour cela qu'il n'existe pas d'anticorps spécifiques de chacune des isoformes. On ne peut donc discriminer ces deux Lf que par leur masse moléculaire et évaluer leur niveau d'expression qu'après séparation électrophorétique. Il apparaît donc plus simple et plus judicieux de doser le niveau d'expression de chacun des transcrits. La transcription à partir des promoteurs alternatifs P1 et P2 permet de distinguer les messagers hLf et Δ Lf, l'exon 1 étant caractéristique du transcrit hLf et l'exon 1 β du transcrit Δ Lf.

Le but de mon travail a donc été de mettre au point un dosage, utilisant une technique de biologie moléculaire, la qPCR Taqman, afin de discriminer et quantifier les transcrits hLf et Δ Lf

II. Moyens mis en œuvre

Nous avons choisi d'utiliser pour cette étude deux cas pathologiques dans lesquels les niveaux d'expression des deux isoformes varient : le cancer et l'inflammation.

Plusieurs équipes dont la nôtre (Benaïssa *et al.*, 2005), ont démontré que les deux isoformes, hLf et Δ Lf, sont sous-exprimées au sein de nombreux tissus cancéreux et cellules tumorales (Siebert et Huang, 1997 ; Liu *et al.*, 2003 ; Benaïssa *et al.*, 2005 ; Goldberg *et al.*, 2005), ceci étant dû à des mutations génétiques, des pertes alléliques du chromosome 3 duquel est issu le gène de la Lf ainsi qu'à une hyperméthylation des deux promoteurs (Teng *et al.*, 2004 ; Iijima *et al.*, 2006). Les techniques utilisées pour ces études faisaient appel à la PCR semi-quantitative et étaient difficilement adaptées à des études cliniques.

Au laboratoire, Fillebeen *et al.* (2001) ont montré que l'expression de la hLf est augmentée dans les cellules microgliales en cas d'inflammation. Cette surexpression a également été constatée dans un contexte inflammatoire dans des cellules provenant d'autres espèces telles que les bovidés (Zheng *et al.*, 2005) et les murinés (Li *et al.*, 2009). Cette régulation positive du gène de la Lf chez ces trois espèces peut être due à la présence d'éléments de réponse aux *stimuli* inflammatoires.

Jusqu'alors, aucune étude n'avait été réalisée concernant une possible régulation de l'expression de la Δ Lf par ce type de *stimuli*. J'ai donc recherché à l'aide des outils informatiques *Genomatix* et *MatInspector* des éléments de réponse aux *stimuli*

inflammatoires tels que STAT3 et cRel/NF κ B au sein des promoteurs P1 et P2. En parallèle, j'ai conçu différents couples d'oligonucléotides ciblant l'exon 1 ou l'exon 1 β afin d'effectuer une quantification par qPCR des transcrits de la Lf. Cette technique ne s'est pas avérée assez sensible et spécifique pour détecter et doser les deux transcrits. J'ai ensuite optimisé le dosage en ayant recours à la technologie qPCR Taqman. J'ai, pour cela, mis au point une série d'oligonucléotides et de sondes Taqman qui ciblent spécifiquement l'exon caractéristique de chaque isoforme (Fig. 1 de l'article). Ce système m'a permis de doser simultanément l'une ou l'autre des isoformes et un gène de ménage, l'HPRT (Hypoxanthine PhosphoRibosylTransferase), afin d'effectuer une quantification relative (le Taqman en triplex ne pouvait être réalisé sur l'appareil de PCR quantitative utilisé au laboratoire).

III. Résultats

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

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

Hoedt, E., Hardivillé, S., Mariller, C., Elass, E., Perraudin, J.-P., et Pierce, A.

Cet article a été publié dans le journal Biometals (*Biometals*, 2010, 23 : 441-452)

Après analyse de l'expression de nos deux isoformes dans un contexte inflammatoire, j'ai pu montrer que même si les promoteurs P1 et P2 contiennent des éléments de réponse putatifs pour des *stimuli* inflammatoires, les transcrits de la hLf sont fortement surexprimés en réponse LPS alors que ceux de la Δ Lf ne le sont pas (Fig. 4 de l'article). J'ai également confirmé la sous-expression ou l'absence des transcrits des deux isoformes dans de nombreuses lignées cancéreuses par rapport aux lignées saines correspondantes.

Par ailleurs, nos résultats mettent en évidence une forte expression des transcrits Δ Lf dans les cellules immunitaires tels que les leucocytes (les transcrits Δ Lf ont une expression soixante fois supérieure à ceux de la hLf) et les lymphocytes (les transcrits Δ Lf ont une expression quinze fois supérieure à ceux de la hLf) (Fig. 2B de l'article). Ainsi, dans le but de purifier la Δ Lf, nous avons isolé les leucocytes et lymphocytes du sang humain (Deligny *et al.*, 2010) puis effectué sur le lysat de la fraction leucocytaire et lymphocytaire, des immunoprécipitations à l'aide d'anticorps dirigés contre la hLf. Après séparation de l'immunoprécipitat sur gel SDS-PAGE, de fines bandes ont été découpées entre 70 et 80 kDa. Les protéines contenues dans les bandes ont été hydrolysées par l'endoprotéase Glu-C (*Staphylococcus aureus Protease V8*) qui est une sérine-protéase qui hydrolyse

sélectivement après les résidus de glutamate et d'aspartate. La Glu-C est la seule protéase capable de discriminer les deux isoformes de la Lf par leur extrémité N-terminale. Notre analyse des peptides par MALDI TOF-TOF ne nous a pas permis de mettre en évidence ces peptides N-terminaux ni d'ailleurs aucun autre peptide appartenant à la Δ Lf. Cette absence de Δ Lf dans nos préparations de leucocytes et de lymphocytes pourrait s'expliquer par une régulation au niveau post-transcriptionnel. En effet, le transcrit de la Δ Lf possède deux cadres ouverts de lecture, un uORF, *upstream Open Reading Frame* présent dans l'exon 1 β directement suivi par un codon stop, ce qui implique un début de la traduction au niveau du pORF, *primary ORF*, présent dans l'exon 2. Chez les eucaryotes, les ribosomes fixent l'extrémité 5' coiffée de l'ARNm et balayent ensuite l'ARNm à la recherche du codon d'initiation. La présence d'un uORF en amont d'un pORF est connue pour perturber l'efficacité de la traduction (Kozak, 2001). En effet, des uORFs sont souvent retrouvés dans des oncogènes et dans des gènes impliqués dans le contrôle de la croissance et la différenciation cellulaire (Sachs et Geballe, 2006). Cette régulation pourrait expliquer pourquoi la Δ Lf est faiblement exprimée et pourquoi nous avons autant de difficultés à l'isoler.

IV. Conclusion

Notre étude démontre l'efficacité de la chimie Taqman dans la discrimination et la quantification des deux isoformes de la Lf. Ces résultats font de la qPCR Taqman une méthode de choix pour quantifier l'une ou l'autre isoforme dans tout modèle cellulaire. Suite à cela, nous avons établi une collaboration avec :

- le Dr. Bo Lonnerdal de l'Université de Davis en Californie, USA, afin de savoir si les transcrits dégradés en présence du micro-ARN, miR-214, que son équipe a mis en évidence sont ceux de la Lf ou de la Δ Lf puisqu'il cible la région 3'UTR commune aux deux transcrits.
- le Dr. Thierry Idziorek de l'Institut de Recherche contre le Cancer de Lille, pour étudier l'expression des deux isoformes dans des biopsies de patients atteints de leucémie myéloïde.

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 / Published online: 14 February 2010
© Springer Science+Business Media, LLC. 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 (Δ Lf) 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 Δ Lf 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 1 β for Δ Lf and a primer pairs/probe set within exon 1 and exon 2 for Lf. In this study, we confirmed that Lf/ Δ Lf Taqman gene expression assay is a powerful tool to investigate the expression of both Lf and Δ Lf transcripts. We also showed that lymphocytes and leukocytes isolated from fresh human blood expressed an extremely high level of Δ Lf 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 Δ Lf expression is not. These results suggest that the NF- κ B/cRel response elements present in the P1 promoter region are functional whereas those present in the P2 promoter region are not and show that Δ Lf is not regulated in inflammatory conditions.

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

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

Abbreviations

Lf	Lactoferrin
Δ Lf	Delta-lactoferrin

LPS	Lipopolysaccharides
FCS	Fetal calf serum
BSA	Bovine serum albumin
PBS	Phosphate buffer saline
NF- κ B	Nuclear factor κ B
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 (Δ Lf) 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 Δ Lf (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 1 β in the Δ Lf 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, Δ Lf 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 Δ Lf, 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 Δ Lf 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 Δ Lf

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- κ B 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 Δ Lf promoter regions.

Downregulation of the expression of Lf and Δ Lf 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 Δ Lf 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 Δ Lf 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% CO₂. 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 Dutcher, 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 10⁶ 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: gatacggcgccacctgtcaa			
Q-PCR	ΔLf primers	ccctcgactgtgtct	60	40	66
		Forward: ctgaaggttgaatctgcagtcaat			
Q-PCR	ΔLf probe	Reverse: aagggttgeaatggcactt	60	40	66
		tggtgctgtaccc			
Q-PCR	HPRT primers	Forward: gccctggcgctgtgatt	60	40	101
		Reverse: ctcagcataatgatttagtatgcaaaa			
RT-PCR	HPRT probe	tgtatgtgaaccaggat	69	16	240
		tgtatgtgaaccaggat			
RT-PCR	GAPDH	Marillier 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

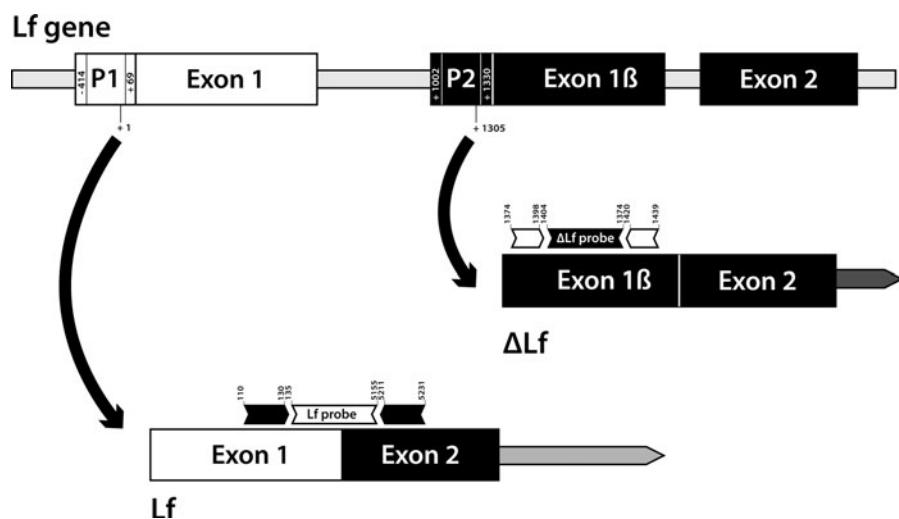
ethidium bromide and image acquisition was performed by UV transillumination using a Gel Doc 1000 system (Bio-Rad).

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 μ l of the QuantiTect Multiplex PCR Kit Mix (Qiagen),

Fig. 1 Diagrammatic 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 1 β 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 $E = 10^{[-1/\text{slope}]}$. 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 side-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_022 517.18, *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 $\Delta\text{Lf}/\text{HPRT}$) 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 ≥ 0.5 confirming strong gene silencing. Lf mRNA was found in the lymphocyte population and not in the leukocyte population. The absence of Lf

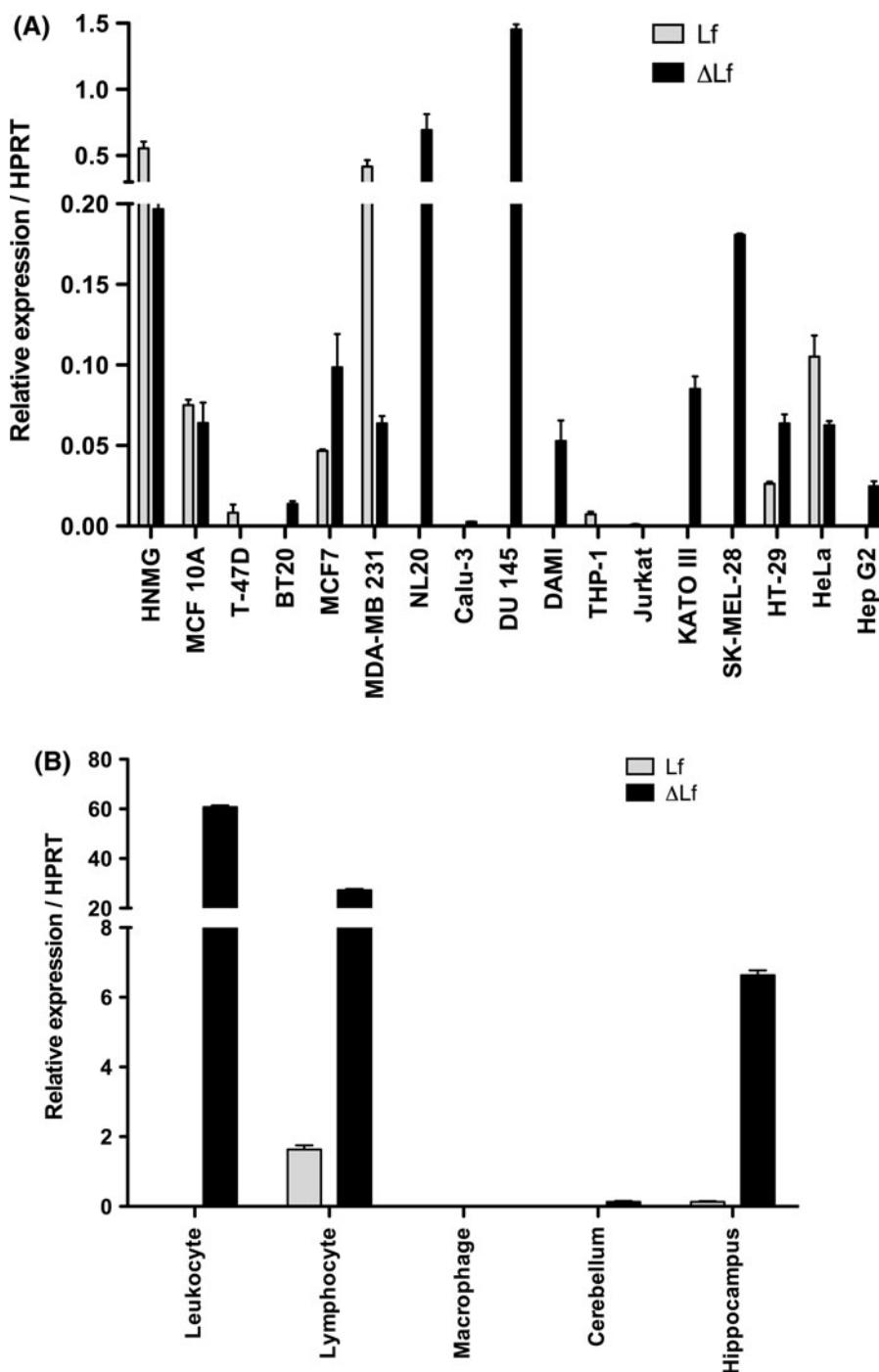
expression in leukocytes, which produce and store it in their secondary granules (Masson et al. 1969), is in accordance with its synthesis occurring only during granulocytic differentiation (Rado et al. 1987).

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 ≥ 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 ≥ 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 (Waslyuk et al. 1993), might be critical for ΔLf expression.

Presence of cRel/NF- κ B/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- κ B, 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-1 β (Karin 1999; Mercurio and Manning 1999) and activation of cRel/NF- κ B transcription factors.

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



The cRel/NF- κ B 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

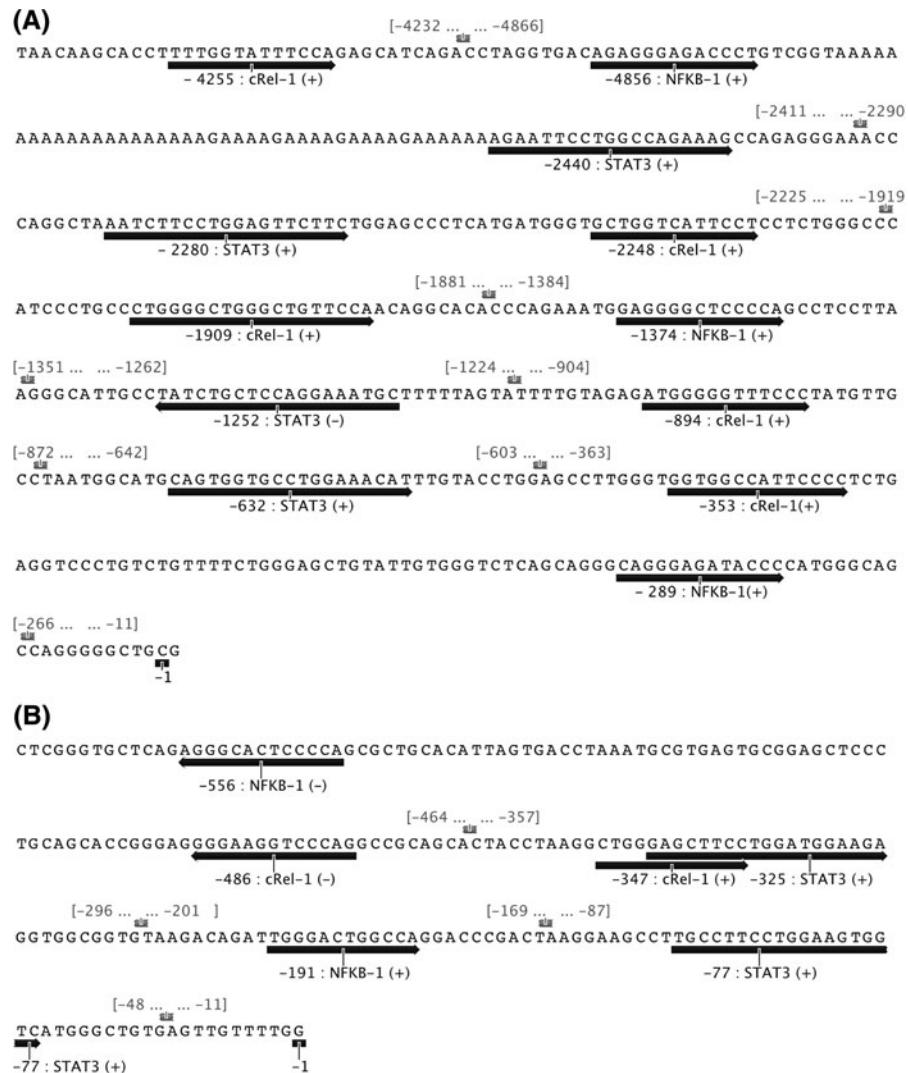
regulates gene expression in response to a variety of stimuli, including cytokines, growth factors, stress, and bacterial and viral infections and in turn controls a number of cellular processes including differentiation, proliferation, and apoptosis.

Therefore, we investigated the Lf gene for the presence of such inflammatory response elements and subsequently checked whether Lf and/or Δ Lf were overexpressed following LPS stimulation. The data presented in Fig. 3 show that potential NF- κ B/cRel/STAT3 promoter modules are present in the regulatory regions of the P1 and P2 promoter regions of the Lf gene. Three NF- κ B 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- κ B 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- κ B 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- κ B, 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- κ B 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 ($100\text{ ng}/10^6\text{ cells}$) was chosen

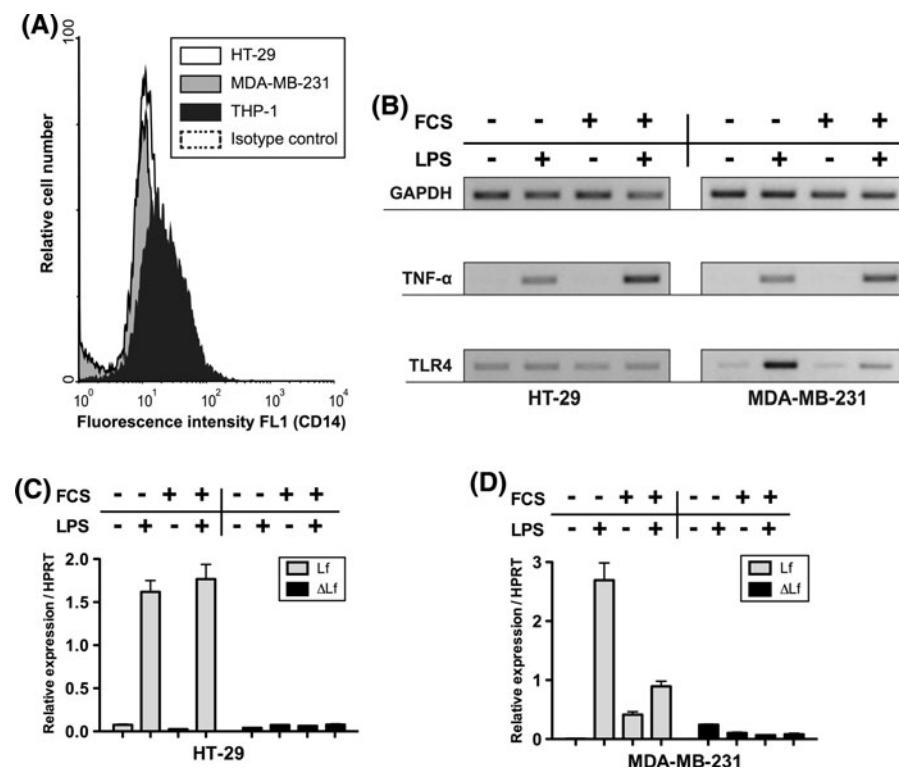
from previous experiments (data not shown) and in accordance to the literature (Li et al. 2009).

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. Δ Lf, 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 Δ Lf 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- κ B/STAT3 response modules are present in the P2 promoter, they are not functional and that Δ Lf 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 Δ Lf 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, Yogeve 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- κ B. *J Clin Investig* 107:241–246
- Barkett M, Gilmore TD (1999) Control of apoptosis by Rel/NF- κ B transcription factors. *Oncogene* 18:6910–6924
- Benäüssa 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
- Er ridge 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
- Fillebein 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- κ B 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- κ B 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, Carpenter 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, Benaïssa M, Hardivillé 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- κ B. *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, Benaïssa 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- κ B 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, Letureq 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
- Wasyluk 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- κ B 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

Deuxième partie

**Découverte d'une nouvelle cible de l'activité
transcriptionnelle de la
delta-lactoferrine**

I. Problématique et objectifs

La Δ Lf a été décrite pour la première fois, par Siebert et Huang, en 1997 comme étant l'isoforme intracellulaire de la Lf. L'expression de son transcrit a été retrouvée dans de nombreux tissus sains mais il est absent dans la plupart des cellules cancéreuses étudiées.

Nous avons voulu connaître l'impact de la réintroduction de la Δ Lf sur le protéome d'une lignée cellulaire immortalisée, les cellules HEK-293. Le travail qui m'a été confié au cours de mon Master Recherche a donc été de réaliser une étude protéomique différentielle *via* l'utilisation d'électrophorèses bidimensionnelles, afin de comparer sans *a priori* les protéomes de cellules exprimant ou non la Δ Lf.

II. Moyens mis en œuvre

Nous avons comparé les variations du protéome des cellules HEK 293 soumis à trois conditions différentes : non traitées, transfectées avec un vecteur d'expression vide, ou transfectées avec le vecteur d'expression de la Δ Lf (3xFLAG- Δ Lf). Des électrophorèses bidimensionnelles des lysats cellulaires correspondant à chaque condition ont été colorées au nitrate d'argent. Après numérisation, leurs images ont été analysées à l'aide du logiciel de quantification des intensités de coloration *ImageMaster 2D Platinum v6*. Les spots confirmés ont été sélectionnés pour une analyse par empreinte peptidique massique par MALDI-TOF ou en LC-MS/MS.

III. Résultats

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

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

Mariller, C., Hardivillé, S., **Hoedt, E.**, Benaïssa, M., Mazurier, J., et Pierce, A.

Cet article a été publié dans le journal Biochimie (*Biochimie*, 2009, 91(01) : 109-22)

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 trois conditions. L'analyse informatique du signal des 214 protéines a révélé une surexpression significativement différente de 14 d'entre elles en présence de Δ Lf. J'ai pu identifier 8 d'entre elles.

Nous avons ensuite cherché à savoir si la surexpression de ces protéines était

directement induite par l'activité transcriptionnelle de la ΔLf . Pour cela, nous avons recherché au niveau des promoteurs des gènes correspondants, l'élément de réponse à la ΔLf (ΔLfRE) et dont la fonctionnalité a été prouvée au laboratoire pour *Skp1*, *Bax* et *Fas* (Tableau V, page 77).

Nous savons que la ΔLf est dotée d'une activité pro-apoptotique et, parmi les protéines identifiées, trois interviennent dans la régulation de l'apoptose. En effet, la surexpression de la protéine *Poly(rC)-binding protein 2* (PCBP 2) conduit à une diminution de la viabilité cellulaire et à une augmentation de l'activité des caspases 3 et 9 entraînant ainsi l'apoptose des cellules qui la produisent (Roychoudhury *et al.*, 2007). Cette protéine n'ayant pas l'élément de réponse à la ΔLf au niveau de son promoteur n'a donc pas une expression directement liée à l'activité de la ΔLf . Quoi qu'il en soit, elle pourrait participer avec la ΔLf à l'induction de la mort cellulaire par apoptose. A l'inverse, la protéine *T complex subunit β* (TCPB) qui est également dépourvue de séquence ΔLfRE a été identifiée comme étant un interactant de la cycline E et comme étant surexprimée lors de la transition G1/S. Son inhibition conduit à une diminution de la viabilité cellulaire, à un arrêt du cycle en G1/S et à la mort cellulaire par apoptose (Coghlin *et al.*, 2006). Cette protéine a donc un rôle important dans la prolifération et la survie cellulaire. Sa surexpression pourrait contrebalancer l'effet pro-apoptotique de la ΔLf et permettre le passage en phase S préalablement bloqué par l'activité de cette dernière. De même, les expériences menées par Lu et Legerski montrent que la surexpression de la protéine *Pre-mRNA processing factor 19* (PRP19) participe également à la survie cellulaire. En effet, elle réduit l'apoptose après exposition des cellules aux agents mutagènes grâce à son activité de réparation de l'ADN (Lu et Legerski, 2007). Nos travaux ne nous ont pas permis de déceler une quelconque surexpression des transcrits de PRP19 en présence de ΔLf (Figure 5 de l'article) bien que cette protéine soit pourvue d'un $\Delta\text{LfRE-like}$ (Tableau V, page 77).

Nous avons également pu identifier 3 autres protéines. Il s'agit de la protéine disulfide isomérase (PDI) qui intervient dans la mise en conformation des protéines et dont le promoteur contient un élément $\Delta\text{LfRE-like}$. La *proteasome subunit β* impliquée dans la dégradation protéolytique et la protéine *vesicle fusing protein NSF* ou *AAA ATPase NSF* qui intervient dans la formation des vésicules intracellulaires. Les promoteurs des gènes correspondants ne contiennent pas de ΔLfRE donc la surexpression de ces protéines ne dépend pas directement de l'activité de la ΔLf .

Enfin, nous avons identifié 3 protéines impliquées dans la maturation et la stabilité des transcrits. La PRP19, citée précédemment pour son rôle dans le maintien de l'intégrité de l'ADN, est également dotée d'un domaine U-box à activité E3 ubiquitine ligase qui est indispensable à l'association des protéines impliquées dans la maturation des transcrits par

épissage (Chan et Cheng, 2005). La protéine « *heterogenous RiboNuclear Protein L* » (hnRNP-L), quant à elle, empêche la maturation des pré-ARNm en se fixant sur des séquences *cis* inhibitrices de l'épissage des exons (Rothrock *et al.*, 2005). Nos expériences n'ont pas montré de surexpression de ses transcrits en présence de ΔLf (Figure 5 de l'article). Il faut noter que cette protéine possède néanmoins un élément similaire et une séquence identique au ΔLfRE . L'absence de transactivation du gène *hnRNP-L* pourrait s'expliquer par l'absence des cofacteurs nécessaires à cette activation. La « *scavenger mRNA decapping enzyme* » (DcpS) hydrolyse la coiffe en 5' des ARNm (Liu et Kiledjian, 2005). Cette protéine voit ses transcrits surexprimés en présence de ΔLf et possède une séquence similaire au ΔLfRE dont nous avons confirmé la fonctionnalité *in vitro* et *in situ* grâce à des expériences de gène rapporteur et d'immunoprecipitation de la chromatine, respectivement (Figure 7 de l'article).

Tableau V : Séquences des différents ΔLfRE classés en fonction de leur efficacité. Ce tableau présente l'alignement des éléments de réponse retrouvés dans les gènes cibles de la ΔLf avec la séquence S1 décrite par He et Furmanski (1995) ainsi que la séquence ΔLfRE consensuelle. 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	
<i>Élément de réponse décrit par He et Furmanski (1995)</i>											
S1	G	G	C	A	C	T	-	T	A/G	C	
ΔLfRE^{Skp1}	G	G	C	A	C	T	G	T	A	C	100
ΔLfRE^{DcpS}		G	C	A	C	T	-	T	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
ΔLfRE	G/A	G	C	A	C	T	(G)	T	A/G	C/G	

IV. Conclusion

Notre étude nous a permis de mettre en évidence une nouvelle cible de l'activité transcriptionnelle de la ΔLf , la protéine DcpS. DcpS est une pyrophosphatase 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 débutant toutes deux par l'hydrolyse de la coiffe 7 méthylguanosine ou 5'-cap (Figure 34, page 78) (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 (Liu et Kiledjian, 2006; Coller et Parker, 2004). 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 être impliquée indirectement dans la régulation de l'épissage ainsi que dans celle de 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.

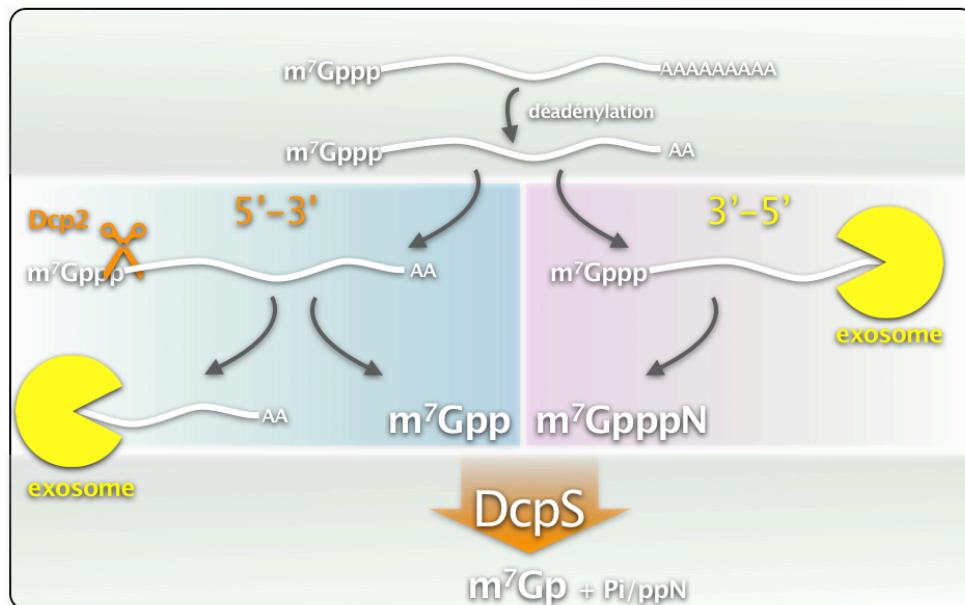


Figure 1 : Modèle de la dégradation des ARNm. Après la déadenylation, les ARNm sont dégradés par l'exosome de 3' en 5' libérant ainsi du m⁷GpppN. L'autre voie de dégradation fait intervenir Dcp2 pour l'hydrolyse de la coiffe libérant du 7mGDP. La dégradation des ARNm de 5' en 3' peut alors être effectuée 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).

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 $\Delta\text{Lf-M4}$ 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 $\Delta\text{Lf-M4}$ -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.

© 2008 Elsevier Masson SAS. All rights reserved.

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).

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 GGCACCTTGC sequence had previously been described [15] and found to be responsible for IL-1 β 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 $\Delta\text{Lf-M4}$ 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 $\Delta\text{Lf-M4}$ 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 (ATCC 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 $\mu\text{g}/\text{ml}$) [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 Dreamfект® 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 × 10⁶ 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) dithiothreitol (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)
<i>RT-PCR</i>				
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'-TCCCTAACATCTGCTCCAT-3' F: 5'-CGCACCCACAACACTT-3'	60	30	464
NSF	S: 5'-ATTGACTCCAACCCCT-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'-GTTTAATGGCTTCTGCT-3'	45	40	393
RPLP0	S: 5'-GATGACCAGCCCAAAGGAGA-3' F: 5'-GTGATGTGCAGCTGATCAAGACT-3'	55	22	101
<i>qRT-PCR</i>				
DcpS	S: 5'-AGACGGGAGATGACTACAGG-3' F: 5'-GTGGGTGTATAACATTCTCGAC-3'	55	40	86
RPLP0	S: 5'-GATGACCAGCCCAAAGGAGA-3' F: 5'-GTGATGTGCAGCTGATCAAGACT-3'	55	40	101
<i>ChIP</i>				
<i>DcpS</i> promoter	S: 5'-CCTATCCCCAGGCTTTCAC-3' F: 5'-GGTCTCAATTCCCTGACCTCG-3'	60	36	136
<i>Albumin</i> promoter	S: 5'-GTGGCTCTGATTGGCTTCTG-3' F: 5'-ACTCATGGAGCTGCTGGTTC-3'	60	36	139
Site-directed mutagenesis				
ΔLf-M4:	Ser10: S: 5'-CGTGGCCCTCCTGTCGCCCTGCATAAAGAGAGA-3' F: 5'-TCTCTTTATGCAGGCGACAGGGAGGGCCACG-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'-GCTTCTAGCCTCAGCCACAGGCTCCGT-3'		manufacturer's instructions	
Plasmid construction				
<i>DcpS</i>	S: 5'-CCTATCCCCAGGCTTTCAC-3' F: 5'-GGTCTCAATTCCCTGACCTCG-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)}/Σ 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 [CH₃OH/H₂O 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 (H₂O/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 C₁₈ PepMap column (Dionex). The flow rate was set at 200 nL/min. Peptides were eluted using a 5–100% linear gradient of solvent B (H₂O/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 μ l 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 μ l 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 μ l of ribonuclease A (20 mg/ml Sigma) and 2 μ l 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 μ l 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.

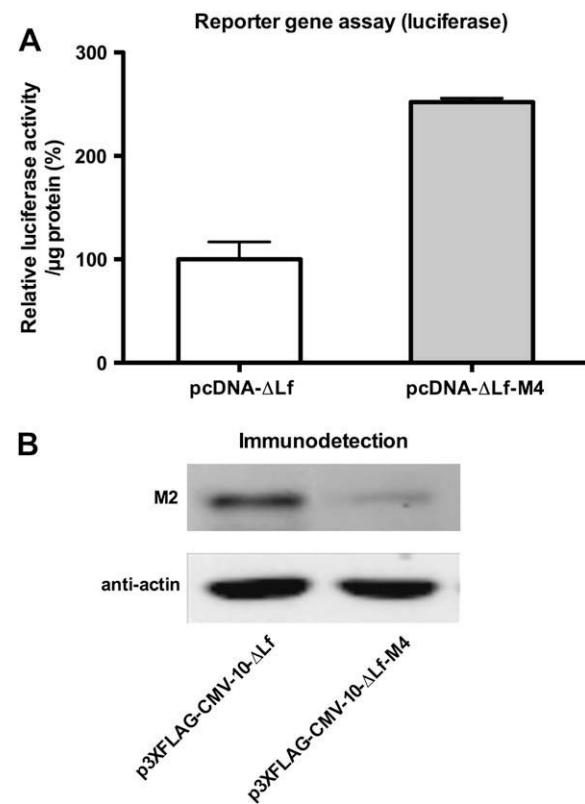


Fig. 1. ΔLf and ΔLf -M4 transactivate transcription from the *Skp1* promoter. Panel A. HEK 293 cells were co-transfected with pGL3-S1^{Skp1} and with either pcDNA- ΔLf or pcDNA- ΔLf -M4 expression vector (200 ng/well) encoding ΔLf or ΔLf -M4 respectively. Cells were lysed 24 h after transfection and 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 p3XFLAG-CMV-10- ΔLf -M4 vectors. Cells were lysed 24 h after transfection and samples (20 μ g 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 endogenous ΔLf . Since *O*-GlcNAc negatively regulates ΔLf activity, mutagenesis of the four *O*-glycosylation sites was performed and the resulting mutant named $\Delta\text{Lf-M4}$. The transcriptional activity and the level of expression of $\Delta\text{Lf-M4}$ 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 $\Delta\text{Lf-M4}$. 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 $\Delta\text{Lf-M4}$ 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 , $\Delta\text{Lf-M4}$) 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 \pm 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 $\Delta\text{Lf-M4}$ class by comparison with the control and ΔLf classes. Fig. 3 summarizes the levels of the differentially

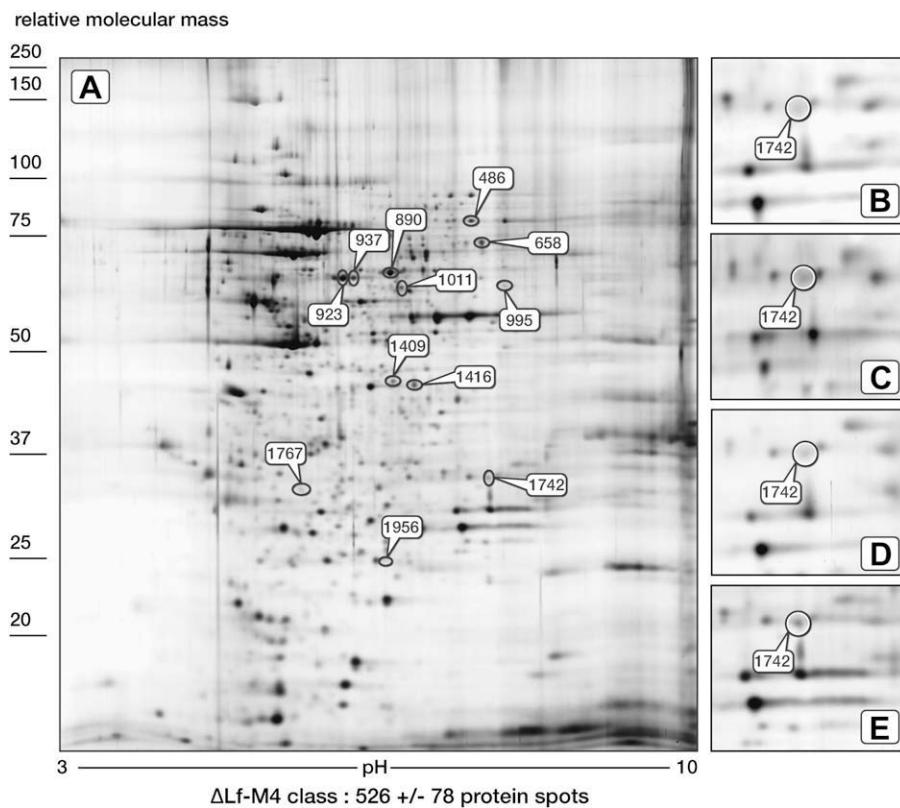


Fig. 2. Proteomic profiling of HEK 293 cells transfected with ΔLf constructs. Panel A: representative 2D image of HEK 293 proteins after transfection by pcDNA- $\Delta\text{Lf-M4}$. Panels B–E: 2D silver-stained portions encompassing DcpS protein from HEK 293 cells. B: not transfected; C: transfected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

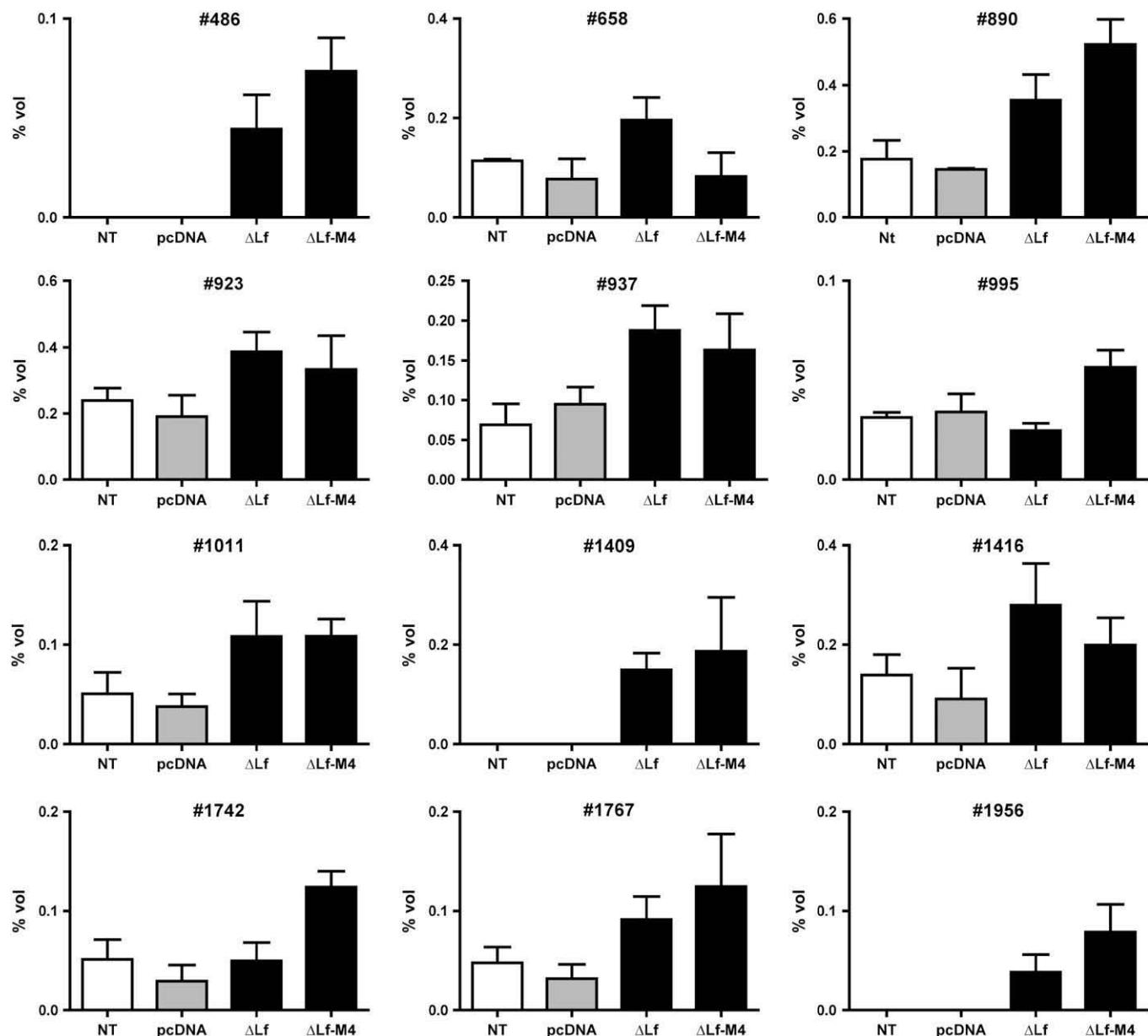


Fig. 3. Modifications in protein levels in Δ Lf or Δ Lf-M4-expressing-cells. HEK 293 cells were not transfected (NT) or transfected with pcDNA3.1 alone, pcDNA- Δ Lf (Δ Lf) or pcDNA- Δ Lf-M4 (Δ Lf-M4). Their total protein content was separated by 2D electrophoresis. Twelve protein spots, of which the intensity varied significantly ($p < 0.05$) were selected and plotted as mean \pm 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- Δ Lf or pcDNA- Δ Lf-M4, whereas two proteins (spots 995 and 1742) were strongly upregulated when transfection occurred with pcDNA- Δ Lf-M4 and one protein (spot 658) was upregulated only in the presence of pcDNA- Δ Lf 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- Δ Lf 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 μ g 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 <i>Gene name</i>	Sequence coverage (%)	Number of peptides
MALDI-TOF					
#486	P46459	Vesicle-fusing ATPase	NSF_HUMAN <i>NSF</i>	29	12
#658	P14866	Heterogeneous nuclear ribonucleoprotein L	HRNPL_HUMAN <i>HRNPL</i>	25	9
#890	P78371	T-complex 1 subunit β	TCPB_HUMAN <i>CCT2</i>	43	20
#923	P30101	Protein disulfide isomerase A3	PDIA3_HUMAN <i>PDIA3</i>	30	16
#1011	Q9UMS4	Pre-mRNA processing factor 19	PRP19_HUMAN <i>PRPF19</i>	24	10
#1742	Q96C86	Scavenger mRNA-decapping enzyme DcpS	DCPS_HUMAN <i>DCPS</i>	21	10
#1956	P49720	Proteasome subunit beta type 3	PSB3_HUMAN <i>PSMB3</i>	41	12
LC-MS/MS					
#658	P14866	Heterogeneous nuclear ribonucleoprotein L	HRNPL_HUMAN <i>HRNPL</i>	19	8
#890	P78371	T-complex 1 subunit β	TCPB_HUMAN <i>CCT2</i>	26	17
#923	P30101	Protein disulfide isomerase A3	PDIA3_HUMAN <i>PDIA3</i>	27	15
#1011	Q9UMS4	Pre-mRNA processing factor 19	PRP19_HUMAN <i>PRPF19</i>	21	10
#1416	Q15366	Poly(rC)-binding protein 2	PCBP2_HUMAN <i>PCBP2</i>	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 10⁶ 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

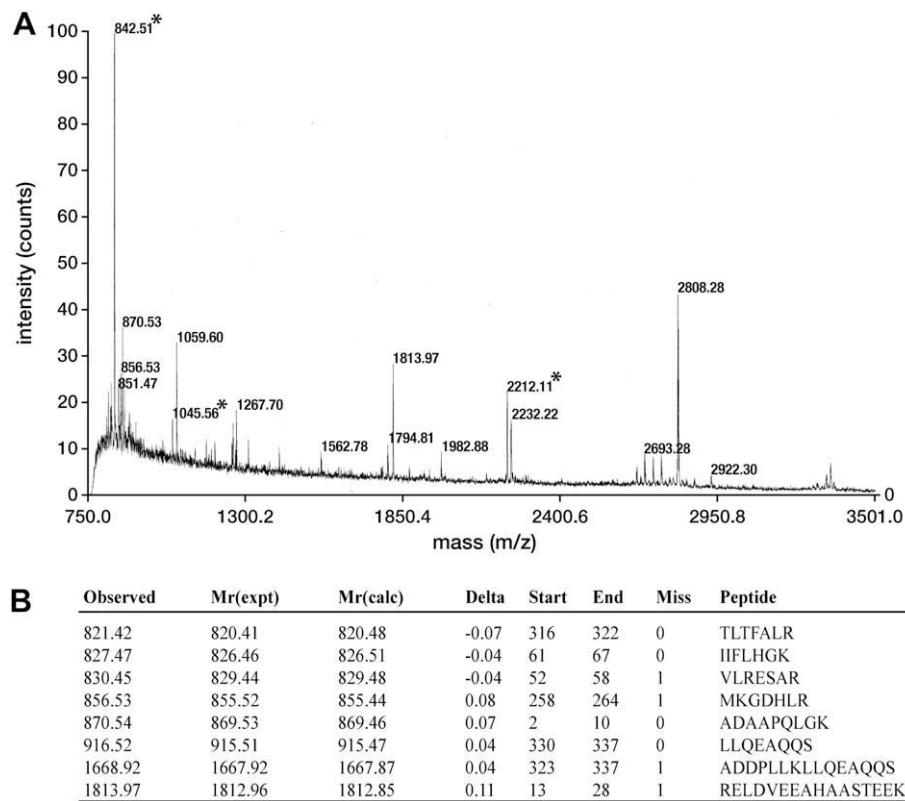


Fig. 4. Panel A. MALDI-TOF peptide-mass fingerprint spectrum of the tryptic digest of spot 1742 protein. *, the trypsin autolytic fragments. Panel B. Proposed peptides were obtained by MASCOT search (MASCOT Peptide Mass Fingerprint: http://www.matrixscience.com/cgi/search_form.pl?FORMVER=2&SEARCH=PMF) 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 Δ Lf 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 Δ Lf 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 Δ Lf 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 Δ Lf and might be targets of its transcriptional activity.

Since Δ Lf 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 Δ Lf 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 Δ Lf isoform in which all the *O*-GlcNAc sites were mutated. The Δ Lf-M4 mutant is, thus, a constitutively active Δ Lf 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 Δ Lf *in vivo*. Our data showed that nearly all the same proteins were overexpressed in Δ Lf- and Δ Lf-M4-expressing HEK293 cells with a higher expression in the latter.

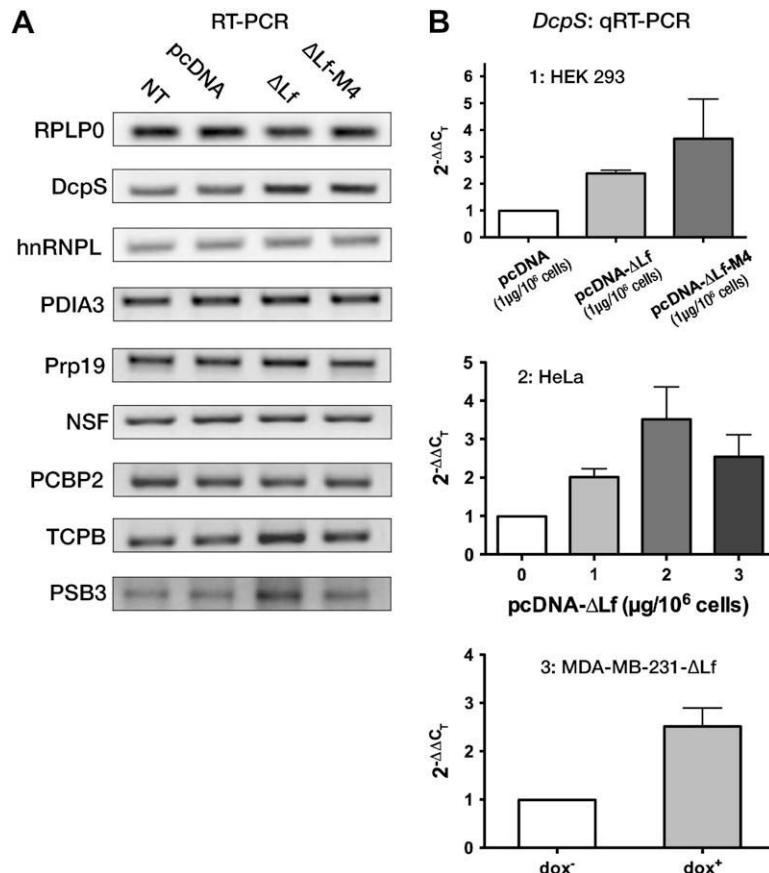


Fig. 5. Panel A. Upregulation of *DcpS* and *TCPB* mRNA in ΔLf and $\Delta\text{Lf}-\text{M}4$ -expressing-HEK 293 cells. HEK 293 cells were either not transfected (NT) or transfected with pcDNA3.1 only (pcDNA), pcDNA- ΔLf (ΔLf) or pcDNA- $\Delta\text{Lf}-\text{M}4$ ($\Delta\text{Lf}-\text{M}4$). Among 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 $2^{-\Delta\Delta C_t}$ of *DcpS* expression compared to RPLP0 expression ($n = 3$). Panel B1. HEK 293 cells transiently transfected with pcDNA, pcDNA- ΔLf or pcDNA- $\Delta\text{Lf}-\text{M}4$. 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
S1 ^{DcpS}	A	G	C	A	C	T	T	G	G	-2228 to -2220

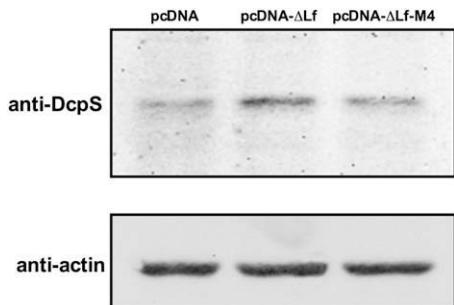


Fig. 6. Immunodetection of DcpS. HEK 293 cells were transfected with pcDNA, pcDNA- Δ Lf or pcDNA- Δ Lf-M4 expression vector (1 μ g/10⁶ cells). 24 h after transfection, total cell extracts were prepared, and samples (30 μ g of protein) were loaded on 10% SDS-PAGE. Western blot detection was done using 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

favors cellular viability [49]. Therefore, these three proteins might participate with Δ Lf 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 spliceosome [50] and its U-box domain is essential for mRNA maturation *in vivo* [43]. The heterogeneous 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 Δ Lf 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 Δ Lf. 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 Δ LfRE 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 Δ LfRE 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 m⁷GMP 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 Δ Lf, 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 Δ Lf-expressing cells.

In conclusion, the proteome profiling technique provided a broad-based and effective approach to identify protein changes induced by Δ Lf and has allowed us to pinpoint DcpS as a new Δ Lf 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 Δ Lf and its

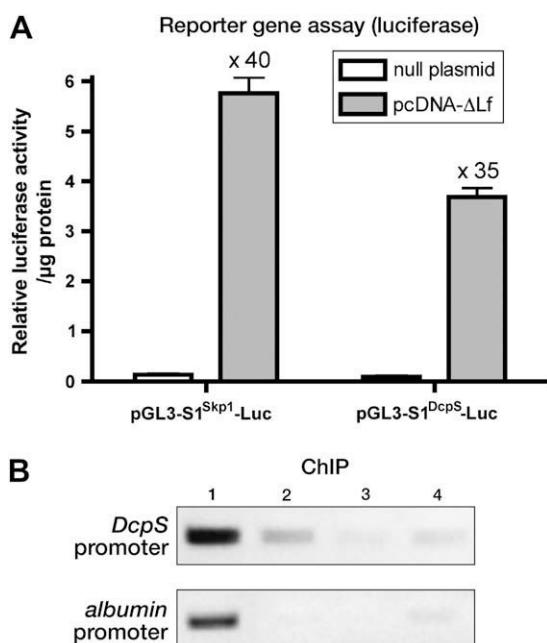


Fig. 7. Δ Lf transactivates transcription from the *DcpS* promoter. Panel A. HEK 293 cells were cotransfected with pGL3-S1^{DcpS}-Luc (A) or pGL3-S1^{Skp1}-Luc (B) constructs (50 ng/well) and with a null plasmid or with pcDNA- Δ Lf expression vector (200 ng/well) encoding Δ Lf, respectively. Cells were 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 \pm SD of triplicates from three independent measurements. Panel B. Δ Lf binds to the *DcpS* promoter *in vivo*. 3XFLAG- Δ Lf/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.

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. Furuchi, 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. Pradelles, 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. Coglin, 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. Liehaber, 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. Guerrero, 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.

Troisième partie

**Impact de la réintroduction de molécules à
activité anti-tumorale,
Lf et Δ Lf, dans les cellules cancéreuses**

MDA-MB-231

I. Introduction

Nous avons voulu étudier le mécanisme d'action des isoformes de la Lf à activité anti-tumorale, en étudiant les conséquences de leur expression sur le protéome de cellules de glande mammaire cancéreuses à caractère invasif, les cellules MDA-MB-231. En effet, les deux isoformes de lactoferrine humaine, hLf et Δ Lf, sont capables de contrôler la progression du cycle cellulaire et l'apoptose, deux mécanismes dont la dérégulation est directement impliquée dans le processus de cancérisation. La Δ Lf et la hLf sont sous-exprimées ou absentes dans le cas de cancer. Elles sont toutes deux des suppresseurs de tumeur potentiels et leur expression est un facteur de bon pronostique dans le cas de cancer du sein. Au laboratoire, les expériences réalisées sur les cellules HEK-293 ont démontré la capacité qu'a la Δ Lf à exercer une activité anti-tumorale. Nous avons aussi vu dans l'introduction du manuscrit que la majorité des études qui ont permis la mise en évidence de l'activité anticancéreuse de la Lf de sécrétion ont été réalisées grâce à l'administration orale de bLf sur des modèles animaux ou dans le cas d'études cliniques. Les études *in vitro* quant à elles ont été réalisées sur des modèles cellulaires hétérologues avec des doses de bLf variant d'un facteur 1 à 100. En conséquence, nous avons décidé de réaliser notre étude dans un système cellulaire unique pouvant indifféremment exprimer la Δ Lf ou être traité par de la hLf exogène. Pour cela nous avons établi des cellules MDA-MB-231 exprimant la Δ Lf de manière stable et inductible. De ce fait, lorsque l'on induit pas l'expression de la Δ Lf, ces cellules peuvent alors être traitées avec de la hLf exogène préalablement purifiée à partir du lait maternel. La quantité de hLf que nous avons choisi d'utiliser pour traiter ces cellules a été décrite comme étant nécessaire et suffisante pour arrêter le cycle des cellules MDA-MB-231 en G1/S (Damiens *et al.*, 1999). Afin d'être à même d'appréhender les modifications globales du protéome induites par la Δ Lf et par la hLf, tout en intégrant un aspect quantitatif rigoureux, nous avons effectué une analyse protéomique quantitative de type SILAC en collaboration avec l'équipe du Dr. Bernard Monsarrat (*IPBS*, Institut de Protéomique et de Biologie Structurale, Toulouse). Cette approche a été choisie car le marquage des protéines est, dans ce cas, un marquage métabolique effectué le plus en amont possible de l'expérience, évitant ainsi les biais de manipulations donc de quantification liés au traitement des échantillons en parallèle. La difficulté majeure a été d'effectuer une analyse SILAC en triplex afin d'intégrer au sein de la même manipulation l'impact des deux isoformes.

À la suite de la première expérience, le protéome s'est avéré n'avoir été que modérément affecté en réponse aux deux isoformes de Lf. Nous avons ensuite validé les surexpressions significatives de protéines par *western-blot* et qPCR. Les sous-expressions seront validées par la suite. Dans un second temps, plutôt que dupliquer l'expérience, nous

avons choisi d'augmenter les doses d'induction afin d'accroître la proportion de protéines dont l'expression est modifiée, avec si possible des ratios de sur- ou sous-expression plus importants. Dans ce but, nous avons, d'une part, transfecté les cellules MDA-MB-231 de manière transitoire à l'aide d'un vecteur d'expression pour la ΔLf et, d'autre part, multiplié par dix la quantité de hLf exogène pour traiter ces cellules, ce qui correspond à la dose nécessaire et suffisante pour induire une apoptose (Lee *et al.*, 2010b).

II. Résultats

A. Identification et quantification du protéome des cellules MDA-MB-231 en réponse aux isoformes de la Lf par l'approche SILAC

1) Marquage des cellules par SILAC et identification des protéines

Nous avons réalisé une analyse différentielle quantitative du protéome des cellules MDA-MB-231 en réponse aux isoformes de la Lf en utilisant le SILAC. Un schéma décrivant la stratégie expérimentale employée est présenté dans la Figure 35A (page 81). Nous avons choisi de réaliser un marquage double, arginine et lysine, pour augmenter le nombre de peptides quantifiables lors de l'analyse par spectrométrie de masse. Le marquage des résidus d'acides aminés Lys et Arg est réalisé avec le ^{13}C et ^{15}N car il est préférentiel à un marquage au deutérium (^2H) qui risque d'induire un décalage de masse du temps de rétention en chromatographie liquide (Zhang *et al.*, 2002). Ainsi, même s'il y a un chevauchement des massifs isotopiques entre [$^{13}\text{C}6$, $^{14}\text{N}2$]-Lysine (Lys 6) et [$^{13}\text{C}6$, $^{15}\text{N}2$]-Lysine (Lys 8), nous pouvons toujours valider la quantification en nous référant aux peptides à arginine qui ont permis l'identification. De plus, l'utilisation d'un spectromètre de masse à haute résolution comme l'Orbitrap permet de distinguer les peptides qui co-éluent et possèdent des *m/z* proches (Makarov *et al.*, 2006). Avant de débuter notre analyse, nous avons aussi vérifié l'incorporation totale des acides aminés marqués. Pour cela, des extraits protéiques provenant les conditions *medium* et *lourdes* ont été analysés par spectrométrie de masse et ont montré une efficacité d'incorporation supérieure à 99%.

Suite au mélange d'un nombre équivalent de cellules pour chacune des trois conditions, nous avons réalisé un fractionnement subcellulaire conduisant à une fraction nucléaire et une fraction cytoplasmique. Le but principal de cette étape est de réduire la complexité des échantillons à analyser en spectrométrie de masse. L'obtention de fractions enrichies a été validée par analyse en *western-blot* à l'aide d'anticorps ciblant des protéines spécifiques de chaque compartiment (Figure 35B, page 81). La mise en évidence de la protéine H2B uniquement dans la fraction nucléaire et la faible immuno-révélation de la protéine GAPDH dans cette même fraction confirment l'efficacité du fractionnement.

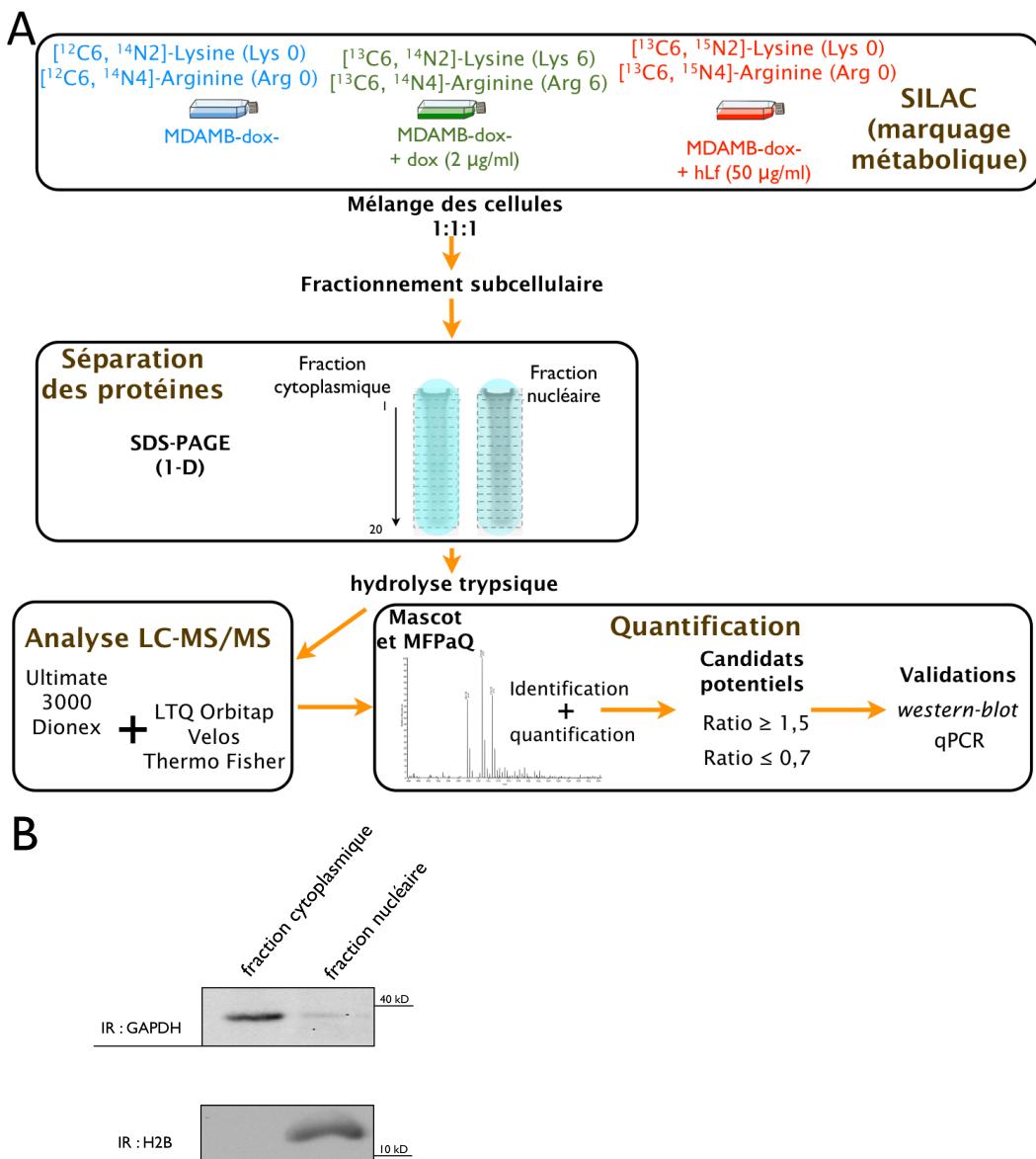


Figure 35 : Stratégie de l'analyse protéomique employée. A : Trois lots de cellules ayant la même origine ont été cultivés dans des milieux de formulation différente. Le milieu *léger* contient des acides aminés naturels [¹²C₆, ¹⁴N₂]-Lysine (Lys 0) et [¹²C₆, ¹⁴N₄]-Arginine (Arg 0), le milieu *medium* contient les acides aminés [¹³C₆, ¹⁴N₂]-Lysine (Lys 6) et [¹³C₆, ¹⁴N₄]-Arginine (Arg 6) et le milieu *lourd* les acides aminés [¹³C₆, ¹⁵N₂]-Lysine (Lys 8) et [¹³C₆, ¹⁵N₄]-Arginine (Arg 10). Après 8 doublements cellulaires afin d'obtenir une incorporation totale des acides aminés, deux fractionnements ont été réalisés. Le premier fractionnement subcellulaire a permis de recueillir une fraction nucléaire et une fraction cytoplasmique. Le second fractionnement permet de séparer les protéines de chacune de ces fractions en fonction de leur masse moléculaire par SDS-PAGE. Chacune des deux pistes est découpée en vingt bandes de même taille qui vont subir une hydrolyse des protéines par la trypsin. Les peptides possèderont alors tous des résidus de lysine ou d'arginine en position C-terminale (excepté le peptide terminal de la protéine). Ces fragments de digestion ont ensuite été séparés par chromatographie en phase inverse sur une nano-LC puis analysés grâce au spectromètre de masse LTQ-Orbitrap Velos. L'analyse bioinformatique a été réalisée via l'utilisation du logiciel Mascot pour l'identification des protéines et le logiciel MFPaQ (Bouyssie et al., 2007) créé au sein de l'IPBS. Ce dernier permet le tri, la validation (élimination des faux-positifs) et la quantification des différents peptides. La validation des résultats de spectrométrie de masse est effectuée par *western-blot* et par qPCR.B : L'efficacité du fractionnement cellulaire a été confirmée par immunorévélation de protéines spécifiques à chaque compartiment. Après mélange des trois conditions au niveau cellulaire, les protéines ont été extraites, séparées sur gel SDS-PAGE puis transférées sur membrane de nitrocellulose. L'immuno-révélation a été effectuée d'une part avec des anticorps dirigés contre la GAPDH, protéine soluble du cytoplasme et d'autre part avec des anticorps dirigés contre l'histone H2B.

L'étude que nous avons réalisée nous a permis d'identifier 5030 protéines au total des deux fractions. Ces résultats viennent compléter l'étude réalisée par l'équipe d'Hubert Hondermarck (Strande *et al.*, 2009) qui avait caractérisé 3481 protéines issues de cellules MDA-MB-231 en utilisant la combinaison gel 1-D et LTQ-Orbitrap de première génération. Réduire la complexité de l'échantillon et utiliser un spectromètre de masse plus performant ont été deux points primordiaux pour identifier un plus grand nombre de protéines avec des couvertures de séquences plus importantes (Lanucara et Evers, 2011; Emmott *et al.*, 2010; Bouyssié *et al.*, 2007).

2) Quantification des protéines et analyse des données

Les protéines identifiées grâce à l'interrogation de la banque de données Uniprot (v. 2010_09) ont été validées automatiquement à l'aide du logiciel MFPAQ v.4.0 (Bouyssié *et al.*, 2007) en fonction du nombre et du score des peptides comme décrit dans le matériel et méthodes. De là, les ratios d'expression supérieurs à 1,5 et inférieurs à 0,7 dans les conditions traitées par rapport à celle non traitée ont été considérés comme significatifs. Ainsi, parmi les protéines identifiées et quantifiées, trois protéines sont surexprimées et douze sous-exprimées en présence de Δ Lf (Tableau VI, page 83) alors que dix protéines sont surexprimées et onze protéines sous-exprimées en présence de hLf (Tableau VII, page 85). Toutes les surexpressions des protéines en réponse à nos deux isoformes on pu être validées par western-blot (Figure 36) exceptées la selénoprotéine H et la highly similar to HIV-1 Rev binding protein pour lesquelles il n'existe pas d'anticorps disponibles ou utilisables en western-blot.

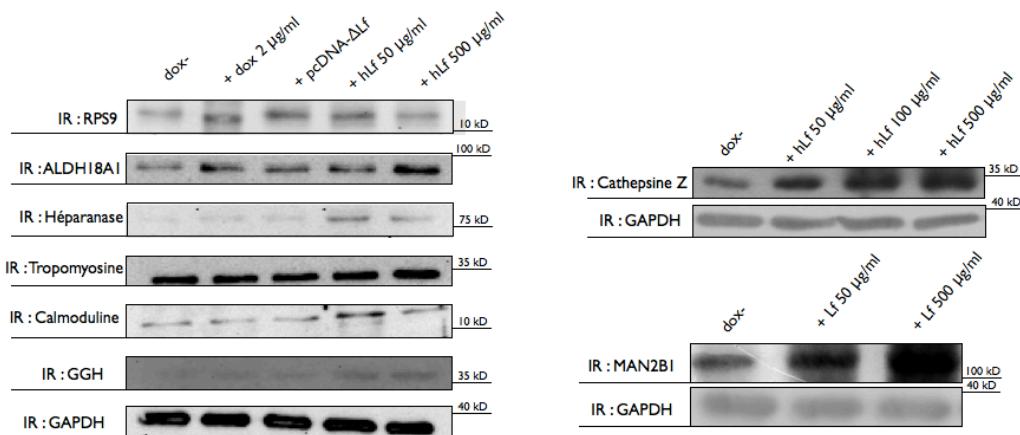


Figure 36 : La surexpression des protéines a été confirmée par immuno-révélation. Les cellules MDA-MB-231 ont été lysées 24 h après le traitement. Les protéines ont été extraites, séparées sur gel SDS-PAGE puis transférées sur membrane de nitrocellulose. L'immuno-révélation a été effectuée avec des anticorps dirigés spécifiquement contre les protéines RPS9, ALDH18A1, héparanase, tropomyosine épithéliale, calmoduline, gamma glutamyl hydrolase (GGH), cathepsine Z, mannosidase cytosolique (MAN2B1) et GAPDH.

Tableau I : Protéines identifiées par LC-MS/MS dans la fraction cytoplasmique dont le ratio d'expression est supérieur ou égal à 1,5 ou inférieur ou égal à 0,7 en réponse à la ΔLf (A) ou en réponse à la hLf (B). Les fonctions imputées aux protéines sont le résultat d'une recherche bibliographique à l'aide de l'outil Pubmed (<http://www.ncbi.nlm.nih.gov/pubmed>).

(A)

Numéro d'accésion	ID	Description	Fonction	Nombre de peptides identifiés	Nombre de peptides séquencés	Pourcentage de couverture (%)	Score	Ratio d'expression
Q8IZQ5	SELH_HUMAN	Sélénoprotéine H (SelH)	Protéine thioredoxin-like impliquée dans l'élimination des ROS par catalyse des peroxydes d'hydrogène (Novoselov <i>et al.</i> , 2007)	12	6	38,5	462,51	1,65
Q05DA8	Q05DA8_HUMAN	UBR7	Proteine E3 ubiquitine-ligase reconnaissant et se fixant aux protéines présentant des résidus N-terminaux spécifiques conduisant à leur ubiquitylation et leur dégradation (Tasaki <i>et al.</i> , 2009)	13	7	35,7	232,42	0,56
Q9NZ23	Q9NZ23_HUMAN	Drug-sensitive protein 1, ARNm	-	18	11	94,9	612,36	0,24
Q8TE01	Q8TE01_HUMAN	DERP12 (<i>Dermal papilla derived protein 12</i>), ARNm	Protéine interagissant avec TRAF6 (<i>TNF receptor associated factor 6</i>) (Ewing <i>et al.</i> , 2007), surexprimée dans le cas de mélanome (Stine <i>et al.</i> , 2011)	5	4	16,6	119,49	0,01
A8K9U8	A8K9U8_HUMAN	ADNC FLJ75480, highly similar to Human lactoferrin	Voir généralités	4	2	6,6	61,26	0,08
B4DUM6	B4DUM6_HUMAN	Endolase	Métalloenzyme impliquée dans la glycolyse (Comi <i>et al.</i> , 2001)	19	10	37,3	1151,16	0,03
B2R4M6	B2R4M6_HUMAN	ADNC, FLJ92148, highly similar to <i>Homo sapiens S100 calcium binding protein A9 (calgranulin B)</i> (S100A9)	Appartient à la famille des protéines liant le calcium, impliquée dans l'organisation du cytosquelette et la prolifération cellulaire (Loser <i>et al.</i> , 2010)	5	5	74,6	89,31	0
Q5DT20	Q5DT20_HUMAN	Hornatin	Membre de la famille des protéines liées au renforcement des enveloppes cellulaires (Henry <i>et al.</i> , 2011)	2	1	1	91,48	0,01
P16104	H2AX_HUMAN	Histone H2A	Protéines de la famille des histones, modification de la chromatine, régulation de la transcription	5	2	49	58,04	0,02
B2R577	B2R577_HUMAN	highly similar to <i>Homo sapiens S100 calcium binding protein A6 (calcyclin)</i> (S100A6)	Appartient à la famille des protéines liant le calcium, impliquée dans l'organisation du cytosquelette, la prolifération et la motilité cellulaire (Lesniak <i>et al.</i> , 2009). Sa surexpression est associée à une diminution de la formation de métastases dans les carcinomes mammaires (Arai <i>et al.</i> , 2008)	6	5	81,1	103,73	0

(B)

Numéro d'accession	ID	Description	Fonction	Nombre de peptides pour l'identification	Nombre de séquences	Pourcentage de couverture	Score	Ratio d'expression
Q15657	Q15657_HUMAN	Tropomyosine (isoforme)	Appartient à la famille des protéines du cytosquelette se liant à l'actine, impliquée dans la morphologie, la division, la motilité cellulaire et le trafic vésiculaire (Schevezov <i>et al.</i> , 2005)	16	11	51,8	171,33	1,54
Q8IZQ5	SELH_HUMAN	Séléprotéine H (SelH)	Protéine <i>thioredoxin-like</i> impliquée dans l'élimination des ROS par catalyse des peroxydes d'hydrogène (Novoselov <i>et al.</i> , 2007)	12	6	38,5	462,51	1,50
B7Z5W1	B7Z5W1_HUMAN	ADNC, FLJ54854, <i>highly similar to Junctional adhesion molecule A</i>	Protéine d'adhésion cellulaire exprimée dans les jonctions serrées des cellules épithéliales, maintient l'intégrité des jonctions serrées (Naik <i>et al.</i> , 2008)	16	10	48,2	336,49	0,56
Q8TEO1	Q8TEO1_HUMAN	DERP12 (<i>Dermal papilla derived protein 12</i>)	Protéine interagissant avec TRAF6 (<i>TNF receptor associated factor 6</i>) (Ewing <i>et al.</i> , 2007), surexprimée dans le cas de mélanome (Stine <i>et al.</i> , 2011)	5	4	16,6	119,49	0,44
Q9NZZ23	Q9NZZ23_HUMAN	<i>Drug-sensitive protein 1</i>	-	18	11	94,9	612,36	0,27
B2MV13	B2MV13_HUMAN	ADNC, FLJ75480, <i>highly similar to Human lactoferrin</i>	Voir généralités	4	2	6,6	61,26	0
B2R577	B2R577_HUMAN	ADNC, FLJ92369, <i>highly similar to Homo sapiens S100 calcium binding protein A6 (calcyclin) (S100A6)</i>	Appartient à la famille des protéines liant le calcium, impliquée dans l'organisation du cytosquelette, la prolifération et la motilité cellulaire (Lesniak <i>et al.</i> , 2009). Sa surexpression est associée à une diminution de la formation de métastases dans les carcinomes mammaires (Arai <i>et al.</i> , 2008)	4	4	64,4	96,85	0,05
B4DPP6	B4DPP6_HUMAN	ADNC, FLJ54371, <i>highly similar to Serum albumin</i>	Protéine sérique	13	11	20,7	113,86	0

Tableau II : Protéines identifiées par LC-MS/MS dans la fraction nucléaire dont le ratio d'expression est supérieur ou égal à 1,5 ou inférieur ou égal à 0,7 en réponse à la ΔIf (C) ou en réponse à la hIf (D). Les fonctions imputées aux protéines sont le résultat d'une recherche bibliographique à l'aide de l'outil Pubmed (<http://www.ncbi.nlm.nih.gov/pubmed>).

(C)		ID	Description	Fonction	Nombre de peptides identifiés	Nombre de peptides séquencés	Pourcentage de couverture (%)	Score	Ratio d'expression
B5MCT8	B5MCT8_HUMAN		<i>Putative uncharacterized protein RPS9</i>	Protéine impliquée dans la biogénèse ribosomale et ayant un rôle anti-apoptotique de protection contre les lésions oxydatives (Kim <i>et al.</i> , 2003)	14	8	43,2	169,43	2,11
Q5T567	Q5T567_HUMAN		<i>Aldéhyde déhydrogénase 18 (family_member_A1) (ALDH18A1)</i>	Protéine impliquée dans la biosynthèse de la proline et de l'ornithine dont l'expression diminue les niveaux de ROS cellulaires (Krishnan <i>et al.</i> , 2008)	47	27	36,9	1343,22	1,85
Q8IZQ5	SELH_HUMAN		<i>Sélénoprotéine H (SelH)</i>	Protéine <i>thioredoxin-like</i> impliquée dans l'élimination des ROS par catalyse des peroxydes d'hydrogène (Novoselov <i>et al.</i> , 2007)	16	6	43,4	884,27	1,58
B3VMW0	B3VMW0_HUMAN		<i>Lactoferrine</i>	Voir généralités	6	3	8,9	47,97	0,4
Q9NZ23	Q9NZ23_HUMAN		<i>Drug-sensitive protein 1</i>	-	6	4	53,7	216,34	0,4
Q2TUW9	Q2TUW9_HUMAN		<i>Lactoferrine</i>	Voir généralités	14	11	20,3	127,25	0,1
Q96IS6	Q96IS6_HUMAN		<i>HSPA8 protein (fragment)</i>	Appartient à la famille des protéines de choc thermique (HSP70), impliquée dans la réponse au stress cellulaire et le repliement des protéines (Garrido <i>et al.</i> , 2003)	18	11	35,3	358,13	0,06
P31151	S10A7_HUMAN		<i>Protéine S100-A7</i>	Appartient à la famille des protéines liant le calcium, elle prononce le stress oxydatif et l'angiogénèse (Shubbar <i>et al.</i> , 2011)	3	3	23,8	137,94	0,32
P06703	S10A6_HUMAN		<i>Protéine S100-A6</i>	Appartient à la famille des protéines liant le calcium, impliquée dans l'organisation du cytosquelette, la prolifération et la motilité cellulaire (Lesniak <i>et al.</i> , 2009). Sa surexpression est associée à une diminution de la formation de métastases dans les carcinomes mammaires (Arai <i>et al.</i> , 2008)	4	4	64,4	96,85	0,08
Q5DT20	Q5DT20_HUMAN		<i>Hornatin</i>	Membre de la famille des protéines liées au S100, fonction probable dans le renforcement des enveloppes cellulaires (Henry <i>et al.</i> , 2011)	7	7	4	214,75	0,01
P02647	AP0A1_HUMAN		<i>Apolipoprotéine A-I</i>	Participe au transport du cholestérol (Pownall et Elinholm, 2006)	3	3	14,6	57,92	0,01

(D)

Numéro d'accession	ID	Description	Fonction	Nombre de peptides pour l'identification	Nombre de séquences	Pourcentage de couverture	Score	Ratio d'expression
Q9UBR2	CATZ_HUMAN	Cathepsine Z	Protéinase lysosomale impliquée dans la migration cellulaire (Obermaier et al., 2006)	15	8	45,9	352,76	2,51
Q92820	GGH_HUMAN	Gamma-glutamyl hydrolase	Enzyme lysosomale agissant comme endo- ou exo-peptidase, impliquée dans le maintien de l'homéostasie des foliates au sein de la cellule (Schneider et Ryan, 2006)	11	8	25,8	126,33	2,16
A8K6A7	A8K6A7_HUMAN	<i>Highly similar to Homo sapiens mannosidase, class 2B, member 1</i> (MAN2B1)	Enzyme lysosomale impliquée dans la dégradation des N-glycoprotéines (Daher et al., 1991)	7	4	50	98,21	2,03
B5MCTR8	B5MCTR8_HUMAN	<i>Putative uncharacterized protein RPS9</i>	Protéine impliquée dans la biogénése ribosomale et ayant un rôle anti-apoptotique de protection contre les lésions oxydatives (Kim et al., 2002)	14	8	43,2	169,43	1,95
B3KQR6	B3KQR6_HUMAN	<i>Highly similar to Heparanase</i>	Enzyme de dégradation des chaînes héparanes sulfatés. Son activité contribue à l'invasion, la formation de métastases (Vlodavsky et al., 1999)(Vlodavsky et al., 1999) et l'angiogénèse (Kelly, 2005)	5	3	21,4	228,13	1,76
Q5T567	Q5T567_HUMAN	Aldéhyde déhydrogénase 18 (family, member A1) (ALDH18A1)	Protéine impliquée dans la biosynthèse de la proline et de l'ornithine dont l'expression diminue les niveaux de ROS cellulaires (Krishnan et al., 2008)	47	27	36,9	987,33	1,46
B4DMSS5	B4DMSS5_HUMAN	<i>Highly similar to HIV-1 Rev-binding protein 2</i>	-	3	2	25,4	76,61	1,51
A8K1M2	A8K1M2_HUMAN	<i>Highly similar to Homo sapiens calmodulin 1</i>	Protéine ayant un rôle de second messager dans la transduction du signal et régulant la prolifération cellulaire (Choi et Husain, 2006)	3	3	16	222,31	1,57
Q8IZQ5	SELH_HUMAN	Sélénoprotéine H (SelH)	Protéine <i>thioredoxin-like</i> impliquée dans l'élimination des ROS par catalyse des peroxydes d'hydrogène (Novoselov et al., 2007)	16	6	43,4	884,27	1,50

Numéro d'accession	ID	Description	Fonction	Nombre de peptides pour l'identification	Nombre de séquences	Pourcentage de couverture	Score	Ratio d'expression
D9YZZV5	D9YZZV5_HUMAN	Tropomyosin1 1 (Alpha), isoforme 4	Appartient à la famille des protéines du cytosquelette se liant à l'actine, impliquée dans la morphologie, la division, la motilité cellulaire et le trafic vésiculaire (Schevezov <i>et al.</i> , 2005)	40	19	57	675,88	1,50
Q96DD3	Q96DD3_HUMAN	ADNC, FLJ54748, highly similar to 2'-oxoglutarate dehydrogenase E1 component, mitochondrial	-	17	12	26,9	124,03	0,59
A6NDA2	A6NDA2_HUMAN	ADNC, FLJ95224, Homo sapiens proline rich 3 (PRR3)	-	11	4	31,2	162,77	0,61
Q96FZ7	CHMP6_HUMAN	Charged multivesicular body protein 6	Protéine impliquée dans le trafic vésiculaire (Lata <i>et al.</i> , 2009)	4	3	17,4	133,69	0,36
Q9NZ23	Q9NZ23_HUMAN	Drug-sensitive protein 1	-	6	4	53,7	216,34	0,38
B3VMW0	B3VMW0_HUMAN	Lactoferrine	Voir généralités	6	3	8,9	47,97	0,27
B4DDT3	B4DDT3_HUMAN	ADNC, FLJ54622, highly similar to Prothrombin	-	4	3	1,6	124,37	0,22
Q96IS6	Q96IS6_HUMAN	Protéine HSP8	Appartient à la famille des protéines de choc thermique (HSP70), impliquée dans la réponse au stress cellulaire et le repliement des protéines (Garrido <i>et al.</i> , 2005)	18	11	35,3	353,13	0,08
P06703	S10A6_HUMAN	Protéine S100-A6	Appartient à la famille des protéines liant le calcium, impliquée dans l'organisation du cytosquelette, la prolifération et la motilité cellulaire (Lesniak <i>et al.</i> , 2009). Sa surexpression est associée à une diminution de la formation de métastases dans les carcinomes mammaires (Arai <i>et al.</i> , 2008)	4	4	64,4	96,85	0,01
Q2TUW9	Q2TUW9_HUMAN	Lactoferrine	Voir généralités	14	11	20,3	127,25	0,11

A. Fonctions cellulaires des protéines sur- ou sous-exprimées en réponse aux Lf

La majeure partie des protéines dont l'expression s'est avérée être modifiée par l'une ou les deux Lf peuvent être associées à trois fonctions cellulaires majeures : la régulation du stress oxydant, la modulation de la migration et de l'adhésion cellulaire et la dégradation des protéines.

1) Stress oxydant

Les espèces réactives de l'oxygène (ROS) sont générées lors du métabolisme oxydatif mitochondrial ou en réponse à des xénobiotiques, cytokines ou une invasion bactérienne. Un excès de ROS ou d'oxydants et des mécanismes de défense cellulaires insuffisants sont à l'origine du stress oxydant. Ce dernier conduit à des lésions macromoléculaires et se manifeste lors de nombreux désordres pathologiques parmi lesquels, le cancer (Ray *et al.*, 2012). Jusqu'à présent, la hLf n'était connue pour réguler le stress oxydant que dans un contexte inflammatoire (Kruzel *et al.*, 2010). Nos résultats montrent que la hLf et la Δ Lf sont toutes deux capables, indirectement, de moduler la réponse au stress oxydant des cellules MDA-MB-231. Trois protéines ayant des fonctions en relation avec le stress oxydant sont surexprimées tant en présence de Δ Lf que de hLf. L'aldéhyde deshydrogénase, ALDH18A1, est une enzyme de détoxicification mitochondriale (Yoshida *et al.*, 1998) qui catalyse la phosphorylation et la réduction du glutamate. L'expression de l'ALDH18A1 est augmentée en réponse au stress oxydant et, bien que le mécanisme ne soit pas réellement connu, cette surexpression protège les cellules de la formation des ROS (Krishnan *et al.*, 2008). La protéine RPS9 fait partie intégrante de la petite sous unité ribosomale. Cependant, dans un contexte pro-oxydant, RPS9 sera surexprimée et se retrouvera présente sous une forme libre possédant une activité intrinsèque qui lui permet de protéger les cellules du stress oxydant (Kim *et al.*, 2003). Enfin, la sélénoprotéine H (SelH) est une thiorédoxine nucléolaire dont la surexpression conduit à la réduction de l'accumulation de superoxydes dans des cellules soumises à une irradiation par les UV (Mendelev *et al.*, 2009; 2011).

La surexpression des deux premières protéines a pu être validée par western-blot (Figure 36, page 82) mais les anticorps dont nous disposons, les seuls présents sur le marché, ne nous ont pas permis de valider la surexpression de la SelH. Après analyse de l'expression des transcrits par qPCR, nous avons pu montrer que seuls les transcrits correspondant à SelH sont surexprimés en réponse aux deux isoformes de Lf (Tableau VIII, page 89).

Tableau I : Validations par *western-blot* et qPCR des protéines surexprimées mises en évidence par l’analyse SILAC.

A : Protéines surexprimées en présence de ΔLf

Numéro d'accesion	Description	Ratio de surexpression	Validation Western-blot	Ratio d'expression des transcrits
B5MCT8	<i>Putative uncharacterized protein RPS9</i>	2,1	oui	0,98
P54886	Delta-1-pyrroline-5-carboxylate synthase (ALDH18A1)	2,09	oui	1,01
Q8IZQ5	Sélénoprotéine H (SELH)	1,59	n.d	1,71

B : Protéines surexprimées en présence de hLf

Numéro d'accesion	Description	Ratio de surexpression	Validation Western-blot	Ratio d'expression des transcrits
B5MCT8	<i>Putative uncharacterized protein RPS9 (RPS9)</i>	1,93	oui	0,95
P54886	Delta-1-pyrroline-5-carboxylate synthase (ALDH18A1)	1,65	oui	0,93
Q8IZQ5	Sélénoprotéine H (SELH)	1,56	n.d	1,53
Q15657	Tropomyosine (isoforme)	1,72	oui	1,96
Q9UBR2	Cathepsine Z	2,54	oui	2,76
Q92820	Gamma-glutamyl hydrolase (GGH)	2,23	oui	2,79
A8K6A7	<i>Highly similar to Homo sapiens mannosidase, alpha, class 2B, member 1 (MAN2B1)</i>	1,97	oui	3,28
D6RAQ1	<i>Highly similar to Heparanase</i>	1,76	oui	2,10
B4DMS5	<i>Highly similar to HIV-1 Rev-binding protein 2</i>	1,61	n.d	1,10
A8K1M2	<i>Highly similar to Homo sapiens calmodulin 1</i>	1,57	oui	1,01

2) Modulation de l’adhésion cellulaire, migration et invasion

Son *et al.* (2002) ont montré l’implication de la hLf dans les phénomènes d’angiogenèse *via* la mise en évidence d’une transactivation *in vitro* du gène *IL-1β* par la hLf conduisant à une augmentation de l’expression des métalloprotéinases (MMP). Il a également été démontré que la hLf était capable d’exercer une activité pro-angiogénique par sa capacité à surexprimer de manière indirecte de nombreuses MMP (Haversen *et al.*, 2000; Suzuki *et al.*, 2002). Récemment, Ha *et al.* (2011) ont montré une augmentation de l’invasion et de la migration des cellules MDA-MB-231 suite au traitement par la hLf.

De manière intéressante, l'étude SILAC nous a permis de mettre en évidence la surexpression de deux protéines dont les fonctions sont reliées à la dégradation de la matrice extracellulaire. Il s'agit l'héparanase et de la cathepsine Z. En effet, l'héparanase, surexprimée en réponse à la hLf est une enzyme impliquée dans la dégradation des protéines de la matrice extracellulaire (MEC), conduisant à l'invasion et à la formation de métastases dans le cas de cancer (Cohen *et al.*, 2006; Vlodavsky *et al.*, 1999). En plus de l'invasion cellulaire, l'héparanase est capable de promouvoir l'angiogénèse et d'accélérer la croissance cellulaire (Kelly, 2005; Cohen *et al.*, 2006). Une autre enzyme de dégradation des protéoglycannes de la MEC, la cathepsine Z, est surexprimée en présence de hLf. Les rôles de cette protéine ne sont pas encore totalement élucidés, néanmoins, l'étude menée par Obermajer *et al.* (2006) montre qu'elle est capable de promouvoir la migration des lymphocytes T. Les données de la littérature impliquent la cathepsine Z dans les phénomènes de migration et invasion. En effet, une forte transcription du gène de la cathepsine Z, localisé dans une région fréquemment amplifiée dans le cas de cancer, est associée à un mauvais pronostic ainsi qu'à la formation accrue de métastases dans le cas de cancer du sein (Hidaka *et al.*, 2000). De plus, l'invalidation de son transcrit diminue fortement la capacité d'invasion des cellules tumorales (Krueger *et al.*, 2005). Récemment, Wang *et al.* (2011) ont mis en évidence une surexpression des transcrits des MMP2, 3 et 9 en réponse à la cathepsine Z, conduisant à l'augmentation du taux de métastases *in vivo* (Wang *et al.*, 2011b). La surexpression de ces protéines en réponse à la hLf a été validée par analyse en *western-blot* (Figure 36, page 82) et les transcrits correspondants sont également surexprimés (Tableau). Les différentes isoformes de la tropomyosine épithéliale sont connues pour être indispensables à la stabilité des microfilaments d'actine du cytosquelette (Schevzov *et al.*, 2005). La surexpression de la tropomyosine alpha isoforme 4 a pu être validée par *western-blot* (Figure 36, page 82) et qPCR (Tableau VIII, page 89). D'autre part, nous avons pu mettre en évidence la sous-expression de la protéine *Junctional adhesion molecule A* en réponse à la réintroduction de la hLf dans les cellules MDA-MB-231. Or, il s'agit d'une protéine d'adhésion cellulaire interagissant avec les intégrines au niveau des jonctions serrées et la perte de son expression dans les cellules MDA-MB-231 conduit à une augmentation de l'invasion cellulaire (Naik *et al.*, 2008).

La protéine S100A6, sous-exprimée en réponse aux deux isoformes de Lf, possède des activités liées à l'adhésion cellulaire et l'invasion. Elle est surexprimée dans de nombreux cas de cancer dont le cancer du sein (Cross *et al.*, 2005) et participe de manière active à la réorganisation du cytosquelette, ce qui implique un rôle clef dans la motilité et la prolifération cellulaire (Leśniak *et al.*, 2009). Parmi les protéines sous-exprimées uniquement en réponse à la Δ Lf, on retrouve deux autres protéines de la famille S100. Les protéines de cette famille sont notamment connues pour réguler les constituants du

cytosquelette et conduire à une motilité cellulaire (Arai *et al.*, 2008). La protéine S100A9 est directement impliquée dans la formation de métastases (Loser *et al.*, 2010). De même, l'expression de la protéine S100A7 a été décrite comme étant inversement proportionnelle à sa capacité à induire la tumorigénèse et la formation de métastases dans les cellules MDA-MB-231. C'est la première étude qui montre que la Δ Lf puisse être impliquée dans des phénomènes de modulation de l'adhésion cellulaire ou de l'invasion. La sous-expression de protéines à activités pro-métastatique et pro-invasive peut être vue comme une réponse de la cellule à la migration induite par les deux isoformes de Lf.

3) Dégradation des protéines

Des protéines dont les fonctions sont en relation avec la dégradation protéolytique ont leur expression modifiée en réponse à l'une ou l'autre isoforme de Lf. En effet, nos résultats mettent en évidence une sous-expression de la protéine UBR7 impliquée dans la dégradation protéasomale *via la N-end rule*, en réponse à la Δ Lf. Au contraire, l'expression de la hLf augmente celle de mannosidase lysosomale qui est une enzyme de dégradation usuelle des N-glycoprotéines. Ces résultats ont pu être validés par western-blot (Figure 36, page 82) et qPCR (Tableau VIII, page 89).

B. Expression des isoformes de Lf

Il est intéressant de constater que, parmi les protéines sous-exprimées, on retrouve la hLf ou Δ Lf (car les peptides ayant conduit à l'identification de la protéine ne permettent pas de discriminer les deux isoformes). Néanmoins, la méthode de quantification des transcrits Lf par qPCR *via la technologie Taqman* que j'ai développée montre une expression des transcrits hLf et Δ Lf augmentée en présence de Δ Lf ou de hLf, à l'inverse des résultats obtenus en SILAC. (Figure 37, page 92).

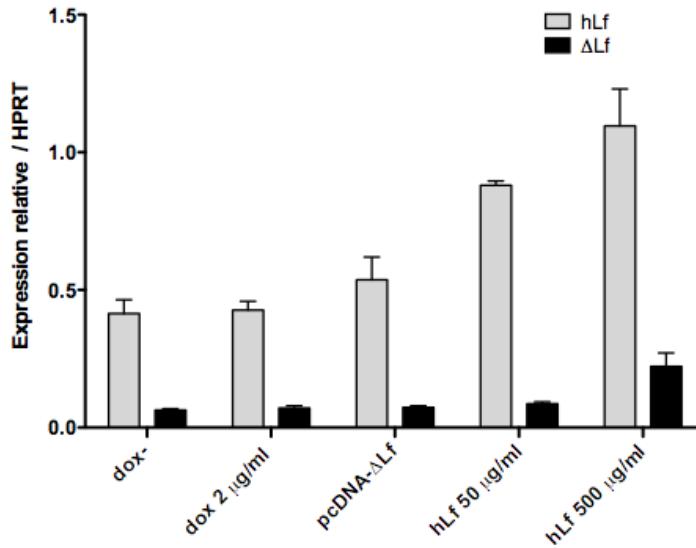


Figure 1: Expression relative des transcrits Lf et Δ Lf dans les cellules MDA-MB-231 dox-, en réponse à différentes doses d'induction de hLf ou Δ Lf. Les valeurs présentées proviennent de trois manipulations indépendantes réalisées en tripliquette.

C. Activité transcriptionnelle des isoformes de la Lf

Comme la Δ Lf est un facteur de transcription, nous avons voulu savoir si la surexpression des protéines ALDH18A1, RPS9 et SelH est directement due à la transactivation des gènes correspondants par la Δ Lf. D'un autre côté, nous avons regardé si la hLf pourrait également être capable de transactiver les gènes correspondants aux protéines surexprimées en sa présence. En effet, He et Furmanski ont été les premiers à démontrer que la hLf est capable de transactiver trois séquences d'ADN *in vitro*, et à montrer que la séquence S1 : GGCACCTT A/G G était la plus efficace. Son *et al.* ont, quant à eux, démontré en 2002 que la hLf est capable de transactiver le gène de l'IL1- β *in vitro*. Récemment, Ha *et al.* ont aussi démontré, avec l'utilisation d'un gène rapporteur et de l'EMSA (*Electrophoretic Mobility Shift Assay*), que la hLf est responsable de la transactivation du gène de l'endotheline 1 grâce à sa fixation sur une séquence identique à S1. Les travaux effectués au laboratoire ont permis de mettre en évidence des nouvelles cibles de l'activité transcriptionnelle de la Δ Lf mais également d'optimiser la séquence cible décrite au préalable par He et Furmanski comme étant la séquence S1 que nous avons appelée Δ LfRE. Nous avons donc recherché la présence d'un Δ LfRE au sein des promoteurs des gènes dont les protéines correspondantes sont surexprimées en réponse à l'une ou l'autre des isoformes de Lf. Tous les promoteurs des gènes correspondants excepté celui de la *Highly similar to HIV-1 Rev binding protein* possèdent au moins une séquence similaire au Δ LfRE (Tableau IX).

Tableau II : Δ LfRE potentiels retrouvés dans les promoteurs des gènes correspondants aux protéines surexprimées. Les nucléotides différents de ceux du Δ LfRE sont en caractère gras.

	Séquence										Localisation	
	1	2	3	4	5	6	7	8		9	10	
Δ LfRE	G/A	G	C	A	C	T	(G)	T	-	A/G	C/G	
ALDH18A1	G	T	C	A	C	T	-	T	-	G	C	- 8208
RPS9	A	G	C	A	C	T	-	T	-	G	G	- 1478
RPS9	A	T	C	A	C	T	-	T	-	G	A	- 1451, -847
SelH	G	G	C	A	C	T	G	T	-	G	G	- 2712
cathepsine Z	A	T	C	A	C	T		T	-	G	A	-4957
GGH	A	T	C	A	C	T		T	-	G	C	- 4962
GGH	A	G	C	A	C	T		T	-	C	A	- 4568
MA2B1 Too1	A	T	C	A	C	T		T	-	C	C	- 2933
MA2B1 Too1	A	G	C	A	C	T		T	-	C	T	- 868
MA2B1 Too8	G	T	C	A	C	T	G	T	-	C	C	- 6390
MA2B1 Too8	G	T	C	A	C	T	-	T	-	C	A	- 4671
héparanase	A	T	C	A	C	T	G	T	-	A	T	- 2865
calmoduline T201	G	T	C	A	C	T	-	T	-	G	A	- 2945
calmoduline T201	A	T	C	A	C	T	-	T	-	G	A	- 2416
calmodulineT202	G	G	C	A	C	T	-	T	-	C	C	- 1012
tropomyosine Too1	A	G	C	A	C	T	G	T	-	T	G	- 4251
tropomyosine T203	A	T	C	A	C	T		T	-	A	G	- 1655
tropomyosine T203	G	G	C	A	C	T		T	-	G	C	- 1423

Il est intéressant de constater que les protéines ALDH18A1, RPS9 et SelH surexprimées en présence de Δ Lf le sont également suite au traitement des cellules MDA-MB-231 par de la hLf exogène. Nous avons pu montrer par qPCR que seuls les transcrits de SelH voient leur expression augmentée en présence de Δ Lf et de Lf (Tableau VIII, page 89). Qui plus est, les ratios de surexpression des transcrits de SelH sont similaires aux ratios de surexpression protéique en réponse à une transfection stable ou transitoire de Δ Lf et quelque soit le type cellulaire étudié (Figure 39, page 94).

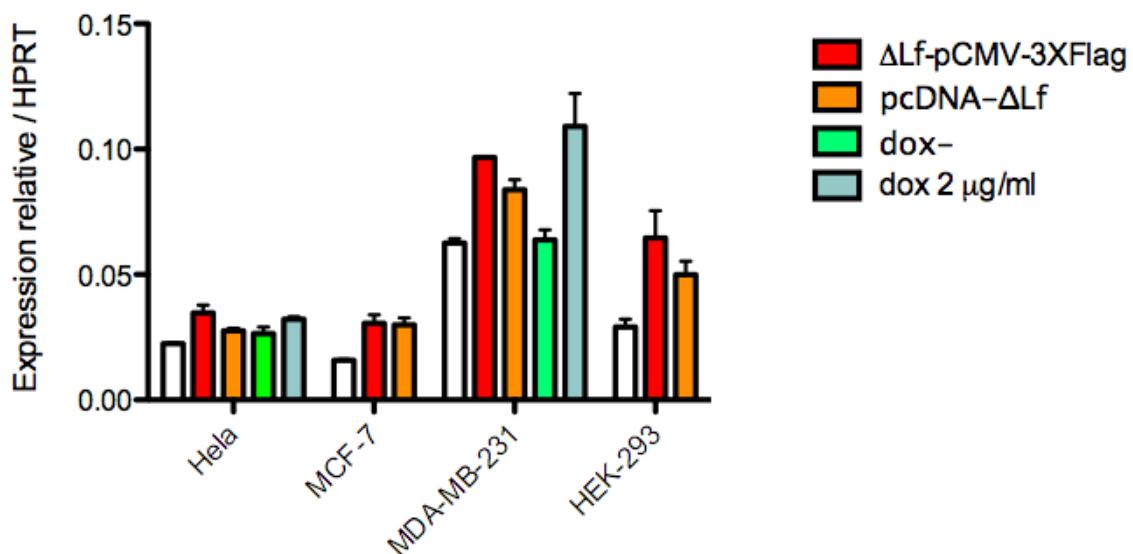


Figure 2 : Expression relative des transcrits SelH dans les cellules HeLa, MCF7, MDA-MB-231 et HEK en réponse à la ΔLf . La ΔLf a été transfectée soit de manière transitoire avec les vecteurs pcDNA- ΔLf ou pCMV10-3XFlag- ΔLf , soit de manière stable suite à l'induction par de la doxycycline (dox 2 $\mu\text{g/ml}$). Les valeurs présentées proviennent de trois manipulations indépendantes réalisées en tripliquette.

Le gène *SelH* possède un ΔLfRE potentiel au sein de son promoteur. La protéine SelH est la seule surexprimée en réponse à la ΔLf et la hLf et dont l'expression des transcrits suit celle de la protéine (Tableau VIII, page 89). Ce gène semble donc être un bon candidat en tant que cible de l'activité transcriptionnelle de nos deux isoformes. Nous avons donc cherché à savoir si Lf et ΔLf sont capables d'activer *in vitro* la transcription du gène *SelH*. Pour cela, nous avons cloné dans le vecteur d'expression pcDNA soit l'ADNC de la ΔLf , soit celui de la hLf ainsi que la partie de la région promotrice du gène *SelH* contenant le ΔLfRE potentiel dans le vecteur pGL3 *enhancer* en amont du gène de la luciférase. Nous avons effectué les expériences de gène rapporteur dans les cellules HEK-293 qui sont très facilement transfectable (efficacité de transcription supérieure à 90 %). Le vecteur pGL3-Skp1 a été utilisé comme contrôle positif de l'activation transcriptionnelle par la ΔLf . Les cellules HEK-293 ont donc été soit transfectées avec le vecteur pGL3-*SelH* ou le vecteur pGL3-Skp1 comme contrôle négatif, soit co-transfектées avec le vecteur pGL3-*SelH* ou le vecteur pGL3-Skp1 et le vecteur pcDNA- ΔLf . Les résultats de mesure de l'activité de la luciférase sont indiqués directement en pourcentage d'efficacité (Tableau X). Les valeurs obtenues montrent que la ΔLf exerce bien directement son activité transcriptionnelle *via* les régions promotrices contenant le $\Delta\text{LfRE}^{\text{Skp1}}$ (Mariller *et al.*, 2007) et le $\Delta\text{LfRE}^{\text{SelH}}$.

Tableau III : Efficacité des éléments de réponse à la ΔLf . Les cellules HEK-293 ont été co-transférées avec 50 ng de vecteur pGL3 contenant le $\Delta\text{LfRE}^{\text{Skp1}}$ ou le $\Delta\text{LfRE}^{\text{SelH}}$ ou le $\Delta\text{LfRE}^{\text{SelH-inv}}$ et 200 ng de vecteur pcDNA- ΔLf . La région promotrice de *Skp1* a été utilisée comme contrôle ayant 100% d'efficacité de transactivation.

	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	
$\Delta\text{LfRE}^{\text{Skp1}}$	G	G	C	A	C	T	G	T	A	C	100
$\Delta\text{LfRE}^{\text{SelH}}$	G	G	C	A	C	T	G	T	G	G	130
$\Delta\text{LfRE}^{\text{SelH-inv}}$	A	G	C	A	C	T	G	T	T	G	130

Nous avons ensuite répété l'expérience pour étudier l'activité transcriptionnelle de la hLf vis-à-vis du $\Delta\text{LfRE}^{\text{SelH}}$. Pour cela, les cellules HEK-293 ont été soit transférées avec le vecteur pGL3-SelH, soit co-transférées avec le vecteur pGL3-SelH et le vecteur pcDNA-hLf. Les résultats de mesure de l'activité de la luciférase sont indiqués directement en *fold* d'activation (Figure 3). La hLf possède une activité de transactivation du gène SelH deux fois supérieure à celle de la ΔLf . Ceci pourrait être du au fait que la ΔLf ne possède qu'une séquence NLS en position C-terminale alors que la hLf possède cette même séquence NLS et un seconde en position N-terminale (¹GRRRR⁵). Cette séquence ¹GRRRR⁵ pourrait être un signal plus efficace pour l'import au noyau.

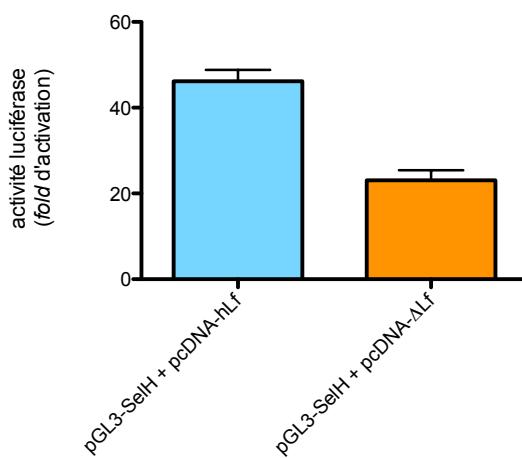


Figure 3 : Les cellules HEK-293 (2×10^5 /puits) ont été transférées pendant 24 h par 50 ng de vecteur pGL3 contenant le promoteur *SelH*, servant de contrôle négatif. En bleu, les cellules HEK 293 ont été co-transférées avec 50 ng de vecteur pGL3 contenant le promoteur *SelH* et 200 ng de vecteur pcDNA-hLf. En orange, les cellules HEK 293 ont été co-transférées avec 50 ng de vecteur pGL3 contenant le promoteur *SelH* et 200 ng de vecteur pcDNA- ΔLf . Les valeurs présentées proviennent de trois manipulations indépendantes réalisées en tripliquette.

Nous avons aussi voulu nous assurer que seul le ΔLfRE identifié dans le fragment cloné était responsable de l'activité transcriptionnelle observée pour la ΔLf . Pour cela, nous avons procédé à la délétion des 5 nucléotides centraux communs à ces RE (GCACT). Les vecteurs

comprenant les Δ LfRE délétés se nomment respectivement : pGL3-dSkp1 et pGL3-dSelH. L'activité transcriptionnelle a ensuite été mesurée, montrant qu'en comparaison avec le contrôle positif (pGL3-DcpS + pcDNA- Δ Lf) (Figure 4o), la suppression des 5 nucléotides centraux des Δ LfRE du gène Skp1 provoque bien l'inhibition de la transactivation du gène de la luciférase par la Δ Lf.

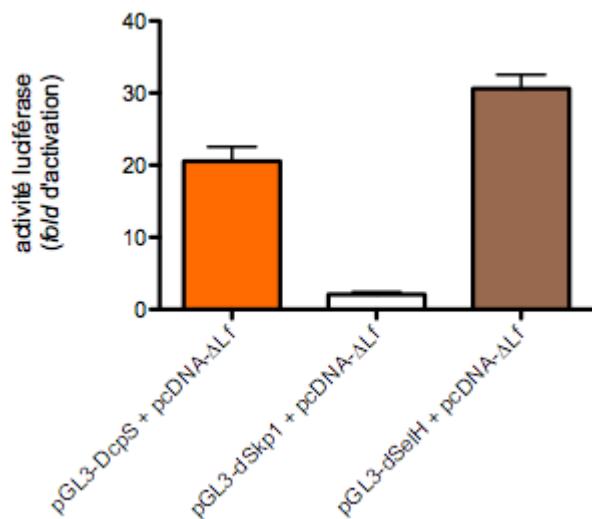


Figure 4 : Les cellules HEK-293 (2×10^5 /puits) ont été transfectées pendant 24 h par 50 ng de vecteur pGL3 contenant le promoteur *DcpS*, *dSkp1* ou *dSelH*, servant de contrôle négatif. En orange, les cellules HEK 293 ont été co-transfектées avec 50 ng de vecteur pGL3 contenant le promoteur *DcpS* et 200 ng de vecteur pcDNA- Δ Lf. En blanc, les cellules HEK 293 ont été co-transfектées avec 50 ng de vecteur pGL3 contenant le promoteur *dSkp1* et 200 ng de vecteur pcDNA- Δ Lf. En marron, les cellules HEK 293 ont été co-transfектées avec 50 ng de vecteur pGL3 contenant le promoteur *dSelH* et 200 ng de vecteur pcDNA- Δ Lf. Les valeurs présentées proviennent de trois manipulations indépendantes réalisées en tripliquette.

Par contre pour le vecteur pGL3-dSelH, la transactivation du gène de la luciférase par la Δ Lf reste effective. L'hypothèse la plus vraisemblable est alors la présence d'un autre Δ LfRE fonctionnel dans la région promotrice clonée. Et en effet, l'expérience présentée dans la Figure 4o, nous a conduit à mettre en évidence une séquence inversée similaire au Δ LfRE (en vert dans le Tableau XI, page 97) dans le promoteur de SelH. Pour certains facteurs de transcription, comme c'est le cas pour p53, l'élément de réponse peut être lu dans les deux sens (pour revue Pan *et al.*, 2010). Cette séquence inversée correspond à AGCACTGTTG comme indiqué dans le Tableau X (page 95). Nous avons donc muté le pGL3-SelH pour obtenir la séquence nucléotidique dans le sens inverse dans le vecteur pGL3 et nous avons ainsi montré qu'elle possède une efficacité de transcription équivalente à la séquence Δ LfRE^{SelH} (Tableau X, page 95). Il ressort de ces différentes expériences que la transactivation du Δ LfRE présent dans le promoteur du gène SelH conduit à l'activité transcriptionnelle la plus importante, suggérant qu'il est le plus proche de l'élément de réponse optimal de la Δ Lf.

Tableau IV : Séquences nucléotidiques utilisées afin de valider les éléments de réponse reconnus par la ΔLf . Les séquences ΔLfRE sont en rouge, la séquence inversée retrouvée dans le promoteur pGL3-SelH est en vert.

	Séquences nucléotidiques clonées dans le promoteur
pGL3-SelH	$-^{2740}\text{CACTAGGAGGAA} \color{red}{CAACAGTGCTCAAATGGATGTCACCGGGCACTGTGGTTCATTATGA}$ GAGCCGGGGAGAGAGGGAACACCAGGATCAGGAATTAAGATGCTAAAGTCT- ²⁶⁴³
pGL3-Skp1	$-^{1095}\text{TGGCCTTCCTGCCAGTCCCATCCCAGA} \color{red}{GGCACTGTACATCTCTGGTTCTCAGGGGG}$ CCCTGTGTGGAAGTATCTTTGTCTCCTGGTGTCAAGGGATATC- ⁹⁸²

D. Poursuite de l'étude par SILAC

Pour la deuxième étude SILAC, les cellules MDA-MB-231 ont été transfectées de manière transitoire avec le vecteur pcDNA- ΔLf et une concentration dix fois plus élevée de hLf (500 µg/ml) a été utilisée pour le traitement exogène. Les résultats de l'analyse bioinformatique n'ont pas encore été exploités dans leur totalité mais nous avons pu mettre en évidence pour l'instant, la surexpression des protéines RPS9 et ALDH18A1 en réponse à la ΔLf et des protéines RPS9, cathepsine Z et *gamma glutamyl hydrolase* en réponse à la hLf avec des ratios supérieurs à 1,5. Par contre, les protéines Skp1, DcpS et Bax, surexprimées en réponse à la ΔLf dans les travaux précédents ont été identifiées mais leur expression n'apparaît pas modifiée en présence de ΔLf .

III. Discussion

Notre projet était consacré à l'étude du protéome des cellules de glande mammaire cancéreuses MDA-MB-231 traitées par la ΔLf ou la hLf qui possèdent toutes les deux des propriétés anti-cancéreuses. Nous avons choisi de mesurer les variations induites par un traitement de 24 h par la hLf ou la ΔLf sur la lignée cellulaire MDA-MB-231. Pour cela, nous avons mis en œuvre une analyse protéomique différentielle quantitative de type SILAC en triplex. Un double marquage sur la lysine et l'arginine a été employé afin d'augmenter le nombre de peptides quantifiables. De même, nous avons eu recours à un premier fractionnement subcellulaire qui nous a permis de scinder notre protéome en fractions cytoplasmique et nucléaire et à un second fractionnement, par électrophorèse 1-D, pour augmenter la couverture de séquence des protéines identifiées et quantifiées. Nous avons aussi pu disposer des performances remarquables du LTQ-Orbitrap Velos qui nous a permis d'identifier plus de 5000 protéines de la lignée cellulaire MDA-MB-231.

L'analyse quantitative a nécessité un travail bioinformatique conséquent avec le logiciel MFPaQ mis au point au sein de l'IPBS à Toulouse dans l'équipe avec laquelle nous collaborons.

Nous avons donc abouti à l'étude la plus exhaustive qu'il soit, quant à l'identification et la quantification des protéines régulées de manière directe ou indirecte par les isoformes de

la Lf sur un même type cellulaire, les cellules de glande mammaire cancéreuses MDA-MB-231. Le nombre de protéines identifiées valide le choix de notre stratégie quant à l'utilisation d'un fractionnement subcellulaire, d'un fractionnement des protéines par SDS-PAGE et de l'utilisation d'un spectromètre de masse à précision de masse élevée et dont la sensibilité permet d'obtenir des couvertures de séquences importantes conduisant à une identification de protéine accrue. Dans un premier temps, nous ne nous sommes intéressés qu'aux protéines surexprimées en réponse à nos deux isoformes, car de manière intéressante, la hLf est capable d'augmenter l'expression des trois protéines surexprimées en réponse à la Δ Lf, les protéines RPS9, ALDH18A1 et SelH, toutes trois impliquées dans la régulation du stress oxydant. Nous avons ainsi pu retrouver les données de la littérature quant à l'implication de la hLf dans le stress oxydant mais, cette fois-ci, dans un contexte autre que l'inflammation. Nous avons également pu mettre en évidence pour la première fois, une implication de la Δ Lf dans la régulation du stress oxydant.

Nous avons aussi identifié d'autres protéines régulées par la hLf ou la Δ Lf et qui sont impliquées dans la motilité et l'invasion cellulaire. Les expériences préliminaires que nous avons effectuées grâce à des chambres d'invasion ont confirmé l'invasivité accrue des cellules MDA-MB-231 en réponse aux deux isoformes de Lf, avec quand même un effet plus prononcée en réponse à la hLf (Figure 41).

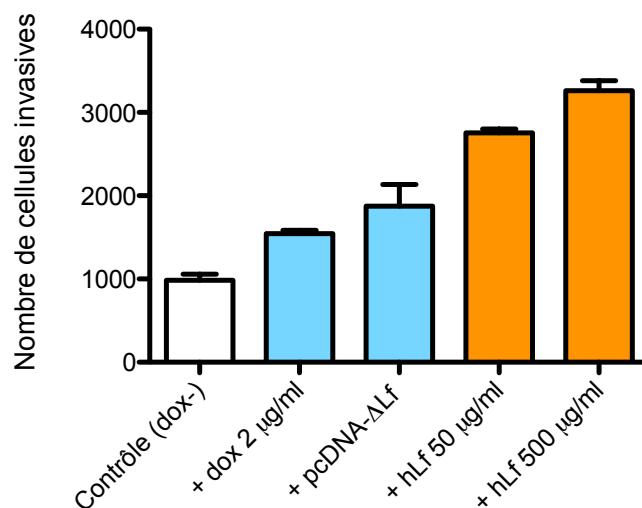


Figure 5 : Histogramme représentant l'effet des isoformes de la Lf sur l'invasion cellulaire par les cellules MDA-MB-231. Les cellules ont été sevrées pendant 24 h et réparties à une densité de 10^5 cellules par puit sur la chambre supérieure de *Matrigel™ Invasion Chambers* (BD Bioscience) et traitées avec de la Δ Lf (en bleu) ou de la hLf (en orange). Après 24 h de traitement, les cellules invasives se trouvant dans la chambre inférieure ont été fixées au paraformaldéhyde et colorées au DAPI afin de les compter. Les résultats proviennent d'une manipulation effectuée en tripliquette.

Nos travaux montrent également que la SelH, qui est la seule protéine surexprimée en réponse aux deux isoformes et dont l'expression des transcrits est également affectée par les

deux isoformes, peut être une cible directe de l'activité transcriptionnelle de la hLf et de la Δ Lf. Les travaux de Legrand *et al.* (2004) ont déjà montré que la hLf était endocytée dans ce type cellulaire de manière clathrine-dépendante et retrouvée au noyau. Toutefois, afin de valider l'hypothèse que le gène *SelH* est une cible de l'activité transcriptionnelle de nos deux isoformes, nous devons chercher à mettre en évidence une fixation *in situ* de chaque isoforme sur l'élément de réponse contenu dans le promoteur du gène *SelH*. Pour cela, des expériences utilisant la technique ChIP seront nécessaires.

D'autre part, l'analyse quantitative effectuée a mis en évidence une sous-expression de Lf alors que les expériences de quantification des transcrits hLf et Δ Lf nous montrent le contraire. De nombreuses hypothèses sont envisageables pour expliquer cette observation. Parmi elles, la régulation post-transcriptionnelle ou une autorégulation de la protéine sont les plus vraisemblables. En effet, l'analyse *in silico* du promoteur du gène *hLf* nous a permis de mettre en évidence une séquence Δ LfRE putative en position -2620 (AGCACTTTGG), suggérant donc que la sous-expression de la Lf ne peut imputée qu'à des mécanismes post-transcriptionnels.

Notre étude nous a permis d'identifier les protéines Bax, Skp1 et DcpS mais aucune variation significative de leur expression n'a pu être détectée dans les cellules MDA-MB-231 en réponse à la Δ Lf, contrairement aux expériences précédentes. Cela pourrait être expliqué par le fait que ces expériences avaient été effectuées par une expression transitoire de Δ Lf dans les cellules HEK-293 ou MCF-7. La transactivation de ces gènes par la Δ Lf pourrait donc être spécifique du type cellulaire employé. De plus, les modifications post-traductionnelles de la Δ Lf et la présence ou l'absence de cofacteurs nécessaires à la transcription de ces gènes pourraient être à l'origine de leur transactivation différentielle.

L'ensemble de nos travaux montre que le protéome des cellules de glande mammaire MDA-MB-231 n'est que peu affecté en réponse aux deux isoformes de Lf. Les résultats, en cours d'exploitation, obtenus avec des doses d'induction plus importantes des deux isoformes ont déjà mis en évidence, d'une part un effet sur le protéome plus important, mais également la confirmation de la surexpression de plusieurs protéines impliquées dans le stress oxydatif, la migration et l'invasion cellulaire.

IV. Matériel et méthodes

A. Mise au point de cellules exprimant la Δ Lf de manière stable et inducible

Le système inducible Tet-On (Clontech) a été utilisé pour exprimer la Δ Lf de manière stable dans les cellules MDA-MB-231 grâce à une induction par la doxycycline. Les cellules ont été cultivées dans le milieu DMEM (*Dubelco's Modified Eagle's Medium*, Thermo Scientific) avec 1 % de L-Glutamine, 10 % de SVF (Sérum de Veau Fœtal) et 500 µg/ml de

génétique. Les cellules ont été incubées à 37 °C sous une atmosphère humide contenant 5 % de CO₂. Une transfection transitoire a d'abord été effectuée à l'aide du lipofectant Dreamfect® (OZ Biosciences) avec le vecteur pTRE contenant le gène de la ΔLf, comme décrit dans Breton *et al.* (2004). Les clones isolés ont été multipliés afin d'obtenir des cellules nommées MDAMB-231 dox-. Pour permettre l'expression de la ΔLf uniquement lors de l'induction à la doxycycline, les cellules ont été cultivées en absence de tétracycline durant toute la période d'induction.

B. Purification de la hLf

La hLf a été purifiée à partir du lait maternel provenant de la banque de lait de l'hôpital Jeanne de Flandres à Lille selon le protocole usuel (Elass *et al.*, 2002). Après purification de la hLf, les endotoxines ont été éliminées par passage sur Detox-gel (*Pierce Chemicals*). Le taux de LPS résiduel dosé par *Limulus amoebocyte lysate assay* (QCL1000; BioWhittaker) était inférieur à 0,09 E.U./μg de hLf.

C. Marquage des cellules par SILAC (Stable Isotope Labelling Amino acid in Cell culture)

Les cellules ont été cultivées dans du milieu DMEM 89985 (Thermo Scientific) déplété en L-Arginine, L-Glutamine et L-Lysine et complété par de 10 % de SVF dialysé (Gibco-Invitrogen). La L-Glutamine (200mM) et la pénicilline (10000 U/ml)/streptomycine (10000 μg/ml) (Gibco-Invitrogen) ont été ajoutées pour obtenir une concentration finale de 1 % (v/v) chacune. La proline (20 mg/ml) a été rajoutée à 1 % (v/v) afin d'éviter l'inter-conversion de résidus d'Arginine en Proline. Les milieux ont ensuite été filtrés (0,45 μm de diamètre). Le marquage métabolique des cellules a été réalisé en cultivant les cellules de la même façon que précédemment mais en présence d'arginine (L-Arg ¹²C₆-¹⁴N₄) et de lysine (L-Lys ¹²C₆-¹⁴N₂) pour la condition *légère*, d'arginine (L-Arg ¹³C₆-¹⁴N₄) et de lysine (L-Lys¹³C₆-¹⁴N₂) pour la condition *medium* ou d'arginine (L-Arg ¹³C₆-¹⁵N₄) et de lysine (L-Lys¹³C₆-¹⁵N₂) pour la condition *lourde*, pendant huit passages avant l'induction. L'incorporation totale des acides aminés au niveau des protéines a été vérifiée par spectrométrie de masse.

D. Traitement des cellules

Les cellules ont été cultivées à raison de 6 x 10⁶ par boîte de 75 cm², 24 h avant le traitement. Un des lots de cellules, cultivé en présence de milieu SILAC *medium* a ensuite été traité pendant 24 h avec 2 μg/ml de doxycycline (Clontech) ou par transfection transitoire du vecteur pcDNA-ΔLf à une concentration de 1 μg/10⁶ cellules afin d'induire l'expression de la ΔLf. Un autre lot, celui cultivé en présence de milieu SILAC *lourd* est traité pendant 24 h avec 50 ou 500 μg/ml de hLf.

E. Fractionnement subcellulaire

24 h après le traitement, les cellules ont été comptées (6 fois par condition) à l'aide d'un microscope optique et d'une cellule de numération de Neubauer. Elles ont ensuite été mélangées à raison de 5×10^6 cellules par condition et rincées au PBS (*Phosphate Buffer Saline*). Les cellules provenant des trois conditions peuvent ensuite être traitées simultanément. Le fractionnement subcellulaire a été effectué selon le protocole de Misawa (2006) que nous avons modifié. Les cellules sont d'abord lysées par 500 µl du tampon A [Tris/HCl 50 mM pH 7,5 ; NaCl 137,5 mM ; Triton X100 0,5 % (p/v) ; Glycérol 10 % (v/v)] en présence d'inhibiteurs de protéases (*Complete*, Thermo) pendant 15 min sur glace. Ce lysat a été centrifugé à 13 000 tr.min⁻¹ (Heraeus, Biofuge 15R₁, HFA 22.2 rotor, 12 000 g) pendant 15 min à + 4 °C, le surnageant obtenu correspondant alors à la fraction cytosolique. Le culot a été lavé deux fois par 500 µl de Tp A (mêmes conditions de centrifugation). Le culot a été ensuite lysé dans 200 µl de tampon B [Tris/HCl 50 mM pH 7,5 ; NaCl 300 mM; Triton X100 0,5 % (p/v) ; Glycérol 10 % (v/v)] par sonication (8 x 4 sec x force 4 avec un sonicateur *Branson 150*). Après centrifugation à 13 000 tr.min⁻¹ pendant 15 min à + 4 °C, le surnageant correspondant cette fois à la fraction nucléaire. Les fractions cytosoliques, et nucléaires ont ensuite été dosés selon la méthode de Bradford avec une gamme étalon de SAB (Sérum Albumine Bovine). La détermination de la pureté de chaque fraction subcellulaire a été effectuée par la recherche de protéines spécifiques de chaque compartiment : H2B pour la fraction nucléaire et GAPDH pour la fraction cytoplasmique. Pour cela, 20 µg de chaque extrait ont été séparés par SDS-PAGE sur un gel homogène à 15 % d'acrylamide. L'électro-transfert et l'immuno-révélation ont été effectués comme décrit dans Mariller *et al.*, (2009).

F. Analyse des échantillons par SDS-PAGE

La réduction des cystéines a été réalisée en diluant 100 µg de chaque fraction dans le tampon de Laemmli (tampon Tris pH 6.8 40 mM ; SDS 2 % (p/v); glycérol 10 % (v/v), bleu de bromophénol 0.001 % (p/v), DTT 25 mM) pendant 30 min à 56 °C. Les cystéines ont ensuite été alkylées dans une solution d'iodoacétamide à 90 mM pendant 30 min à température ambiante et à l'obscurité.

Les extraits protéiques de chacune des deux fractions ont ensuite été séparés sur gel de polyacrylamide NuPAGE® Bis-Tris 4-12% (Invitrogen) et révélés par coloration au bleu de coomassie. Chaque piste a été découpée en 20 bandes homogènes qui ont ensuite été lavées par des solutions successives de tampon bicarbonate d'ammonium 100 mM et tampon bicarbonate d'ammonium / acetonitrile (50/50) pendant 15 min à 37° C chacune, avant d'être séchées sous vide dans un concentrateur rotatif sous vide de type Speed-Vacuum®.

G. Digestion trypsique et extraction des peptides

L'hydrolyse des protéines contenues dans chaque bande de gel est réalisée grâce à 50 µl

de bicarbonate d'ammonium 50 mM contenant 20 µg/ml de trypsine (*Sequencing Grade Modified Trypsin V511A*, Promega). Les tubes ont été ensuite incubés à 37 °C pendant la nuit.

Le surnageant de digestion de chaque tube a été prélevé et conservé, puis les morceaux de gel ont été incubés pendant 15 min à 37 °C sous agitation dans du bicarbonate d'ammonium 50 mM et soumis aux ultrasons pendant 5 min. Le nouveau surnageant a été récupéré et ajouté au premier. L'élution des peptides contenus dans le gel a été réalisées en trois étapes. Une première élution est obtenue par une solution d'acétonitrile à 2,5 % d'acide formique pendant 15 min à 37 °C sous agitation. Chaque échantillon a été soumis aux ultrasons à nouveau pendant 5 min et le surnageant a été additionné aux précédents. Une seconde extraction a été effectuée deux fois par une solution contenant 45 % d'acétonitrile et 10 % d'acide formique pendant 15 min à 37° C sous agitation. Les éluats ont ensuite été séchés sous vide et repris par 14 µl d'une solution d'acétonitrile à 10 % contenant 0,05 % acide trifluoroacétique.

H. Analyse par LC-MS/MS

Chaque échantillon a été analysé par nanoLC-MS/MS grâce à l'*UltiMate 3000 LC system* (Dionex) couplé au spectromètre de masse LTQ-Orbitrap Velos ETD (Thermo Fisher Scientific). 5 µl de mélange peptidique de chaque échantillon sont injectés grâce à un système automatique de prélèvement d'échantillons sur une colonne capillaire de pré-concentration à polarité inversée (C₁₈, 300 µm d.i x 5 mm, Dionex), équilibrée dans une solution contenant 5 % d'acetonitrile et 0.05 % d'acide trifluoroacétique à un débit de 20 µl/min. Après 5 min de dessalage, la pré-colonne est basculée en ligne avec une colonne capillaire analytique réalisée au laboratoire (ReproSil-Pur® C18-AQ, 75 mm d.i × 15 cm, 3 µm) équilibrée dans 95 % de solvant A (acétonitrile 5 %, acide trifluoroacétique 0,2 %) et 5 % de solvant B (acétonitrile 80 %, acide trifluoroacétique 0,2 %). Les peptides sont élués de la colonne analytique vers le spectromètre de masse à l'aide d'un gradient linéaire de 0 % à 50 % de solvant B pendant 105 min à un débit de 300 nl/min.

L'acquisition des données de la LTQ-Orbitrap est réalisée en continu par le logiciel XCalibur (v. 2.1). Lors du pré-scan, les données de MS sont accumulées dans l'Orbitrap pour une gamme dynamique de 350 à 2000 *m/z* avec une résolution de 60000. Les 20 ions les plus abondants sont sélectionnés pour être fragmentés en MS/MS par CID. Les fragments résultants sont analysés dans la trappe linéaire (LTQ). Une fenêtre dynamique d'exclusion de 60 s a été utilisée pour éviter la sélection répétitive de mêmes ions.

I. Traitement bioinformatique des données

Le logiciel Mascot (v.2.3.01) a été utilisé pour extraire automatiquement les listes de masses sous format de fichier .raw. Dans la banque de donnée Uniprot v. 2010_09 (Swiss-Prot et TrEMBL), l'espèce *Homo sapiens* a été sélectionnée puis la recherche a été effectuée

en tenant compte des paramètres résultant de la préparation des échantillons : la carbamidométhylation des cystéines a été indiquée comme étant une modification fixe des peptides ; l'oxydation des méthionines, l'acétylation en N- terminale et le marquage des résidus de lysine et arginine en $^{13}\text{C}_6$ et/ou $^{15}\text{N}_4$ ont été indiqué comme étant des modifications variables des peptides pour toutes les recherches Mascot. Un site de clivage manqué par la trypsine a été autorisé, la tolérance de masse en MS et MS/MS est fixée à 5 ppm et 0,8 Da respectivement et l'appareil utilisé a été spécifié comme étant l'*ESI-Trap*.

Le logiciel MFPaQ v.4.0 (Bouyssié *et al.*, 2007), développé à l'IPBS au sein de l'équipe de Bernard Monsarrat a été utilisé pour valider et trier automatiquement les protéines selon les scores des peptides ayant conduit à leur identification. Les protéines ont été validées lorsqu'elles atteignaient les critères suivants : identification avec au moins un peptide de rang 1 et de score Mascot supérieur à 50 (ce qui correspond à une *p-value* inférieure à 0,001) ou au moins 2 peptides de rang 1 et de score Mascot supérieur à 33 (ce qui correspond à une *p-value* inférieure à 0,05).

J. Analyse par Western-blot

L'électro-transfert et l'immunorévélation pour la validation des protéines dont l'expression apparaît modifiée lors de l'analyse protéomique ont été effectués comme décrit dans Mariller *et al.* (2009). Les anticorps utilisés et les conditions expérimentales qui s'y rapportent sont reportés dans le **Tableau V**.

Tableau V : Anticorps et conditions expérimentales utilisées pour l'immuno-révélation des protéines dont l'expression est augmentée en réponse aux isoformes de lactoferrine.

Anticorps primaires	Référence	Origine
H2B	Cell Signaling	Monoclonal de lapin
GAPDH	Santa Cruz	Polyclonal de souris
ALDH18A1	Santa Cruz	Polyclonal de chèvre
RPS9	Santa Cruz	Polyclonal de chèvre
SelH	Santa Cruz	Polyclonal de chèvre
Cathepsine Z	Santa Cruz	Polyclonal de chèvre
Héparanase	Abcam	Polyclonal de lapin
Gamma Glutamyl hydrolase	Santa Cruz	Polyclonal de chèvre
Tropomyosine	Abcam	Monoclonal de souris
MAN2B1	Abnova	Polyclonal de souris
Calmoduline	Abcam	Polyclonal de lapin
Intégrine alpha-6	Abcam	Polyclonal de lapin
Anti-Ig de lapin	Amersham	Polyclonal de chèvre
Anti-Ig de souris	Amersham	Polyclonal de mouton
Anti-Ig de chèvre	Sigma	Polyclonal de lapin

K. Quantification de l'expression des ARNm par qPCR

L'extraction des ARN totaux et la PCR quantitative ont été réalisées comme décrit dans Mariller *et al.* (2009). Les séquences des couples d'amorces utilisées sont référencées dans le **Tableau VI** (page 104).

Tableau VI : Oligonucléotides utilisés pour la qPCR

Gène cible	Séquence de l'amorce sens	Séquence de l'amorce anti-sens	Tm (°C)
ALDH18A1	CTGAGCGCTCCCCACTGCAAA	GACGTAACCTCTGACCTGGTTGGC	53
RPS9	GCGATGACGCCCATGAGGCTA	GCGATGACGCCCATGAGGCTA	55
SelH	GCTTCCAGTAAAGGTGAACCCG	ACCCAATCTCCCT ACGACAGG	55
Cathepsine Z	ACTGGCGCAATGTGGATG	GATGGCGTGGCACTCTTG	57
Héparanase	CTCCTTCCGCCCTTTCC	ACCAAATACTGCATGTTCTCACTT	55
Gamma Glutamyl hydrolase	GAGTTTATTCAACAATGGAAGGATATAAG	TGGGAAATGCCATCCAAATT	69
Tropomyosine	AGCGGGCTGAGCTCTCAGAAG	TGCGAGTACTTATCCTCTGCAGCC	54
MAN2B1	CCGTGAGCGGCAATCCAAACA	CCTGTCCGGTGAGTCTGGGG	54
Calmoduline	TGAGACGTGGCCCCACTGGT	GGCATGCCACCAGGGTGTGT	66
Highly similar to HIV-1 Rev-binding protein 2	GTCTGGTGCTGCTGCGAGGG	CCGGCAGCTCCAGCAAGTGG	68

Discussion & conclusion

Au cours de ma thèse, les travaux effectués ont permis de discriminer les deux isoformes par leurs transcrits et de montrer que la transcription de la hLf peut être régulée dans un contexte inflammatoire (Hoedt *et al.*, 2010). D'autre part, j'ai pu mettre en place deux études protéomiques. Une première basée sur l'utilisation d'électrophorèses bidimensionnelles (2-D) et l'identification des protéines surexprimés en réponse à la ΔLf par spectrométrie de masse, nous a permis de mettre en évidence une nouvelle cible de l'activité transcriptionnelle de la ΔLf, la pyrophosphatase DcpS (Mariller *et al.*, 2009). En parallèle, nous avons pu montrer que l'activité transcriptionnelle de la ΔLf est hautement régulée par des modifications post-traductionnelles, notamment la O-GlcNAcylation (Hardivillé *et al.*, 2010). Pour s'affranchir des limitations de la 2-D, la seconde analyse protéomique a été basée sur une approche SILAC et l'utilisation d'un spectromètre de masse de haute performance. Elle nous a permis d'identifier et de quantifier grâce au logiciel MFPaQ un nombre considérable de protéines. Une activité de facteur de transcription est de plus en plus attribuée à la Lf de sécrétion et notre dernière étude nous a permis de mettre en évidence la sélénoprotéine H en tant que cible potentielle de l'activité transcriptionnelle de la ΔLf et de la hLf (troisième partie des travaux personnels).

I. Discrimination des isoformes de la Lf

hLf et ΔLf possèdent la même séquence primaire, mis à part les 45 premiers résidus d'acides aminés de la hLf délétés pour la ΔLf. Ne possédant pas d'anticorps assez sensibles pour discriminer les deux isoformes, nous avons choisi d'utiliser une approche de biologie moléculaire utilisable en clinique, qui permet de les différencier grâce à l'exon spécifique duquel est issu leur transcrit respectif. Nous avons donc mis au point des sondes et oligonucléotides qui ciblent spécifiquement l'exon 1 ou l'exon 1 β pour quantifier grâce à la technologie Taqman les transcrits hLf et ΔLf respectivement. Cette méthodologie nous a permis de confirmer la sous-expression des deux isoformes dans des lignées cellulaires cancéreuses par rapport aux lignées saines correspondantes. D'autre part, bien que les promoteurs des deux transcrits possèdent des éléments de réponse à des *stimuli* inflammatoires, seule l'expression de la hLf peut être régulée dans un contexte inflammatoire. Nous avons également souligné une forte expression des transcrits ΔLf dans les cellules leucocytaires et lymphocytaires, confirmant ainsi les travaux de l'équipe de Christina Teng qui avait montré que le proto-oncogène c-Ets est préférentiellement exprimé dans les cellules lymphoïdes où il active le promoteur P2 qui possède de nombreux sites de fixation à Ets (Liu *et al.*, 2003). Auparavant, les nombreuses tentatives de purification de la ΔLf non étiquetée en vue de sa caractérisation structurale ont été un échec. Une régulation post-

transcriptionnelle pourrait être à l'origine d'un faible nombre de copies de protéines ΔLf même dans des cellules où les transcrits y sont fortement représentés. En effet, les travaux effectués au laboratoire ont montré que la ΔLf possède une demi-vie très courte (Hardivillé *et al.*, 2010), ce qui n'est pas étonnant puisqu'une molécule à activité de facteur de transcription capable de réguler la progression du cycle cellulaire doit avoir une expression transitoire induite par des *stimuli* environnementaux et ne pas être constamment présente, qui plus est en grande quantité. Récemment, Liao *et al.*, ont démontré que l'expression endogène de la hLf est sous régulée par le miR-214 qui se fixe sur la région 3' non transcrive (3'UTR, 3'UnTranslated Region) de l'ARNm de la hLf. Puisque les transcrits Lf et ΔLf possèdent la même région 3'UTR, nous sommes actuellement entrain de chercher à savoir si le miR-214 cible également le 3'UTR de la ΔLf , ce qui expliquerait sa faible expression. Nous avons malgré tout essayé d'isoler la ΔLf à partir des cellules leucocytaires purifiées du sang humain. Malheureusement en plus de la faible quantité de ΔLf présente nous nous sommes heurtés à un double problème, l'optimisation de l'activité de l'endoprotéinase Glu-C, seule enzyme capable de distinguer la hLf de la ΔLf par la libération de peptides N-terminaux spécifiques de chaque isoforme et l'utilisation du MALDI TOF-TOF comme spectromètre de masse. Une analyse protéomique utilisant une approche type AQUA pour la synthèse d'un peptide spécifique de la ΔLf , couplée à une analyse en spectrométrie de masse de type MRM présenterait un atout considérable pour une mise en évidence de la ΔLf .

II. ΔLfRE , élément de réponse à la Lf et à la ΔLf

La ΔLf exerce son activité de facteur de transcription en se fixant *in vivo* sur une séquence spécifique d'ADN appelée élément de réponse, ΔLfRE . Le ΔLfRE est une séquence homologue à la séquence S1 décrite par He et Furmanski (1995) pour la hLf. Les différents travaux effectués au laboratoire ont permis de montrer que les gènes *Skp1* (Mariller *et al.*, 2007), *DcpS* (Mariller *et al.*, 2009), *Bax* et *Fas* (Hardivillé, communication personnelle) sont directement régulés par la ΔLf . Pour cela, nous avons principalement validé des ΔLfRE issus des régions promotrices des gènes *Skp1*, *DcpS*, *Bax*, *Fas* et *SelH* en tant qu'unité transactivatrice. Les séquences promotrices de ces gènes présentent toutes une séquence commune [G/A]GCACT(G)T[A/G][C/G] et la mutation des cinq bases centrales (GCACT) conduit à une inhibition de la transactivation de ces gènes. Nous avons proposé cette séquence aux responsables du logiciel Genomatix (<http://www.genomatix.de>) qui l'ont inscrite dans leur base de recherche MatInspector sous le nom de « *lactoferrin motif* ». Pour l'instant, la pertinence de recherche de ce motif n'est pas optimale puisqu'aucun poids n'est imposé par base du consensus. L'activation de la transcription via l'élément ΔLfRE présent dans le promoteur de *SelH* est hautement efficace suggérant que la séquence de ce ΔLfRE

devrait être très proche de la séquence consensus optimale (Tableau XIV). L'efficacité de l'élément de réponse présent dans le promoteur de *Skp1* ne semble pas être affectée par la présence d'une adénine en position 9 et d'une cytosine en position 10. Par contre, l'absence de guanine en position 7 semble préjudiciable puisque les efficacités transcriptionnelles sont moins élevées. La thymine en position 8 est indispensable puisque son remplacement par une adénine dans le ΔLfRE de *Fas* conduit à une activité transcriptionnelle extrêmement faible. En outre, nous avons pu démontrer que la séquence ΔLfRE en position inversée est fonctionnelle comme on peut le voir pour celle de nombreux facteurs de transcription. La mutation séquentielle de chaque base du $\Delta\text{LfRE}^{\text{SelH}}$ ou du $\Delta\text{LfRE}^{\text{Skp1}}$ suivie de la mesure de l'efficacité transcriptionnelle des mutants obtenus devraient nous permettre de pondérer chaque nucléotide et par une approche bioinformatique ultérieure de rechercher de nouveaux gènes cibles de la ΔLf . En parallèle, il serait intéressant également d'effectuer un séquençage à haut débit des segments d'ADN sur lesquels se fixe la ΔLf (ChIP-Seq, *Chromatin ImmunoPrecipitation Sequencing*) pour connaître les différentes séquences cibles de la ΔLf . Il nous faudra par la suite déterminer si la fixation de la ΔLf à ces séquences conduit à une activation ou une répression transcriptionnelle. En effet, des études récentes réalisées dans notre groupe montrent que la ΔLf est la cible de nombreuses modifications post-traductionnelles dont certaines sont connues pour réguler la transactivation ou la répression transcriptionnelle.

Tableau VII : Efficacité respective des éléments de réponse à la ΔLf par rapport à celle du ΔLfRE présent dans le promoteur du gène *Skp1*.

ΔLfRE	Séquence										Efficacité en %
	1 G/A	2 G/T	3 C	4 A	5 C	6 T	7 (G)	8 T	9 A/G	10 C/G	
$\Delta\text{LfRE}^{\text{Skp1}}$	G	G	C	A	C	T	G	T	A	C	100
$\Delta\text{LfRE}^{\text{SelH}}$	G	G	C	A	C	T	G	T	G	G	130
$\Delta\text{LfRE}^{\text{SelH-inv}}$	A	G	C	A	C	T	G	T	T	G	130
$\Delta\text{LfRE}^{\text{DepS}}$	A	G	C	A	C	T	-	T	G	G	90
$\Delta\text{LfRE}^{\text{Bax}}$	G	G	C	A	C	T	-	T	A	T	60
$\Delta\text{LfRE}^{\text{Fas}}$	G	G	C	A	C	T	-	A	A	C	5

He et Furmanski (1995) ont été les premiers à proposer une activité de facteur de transcription pour la hLf. Ils ont montré qu'elle est capable d'interagir *in vitro* avec des séquences spécifiques d'ADN dont une, la séquence S1 qui s'est avérée être fonctionnelle. Depuis, différents auteurs ont montré que la hLf est capable de transactiver les gènes *IL-1 β* (Son *et al.*, 2002), *NFATc1*, *Nuclear Factor Activated T cells* (Naot *et al.*, 2011) et *endotheline-1* (Ha *et al.*, 2011) *via* un élément de réponse identique à la séquence S1. Nos travaux montrant que ΔLf et hLf sont toutes deux capables de conduire à la surexpression de la sélénoprotéine H *via* le même élément de réponse suggèrent que ces deux isoformes

pourraient avoir les mêmes gènes cibles. Puisque les deux isoformes partagent le même élément de réponse, il serait intéressant de savoir ce qu'il adviendrait suite à la réintroduction simultanée des deux isoformes dans une même cellule.

L'activité facteur de transcription est liée à la présence du facteur impliqué dans le noyau. La polémique autour de l'activité de facteur de transcription de la hLf concerne davantage l'adressage au noyau d'une Lf d'origine exogène. En effet, après avoir été endocytée, la Lf est retrouvée dans l'endosome avant d'être dégradée dans les lysosomes. Néanmoins une faible proportion de cette hLf a été observée dans le noyau de certaines cellules comme dans celui des cellules MDAMB-231 sans que l'on connaisse réellement le mécanisme de transport mis en jeu (Legrand *et al.*, 2004). Le transfert au noyau dépend également de la reconnaissance par des importines d'une séquence de domiciliation nucléaire. La hLf, en plus de la séquence NLS commune à toutes espèces de Lf, possède un second signal de localisation nucléaire qui correspond aux 5 premiers résidus d'acides aminés ¹GRRR⁵. La présence de ces deux signaux de localisation nucléaire pourrait expliquer un transfert plus efficace de la hLf au noyau et donc une meilleure transactivation du gène *SelH* par la hLf endogène par rapport à la ΔLf. En effet, cette étude a été réalisée avec une hLf amputée de la séquence signal donc une hLf intracellulaire. Par la suite, il nous faudra comparer cette efficacité de transactivation avec celle de la hLf exogène. Les résultats obtenus pourraient nous renseigner quant à l'efficacité de l'adressage nucléaire de la hLf selon le type cellulaire étudié et donc selon le type de récepteurs responsables de l'endocytose.

Puisque les modifications post-traductionnelles contrôlent étroitement l'activité et la stabilité de différents facteurs de transcription dont la ΔLf, il nous faudra étudier le profil de ces modifications sur la hLf après endocytose et transfert au noyau afin de voir si elles sont présentes. La hLf étant une N-glycosylprotéine, il serait intéressant de connaître le profil de N-glycosylation de la hLf exogène une fois cette dernière transportée au noyau.

III. Régulation de l'activité transcriptionnelle de la ΔLf

L'adressage au noyau de la ΔLf est une condition nécessaire mais pas suffisante pour qu'elle soit pleinement active. En effet, nous avons montré que la ΔLf est la cible de différentes modifications post-traductionnelles et que chacune d'elle est susceptible de réguler sa stabilité et/ou son activité transcriptionnelle. Nous avons notamment pu montrer que la modification par O-GlcNAcylation du résidu de S¹⁰ est associée à une augmentation de la stabilité de la ΔLf, grâce à une inhibition de la dégradation protéasomale. D'autre part, cette modification régule négativement l'activité transcriptionnelle de la ΔLf alors que la phosphorylation du même résidu l'active (Hardivillé *et al.*, 2010). En outre, nous avons montré que la ΔLf pouvait être SUMOylée sur les résidus de K¹³, K³⁰⁸, K³⁶¹ et K³⁹¹ et la

recherche *in silico* de motifs acétylables nous a prédit une acétylation de la K³. Nos expériences récentes d'immunoprecipitation de la ΔLf, immunorévélée à l'aide d'anticorps dirigés contre les lysines acétylées nous ont montré que la ΔLf est également acétylée. Il est maintenant indispensable de muter la lysine en position 3 en arginine pour savoir s'il s'agit de la seule cible des acétyl-transférases. Les travaux actuels du groupe cherchent à mettre en évidence une implication potentielle de ces modifications post-traductionnelles dans la régulation de l'activité transcriptionnelle de la ΔLf et/ou sa stabilité et à étudier comment s'établit le dialogue entre ces différentes modifications post-traductionnelles. En effet, ces modifications sont essentiellement concentrées au niveau de deux domaines clef de la ΔLf : le domaine N-terminal avec le résidu potentiel de lysine acétylable (K³), la S¹⁰ qui peut être phosphorylée ou GlcNAcylée et la Lys¹³ SUMOylable et la région centrale avec le DBD2, la séquence PEST susceptible d'être phosphorylée ou GlcNAcylée, les K³⁰⁸ et K³⁶¹ et K³⁹¹ SUMOylables et les K³⁷⁹ et K³⁹¹ cibles de l'ubiquitine.

L'étude de la relation entre ces différentes modifications post-traductionnelles est d'autant plus intéressante qu'il a été démontré qu'elles régulent le trafic, la stabilité, la fixation à l'ADN, le recrutement de partenaires ou la sélectivité des gènes transactivés donc l'activité des facteurs de transcription (Benayoun et Veitia, 2009). SUMOylation et acétylation ont été décrites comme étant deux modifications mutuellement exclusives alors que SUMOylation et phosphorylation peuvent être agonistes ou antagonistes selon le substrat. Ainsi, l'acétylation de la ΔLf pourrait conduire à une activation de la transcription (Glozak *et al.*, 2005) tout comme la phosphorylation alors que la O-GlcNAcylation (Hardivillé *et al.*, 2010) et/ou la SUMOylation conduiraient à la répression de la transcription (Gill, 2005). Ces quatre modifications post-traductionnelles présentes en N-terminal pourraient ainsi contrôler très finement l'activité transcriptionnelle de la ΔLf.

Par ailleurs, la SUMOylation en entrant en compétition avec l'ubiquitylation permet de stabiliser le facteur de transcription cible (Lin *et al.*, 2003; Desterro *et al.*, 1998). Ainsi la ΔLf dont le résidu de K³⁹¹ est à la fois cible de l'ubiquitylation et de la SUMOylation pourrait avoir sa demi-durée de vie régie par ces modifications. Sachant que la séquence PEST peut également être phosphorylée ou GlcNAcylée, le dialogue entre ces différentes modifications pourrait être responsables de la stabilité de la ΔLf. Il ne faut cependant pas oublier le rôle clef de la S¹⁰ sur l'ubiquitylation de la ΔLf confirmant que le domaine N-terminal a un rôle crucial dans la régulation à la fois de l'activité et de la stabilité de la ΔLf.

Le décodage des modifications post-traductionnelles et de leurs interconnections permettra de décrypter les modalités d'action de la ΔLf. Ainsi, sur la base de nos résultats et des données de la littérature nous proposons un schéma qui tente d'expliquer comment ces différentes modifications post-traductionnelles peuvent contrôler l'activité et la stabilité de la ΔLf (Figure 42, page 110).

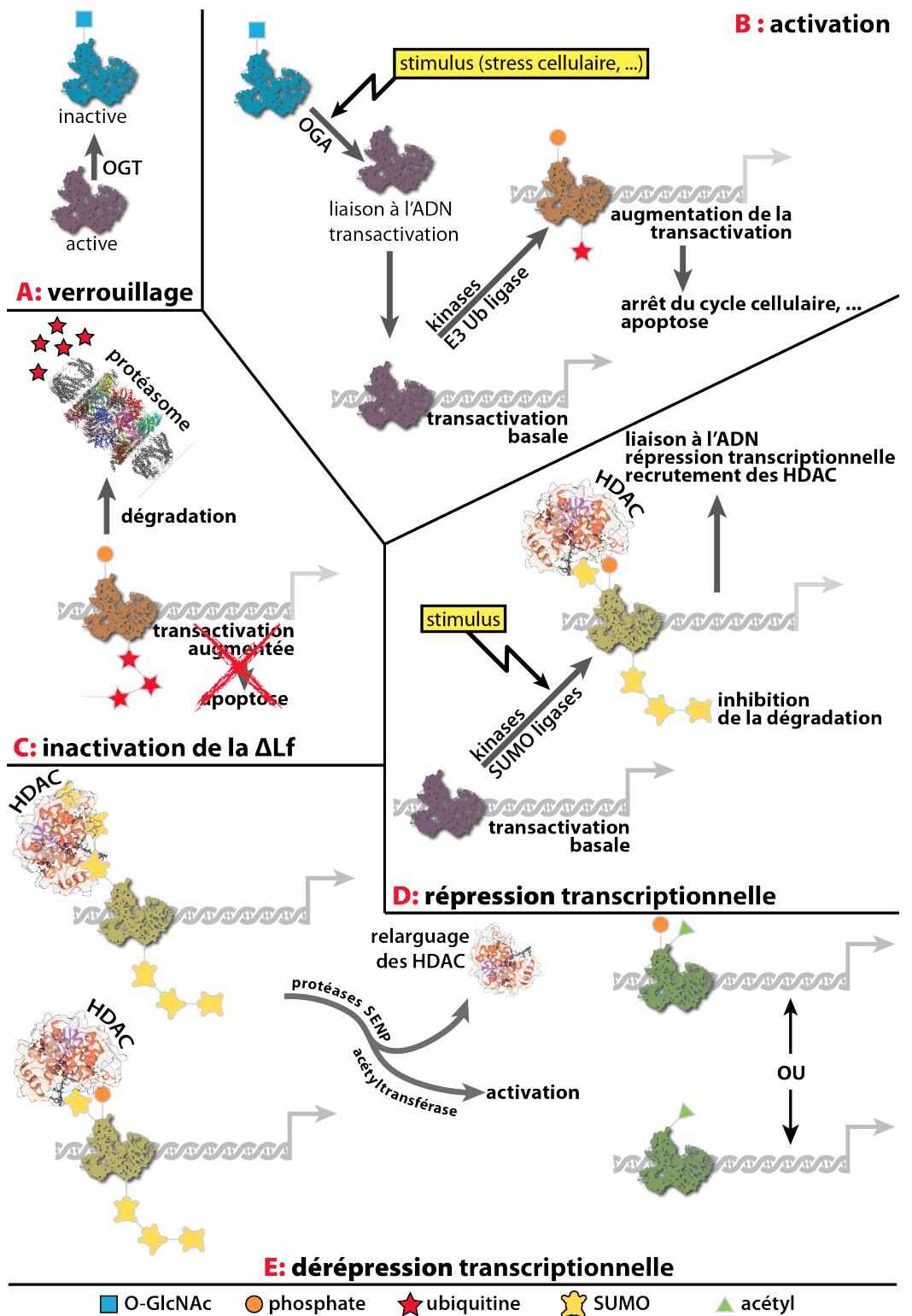


Figure 6 : Illustration de la régulation de l'activité transcriptionnelle et de la stabilité de la ΔLf par les modifications post-traductionnelles (d'après Mariller *et al.*, 2012). A : La ΔLf est présente dans la cellule sous une forme glycosylée inactive. En réponse à un *stimulus* comme le stress cellulaire, la O-GlcNAcase déglycosyle la ΔLf , rendant ainsi le résidu de S¹⁰ accessible aux kinases. La phosphorylation du résidu S¹⁰ potentialisera la phosphorylation de la séquence PEST, l'ubiquitinylation, la fixation de la ΔLf à l'ADN et la transactivation de ses gènes cibles. C : suite à l'activation transcriptionnelle, la poly-ubiquitination de la ΔLf devrait la conduire à la dégradation protéasomale. La seconde régulation pourrait provenir d'une répression de l'activité transcriptionnelle de la ΔLf via sa SUMOylation. La SUMOylation pourrait être à l'origine d'un recrutement de co-répresseurs tels que les HDAC : Histones DéAcétylases qui condensent la chromatine et empêchent la transcription. E : la DéSUMOylation par les SENP protéases et l'acétylation pourrait conduire à une dérépression de la transcription.

IV. Les lactoferrines et le stress oxydant

La capacité qu'a la hLf à chélater le fer lui confère des propriétés anti-oxydantes (Kruzel *et al.*, 2010). Nos résultats ont permis de montrer que la hLf est également capable d'exercer cette activité en régulant l'expression de protéines telles que l'aldéhyde deshydrogénase 18 A1, la protéine ribosomale RPS9 et la sélénoprotéine H, dont la surexpression est associée à la protection des cellules contre le stress oxydant. Les nombreuses études qui ont montré que l'oxygénéation et les niveaux de ROS sont étroitement liés (Choksi *et al.*, 2011) nous ont encouragés à chercher si la réduction des radicaux libres qu'entraîneraient ces protéines à activité anti-oxydante ne serait pas à l'origine d'un état d'hypoxie. Nous nous sommes donc intéressés à l'expression du facteur de transcription HIF-1 (*Hypoxia Inducible Factor-1*), marqueur de l'état d'hypoxie. En effet, en condition de normoxie, la protéine HIF-1 est très instable car elle est la cible du complexe E3 à activité ubiquitine ligase qui induit sa dégradation protéasomale. En condition d'hypoxie, HIF-1 est transloquée au noyau et en présence de ses cofacteurs, transactive les gènes qui possèdent des éléments de réponse à l'hypoxie (HRE, *Hypoxia Response Element*). Nous avons recherché *in silico* la présence de séquences HRE au sein des promoteurs P1 et P2 desquels sont issus les transcrits hLf et ΔLf respectivement. Seul le promoteur P1 possède deux séquences HRE putatives en positions -6772 et -4708. De manière intéressante, la quantification des transcrits HIF-1 par qPCR suite au traitement des cellules MDA-MB-231 par de la hLf ou de la ΔLf montre que seule la réintroduction de hLf conduit à la surexpression des transcrits HIF-1 d'un facteur deux. Ainsi la hLf exogène après interaction avec son ou ses récepteurs de surface pourrait déclencher une réponse anti-oxydante mais aussi après internalisation et transport au noyau transactiver certains gènes cibles comme le gène *SelH*. Afin d'éviter les effets délétères d'un état de type « hypoxique » les cellules sont capables d'exploiter des voies de survie *via* l'activation de facteurs de transcription tels que HIF-1. De manière tout à fait surprenante la hLf pourrait exercer un rétrocontrôle du processus anti-oxydant qu'elle déclenche en conduisant à la surexpression du gène *HIF-1*. Ces résultats sont extrêmement importants, qui plus est dans un contexte de cellules tumorales, et demandent à être confortés.

V. Les lactoferrines et l'invasion cellulaire ?

Depuis 1994, grâce aux travaux de Bezault *et al.*, nous savons que la Lf possède une activité anti-métastatique. Suite à cela, les nombreuses études réalisées, notamment sur des modèles animaux de carcinogénèse chimique ont confirmé cette propriété. Cependant, Kim et collaborateurs ont montré que la Lf est capable d'augmenter l'expression des métalloprotéinases MMP suite à la transactivation du gène *IL-1β*. Les MMP sont des enzymes protéolytiques dont le mécanisme d'action consiste à dégrader les protéines de la matrice

extracellulaire. La surexpression de MMP est souvent associée à l'invasion cellulaire et à une augmentation du nombre de métastases. Ces travaux montrent donc que la Lf est capable de promouvoir la motilité cellulaire de manière indirecte, ce qui va à l'encontre d'une inhibition de la formation de métastases. Récemment, Ha *et al.*, ont démontré que le traitement par de la hLf de plusieurs lignées cellulaires de glande mammaire cancéreuse, parmi lesquelles les cellules MDA-MB-231, augmente leurs capacités de migration et d'invasion. La surexpression des protéines héparanase et cathepsine Z en réponse à la hLf a été mise en évidence grâce à l'analyse protéomique quantitative que nous avons effectuée. L'activité enzymatique de ces deux protéines conduit à la dégradation des protéoglycans de la MEC. Or il a été démontré que les cellules MDA-MB-231 possèdent à leur surface une forte proportion de protéoglycans qui piégent la hLf grâce à des interactions de faible affinité (Rochard *et al.*, 1992). Il serait donc intéressant de savoir si la présence de Lf en surface des cellules MDAMB-231 pourrait contribuer à la dégradation des protéoglycans qui la piégent et par voie de conséquence favoriser son interaction avec ses récepteurs de haute affinité. Il pourrait s'agir d'un mécanisme contrôlant la fixation à ses récepteurs spécifiques, l'induction de la signalisation cellulaire associée et son endocytose.

VI. Les lactoferrines, régulation du cycle cellulaire et apoptose

Les expériences réalisées au laboratoire ainsi que celles relatées par la littérature ont mis en évidence une activité anticancéreuse des deux isoformes de la Lf, *via* un arrêt du cycle en G1/S et l'induction de l'apoptose.

L'analyse protéomique différentielle réalisée suite à l'utilisation de gels 2-D et identification des protéines par spectrométrie de masse ne concernait que la ΔLf, transfectée transitoirement dans les cellules HEK293 et, parmi les protéines identifiées, la protéine Poly(rC)-binding protein 2 (PCBP 2) possède des activités pro-apoptotiques (Roychoudhury *et al.*, 2007). A l'inverse, les protéines T complex subunit β (TCPB) et PRP19 (Lu et Legerski, 2007) participent toutes deux à la survie cellulaire. Leur surexpression pourrait contrebalancer l'effet pro-apoptotique de la ΔLf et permettre le passage en phase S préalablement bloqué par l'activité de cette dernière.

Par contre, l'analyse protéomique différentielle quantitative réalisée par le biais du SILAC ne nous a pas permis de mettre en évidence de surexpression de protéines à activité pro-apoptotique ou impliquées dans la progression du cycle cellulaire. L'analyse des protéomes traités par des doses plus importantes de nos isoformes sera peut-être plus fructueuse en ce sens puisque toutes les études réalisées à ce jour concernant la régulation du cycle et de l'apoptose ont été réalisées suite à une transfection transitoire de ΔLf (Mariller *et al.*, 2007 ; Hardivillé, communication personnelle) et grâce au traitement par au moins

500 µg/ml de Lf (Son *et al.*, 2010). Néanmoins, Actor *et al.*, ont montré en 2009 que le prétraitement avec de la hLf (125 ou 250 µg/ml) de monocytes humains soumis à un agent apoptogène empêche la maturation de la caspase 3 (Actor *et al.*, 2009). De manière intéressante, il semble que deux protéines dont l'expression est augmentée en réponse aux deux isoformes de Lf possèdent des propriétés anti-apoptotiques. En effet, les travaux réalisés par Lindström et Nistér montrent que l'invalidation de l'expression de RPS9 par l'utilisation de *siRNA* conduit à l'induction de l'apoptose dans les cellules de carcinome du col de l'utérus. L'expression de RPS9 serait donc plutôt associée à des effets anti-apoptotiques (Lindström et Nistér, 2010). De la même manière, la surexpression de la sélénoprotéine H humaine dans des cellules neuronales murines les protège de la mort cellulaire induite par une irradiation aux UVB en bloquant l'apoptose induite par la voie mitochondriale et en activant des signaux de survie cellulaire (Mendelev *et al.*, 2009). Il semblerait donc que les deux isoformes de la Lf puissent avoir des effets pro- ou anti-apoptotiques selon la dose de traitement.

Annexes

O-GlcNAcylation/Phosphorylation Cycling at Ser¹⁰ Controls Both Transcriptional Activity and Stability of Δ-Lactoferrin*

Received for publication, November 2, 2009, and in revised form, March 25, 2010. Published, JBC Papers in Press, April 19, 2010, DOI 10.1074/jbc.M109.080572

Stéphan Hardivillé, Esthelle Hoedt, Christophe Mariller, Monique Benäissa, and Annick Pierce¹

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

Δ-Lactoferrin (**ΔLf**) is a transcription factor that up-regulates *DcpS*, *Skp1*, and *Bax* genes, provoking cell cycle arrest and apoptosis. It is post-translationally modified either by O-GlcNAc or phosphate, but the effects of the O-GlcNAc/phosphorylation interplay on **ΔLf** function are not yet understood. Here, using a series of glycosylation mutants, we showed that Ser¹⁰ 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 ³⁹¹KSQQSSDPDPNCVD⁴⁰⁴ sequence as a functional PEST motif responsible for **ΔLf** degradation and defined Lys³⁷⁹ as the main polyubiquitin acceptor site. We next investigated the control of **ΔLf** transcriptional activity by the O-GlcNAc/phosphorylation interplay. Reporter gene analyses using the *Skp1* promoter fragment containing a **ΔLf** response element showed that O-GlcNAcylation at Ser¹⁰ negatively regulates **ΔLf** transcriptional activity, whereas phosphorylation activates it. Using a chromatin immunoprecipitation 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 Ser¹⁰ by aspartate. Moreover, using double chromatin immunoprecipitation 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 Ser¹⁰ 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)² 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).

Because numerous proteins, such as transcription factors, signaling components, and metabolic enzymes are modified, O-GlcNAcylation is 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 C-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 down-regulation (7, 8). OGA may favor gene transcription, not only by reducing the level of glycosylation 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 proteasome 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 (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 toward 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

* This work was supported in part by the CNRS UMR 8576 (Unité de Glycobiologie Structurale et Fonctionnelle), the Institut Fédératif de Recherche 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").

¹ To whom correspondence should be addressed. Tel.: 33-3-20-33-72-38; Fax: 33-3-20-43-65-55; E-mail: annick.pierce@univ-lille1.fr.

² The abbreviations used are: OGT, O-GlcNAc transferase; **ΔLf**, Δ-lactoferrin; Lf, lactoferrin; OGA, O-GlcNAc hydrolase; Ub, ubiquitin; **ΔLfRE**, **ΔLf**

response element; OA, okadaic acid; GlcNH₂, glucosamine; O-GlcNAc, β-O-linked N-acetylglucosamine; O-GlcNAc/P, O-GlcNAc/phosphorylation; ChIP, chromatin immunoprecipitation; re-ChIP, double ChIP; WT, wild type; qPCR, quantitative PCR; HA, hemagglutinin.

O-GlcNAc/P Interplay Regulates ΔLf Stability and Activity

mechanism and high molecular diversity and fine tunes protein interactions and functions, may also target Δ -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 lactoferrin (L f) gene (29), and the use of the first available AUG codon in frame produces an alternative N terminus. Thus, compared with L f , its secretory counterpart, ΔLf is a 73-kDa cytoplasmic protein able to enter the nucleus (30). Potential DNA-binding domains have been suggested for L f , 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 because we recently linked ΔLf overexpression to up-regulation of the *Bax* promoter and apoptosis (35). Because 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 O-GlcNAc/phosphorylation sites were found at Ser¹⁰, Ser²²⁷, Ser⁴⁷², and Thr⁵⁵⁹, and their mutation led to a constitutively active $\Delta\text{Lf}^{\text{M4}}$ mutant (34). Here, we map the major O-GlcNAc/P site to Ser¹⁰, the PEST sequence (amino acids 391–404), and the main poly-Ub site to Lys³⁷⁹. We also report that O-GlcNAcylation at Ser¹⁰ down-regulates ΔLf transcriptional activity and up-regulates its stability by abrogating Ub-mediated proteolysis, whereas phosphorylation activates both transcription and degradation.

EXPERIMENTAL PROCEDURES

Cell Culture, Reagents, and Transfection—HEK 293 cells (ATCCRL-1573) were grown in monolayers and transfected as described (2 μg of DNA for 2×10^6 cells) (30) using DreamFectTM (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 Corp., 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), Ser(P) motif (rabbit polyclonal antibodies, Millipore), and actin (goat polyclonal antibodies I-19, Santa Cruz Biotechnology, Inc. (Santa Cruz, CA)) were used for immunofluorescence, immunoprecipitation, and/or immunoblotting.

Immunofluorescence and Microscopy—HEK 293 cells were transfected by ΔLf C-terminal fused GFP expression vector 24 h prior the 4',6-diamidino-2-phenylindole (Sigma) staining. The p ΔLf -N-EGFP vector was kindly provided by Dr. C. Teng (National Institutes of Health, Research Triangle Park, NC). Immunofluorescence and microscopy were performed as described (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 $\times 40$ 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 minikit (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 described (30). The primer pairs used for the detection of ΔLf (forward, 5'-AAGCCAGTGGACAAGTTCA-3'; reverse, 5'-GCTTGTTGGGATTGTAGTT-3'; annealing temperature, 55 °C), ribosomal protein, large, P0 (34), and hypoxanthine-guanine phosphoribosyltransferase (forward, 5'-GACCAGTCACAGGGGACAT-3'; reverse, 5'-AACACTTCGTGGGGTCCTTTC-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 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, The John Hopkins University School of Medicine (Baltimore, MD)) 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 1. The constructs in which several sites were mutated were done sequentially. Following sequencing, the HindIII-NotI 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 (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 described (30) and expressed as a percentage; 100% corresponds to the relative luciferase activity of $\Delta\text{Lf}^{\text{WT}}$. Basal luciferase expression was assayed using a null vector and was determined for each condition (OGT, okadaic acid (OA), and glucosamine (GlcNH₂)) at each concentration.

O-GlcNAc/P Interplay Regulates ΔLf Stability and Activity

TABLE 1

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'-CGTGGCCCTCCTGTCGCCATCAAAGAGAGA-3' R: 5'-TCTCTTTATGCAGGCACAGGAGGGCCACG-3'
ΔLf ^{S227A}	S227A	F: 5'-CCCGGGTCCCTGCTCATGCCGTTG-3' R: 5'-CAACGGCATGAGCAGGGACCCGGG-3'
ΔLf ^{S472A}	S472A	F: 5'-GTGCCCTGGGGCTGACCGAGAT-3' R: 5'-ATCTCGGGTCAGCCCAGGGGCAC-3'
ΔLf ^{T559A}	T559A	F: 5'-ACGGAAGCCTGAGGCTAGAACG-3' R: 5'-GCTCTAGCCTCAGCCACAGGCTCCGT-3'
ΔLf ^{S10D}	S10D	F: 5'-GTGCCCTCCTGTCGACTGCATAAAAGAGAG-3' R: 5'-CTCTCTTATGCAGTCGACAGGAGGGCAC-3'
ΔLf ^{S16D}	S16D	F: 5'-CTGCATAAAGAGAGACGACCCATCCAGTGTATC-3' R: 5'-GATACACTGGATGGGTCGTCAGTCTTTATGCAG-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'-CAGAGAACTACAAAGCCAACAAAGCAGTG-3' R: 5'-CACTGTTGTTGGGCTTTGTAGTTCTCTG-3'
ΔLf ^{S395A}	S395A	F: 5'-CTACAAATCCCAACAAGCCAGTGACCCCTGATCC-3' R: 5'-GGATCAGGGTCACTGGCTTGTGGGATTGTAG-3'
ΔLf ^{S396A}	S396A	F: 5'-CAAATCCCAACAAAGCGCTGACCCCTGATCCTAAC-3' R: 5'-GTTAGGATCAGGGTCAGCGTTGTGGGATTG-3'
ΔLf ^{PEST}	S392A, S395A, S396A	
ΔLf ^{M4-S392A}	S10A, S227A, S392A, S472A, T559A	
ΔLf ^{M4-S395A}	S10A, S227A, S395A, S472A, T559A	
ΔLf ^{M4-S396A}	S10A, S227A, S396A, S472A, T559A	
ΔLf ^{M4-PEST}	S10A, S227A, S392A, S395A, S396A, S472A, T559A	
ΔLf ^{K379A}	K379A	F: 5'-CTGGCAGAGAACTACGCATCCAAACAAAGCAG-3' R: 5'-CTGTTGTTGGGATGCGTAGTTCTCTGCCAG-3'
ΔLf ^{K391A}	K391A	F: 5'-GTGTACACTGCGATGCGATGTGGTTGGTGC-3' R: 5'-GCACCAAACACATGCGCATGCGATGTACAC-3'
ΔLf ^{KK}	K379A, K391A	
ΔLf ^{M4-K379A}	S10A, S227A, K379A, S472A, T559A	
ΔLf ^{M4-K379A}	S10A, S227A, K391A, S472A, T559A	
ΔLf ^{Δ1-14}	Deletion of amino acids 1 to 14	F: 5'-AAGCTTATGGACTCCCCATCCAGTGTATCC-3' R: 5'-GGATACACTGGATGGGGAGTCCATAAGCTT-3'

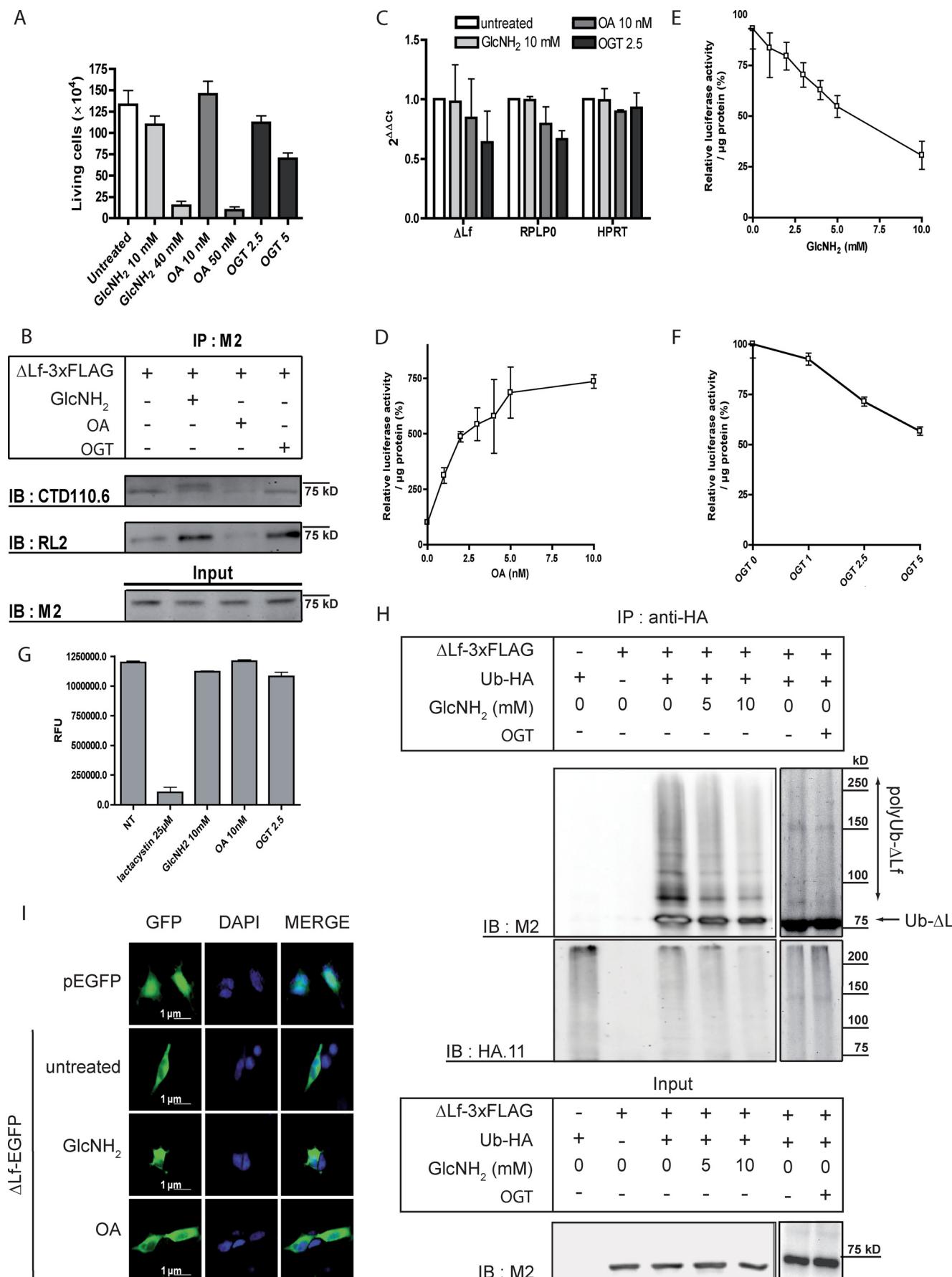
^a Single-letter amino acid codes are used.

Each experiment represents at least three sets of independent triplicates.

In Vivo DNA Binding Assays—Chromatin immunoprecipitation (ChIP) and double ChIP (re-ChIP) assays were performed as described (34, 37) with some modifications introduced for re-ChIP. Briefly, ChIP complexes (8×10^6 cells) were immunoprecipitated with M2, RL2, HA tag, or anti-Ser(P) antibodies all used at 1:250 and twice eluted with 100 μ l of 1 mM dithiothreitol for 30 min at 37 °C. After centrifugation, pooled eluted fractions were diluted 40 times to reduce the dithiothreitol concentration to 25 μ M with ChIP dilution buffer and then immunoprecipitated with either M2 or rabbit anti-IgG (GE Healthcare) or without antibodies. Ampli-

fication conditions of *Skp1* and *albumin* promoters were as described (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 μ l of purified 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; hybridization, 30 s at 55 °C; elongation, 30 s at 72 °C) in a thermocycler Mx4000 (Stratagene). Data presented in Fig. 5D are expressed as a percentage of input.

O-GlcNAc/P Interplay Regulates ΔLf Stability and Activity



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 20 S proteasome inhibitor. Fluorescence data were collected using a Tristar multimode microplate reader LB 941 (Berthold Technologies) using 380-nm excitation and 460-nm emission filters.

Immunoblotting and Immunoprecipitation—Proteins were extracted from frozen cell pellets in radioimmune precipitation buffer as described (30). For direct immunoblotting, samples mixed with 4× 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 4× 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-Ser(P), 1:500; RL2, 1:2000; and anti-actin, 1:10,000) for 2 h and with either secondary anti-IgG antibodies conjugated to horseradish peroxidase (GE Healthcare) or secondary anti-IgM antibodies conjugated to horseradish peroxidase (Chemicon International) for 1 h before detection by chemiluminescence (ECL Advance, GE Healthcare). 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 version 4.1 software (Bio-Rad), and acquisition was carried out with a GS710-calibrated densitometer (Bio-Rad). M2 densitometric values were normalized to actin and expressed as $R = D^{\text{M2}}/D^{\text{act}}$. The -fold stability (X) is expressed as this ratio related to the wild type ratio and to the t_0 value as follows for $\Delta\text{Lf}^{\text{PEST}}$. $X = \frac{\text{PEST}}{\text{PEST}} R_t / \frac{\text{PEST}}{\text{PEST}} R_{t_0} / \frac{\text{WT}}{\text{WT}} R_t / \frac{\text{WT}}{\text{WT}} R_{t_0}$. All of the statistical analyses were done using Origin® 7 software (OriginLab Corp.). Means were statistically analyzed using the *t* test or analysis of variance, and differences were 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 GlcNH₂ or overexpressing OGT exhibit enhanced levels of protein O-GlcNAcylation (39). On the other hand, the use of 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 HEK 293 cells possess rapid, inducible O-GlcNAc/P mechanisms at the OGT, GlcNH₂ 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 GlcNH₂ and 10 nM 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/10⁶ cells) and OGT (2.5 μg of DNA/10⁶ cells) expression vectors did not significantly affect cell viability (Fig. 1A). Fig. 1B shows that ΔLf is indeed sensitive to OA and GlcNH₂ or OGT but with opposite effects. Treatment with OA led to decreased ΔLf glycosylation, whereas treatment with GlcNH₂ 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 favors phosphorylation, decreases the ΔLf glycosylation level, suggesting that glycosylation site(s) may exist in balance with phosphorylation site(s). Because RNA polymerase II activity is also controlled by this interplay, we next verified that transcription of ΔLf ; ribosomal protein, large, P0; or hypoxanthine-guanine phosphoribosyltransferase was indeed not altered under OGT, GlcNH₂, 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, GlcNH₂, or OGT we used.

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

FIGURE 1. O-GlcNAc/P interplay regulates ΔLf transcriptional activity. HEK 293 cells were incubated with GlcNH₂ or 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 10⁴ HEK 293 was assayed 24 h after GlcNH₂ or OA treatment or after transfection with pcDNA-OGT at 2.5 or 5 μg of DNA/10⁶ cells ($n = 9$). *B*, ΔLf O-GlcNAcylation status. Treated and untreated 3xFLAG- ΔLf -expressing HEK 293 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 GlcNH₂ and OA treatment or OGT overexpression. Poly(A)⁺ RNA was purified from total RNA of ΔLf -expressing cells treated with OGT, GlcNH₂, or OA 24 h after transfection and assayed using real-time PCR. RPLPO and hypoxanthine-guanine phosphoribosyltransferase are internal controls ($n = 3$). *D–F*, ΔLf transcriptional activity is modulated by OGT, GlcNH₂, or OA treatment. Cells were co-transfected with pcDNA- ΔLf and pGL3-S1^{Skp1}-Luc and incubated with GlcNH₂ or OA or co-transfected with pcDNA- ΔLf , pcDNA-OGT, and pGL3-S1^{Skp1}-Luc for 24 h prior to lysis. Relative luciferase activities are expressed as described under “Experimental Procedures” ($n \geq 9$). *G*, the endopeptidolytic activity of the proteasome is not altered under OGT, GlcNH₂, 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 GlcNH₂-sensitive. Cells were co-transfected with or without 3xFLAG-tagged ΔLf (ΔLf -3xFLAG) and Ub-HA vectors, treated or not with GlcNH₂, 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 (top and bottom) or HA.11 (middle) antibodies. *I*, ΔLf traffic is not affected by GlcNH₂ or OA treatment. Cells were transfected with pEGFP empty or pEGFP- ΔLf vector and incubated with GlcNH₂ or OA. Fluorescent microscopy was performed after 4',6-diamidino-2-phenylindole (DAPI) staining. Error bars, S.D. *IB*, immunoblot; *IP*, immunoprecipitation.

O-GlcNAc/P Interplay Regulates ΔLf Stability and Activity

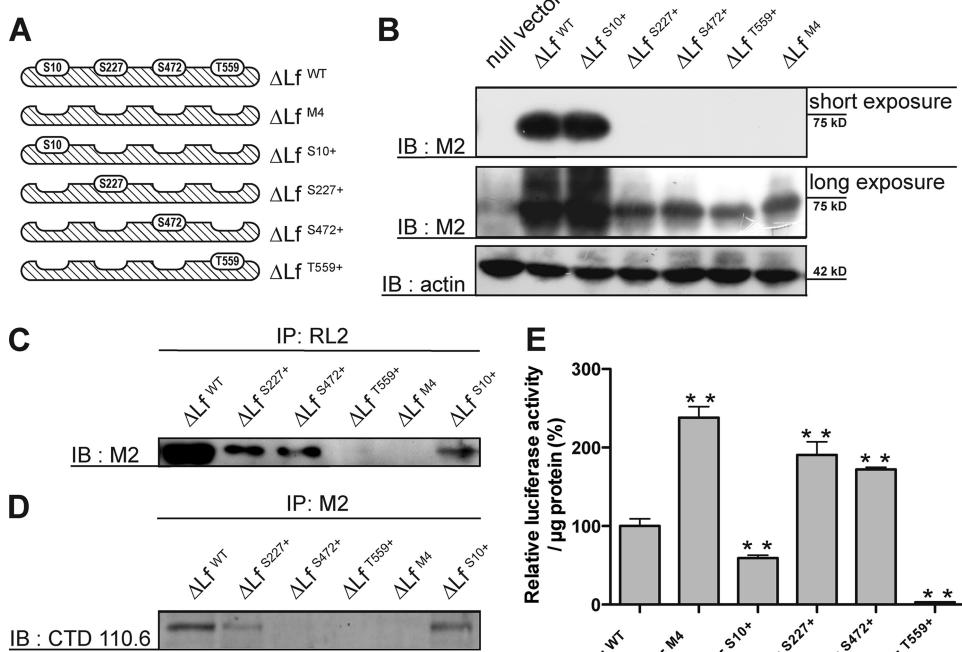


FIGURE 2. Post-translational modifications of Ser¹⁰ modulate ΔLf transcriptional activity and stability. *A*, schematic representation of ΔLf glycosylated mutant constructs. *B*, expression of 3xFLAG- $\Delta\text{Lf}^{\text{WT}}$ and its O-GlcNAcylated mutants. $\Delta\text{Lf}^{\text{WT}}$ and mutant 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* and *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 CTD 110.6. *E*, relative luciferase activity of ΔLf and its mutants. Cells were co-transfected with pGL3-S1^{Skp1}-Luc reporter vector and pcDNA- ΔLf ($\Delta\text{Lf}^{\text{WT}}$) vector or the O-GlcNAc mutant constructs. Relative luciferase activities are expressed as described under “Experimental Procedures” ($n \geq 9$; **, $p < 0.01$). Error bars, S.D. IB, immunoblot; IP, immunoprecipitation.

the ΔLf RE known to be highly transactivated by ΔLf (30). ΔLf transcriptional activity increased in line with OA concentration (Fig. 1*D*), whereas it decreased in a dose-dependent manner in the presence of GlcNH₂ (Fig. 1*E*). Thus, when phosphorylation was augmented, transactivation was increased 6–7-fold compared with 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. 1*F*).

Because, 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, GlcNH₂, or OGT did not disturb the proteasome pathway. As shown in Fig. 1*G*, these treatments did not alter or exacerbate proteasomal degradation compared with the untreated and lactacystin-treated conditions. Fig. 1*H* shows that a ladder of polyubiquitinated ΔLf forms is visible (*top*, 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 GlcNH₂ treatment (Fig. 1*H*, *top*, lanes 4 and 5) and after OGT overexpression (Fig. 1*H*, *top*, lane 7). Equivalent loadings of Ub-HA protein (Fig. 1*H*, *middle*) and ΔLf (Fig. 1*H*, *bottom*) were confirmed by immunoblotting. These data demonstrated that ΔLf is more stable in an environment favoring 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. 1*I*, *panel 2*). Here, we showed that the subcellular localization of ΔLf -GFP was not modified with either GlcNH₂ or OA (Fig. 1*I*, *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 Ser¹⁰—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 labeling 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 $\Delta\text{Lf}^{\text{S10}+}$, $\Delta\text{Lf}^{\text{S227}+}$, $\Delta\text{Lf}^{\text{S472}+}$, and $\Delta\text{Lf}^{\text{T559}+}$, respectively (Fig. 2*A*).

$\Delta\text{Lf}^{\text{M4}}$ (34) and $\Delta\text{Lf}^{\text{WT}}$ were used as controls. ΔLf and its glycosylation mutants were then expressed in HEK 293 cells, and their levels of expression were compared. Fig. 2*B* shows that the $\Delta\text{Lf}^{\text{S10}+}$ mutant was expressed at the same level as $\Delta\text{Lf}^{\text{WT}}$ (short exposure time) in contrast to the other mutants (long exposure time). $\Delta\text{Lf}^{\text{S227}+}$ and $\Delta\text{Lf}^{\text{S472}+}$ were slightly more expressed than were $\Delta\text{Lf}^{\text{M4}}$ and $\Delta\text{Lf}^{\text{T559}+}$, which were both feebly expressed. These data suggest that the post-translational modifications present on Ser¹⁰ 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. Because ΔLf mutants are feebly produced, we first immunoprecipitated ΔLf -expressing cell lysates with RL2 in order to accumulate enough O-GlcNAcylated material (Fig. 2*C*). A reverse immunoprecipitation was then performed using the M2 antibody in order to specifically immunoprecipitate ΔLf or its glycovariants (Fig. 2*D*). Fig. 2*C* shows that ΔLf was effectively glycosylated, whereas $\Delta\text{Lf}^{\text{M4}}$ was not, confirming that no other O-GlcNAc sites are present on the protein. $\Delta\text{Lf}^{\text{S10}+}$, $\Delta\text{Lf}^{\text{S227}+}$, and $\Delta\text{Lf}^{\text{S472}+}$ mutants were glycosylated, whereas $\Delta\text{Lf}^{\text{T559}+}$ was not (Fig. 2*C*). The reverse immunoprecipitation of the cell lysates with M2 antibody followed by O-GlcNAc immunodetection with the CTD 110.6 antibody (Fig. 2*D*) confirmed that ΔLf and its $\Delta\text{Lf}^{\text{S10}+}$ mutant were glycosylated, whereas $\Delta\text{Lf}^{\text{M4}}$ and $\Delta\text{Lf}^{\text{T559}+}$ were not. The O-GlcNAcylated signals corresponding to $\Delta\text{Lf}^{\text{S227}+}$ and $\Delta\text{Lf}^{\text{S472}+}$ mutants that were effec-

O-GlcNAc/P Interplay Regulates ΔLf Stability and Activity

tively 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 Ser⁴⁷² site was not fully glycosylated. However, we cannot exclude the possibility that performing the immunoprecipitation first with RL2 antibody may favor immunoprecipitation of O-GlcNAcylated ΔLf partners. Nevertheless, both immunoprecipitations confirmed without ambiguity that ΔLf^{S10+} , ΔLf^{S227+} , and $\Delta\text{Lf}^{\text{WT}}$ mutants were glycosylated, whereas ΔLf^{T559+} and ΔLf^{M4} were not.

We next assayed the transcriptional activity of the mutants compared with 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 (Fig. 1, E and F). We compared the activity of mutants in which only one glycosylation site was preserved with 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 with 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 because 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 Thr⁵⁵⁹. O-GlcNAc/P modifications on the Ser¹⁰ site led to a 5-fold inhibition of ΔLf transcriptional activity compared with ΔLf^{M4} and 2-fold inhibition compared with $\Delta\text{Lf}^{\text{WT}}$ (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 Ser¹⁰ 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 (Amino Acids 384–404) and Lys³⁷⁹—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 Ser¹⁰ environment but indicated one at the C terminus with three nearly contiguous Ser (Ser³⁹², Ser³⁹⁵, and Ser³⁹⁶) and two flanking Lys (Lys³⁷⁹ and Lys³⁹¹) residues as potential targets either for kinase/OGT or Ub ligase,

TABLE 2
Alignment of ΔLf PEST sequence to lactoferrin sequences from different species

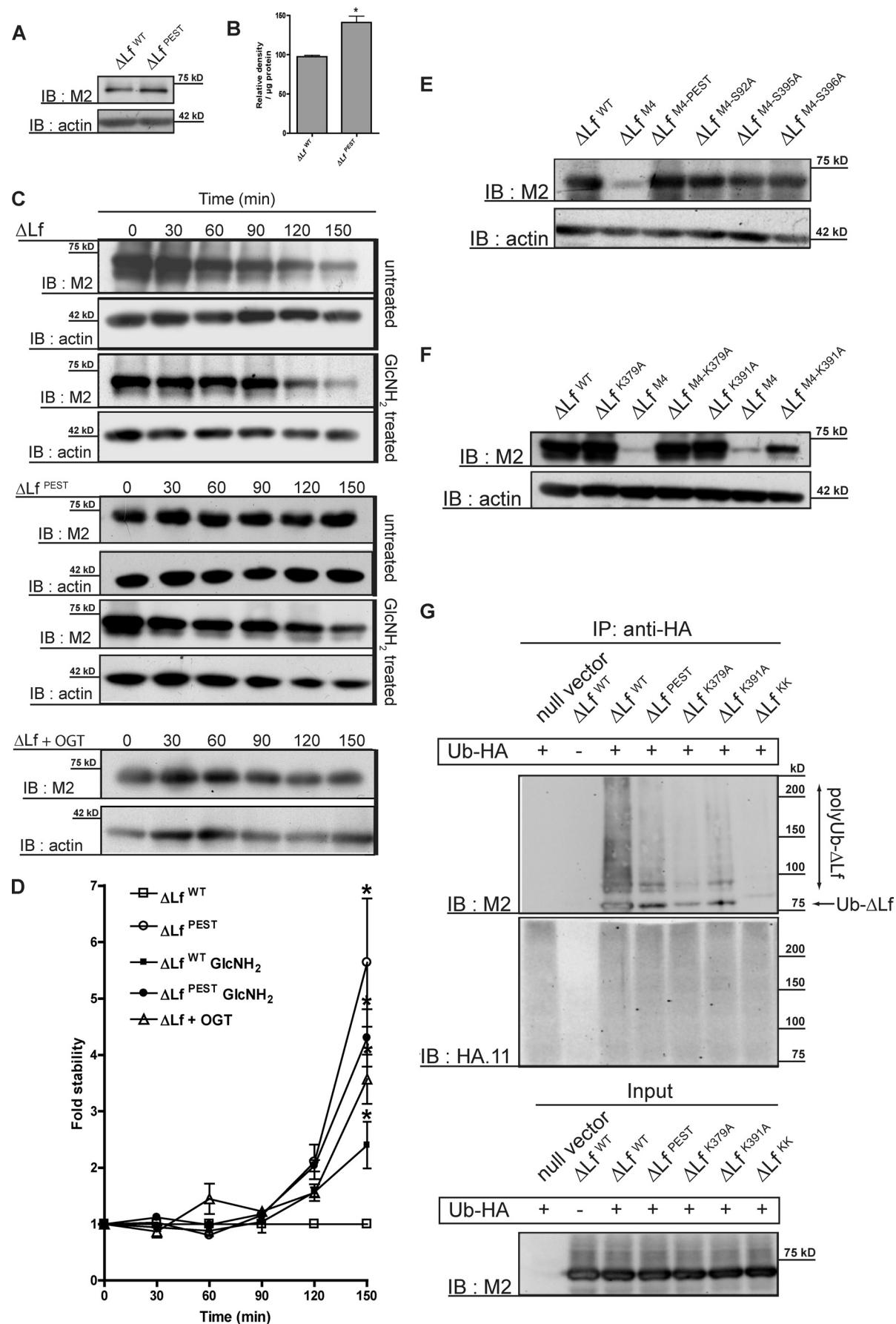
Species	Sequences ^a	Accession number / ref
Delta-lactoferrin		
<i>Homo sapiens</i>	<u>379</u> KCGLVPVLAENYK <u>SQOS</u> S DPDPNCVDRP <u>406</u>	(25)
Lactoferrin		
<i>Homo sapiens</i>	423 KCGLVPVLAENYK SQOS S DPDPNCVDRP <u>450</u>	Q5EK51
<i>Sus scrofa</i>	419 KCGLVPVLAENQK <u>SRQ</u> SSSD --CVHRP <u>444</u>	P14632
<i>Mus musculus</i>	422 KCGLVPVLAENQK <u>SSK</u> S NGLD--CVNRP <u>447</u>	P08071
<i>Equus caballus</i>	410 KCGLVPVLAENQK <u>SQNS</u> SNAPD--CVHRP <u>435</u>	O77811
<i>Bubalus bubalis</i>	423 KCGLVPVLAENRK <u>SSKHS</u> SSLD--CVLRP <u>448</u>	O77698
<i>Bos taurus</i>	423 KCGLVPVLAENRK <u>SSKHS</u> SSLD--CVLRP <u>448</u>	P24627
<i>Capra hircus</i>	423 KCGLVPVMAENRK <u>SSKYS</u> SSLD--CVLRP <u>448</u>	Q29477
<i>Ovis aries</i>	423 KCGLVPVMAENRE <u>SSKYS</u> SSLD--CVLRP <u>448</u>	NP_001020033
<i>Rattus norvegicus</i>	422 KCGLVPVLAEIQK <u>SPN</u> S NGSD--CVDRP <u>447</u>	NP_001100334
<i>Camelus dromedarius</i>	423 KCGLVPVLAESQQ <u>SPES</u> SSLD--CVHRP <u>448</u>	Q9TUM0

^a Single-letter amino acid codes are used; the PEST sequence of ΔLf predicted using the PESTfind software (EMBnet Austria) is in boldface type; italic boldface 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.

respectively. Alignment of Lf sequences from other species to this PEST motif shows that the locus is conserved (Table 2).

We evaluated the functionality of the PEST sequence using a $\Delta\text{Lf}^{\text{PEST}}$ 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 with WT (Fig. 3, A and B). To measure the ΔLf turnover rate indirectly, we performed incubations (0–150 min) with cycloheximide, a potent inhibitor of *de novo* protein synthesis (45, 46). The ΔLf content of HEK cells transfected with either $\Delta\text{Lf}^{\text{WT}}$ or $\Delta\text{Lf}^{\text{PEST}}$ constructs was analyzed following the addition of cycloheximide (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 cycloheximide treatment (Fig. 3C, panel 1). Mutation of the Ser residues in the PEST sequence conferred stability on ΔLf (Fig. 3C, panel 5). GlcNH₂ treatment of HEK cells transfected with either $\Delta\text{Lf}^{\text{WT}}$ or $\Delta\text{Lf}^{\text{PEST}}$ constructs was also performed (Fig. 3C, panels 3 and 7, respectively), and OGT was coexpressed with $\Delta\text{Lf}^{\text{WT}}$ (Fig. 3C, panel 9). Actin, which is stable under the same experimental conditions, was used as an internal control (Fig. 3C, panels 2, 4, 6, 8, and 10). Densitometric data are expressed as -fold stability as described under “Experimental Procedures” (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. GlcNH₂ treatment or OGT overexpression led to a 2- or 3-fold increase in $\Delta\text{Lf}^{\text{WT}}$ stability, respectively, compared with controls, visible at 90 min, confirming that increasing O-GlcNAcylation protects ΔLf from degradation. However, when $\Delta\text{Lf}^{\text{PEST}}$ -expressing cells were submitted to the GlcNH₂ treatment, the stability of $\Delta\text{Lf}^{\text{PEST}}$ was not significantly different in the presence or absence of GlcNH₂ 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 because, if they were independent, a greater stability of the $\Delta\text{Lf}^{\text{PEST}}$ isoform should be visible in the presence of GlcNH₂.

O-GlcNAc/P Interplay Regulates ΔLf Stability and Activity



O-GlcNAc/P Interplay Regulates ΔLf Stability and Activity

We next studied the invalidation of the PEST sequence on the $\Delta\text{Lf}^{\text{M}4}$ mutant. This mutant is detected at low levels in transfected cells, indicating that it is either feebly expressed or rapidly degraded (Fig. 2B). Fig. 3E shows that this invalidation increased $\Delta\text{Lf}^{\text{M}4}$ 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 2). 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.DTU; available on the World Wide Web) also emphasized these three Ser residues as kinase targets, albeit with a higher score for Ser³⁹⁶ (Ser³⁹², 0.766; Ser³⁹⁵, 0.789; Ser³⁹⁶, 0.977).

We then studied whether ΔLf -mediated ubiquitination occurs predominantly through Lys³⁷⁹- or Lys³⁹¹-linked chains by constructing a series of mutants in which residue 379 or 391 was mutated to Ala either in ΔLf or $\Delta\text{Lf}^{\text{M}4}$. The K379A mutation led to a slightly increased expression level of ΔLf and completely restored stability to $\Delta\text{Lf}^{\text{M}4}$ compared with controls, whereas the K391A mutation had no effect on the ΔLf expression level and only slightly increased expression of $\Delta\text{Lf}^{\text{M}4}$ (Fig. 3F). This result confirms that the flanking Lys³⁷⁹, which is highly conserved among species (Table 2), is involved in ΔLf turnover and suggests that it is the major poly-Ub acceptor site. We next verified that ubiquitination of ΔLf is indeed Lys³⁷⁹-linked. Fig. 3G shows that polyubiquitination was strongly visible on $\Delta\text{Lf}^{\text{WT}}$ (*top, lane 3*), was lower on $\Delta\text{Lf}^{\text{PEST}}$ (*lane 4*) and $\Delta\text{Lf}^{\text{K391A}}$ mutants (*lane 6*), was poorly visible on $\Delta\text{Lf}^{\text{K379A}}$ (*lane 5*), and was not at all visible on the double mutant $\Delta\text{Lf}^{\text{KK}}$ (*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 Lys³⁷⁹ residue corresponds to the main Ub ligase target and that Lys³⁹¹ corresponds 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 Ser¹⁰ Protects ΔLf from Polyubiquitination—To determine whether ΔLf protein stability was controlled via an O-GlcNAc/P switch on Ser¹⁰, other mutants were constructed (Fig. 4A), such as the $\Delta\text{Lf}^{\text{S}10\text{A}}$ mutant that has only the Ser¹⁰ residue mutated and the $\Delta\text{Lf}^{\text{S}10\text{D}}$ mutant in which an Asp residue was introduced in place of Ser in order

to mimic constitutive phosphorylation as described previously (47, 48). Immunoblotting of the different mutants with M2 antibodies is presented in Fig. 4B. $\Delta\text{Lf}^{\text{S}10\text{A}}$ and $\Delta\text{Lf}^{\text{S}10\text{D}}$ had expression levels similar to that of WT, whereas $\Delta\text{Lf}^{\text{S}10\text{D}}$ 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 Ser¹⁰, $\Delta\text{Lf}^{\text{S}10\text{A}}$ was expected, like the $\Delta\text{Lf}^{\text{M}4}$, $\Delta\text{Lf}^{\text{T}559+}$, $\Delta\text{Lf}^{\text{S}227+}$, and $\Delta\text{Lf}^{\text{S}472+}$ mutants (Fig. 2B), to be less stable than WT. In $\Delta\text{Lf}^{\text{S}10\text{A}}$, only Ser¹⁰ is mutated. Therefore, the stability of $\Delta\text{Lf}^{\text{S}10\text{A}}$ might be due to the other sites, which could be used as “protecting sites” in the absence of Ser¹⁰.

The turnover of these different Ser¹⁰ mutants compared with WT and actin (internal control) is shown in Fig. 4C. Differences in the steady state levels of ΔLf and $\Delta\text{Lf}^{\text{S}10\text{D}}$ mutant were readily apparent around 30–60 min and strongly visible after 90 min (Fig. 4C, *panels 1 and 5*). Invalidation of the Ser¹⁰ site in $\Delta\text{Lf}^{\text{S}10\text{A}}$ resulted in a markedly prolonged half-life (Fig. 4C, *panel 3*). Comparable results were obtained when cells expressing $\Delta\text{Lf}^{\text{S}10\text{D}}$ were cultured in the presence of GlcNH₂, confirming the crucial role of O-GlcNAcylation in ΔLf stability (Fig. 4C, *panel 7*). Interestingly, $\Delta\text{Lf}^{\text{S}10\text{D}}$ had a faster turnover rate compared with WT (*panel 9*), indicating that mimicking phosphorylation at this locus triggers degradation. Fig. 4D summarizes the densitometric data of the ΔLf immunoblots expressed as -fold stability, as described under “Experimental Procedures.” $\Delta\text{Lf}^{\text{S}10\text{D}}$, which could be either 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 Ser¹⁰ to Ala also led to ΔLf stability (3.5-fold compared with 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 Ser¹⁰ in the $\Delta\text{Lf}^{\text{S}10\text{D}}$ mutant shortened its half-life. From these experiments, we conclude that phosphorylation at Ser¹⁰ accelerates ΔLf degradation, whereas O-GlcNAcylation at Ser¹⁰ 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 Ser¹⁰ phosphorylation. Fig. 4E shows that polyubiquitination was marked on $\Delta\text{Lf}^{\text{WT}}$ and $\Delta\text{Lf}^{\text{S}10\text{D}}$ (*top, lanes 3 and 4*), whereas it was reduced on $\Delta\text{Lf}^{\text{S}10\text{A}}$ (*lane 5*). Control levels of ubiquitination and ΔLf expression are shown in the *middle* and *bottom panels* (Fig. 4E). Unfortunately, the high turnover rate of the

FIGURE 3. Ub-dependent ΔLf degradation is mediated through a PEST sequence at the C terminus and Lys³⁷⁹ and is inhibited by O-GlcNAcylation. A and B, deletion of the PEST sequence slightly increases ΔLf stability. HEK 293 cells were transfected with either $\Delta\text{Lf}^{\text{WT}}$ or $\Delta\text{Lf}^{\text{PEST}}$ 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 $\Delta\text{Lf}^{\text{WT}}$ or $\Delta\text{Lf}^{\text{PEST}}$ and GlcNH₂-treated or cotransfected with the OGT-construct or not and then incubated with fresh medium supplemented by 10 $\mu\text{g}/\text{ml}$ cycloheximide 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 under “Experimental Procedures.” *, $p < 0.05$. E, the three Ser residues of the PEST sequence are equivalent. Mutation of Ser residues was done on the 3xFLAG- $\Delta\text{Lf}^{\text{M}4}$ 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 Lys³⁷⁹ rather than of Lys³⁹¹ inhibits degradation. 3xFLAG- ΔLf and 3xFLAG- $\Delta\text{Lf}^{\text{M}4}$ constructs were used as template to obtain Lys³⁷⁹ and Lys³⁹¹ 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, Lys³⁷⁹ 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 a 10 μM concentration 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 (*top and bottom*) or with HA.11 (*middle*) antibodies. Error bars, S.D. IB, immunoblot.

O-GlcNAc/P Interplay Regulates ΔLf Stability and Activity

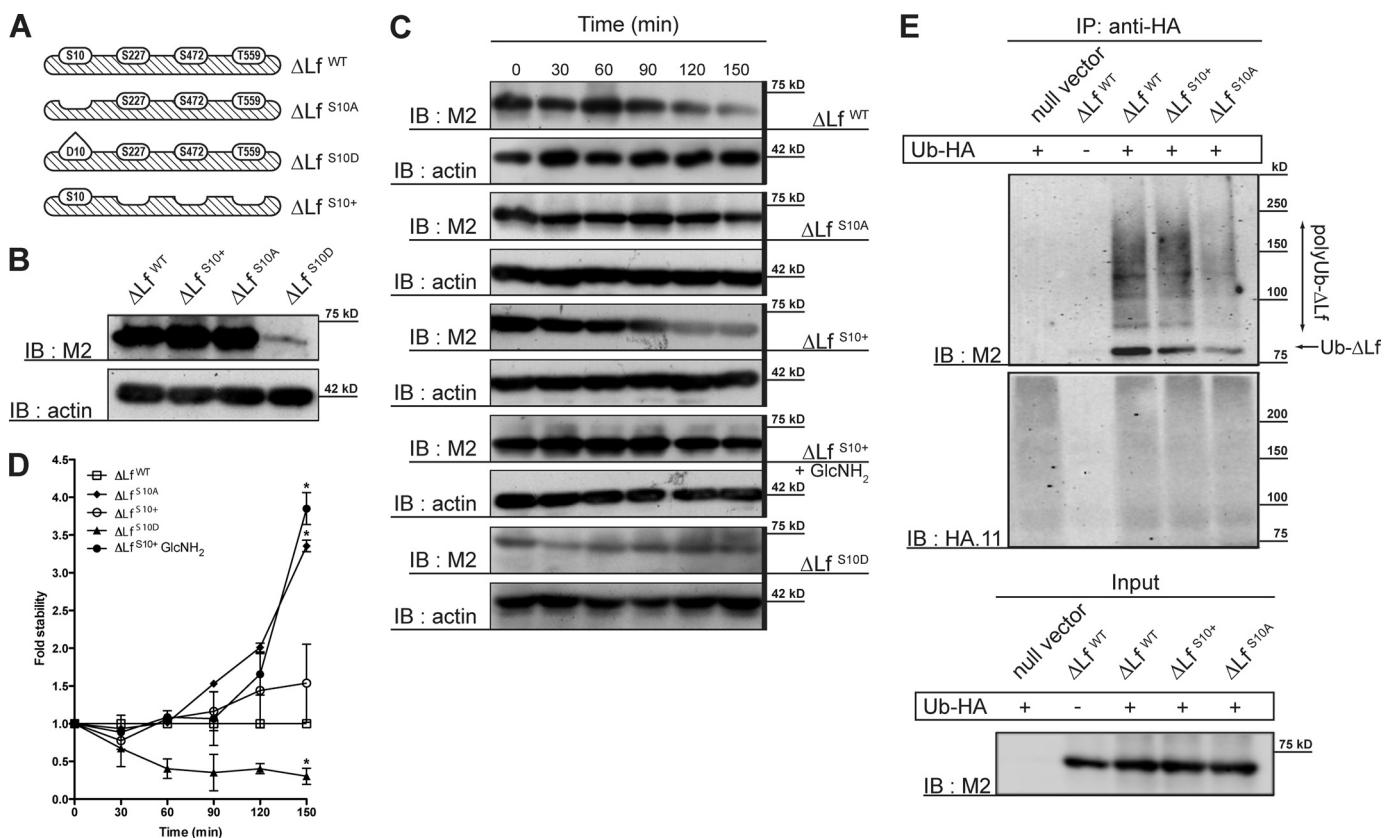


FIGURE 4. Ser¹⁰ posttranslational modifications control ΔLf turnover; O-GlcNAc confers stability, whereas phosphorylation triggers degradation. A, schematic representation of the Ser¹⁰ mutant constructs. B, total protein extract of HEK 293 cells expressing 3xFLAG- ΔLf or its Ser¹⁰ mutant constructs was immunoblotted with M2 or anti-actin antibodies. C, O-GlcNAc posttranslational modification at Ser¹⁰ increases ΔLf stability. Transfected cells were incubated 24 h posttransfection with fresh medium supplemented with 10 μ g/ml cycloheximide. ΔLf^{WT} and ΔLf^{S10+} transfected cells were incubated with or without GlcNH₂. Cell lysates were immunoblotted with M2 or anti-actin antibodies. D, the graph data are expressed as -fold stability as described under “Experimental Procedures” ($n = 5$, *, $p < 0.05$). E, invalidation of Ser¹⁰ protects against Ub-dependent degradation. Cells were transiently co-transfected with or without the 3xFLAG-tagged ΔLf or Ser¹⁰ mutant constructs and the Ub-HA expression vector for 24 h and then incubated with 10 μ M 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 (top and bottom) or with HA.11 antibodies (middle). Error bars, S.D. IB, immunoblot; IP, immunoprecipitation.

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 Lys³⁷⁹. We also demonstrated that the degradation process is regulated via the O-GlcNAc/P interplay, which targets Ser¹⁰. 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 Ser¹⁰ favors phosphorylation of the PEST sequence, whereas O-GlcNAcylation of Ser¹⁰ prevents it.

Phosphorylation at Ser¹⁰ Controls ΔLf Transcriptional Activity—Because 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-Ser(P) antibody, we studied the phosphorylation status of ΔLf (Fig. 5A, left). Phosphatase treatment markedly abrogated the phosphorylation signal, confirming antibody specificity (right). Immunoblotting (Fig. 5A) showed that ΔLf and its Ser¹⁰ mutants exist as phosphoforms. The decreased phosphorylation signals observed under GlcNH₂ 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 with control (lane 5) strongly suggests that the O-GlcNAc/phosphorylation interplay targets the Ser¹⁰ site. However, because 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 Ser¹⁰ controls ΔLf transcriptional activity. Fig. 5B shows that, compared with 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} was increased 4.5–5-fold. The prevention of glycosylation of Ser¹⁰ favored transcription, suggesting that O-GlcNAcylation at this site inhibits ΔLf transcriptional activity. Mimicking phosphorylation at Ser¹⁰ rendered ΔLf even more active than ΔLf^{M4} (Fig. 2E) and strongly suggests that the presence of a phosphate group on this site favors transactivation (Fig. 5B). This result reinforces the status of ΔLf^{S10D} as a constitutive phosphorylated mutant. Because Ser¹⁰ is present in a basic environment (¹MRKVVRGPPVSCIKR¹⁴) within a putative truncated DNA-binding domain, we constructed a ΔLf^{A1-14} mutant in which

O-GlcNAc/P Interplay Regulates ΔLf Stability and Activity

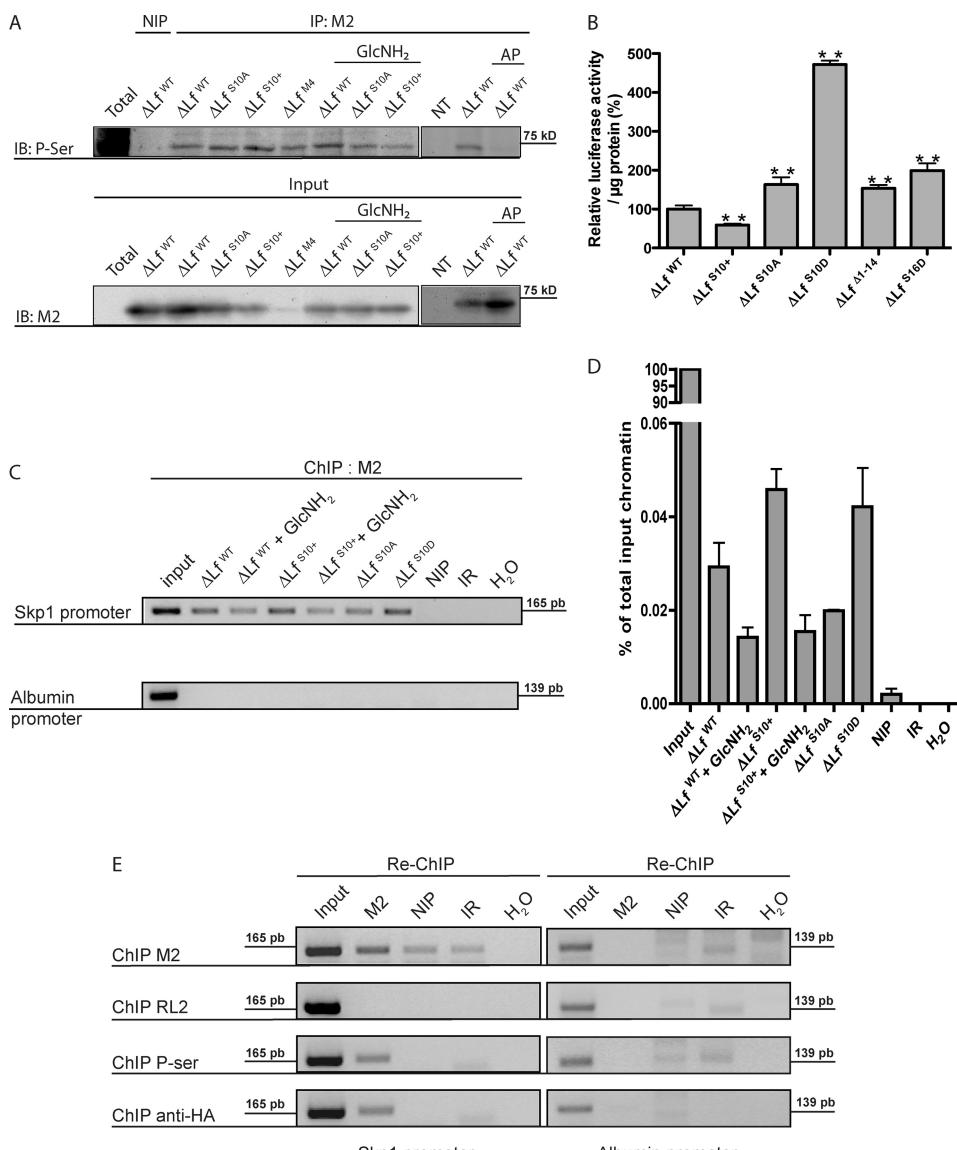


FIGURE 5. **O-GlcNAcylation at Ser¹⁰ negatively controls DNA binding and ΔLf transcriptional activity.** A, ΔLf is a phosphorylated protein. HEK 293 cells were transfected by different ΔLf constructs in the presence or not of GlcNH₂ for 24 h prior to lysis. M2 immunoprecipitates were immunoblotted with anti-Ser(P), and as a loading control, input was immunoblotted with M2 (left). Phosphatase treatment (1 unit/IP) confirms anti-phosphate antibody specificity (AP, right). 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- $\Delta\text{Lf}^{\text{WT}}$ or Ser¹⁰ mutant constructs. Relative luciferase activities are expressed as described under “Experimental Procedures” ($n \geq 9$; **, $p < 0.01$). C and D, O-GlcNAcylation inhibits DNA binding. The *in vivo* binding of ΔLf and its Ser¹⁰ mutants to the Skp1 promoter fragment was examined in HEK 293 cells treated or not with GlcNH₂ ($n = 3$). 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 ΔLf RE 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 (irrelevant; 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-Ser(P) or anti-HA antibodies. Then, prior to reversal of protein-DNA cross-linking, the chromatin fragments were subjected to reprecipitation using M2, irrelevant antibody, or no antibodies ($n = 3$). Error bars, S.D. IB, immunoblot; IP, immunoprecipitation.

the first 14 amino acid residues were deleted. Surprisingly, this deletion did not affect ΔLf transcriptional activity (Fig. 5B), suggesting that the ΔLf DNA-binding domain must be located at the hinge region (31, 32). Because O-GlcNAcylation and phosphorylation might occur on neighboring sites, we screened the vicinity of Ser¹⁰ and identified Ser¹⁶ that might be used as a replacement target by kinases. We therefore constructed a

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 $\Delta\text{Lf}^{\text{S10+}}$ and $\Delta\text{Lf}^{\text{S10D}}$ was twice as high as that of WT. Treatment with GlcNH₂ led to a 0.5-fold promoter site occupancy compared with WT, confirming that favoring GlcNAcylation prevents DNA binding.

$\Delta\text{Lf}^{\text{S16D}}$ 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 $\Delta\text{Lf}^{\text{S10D}}$ and does not take over when the major acceptor site is invalidated, confirming the key role of Ser¹⁰.

Because ΔLf transcriptional activity is altered by O-GlcNAcylation at Ser¹⁰ and 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 ChIP assay we investigated the binding of the different Ser¹⁰ mutants compared with WT. As shown in Fig. 5C, specific ChIP PCR products were detected for each mutant. It is interesting to note that the PCR product signals for $\Delta\text{Lf}^{\text{WT}}$ and $\Delta\text{Lf}^{\text{S10+}}$ were equivalent, whereas treatment with GlcNH₂ led to a weaker signal for both, suggesting that fewer promoter sites were occupied. Because $\Delta\text{Lf}^{\text{WT}}$ and $\Delta\text{Lf}^{\text{S10+}}$ were equivalently expressed (Fig. 4B) even under GlcNH₂ treatment (Fig. 1, B and C, respectively), we suggest that glycosylation inhibits binding to DNA and that among the ΔLf intracellular pool, only the Ser¹⁰ phosphoforms bind ΔLf RE. These results were confirmed by the detection of a PCR product signal comparable with that of WT for $\Delta\text{Lf}^{\text{S10D}}$, which was poorly expressed but extremely active (Fig. 4B), suggesting that a large proportion of $\Delta\text{Lf}^{\text{S10D}}$ binds ΔLf RE (Fig. 5C). The detection of a weaker signal for $\Delta\text{Lf}^{\text{S10A}}$, which was expressed similarly to WT, shows that without phosphorylation and glycosylation at Ser¹⁰, ΔLf still binds DNA, but its capacity to occupy promoter sites is reduced.

O-GlcNAc/P Interplay Regulates ΔLf Stability and Activity

In addition, we performed a 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 a 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 binds the *Skp1* promoter segment with close proximity *in vivo*, whereas glycosylated ΔLf or ΔLf associated with glycosylated proteins does not. Because ΔLf is ubiquitinated at Lys³⁷⁹ and phosphorylated at Ser¹⁰, we suggest that these two post-translational modifications might be concomitantly 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-Ser¹⁰ or the Ub-Lys³⁷⁹ or poly-Ub-Lys³⁷⁹ 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 Ser¹⁰ 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 eIF2a-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 Ser¹⁰ is O-GlcNAcylated. Moreover, we showed that Ser¹⁰ is not present within a phosphodegron, which is a recognition signal for Ub ligases. The ΔLf degradation motif (³⁹¹KSQQSSDPDPNCVD⁴⁰⁴) 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 moieties. YinOYang 1.2 server predictions indicated Ser²⁹² and Ser³⁹⁵ as OGT targets but with low scores. The $\Delta\text{Lf}^{\text{M4}}$ isoform is not glycosylated, suggesting that no further

glycosylation sites are present, but we cannot exclude glycosylation of the PEST motif only when Ser¹⁰ is glycosylated or the possibility that the $\Delta\text{Lf}^{\text{M4}}$ isoform, which is extremely unstable, exists only as a phosphorylated PEST isoform.

We next investigated Ub targets by mutating lysine residues neighboring the PEST motif and demonstrated that ΔLf ubiquitination occurs on Lys³⁷⁹ and Lys³⁹¹ with Lys³⁷⁹ as the main target. The $\Delta\text{Lf}^{\text{KK}}$ double mutant was devoid of Ub, confirming that only these two residues are involved. The formation of Ub ladders observed with $\Delta\text{Lf}^{\text{K379}}$ and $\Delta\text{Lf}^{\text{K391}}$ also revealed that, despite the possibility of its multimonoubiquitination, ΔLf undergoes polyubiquitination. Unexpectedly, the $\Delta\text{Lf}^{\text{PEST}}$ 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 anaphase promoting complex/cyclosome, might be involved. Interestingly, ΔLf possesses a ⁴⁷⁵RSNLCAL⁴⁹¹ sequence, which may behave as a potential RXXLXX(L/I/V/M) D-box motif (ELM D-box entry), the target of anaphase promoting complex/cyclosome (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. Although 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 Ser¹⁰ as in the $\Delta\text{Lf}^{\text{S10A}}$ mutant was accompanied by a decrease in polyubiquitination, suggesting that this modification of $\Delta\text{Lf}^{\text{S10+}}$ and $\Delta\text{Lf}^{\text{WT}}$ only occurs on the phosphoforms. Phosphorylation at Ser¹⁰, 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 Ser¹⁰.

The O-GlcNAc/P interplay also modulates transcriptional activity. O-GlcNAcylation directly activates FoxO1 (52), p53 (53, 54), and NF- κ B (55) and Sp1 indirectly via cofactors (14, 15), whereas it inhibits c-Myc (24) and mouse estrogen receptor β (49). In this work, we have demonstrated that GlcNAcylation inhibits ΔLf transcriptional activity, whereas phosphorylation activates it, and that Ser¹⁰ is central to this regulation. An absence of modification at Ser¹⁰ leads to gene transactivation, whereas phosphomimetic increases it, confirming the inhibitory role of glycosylation. Because the expression of $\Delta\text{Lf}^{\text{S10+}}$ is much greater than that of $\Delta\text{Lf}^{\text{S10D}}$, 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 because 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

O-GlcNAc/P Interplay Regulates ΔLf Stability and Activity

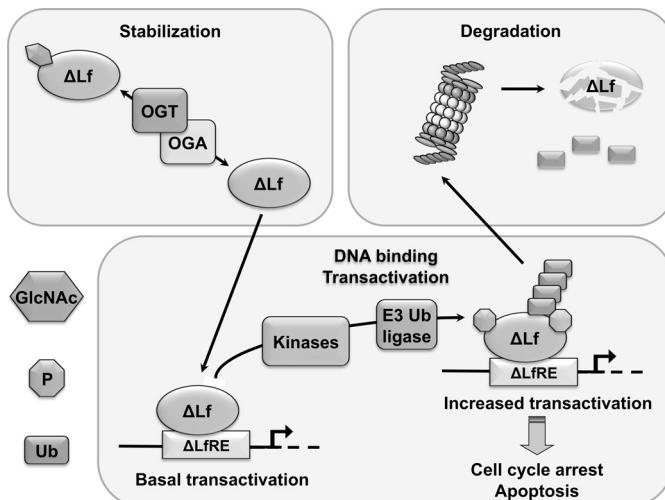


FIGURE 6. Illustration of the regulation of ΔLf activity and stability by the O-GlcNAc/P interplay. ΔLf exists as an inactivated and stable Ser¹⁰-O-GlcNAcylated isoform pool in cells. Depending on cellular events, such as transcription, translation, cell cycle progression, or cell signaling, or in the case of cell stress/injury, OGA will trigger ΔLf deglycosylation, unmasking the Ser¹⁰ hydroxyl group. Deglycosylated ΔLf can bind DNA and transactivate basal transcription. Phosphorylation at Ser¹⁰ by an appropriate kinase leads to a strong amplification of the process and mediates PEST phosphorylation, poly-Ub at Lys³⁷⁹, and ΔLf degradation. P, phosphate.

GlcNH₂ or OA treatment. But even if the nucleocytoplasmic traffic is not governed by the O-GlcNAc/P interplay, because 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 occurred with the unmodified or the phosphomimetic Ser¹⁰ isoform but decreased when O-GlcNAcylation was increased, suggesting that the glycoform is unable to bind DNA. $\Delta\text{Lf}^{\text{S10A}}$ 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 Ser¹⁰ inhibits DNA binding, whereas phosphorylation favors it and promotes transactivation.

The O-GlcNAc/P content fluctuates during cell cycle progression. A recent study showed that increasing O-GlcNAc levels induces a slowing down of both S and G₂/M phases, whereas a reduced O-GlcNAc level impairs the G₁/S checkpoint transi-

tion (63). Because temporal control of Ub-proteasome-mediated protein degradation is critical for normal G₁ 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 G₁/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 down-regulating 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-proteasome-mediated degradation of cell cycle regulators. Furthermore, we demonstrated that ΔLf is itself ubiquitinylated; 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 preubiquitination and phosphorylation as a transcriptional activation signal before being degraded as a polyubiquitinated isoform.

Acknowledgments—We thank Dr. R. J. Pierce (Centre d'Infection et d'Immunité de Lille, INSERM U1019, CNRS UMR 8024, Institut Pasteur de Lille) and Prof. T. Lefebvre (UMR 8576 CNRS, Université des Sciences et Techniques de Lille) for reviewing the manuscript.

REFERENCES

- Hart, G. W., Haltiwanger, R. S., Holt, G. D., and Kelly, W. G. (1989) *Ciba Found. Symp.* **145**, 102–112, discussion 112–108
- Hanover, J. A. (2001) *FASEB J.* **15**, 1865–1876
- O'Donnell, N., Zachara, N. E., Hart, G. W., and Marth, J. D. (2004) *Mol. Cell. Biol.* **24**, 1680–1690
- Lefebvre, T., Dehennaut, V., Guinez, C., Olivier, S., Drougas, L., Mir, A. M., Mortuaire, M., Vercoutter-Edouart, A. S., and Michalski, J. C. (2010) *Biochim. Biophys. Acta* **1800**, 67–79
- Comer, F. I., and Hart, G. W. (1999) *Biochim. Biophys. Acta* **1473**, 161–171
- Hart, G. W. (1997) *Annu. Rev. Biochem.* **66**, 315–335
- Roos, M. D., and Hanover, J. A. (2000) *Biochem. Biophys. Res. Commun.* **271**, 275–280
- Yang, X., Zhang, F., and Kudlow, J. E. (2002) *Cell* **110**, 69–80
- Toleman, C., Paterson, A. J., Whisenhunt, T. R., and Kudlow, J. E. (2004) *J. Biol. Chem.* **279**, 53665–53673
- Wells, L., Kreppel, L. K., Comer, F. I., Wadzinski, B. E., and Hart, G. W. (2004) *J. Biol. Chem.* **279**, 38466–38470
- Hart, G. W., Housley, M. P., and Slawson, C. (2007) *Nature* **446**, 1017–1022
- Slawson, C., Housley, M. P., and Hart, G. W. (2006) *J. Cell. Biochem.* **97**, 71–83
- Zachara, N. E., and Hart, G. W. (2004) *Trends Cell Biol.* **14**, 218–221
- Han, I., and Kudlow, J. E. (1997) *Mol. Cell. Biol.* **17**, 2550–2558
- 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
- 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
- Cheng, X., Cole, R. N., Zaia, J., and Hart, G. W. (2000) *Biochemistry* **39**, 11609–11620

O-GlcNAc/P Interplay Regulates ΔLf Stability and Activity

18. Rechsteiner, M., and Rogers, S. W. (1996) *Trends Biochem. Sci.* **21**, 267–271
19. Sümegei, M., Hunyadi-Gulyás, 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 Acids 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.*, in press
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. Métivier, R., Penot, G., Hübner, 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* **20**, 4524–4530
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., Bussière, 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., Guérardel, C., Deltour, S., Martin-Soudant, N., Sloomiany, 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

REVIEW / SYNTHÈSE

Delta-lactoferrin, an intracellular lactoferrin isoform that acts as a transcription factor¹

Christophe Mariller, Stephan Hardivillé, Esthelle Hoedt, Isabelle Huvent, Socorro Pina-Canseco, and Annick Pierce

Abstract: Delta-lactoferrin (ΔLf) is a transcription factor of which the expression is downregulated in cancer. It is a healthy tissue marker and a high expression level of its transcripts was correlated with a good prognosis in breast cancer. ΔLf results from alternative promoter usage of the *hLf* gene leading to the production of 2 isoforms with alternative N-termini: lactoferrin, which is secreted, and ΔLf , its nucleocytoplasmic counterpart. ΔLf possesses antiproliferative properties and induces cell cycle arrest. It is an efficient transcription factor interacting in vivo via a ΔLf response element found in the *Skp1*, *Bax*, *DcpS*, and *SelH* promoters. Since ΔLf possesses different target genes, modifications in its activity or concentration may have crucial effects on cell homeostasis. Posttranslational modifications modulate ΔLf transcription factor activity. Our earlier investigations showed that *O*-GlcNAcylation negatively regulates ΔLf transcriptional activity, whilst inhibiting its ubiquitination and increasing its half-life. On the other hand, phosphorylation potentiates ΔLf transcriptional activity. Recently, we showed that ΔLf is also modified by SUMOylation. Therefore, cooperation and (or) competition among SUMOylation, ubiquitination, phosphorylation, and *O*-GlcNAcylation may contribute to the establishment of a fine regulation of ΔLf transcriptional activity depending on the type of target gene and cellular homeostasis.

Key words: delta-lactoferrin, prognostic factor, antitumoral agent, transcription factor, posttranslational modifications.

Résumé : La delta-lactoferrine (ΔLf) est un facteur de transcription dont l'expression est régulée à la baisse dans le cancer. Elle constitue un marqueur du tissu sain et le haut niveau d'expression de ses transcrits est corrélé à un bon pronostic dans le cancer du sein. La ΔLf résulte d'une utilisation alternative du promoteur du gène *hLf* qui conduit à la production de deux isoformes possédant des extrémités N-terminales alternatives : la lactoferrine, qui est secrétée, et la ΔLf , la contrepartie nucléocytoplasmique. La ΔLf possède des propriétés anti-prolifératives et induit l'arrêt du cycle cellulaire. C'est un facteur de transcription efficace qui interagit in vivo avec un élément de réponse au ΔLf trouvé dans les promoteurs de *Skp1*, *Bax*, *DcpS* et *SelH*. Puisque ΔLf possède différents gènes cibles, des modifications de son activité ou de sa concentration peuvent avoir des incidences importantes sur l'homéostasie cellulaire. Des modifications post-traductionnelles modulent l'activité du facteur de transcription ΔLf . Nos recherches antérieures ont montré que la *O*-GlcNAcylation régule négativement l'activité transcriptionnelle de ΔLf , pendant qu'elle inhibe son ubiquitination et augmente sa demi-vie. D'un autre côté, la phosphorylation stimule l'activité transcriptionnelle de la ΔLf . Nous avons montré récemment que la ΔLf est aussi modifiée par SUMOylation. En conséquence, la coopération et (ou) la compétition entre la SUMOylation, l'ubiquitination, la phosphorylation et la *O*-GlcNAcylation peut contribuer à établir une régulation fine de l'activité transcriptionnelle de la ΔLf selon le type de gènes cibles et l'homéostasie cellulaire.

Mots-clés : delta-lactoferrine, facteur pronostique, agent anti-tumoral, facteur de transcription, modifications post-traductionnelles.

[Traduit par la Rédaction]

Received 31 July 2011. Revision received 20 September 2011. Accepted 28 September 2011. Published at www.nrcresearchpress.com/bcb on .

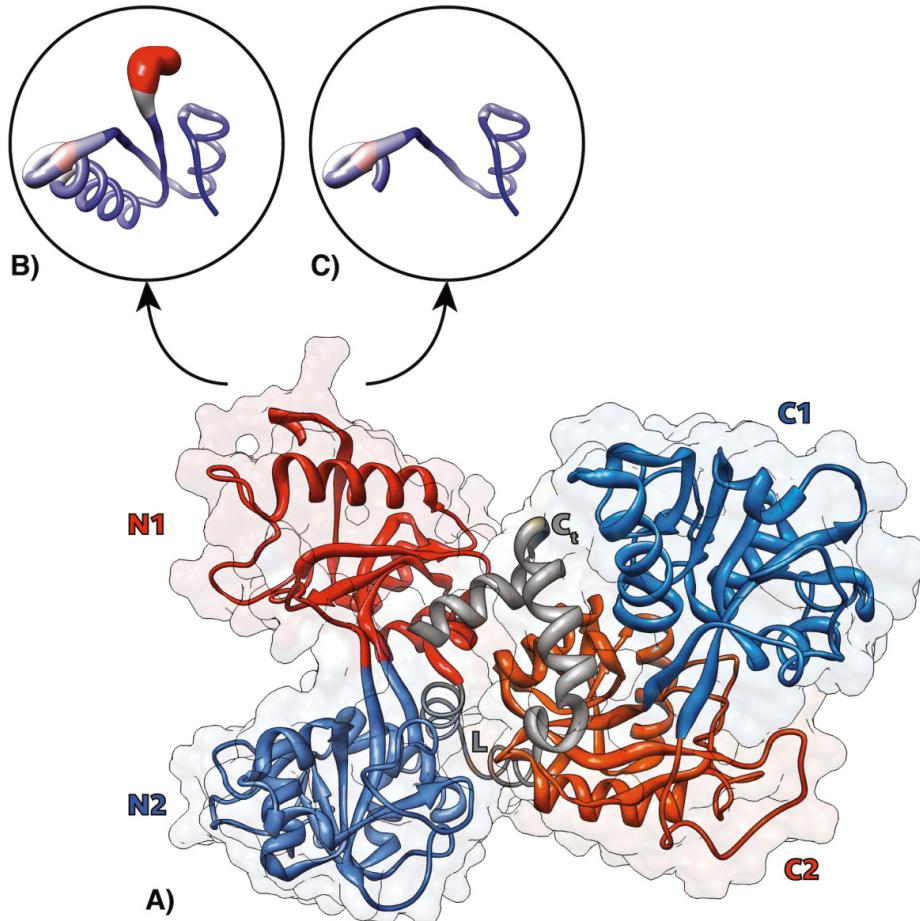
Abbreviations: Abbreviations: ΔLf , delta-lactoferrin; ΔLf RE, ΔLf response element; Ub, ubiquitin; *O*-GlcNAc, *O*-N-acetyl-glucosamine; and SUMO, small ubiquitin modifier.

C. Mariller, S. Hardivillé, E. Hoedt, I. Huvent, S. Pina-Canseco, and A. Pierce. Unité de Glycobiologie Structurale et Fonctionnelle, UMR 8576 CNRS-Université des Sciences et Technologies de Lille, IFR 148, 59655 Villeneuve d'Ascq, France.

Corresponding author: Christophe Mariller (e-mail: christophe.mariller@univ-lille1.fr).

¹This article is part of a Special Issue entitled Lactoferrin and has undergone the Journal's usual peer review process.

Fig. 1. Polypeptide folding of Δ -lactoferrin (Δ Lf). (A) The structure of Δ Lf has been created using coordinates from Lf crystallographic data (Anderson et al. 1987) and plotted using UCSF Chimera 1.5.3 (Pettersen et al. 2004). The surface displays a solvent-excluded molecular surface with a contact probe radius of 1.4 Å. (B) N-terminal end of Lf (residues 1–58 in human Lf) with a diameter proportional to the B-factor. (C) N-terminal end of Δ Lf (residues 1–33 in human Δ Lf) with the same representation.



Introduction

Delta-lactoferrin (Δ Lf) is a truncated intracellular isoform of lactoferrin (Lf) that was described as a transcript for the first time in 1997 in human tissue (Siebert and Huang 1997). In this review, we will summarize our current understanding of the Δ Lf structure function relationship. Its restricted expression and regulation of processes such as cell cycle progression, apoptosis, and mRNA decay suggest that Δ Lf may belong to the tumor suppressor family.

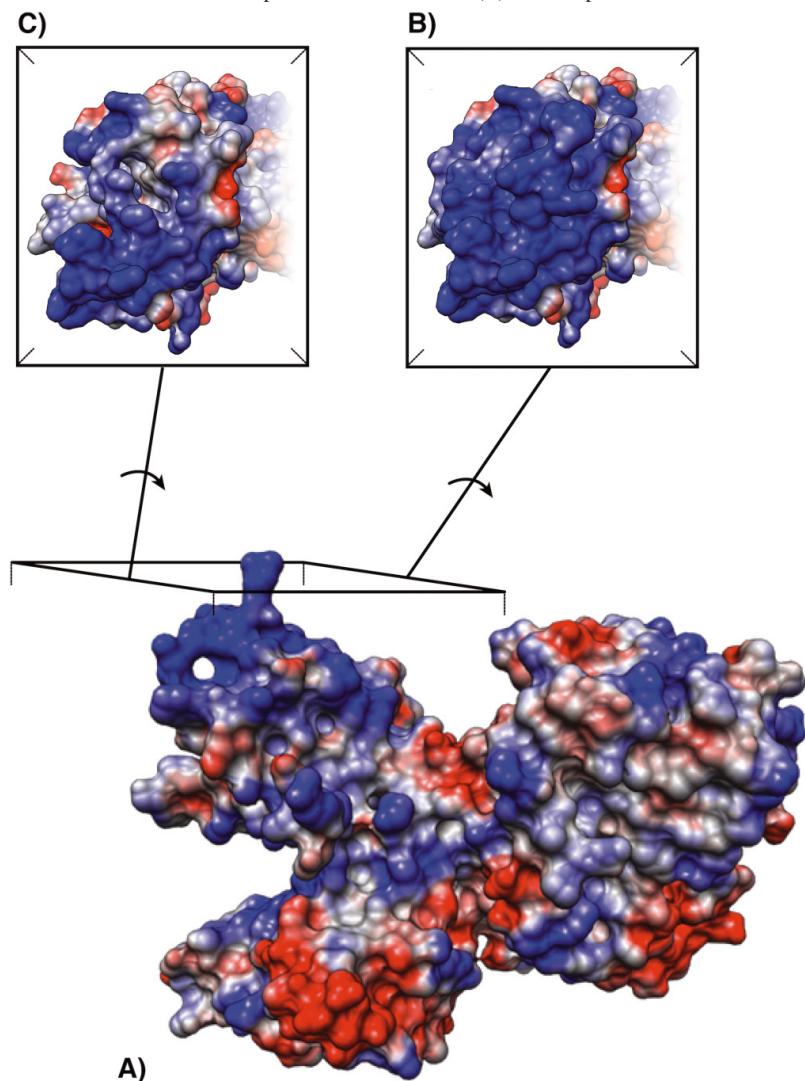
The *Lf* gene is present in mammals (Baldwin 1993) and derives from a duplication of the transferrin gene. It is mapped to human chromosome 3 at 3q21.3 (McCombs et al. 1988), has a total length of 29 kbp (Siebert and Huang 1997; Kim et al. 1998), and is fragmented into 17 exons from 48 to 210 bp and 16 introns of 300 bp to 3.3 kbp. It possesses 2 promoter regions with the P2 promoter located in the first intron of the *Lf* gene (Liu et al. 2003). Activation from the P1 promoter leads to the synthesis of the secreted Lf and from the alternative P2 promoter to Δ Lf. Therefore, there are 2 transcripts derived from the human *Lf* gene, Lf mRNA and Δ Lf mRNA. Both messengers have in common exons 2 to

17 and differ only in their first exon (Siebert and Huang 1997; Liu et al. 2003). Indeed, the use of the P2 promoter leads to the replacement of exon 1 by exon 1 β , longer by ~100 bp. Translation begins at exon 2 and leads to a protein identical to Lf that lacks the leader sequence and the first 26 amino acid residues (Siebert and Huang 1997). Δ Lf is, therefore, a 73 kDa intracellular protein in which the Lf N-terminus, involved in binding to receptors and in bacterial killing functions (Ward et al. 2005; Legrand et al. 2008), is removed, explaining why Δ Lf acts differently from Lf and is not involved in antimicrobial activities and immunomodulation. Until now Δ Lf expression seems specifically restricted to humans but we are currently investigating its presence in other species (S. Hardivillé, personal communication 2011).

Structural features of Δ Lf

The 3-D structure of Δ Lf remains to be established, but, as its sequence is identical to the secretory Lf minus the 25 N-terminal amino acid residues, many of the structural features of Δ Lf can be speculated on the basis of the knowledge acquired from the Lf 3-D structure.

Fig. 2. Electrostatic potential surface of ΔLf . (A) The electrostatic potential surface of ΔLf has been calculated by solving the Poisson-Boltzmann equation using DelPhi (Rocchia et al. 2002). Surfaces are colored according to electrostatic projections at a level of +10 (blue) and -10 (red) kT/e , where k is Boltzmann's constant, T is the absolute temperature, and e is the magnitude of the electron charge. (B) Magnification of the N-terminal domain of Lf electrostatic potential surface of Lf. (C) Same representation but for the N-terminal domain of ΔLf .



The structure of Lf was first determined by Anderson et al. (1987). Its unique polypeptide chain contains 691 amino acids folded in 2 globular lobes. These 2 lobes result from an internal duplication, and the N-terminal lobe (residues 1–333) shares ~40% sequence identity with the C-terminal lobe (residues 345–691). The linker between the 2 lobes is constituted by a short alpha helix (residues 334–344). Lf itself also shares ~60% sequence identity with hTf. Lf is also characterized by its high pI (>9) reflected by its electrostatic potential surface distribution. Positive charges are mainly clustered at the N-terminal end (specifically residues 1–7 and 13–30 of the first helix) (Baker 2005) and in the interlobe region (Baker et al. 2002). This region of the protein is a proposed binding site for DNA, lipopolysaccharides, heparin, and glycosaminoglycans (Mann et al. 1994; van Berkel et al. 1997). Lf is also well known for its iron-binding ability. It was

shown to reversibly bind iron as Fe^{3+} with a K_D of $\sim 10^{-22}$ mol/L. This binding is dependent on the synergistic action of one carbonate anion with the amino acids of the binding site (Aisen and Leibman 1972; Anderson et al. 1989). Fe^{3+} binding and release is associated with the closing or opening of the 2 domains of each lobe over the bound Fe^{3+} . This release is triggered by lowering the pH to 3–4.

As they share identical sequences, Lf and ΔLf are assumed to exhibit a similar conformation. The main difference between them resides in the absence of the first 26 amino acids that are of great significance in the structure and properties of Lf. The N-terminal residues 1–5 of Lf extend out from the surface of the protein and have been demonstrated to be highly flexible (Fig. 1, panel B) (Baker and Baker 2009). This N-terminal end also contains a large proportion of basic residues (Fig. 2, panel B) mainly located in the first 5 resi-

Fig. 3. Schematic diagram of domains and functional motifs/sites of ΔLf . Schematic diagram of ΔLf with functional motifs/sites. PEST, rapid degradation domain; and NLS, nuclear localization signal. *O*-GlcNAcylation sites are indicated by blue boxes and ubiquitination sites by red boxes. For the acetylation and SUMOylation sites and the PEST sequence, the target amino acid is outlined in red. The numbering of the amino acid residues corresponds to human ΔLf .

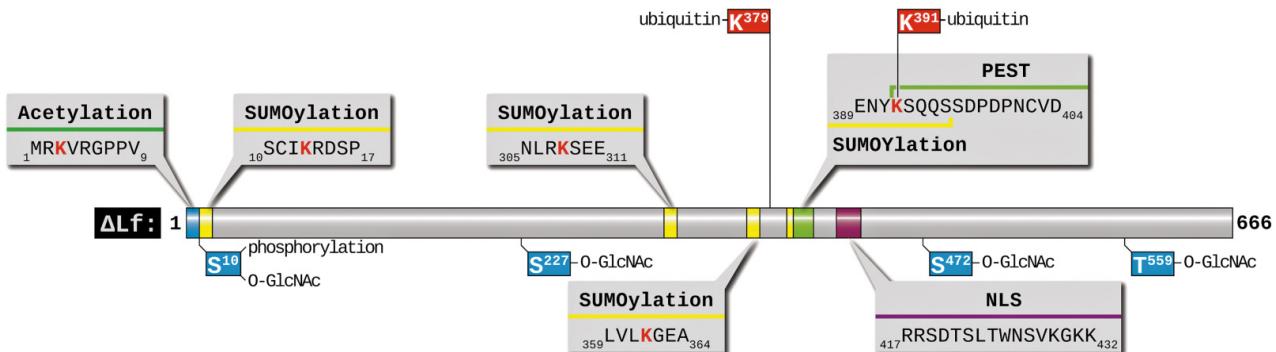


Fig. 4. Illustration of the regulation of ΔLf activity and stability by posttranslational modifications (PTMs). (A) ΔLf exists in cells as a stable but inactive glycoform waiting for a stimulus to be activated. (B) Under stimuli such as cell stress or injury, *O*-GlcNAcase deglycosylates ΔLf and renders Ser¹⁰ accessible to kinases. Phosphorylation of Ser¹⁰ will trigger phosphorylation of the PEST motif, ubiquitination, binding to DNA, and transactivation of ΔLf target genes. (C) Following activation of transcription, polyubiquitination will occur triggering ΔLf to proteasomal degradation to inactivate it. (D) A second layer of control might be provided by blocking activation of ΔLf through the activation of the SUMO pathway leading to repression of ΔLf transcriptional activity. SUMOylation then triggers the recruitment of corepressors such histone deacetylases (HDACs) that condense chromatin and prevent transcription. (E) DeSUMOylation by SENP proteases and acetylation might lead to derepression of the transcription.. OGA, *N*-acetyl- β -D-glucosaminidase; and OGT, *O*-GlcNAc transferase.

dues and the C-terminal end of helix 1 (residues 27–30). This strong basic characteristic of the Lf N-terminal end is greatly reduced in the case of ΔLf (Fig. 2A) but the surface of the latter still remains largely basic. Moreover, despite lacking the N-terminal end of Lf, ΔLf remains a basic protein with a predicted PI of 8.19. Furthermore ΔLf should retain its iron-binding capability and be able to scavenge iron within cells.

Multiple structural aspects of ΔLf can be predicted by bioinformatic tools. First, ΔLf is an intracellular protein and unlike Lf, it does not undergo typical *N*- or *O*-glycosylation. However, many intracellular potential posttranslational modifications (PTMs) can be postulated for ΔLf using different in silico prediction methods, most of which have been experimentally confirmed. Each characteristic presented here is further discussed in the relevant sections. The SUMOsp v. 2.04 software (<http://sumosp.biocuckoo.org/prediction.php>) has been employed to detect putative SUMOylation sites on ΔLf . Four potential sites have been detected (Fig. 3) at positions SCIK¹³RDSP, NLRK³⁰⁸SEE, LVLK³⁶¹GEA, and ENYK³⁹¹SQ¹ in which each lysine residue is potentially a target for SUMO ligases. Mutation of each SUMO site fails to abolish SUMOylation showing that all 4 sites can be SUMOylated (E. Hoedt, personal communication 2011). We are currently studying the role of SUMOylation on ΔLf transcriptional activity and stability (see Regulation of ΔLf activity). In the same way, acetylation sites were screened with the Bdmpail webserver (<http://Bdmpail.biocuckoo.org/prediction.php>). One site was located at Lys3 (MR³KVRGPPV) but remains to be validated. (see Regulation of ΔLf activity). The YinYang 1.2 server (Gupta and Brunak 2002) has high-

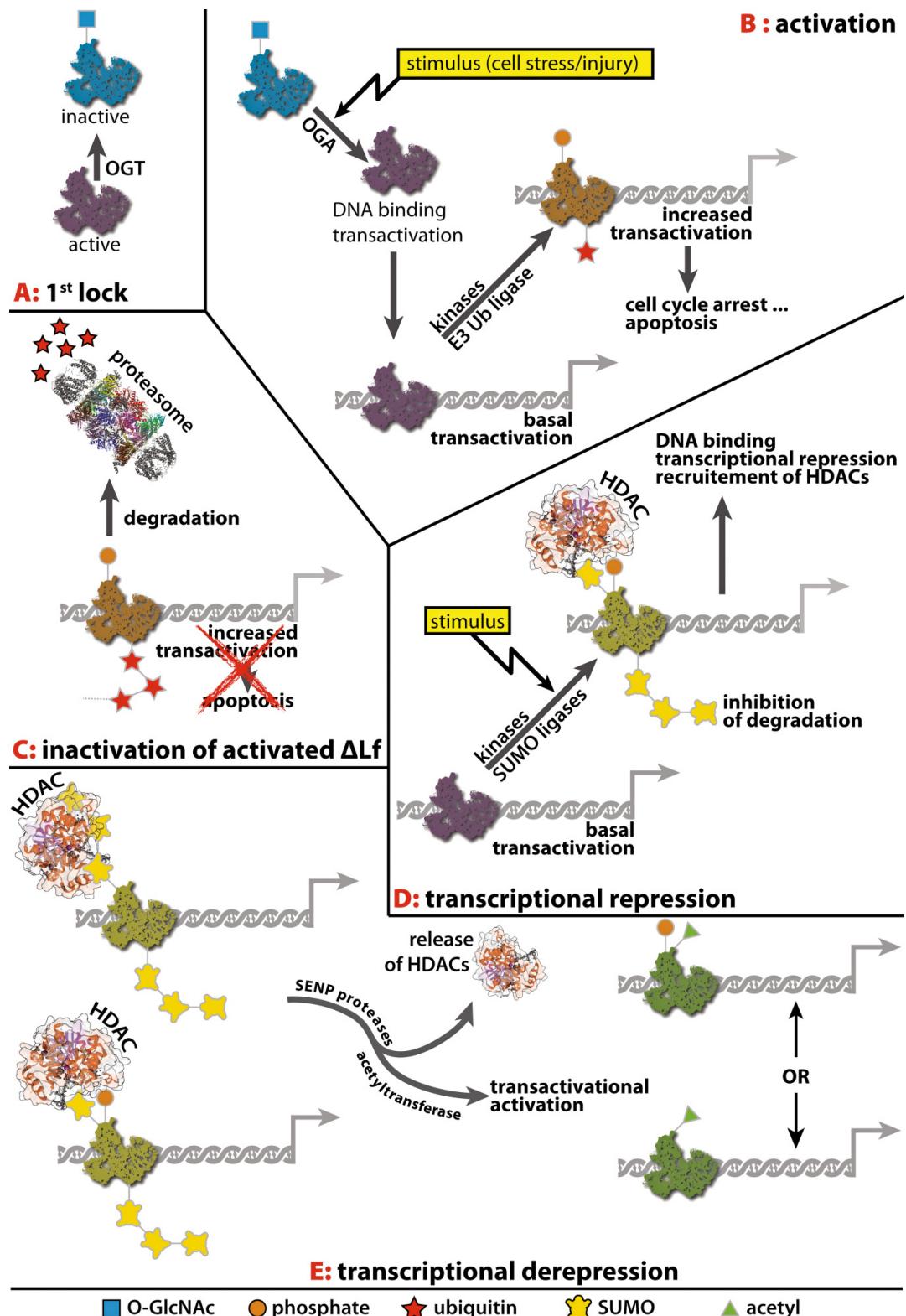
lighted 4 potential *O*-GlcNAcylation sites in the ΔLf sequence at positions 10, 227, 472, and 557 (Fig. 3). Due to its highly versatile and intricate nature, phosphorylation is discussed in its context in the corresponding paragraphs (see Regulation of ΔLf turnover and Regulation of ΔLf activity).

Likewise, several parts of the ΔLf sequence can be singled out for their importance in the properties of ΔLf . Two sequences in the C-terminal lobe are of the utmost importance for understanding the mode of action of ΔLf . The first is a bipartite putative nuclear localization signal (NLS) located at position 417–432 (RRSDTSLTWNSVKGKK) and first predicted using the PSORT II server (<http://psort.hgc.jp/form2.html>) and the subprogram nucdisc (Fig. 3). This sequence is similar to the functional motif found in Rb (Efthymiadis et al. 1997) and IL-5 (Jans et al. 1997) and has been proven to be effective in locating ΔLf to the nucleus (see Transcriptional activity of ΔLf) (Breton et al. 2004). This consensus sequence is conserved in Lfs from different species, such as the cow, mouse, pig, horse, and goat (Mariller et al. 2007). The second is a rapid degradation domain (PEST sequence) detected with the module pestfind from the EMBOSS suite (<http://emboss.bioinformatics.nl/cgi-bin/emboss/pestfind>) and is located at position 391–405 (PVLAENYKSQQSSDPDPNCVD) (Fig. 3). (see Regulation of ΔLf turnover).

Transcriptional activity of ΔLf

Several observations suggest that ΔLf possesses transcriptional activity. Such activity has already been described for

¹Positions in the Lf and ΔLf sequences are numbered from 1 at their respective N-termini, therefore the sequence numbering between Lf and ΔLf is shifted by 26 residues.



Lf. The GRRRR sequence (position 1–5) of Lf, with its 4 consecutive arginine residues, has been shown to be a NLS (Penco et al. 2001). In addition, specific binding of Lf to

DNA leading to transcriptional activation has also been demonstrated (He and Furmanski 1995). Two consensus sequences, S1 and S2, were identified: GGCACCTT(G/A)C and

Table 1. ΔLfRE sequences found in the promoters of genes directly regulated by ΔLf compared with the S1 sequence described by He and Furmanski (1995).

Promoter	ΔLfRE										Location*	Acc. No.
	1	2	3	4	5	6	7	8	9	10		
S1	G	G	C	A	C	T	—	T	A/G	C		
<i>Consensus</i>	G	G	C	A	C	T	—	T	A/G	C/G		
<i>Skp1</i>	G	G	C	A	C	T	G	T	A	C	-1067 to -1058	AC007199
<i>Bax</i>	G	G	C	A	C	T	—	T	A	T	-155 to -146	NW_001838497
<i>Fas</i>	G	G	C	A	C	T	—	A	A	C	-1013 to -1005	NT_00059
<i>DcpS</i>	A	G	C	A	C	T	—	T	G	G	-2228 to -2220	NT_033899.7
<i>SelH</i>	G	G	C	A	C	T	G	—	G	G	-2712 to -2704	NT_168190.1

*Location is relative to the transcription start.

ACTACAGTCTACA, respectively (Table 1). However, Lf is a secreted protein and, therefore, has to access the nucleus after endocytosis and before its transfer to the lysosome. Nevertheless, Lf has been shown to interact with surface nucleolin and partially reaches the nucleus after endocytosis (Legrand et al. 2004). The GRRR motif is absent from the ΔLf sequence but the bipartite NLS (position 417–432) is present. Immunostaining and confocal microscopy have shown that the subcellular distribution of ΔLf is mainly cytoplasmic but also nuclear (Breton et al. 2004). A ΔLf -GFP fusion protein has also been used to demonstrate the nuclear localization of ΔLf (Liu et al. 2003; Goldberg et al. 2005). Furthermore, disruption of both basic amino acid sites in the bipartite NLS (⁴¹⁷R_{RS}DTSLTWNSVKGKK⁴³²) abolishes the ΔLf nuclear traffic (Mariller et al. 2007). Taken together, the facts that ΔLf shows nuclear localization and is the intracellular counterpart of Lf that is able to specifically bind DNA means that ΔLf could exhibit transcriptional activity. This has been demonstrated several times with separate approaches.

Previous studies by DNA arrays of gene expression have not revealed any significant modification induced by ΔLf expression using cell lines such as mouse embryo fibroblasts, macrophages, or colon carcinoma (Co26) cells (Goldberg et al. 2005). However, such observations do not disprove the ability of ΔLf to exhibit transcriptional activity when expressed in the right cellular context and (or) in a homologous system.

ΔLf expression was shown to induce cell cycle arrest of stable inducible MCF7 cells (Breton et al. 2004). Using gene expression macroarrays with 23 different genes involved in the regulation of G1/S phase progression, S phase kinase-associated protein (Skp1) overexpression was observed. This overexpression was shown to be directly under the control of transcriptional activation by ΔLf via a ΔLf response element (ΔLfRE) very similar to the S1 sequence described for Lf (Table 1) (Mariller et al. 2007). Skp1 is a highly conserved ubiquitous eukaryotic protein belonging to the SCF complex (Bai et al. 1996). The latter has 4 components including Skp1, Cullin, Rbx1, and an F-box protein acting as a receptor for target proteins, whilst Skp1 acts as an adaptor between Cullin and the variable F-box proteins. SCF controls the ubiquitination of several cell cycle regulators, and at the G1/S transition, it is involved in the recruitment of cyclin E, cyclin A, p21, and p27, leading to their degradation by the proteasome (Zhang et al. 1995; Carrano et al. 1999; Tsvetkov et al. 1999). All these observations underline the key role played

by Skp1 at both the G1/S and G2/M transitions and corroborate the involvement of ΔLf in the regulation of cell cycle progression.

From this illustration of the transcriptional capability of ΔLf , several approaches were used to find other genes that are under its direct control. As ΔLf overexpression leads to cell cycle arrest in S phase (Breton et al. 2004) and to cell death, transcriptional regulation of the apoptosis genes by ΔLf has also been investigated. An in silico search has highlighted 2 ΔLfRE in *Bax* and *Fas* gene promoters and both of them have been proven to be directly regulated by ΔLf via their ΔLfRE (Table 1) leading to alteration of the Bax/Bcl-xL and Bax/Bcl-2 ratios. Consequently, ΔLf , by modifying the balance between Bcl-2 members at the mitochondrial membrane, leads to cytochrome *c* release and in turn to the cleavage of caspase-9. This demonstrates the proapoptotic effect of ΔLf by the activation of the intrinsic mitochondria death pathway (S. Hardivillé, personal communication 2011).

Proteomic approaches were also useful to find genes transactivated by ΔLf . A differential proteomic approach using 2-D gel electrophoresis/mass spectrometry with HEK 293 cells transiently expressing ΔLf has revealed the direct transactivation of the scavenger decapping enzyme (*DcpS*) gene (Mariller et al. 2009). The promoter of *DcpS* contains a sequence 80% identical to the S1 sequence previously described by He and Furmanski (1995) (Table 1). *DcpS* is a nucleocytoplasmonic shuttling protein that is predominantly observed in the nucleus (Shen et al. 2008), and its scavenger decapping enzyme activity is involved in cap nucleotide metabolism. It catalyzes the release of m7GMP ensuring that no excess unhydrolyzed cap accumulates and positively modulates translation and cap-proximal intron splicing by controlling eIF4E (Liu et al. 2004) and Cbp20 (Shen et al. 2008) availability. It may also perform the cap removal from aberrant nuclear transcripts and thereby may improve the quality control of mRNA biogenesis (Frischmeyer et al. 2002). The expression of *DcpS* is crucial for ensuring mRNA turnover and quality.

Proteomic studies have been carried out using stable isotope labeling by amino acids in cell culture (SILAC). This methodology allows a much more exhaustive determination of protein expression profile modifications. Experiments were carried out using MDA-MB-231 cells stably expressing ΔLf . Among the proteins that are under- or overexpressed in response to ΔLf expression, the selenoprotein H (SelH) particularly stands out and its gene promoter presents a ΔLfRE that is directly responsible for its overexpression (Table 1). SelH is 1 of the 25 so far identified selenoproteins with no

sequence homology to functionally characterized proteins. It is a thioredoxin fold-like protein in which a conserved CXXU motif (cysteine separated by 2 other residues from selenocysteine) corresponds to the CXXC motif in thioredoxins (Novoselov et al. 2007).

The ΔLf RE

The first evidence of Lf interacting in vivo with DNA was provided in 1992 (Garré et al. 1992). Its transcriptional activity was demonstrated by the inhibition of granulocyte macrophage colony-stimulating factor gene expression (Penco et al. 1995) and the use of a reporter gene transfection assay (He and Furmanski 1995; Penco et al. 1995). The specificity of Lf DNA binding was established and 3 consensus sequences were highlighted, 2 of them being functional in a reporter gene transfection assay (He and Furmanski 1995). More recently, Lf has been shown to activate the transcription of the *IL1- β* gene synergistically with phorbol 12-myristate 13-acetate activation (Son et al. 2002). Despite the uncertainty about Lf nuclear localization, these response elements for Lf were used as framework to investigate the promoters of genes under the control of ΔLf .

Response elements are overriding keys to gene regulation. Their size is usually between 4 and 20 nucleotides and they can be located near the initiation site or further upstream (Pan et al. 2010). ΔLf RE are not exceptional in this respect. Previous studies on the transcriptional activity of ΔLf have permitted the identification of a consensus sequence for the ΔLf RE (Table 1), GGCACCTT(A/G)(C/G), which is very close to the S1 sequence described by He and Furmanski (1995) for Lf. The base positions 1, 2, and 5 (G,G, and C, respectively) are essential for transactivation with position 1 being less critical. These positions are well conserved in the consensus ΔLf RE with the exception of position 1 in the *DcpS* ΔLf RE where G is substituted by A (Table 1). Deletions of the nucleotides from positions 4 to 9 abolished transactivation by ΔLf with the *Skp1*, *Fas*, and *DcpS* gene promoters but not for the *Bax* gene promoter, suggesting the existence of another ΔLf RE in this case (Mariller et al. 2007; S. Hardivillé, personal communication 2011). The direct interaction of ΔLf with the ΔLf RE was confirmed for *Skp1*, *Bax*, *Fas*, and *DcpS* using ChIP experiments (Mariller et al. 2007, 2009; S. Hardivillé, personal communication 2011).

Two potential DNA-binding domains (DBD) have been proposed for Lf, mainly due to their strong concentration of positive charges. The first is located at the N-terminal end of the protein and consists of the first 5 residues (GRRRR) and the C-terminal end of helix 1 (residues 27–30) (van Berkel et al. 1997; Baker and Baker 2009). This is supported by the demonstration that the N-terminal 98 amino acids of the protein are for the transactivation of the *IL-1 β* gene (Son et al. 2002). The second potential DBD could involve the interlobe region of the protein, which is also rich in positive charges (Baker et al. 2002). However, the positive charge of the N-terminal end is greatly reduced in the case of ΔLf and the deletion of the first 14 residues of ΔLf does not result in the loss of its transcriptional activity (Hardivillé et al. 2010).

The vast proportion of transcription factors act as dimers and bind a response element split into 2 hemi-sites (Pan et

al. 2010), but currently, there is no evidence for such a ΔLf RE organization or for ΔLf acting as a dimer.

Regulation of ΔLf expression

ΔLf expression provokes antiproliferative effects and induces cell cycle arrest in S phase. It is also at the crossroads between cell survival and cell death. Since ΔLf has several crucial target genes and may act as a tumor suppressor, modifications in its concentration or activity may have marked consequences for cell homeostasis and its transcriptional activity should be strongly regulated. Indeed ΔLf expression is regulated at each step from transcription, via the processing of its mRNA to translation. In addition, when ΔLf is finally expressed, the control of its activity and stability is realized by the interplay between numerous PTMs.

Regulation at the transcriptional level

Potential constitutive and inducible regulatory elements are both present at the P2 promoter/enhancer region. Comparison of these elements with those present at the P1 promoter/enhancer region suggests that the *Lf* and ΔLf promoters are differentially transactivated (Teng et al. 2002). The upstream putative regulatory sequences of the P2 promoter were identified by Siebert and Huang (1997) and confirmed by the team of Cristina Teng (Liu et al. 2003). An inverted TATA box is located at -29 nucleotides from the initiation start site. Putative regulatory elements specific to mammals have been identified including an *Sp1* consensus sequence, a glucocorticoid response element (GRE), an enhancer binding sequence already found in *NF- κ B* genes, a GA box found in the *c-Myc* gene, a NF-E2 box, a CCAAT-binding protein site, an IRF (interferon regulatory factor (IRF) binding site, and binding sequences for transcription factors such as XRE (xenobiotic response element), 1/2 HRE (hormone response element half site), and AML-1a (acute myeloid leukemia-1). Three functional Ets response elements were also found in the vicinity of the TATA box (Liu et al. 2003). Potential *NF- κ B/cRel* response elements were found on both *Lf* promoters but only those present upstream of the P1 promoter were functional, confirming that the P2 promoter is not activated under inflammatory conditions (Hoedt et al. 2010). Liu et al. (2003) studied the differential expression of CAT reporter genes following activation by either P1 or P2 promoter/enhancer regions and demonstrated that Lf and ΔLf expressions are regulated differently in different cell lines, the P2 promoter being highly active in lymphoid cells.

Evaluation of ΔLf gene expression levels required discrimination between Lf and ΔLf mRNAs and has been realized using either RT-PCR (Siebert and Huang 1997; Liu et al. 2003) or TaqMan qPCR (Hoedt et al. 2010). It has been shown that ΔLf transcripts are expressed in all normal tissues suggesting a ubiquitous expression (Siebert and Huang 1997). Indeed, these transcripts are found in varying amounts in the mammary gland, ovaries, uterus, placenta, testis, prostate, adult and fetal liver, small intestine, pancreas, thymus, leukocytes, kidney, spleen, fetal and adult brain, lungs, and skeletal muscle (Siebert and Huang 1997). Moreover, ΔLf mRNA is overexpressed in leukocytes and lymphocytes by nearly 300- and 150-fold, respectively, compared with normal human mammary gland (Hoedt et al. 2010). In many normal

tissues, Lf mRNAs coexist with Δ Lf mRNAs but are expressed at different levels (Siebert and Huang 1997; Benäïssa et al. 2005; Goldberg et al. 2005). In addition to a selective activation by transcription factors of each promoter, differential expression of both isoforms might also be due to a different level of promoter methylation. Thus, the P1 promoter was found to be undermethylated in numerous tissues but more specifically in leukocytes, spleen, and kidney. The P2 promoter was found to be hypermethylated in colon, lung, muscle, heart, spleen, and kidney while it was hypomethylated in normal leukocytes (Teng et al. 2004) confirming the data from Hoedt et al. (2010). Therefore, it is likely that transcription at P2 is induced by different mechanisms in different cellular environments leading to a modulation of its transcriptional activity.

Since Δ Lf is a potential antitumor agent, its level of expression has also been determined in cases of neoplastic transformation. Δ Lf transcripts have not been detected or not detected at a feeble level in tumor cells whereas they were present in all normal cells (Siebert and Huang 1997; Liu et al. 2003). In human breast cancer biopsies, Δ Lf expression was downregulated (Benäïssa et al. 2005) while it was reported that Δ Lf was equally expressed in lung adenocarcinoma and adjacent healthy tissues (Goldberg et al. 2005). Thus, Δ Lf expression is regulated differently depending on cell type and the normal or malignant phenotype of these tissues. This downregulation was also described for Lf suggesting that there is a global deregulation of *Lf* gene expression in the case of cancer. Downregulation or silencing of the expression of many genes is a characteristic of tumor suppressor genes and is mainly due to hypermethylation of their promoters. Thus, the aberrant DNA methylation of CpG islands is involved in the extinction of the expression of anti-oncogenes (Herman and Baylin 2003). Therefore, the degree and location of the methylation sites of the *Lf* gene promoters have been studied. Teng et al. (2004) studied the methylation profile of 42 CpG sites within the P1 and P2 promoter regions and showed that they are differentially methylated among tissues. Moreover, they found that most CpG sites were methylated at a higher level in leukemic cells from leukemia patients compared with normal leukocytes from healthy volunteers. In addition to gene silencing, the *Lf* gene is present within the 3q21 chromosomal region that undergoes various rearrangements in cases of cancer and particularly in cases of leukemia (Testoni et al. 1999; Lahortiga et al. 2004; McGrattan et al. 2010).

Δ Lf may therefore serve as a tumor cell marker by virtue of its absence (Siebert and Huang 1997; Liu et al. 2003). A study conducted in our laboratory reported a positive correlation between the expression profiles of Δ Lf mRNA in 99 biopsies of mammary gland cancer and the survival of patients. This retrospective study has established a threshold beyond which no recurrence was diagnosed and this after a follow-up of 4000 days. Thus, Δ Lf expression might be a good prognosis parameter, high concentrations being associated with longer overall survival (Benäïssa et al. 2005).

Regulation at the posttranscriptional level

Posttranscriptional regulation may occur to control Δ Lf expression. miRNAs are posttranscriptional regulators that bind to complementary sequences on target mRNA tran-

scripts, usually resulting in either translational repression or target degradation leading to gene silencing (Ambros 2004). They have a widely recognized role in development, differentiation, growth, apoptosis, and stress responses (Baltimore et al. 2008; Carthew and Sontheimer 2009; Garofalo et al. 2010; Melino and Knight 2010). Recently, Liao et al. (2010) reported that miR-214 downregulates endogenous expression of Lf by binding to the 3'-untranslated region (UTR) of the Lf mRNA. Since Lf and Δ Lf transcripts both possess the same 3'UTR sequence, we are currently investigating whether miR-214 may also target Δ Lf 3'UTR leading to the downregulation of its expression. Moreover, a growing number of studies indicate that deregulated miRNAs have a key role in cancer development, acting both as oncogenes and as tumor suppressor genes. It will be interesting in parallel to assay Δ Lf expression level in the presence of miR-214 and to investigate the miR-214 expression level in tumor cells compared with their normal counterparts. An increased expression of miR-214 expression might correspond to a second level of regulation of Δ Lf expression.

Regulation at the translational level

A third level of regulation may appear during translation. Translational regulation is often controlled by short sequence elements in the UTRs of mRNA. One such element is the upstream open reading frame(uORF) that is defined by a start codon in the 5' UTR that is out-of-frame with the main coding sequence. uORFs influence the protein expression of thousands of mammalian genes (Calvo et al. 2009). They occur mainly in oncogenes and genes involved in the control of cellular growth and differentiation (Morris and Geballe 2000) and correlate with significantly reduced protein expression from the downstream ORF known as the primary ORF (pORF) (Kozak 2002). The Δ Lf transcript effectively possesses 2 open reading frames, uORF and a pORF. The uORF present in exon 1 β is immediately followed by a stop codon triggering translation at pORF in exon 2. Because eukaryotic ribosomes usually load on the 5' cap of mRNA and scan for the presence of the first AUG start codon, uORFs disrupt the efficient translation of the pORF, possibly explaining why Δ Lf is so feebly expressed.

Regulation of Δ Lf turnover

Δ Lf possesses a degradation motif (PEST; ³⁹¹KSQQSSDPDPNCVD⁴⁰⁴) at the C-terminus that is phylogenetically conserved among Lfs (Hardivillé et al. 2010). Phosphorylation at serine or threonine residues and ubiquitination, often of the flanking lysine residues, usually triggers proteasomal degradation (Rogers et al. 1986; Rechsteiner and Rogers 1996). Thus, the difficulty to produce Δ Lf even when overexpressed suggests that it is very unstable and that the PEST sequence is fully functional. We therefore investigated Δ Lf ubiquitination and showed that lysine residues positioned at 379 and 391 were both targeted by E3-Ub ligases with K³⁷⁹ being the main polyubiquitin acceptor site (Fig. 3). Moreover, mutation of these 2 lysine residues or of the 3 serine residues (S³⁹², S³⁹⁵, and S³⁹⁶) within the PEST motif strongly increased the Δ Lf half-life. We also showed that the nearly adjacent serine residues are equivalent targets of kinases (Hardivillé et al. 2010).

PEST sequences usually allow the interaction of proteins with the SCF (Skp1/Cullin-1/F-box) Ub–ligase complex, which directs their polyubiquitylation at the G1/S transition. The antiproliferative activity of ΔLf leading to cell cycle arrest at S phase (Breton et al. 2004) and the upregulation of Skp1 expression might both be downregulated by the activation of the ΔLf PEST sequence prior to its degradation. Unexpectedly, the PEST mutant still shows a residual polyubiquitylation suggesting that other domains of ΔLf may be involved in this process (Hardivillé et al. 2010). Analysis of the primary sequence of ΔLf points out the presence of 2 putative D-boxes, D-Box1 ($^{108}\text{RPFL}^{111}$) and D-Box2 ($^{475}\text{RSNL}^{478}$), conserved among species. These sequences are known to interact with the APC/C Ub–ligase complex that directs protein degradation at G2/M (Linder et al. 1989; Hames et al. 2001). Therefore, if ΔLf still remains at the end of the S phase, activation of its D-boxes may lead to its degradation at G2/M. Further work has to be done to confirm the functionality of these 2 D-boxes.

Increasing evidence has linked *O*-GlcNAcylation to a lower proteasomal susceptibility and in the next paragraph we will show that the stability of ΔLf , like that of many transcription factors (Han and Kudlow 1997; Cheng and Hart 2000; Yang et al. 2001), is controlled by this modification.

Regulation of ΔLf activity

Most of the proteins that are translated from mRNA undergo chemical modifications known PTMs before becoming functional in different body cells. So far, more than 200 PTMs have been characterized that modulate the function of most eukaryote proteins by altering their activity state, localization, turnover, and interactions with other proteins. Therefore, we investigated which PTMs may control ΔLf activity and stability and found that *O*-GlcNAcylation, phosphorylation, and ubiquitination were involved (Hardivillé et al. 2010). Recently, we have started to investigate other PTMs such as SUMOylation and acetylation. Transcription factors are often regulated by combinations of different PTMs that might act as a molecular barcode (Benayoun and Veitia 2009). We are currently investigating the crosstalk between them to decrypt the ΔLf code responsible for the control of its activity. The ΔLf code should both provide a further layer of regulation of ΔLf stability and activity and lead to highly specific cellular outputs.

First of all, we evaluated the possibility that *O*-GlcNAcylation may act as a potent regulator. *O*-linked- β -N-*O*-GlcNAc (*O*-GlcNAc) is a dynamic and reversible intracellular PTM controlled by *O*-GlcNAc transferase (OGT) that attaches a single *N*-acetylglucosamine to serine or threonine residues and *N*-acetyl- β -D-glucosaminidase (OGA) that removes it (Hart et al. 1989). In many *O*-GlcNAcylated proteins, this modification competes directly with phosphate for the same or nearby serine or threonine residues (Comer and Hart 2001; Kamemura and Hart 2003; Yang et al. 2006). This *O*-GlcNAc/P interplay leads to rapid response mechanisms controlling development, cell cycle, apoptosis, and cell architecture via the control of key actors involved in cellular processes such as transcription, translation, cell signalling, cell trafficking, or protein degradation (Kamemura and Hart 2003; Yang et al. 2006). Alterations of the *O*-GlcNAc profile appear in neurodegenerative disorders, diabetes, and cancer,

and in the case of breast cancer, a close relationship between *O*-GlcNAc and tumorigenesis may exist (reviewed in Lefebvre et al. (2010)). Recently, we showed using glycosylation mutants, inhibitors and (or) activators of the hexosamine biosynthesis pathway that ΔLf nucleocytoplasmic traffic was not altered when *O*-GlcNAc content was modified in cells and that *O*-GlcNAcylation negatively regulates ΔLf transcriptional activity while phosphorylation activates it (Hardivillé et al. 2010). The directed mutagenesis of the *O*-GlcNAc sites led to a constitutively active ΔLf isoform ($\Delta\text{Lf-M4}$) with increased pro-apoptotic effects compared with the wild type. Using a series of different glycosylation mutants in which only 1 *O*-GlcNAcylation site was preserved, we showed that *O*-GlcNAcylation at Ser¹⁰ was crucial and negatively regulated ΔLf activity. Using a ChIP assay, we showed that *O*-GlcNAcylation of ΔLf inhibits its DNA binding. Deglycosylation leads to DNA binding and transactivation of the ΔLfRE present in the *Skp1* promoter at a basal level. Basal transactivation was markedly enhanced when phosphorylation was mimicked at Ser¹⁰ by aspartate. We next demonstrated that *O*-GlcNAcylation inhibits ΔLf polyubiquitination and increases its half-life. Thus, this PTM on the Ser¹⁰ residue controls ΔLf turnover. Moreover, using Re-ChIP assays, we showed that the ΔLf transcriptional complex binds to the ΔLfRE and is phosphorylated and (or) ubiquitinated, suggesting that ΔLf transcriptional activity and degradation are concomitant events. Collectively, our results indicate that reciprocal occupancy of Ser¹⁰ by either *O*-phosphate or *O*-GlcNAc coordinately regulates ΔLf stability and transcriptional activity. Taken together, these results suggest that in cells, the pool of ΔLf may exist under a stable but not functional *O*-GlcNAc isoform (Fig. 4A). Since the level of *O*-GlcNAc changes during the cell cycle or is altered in tumorigenesis, deglycosylated ΔLf will become the target of kinases leading to its activation and polyubiquitination (Fig. 4B). ΔLf will then transactivate its target genes prior to being degraded by the proteasome (Fig. 4C)(Hardivillé et al. 2010).

We further investigated whether ΔLf was modified by SUMOylation. And so far, 4 sites have been detected at positions 13, 308, 361, and 391. SUMO stands for small ubiquitin-like modifier, a ubiquitin family member (Johnson 2004; Kirkin and Dikic 2007). SUMOylation is a widespread mechanism for rapid and reversible changes in protein function. It is highly dynamic involving SUMO-conjugation and SUMO-deconjugation pathways. SUMOylation mainly targets nuclear factors and regulates numerous cellular processes such as nuclear transport, transcription, and DNA repair (reviewed in Gareau and Lima 2010 and Hannoun et al. 2010). In most cases SUMOylation triggers transcriptional repression. The covalent attachment of SUMO to transcription factors inhibits transcription by recruiting transcriptional corepressors, such as histone deacetylases (Gill 2005). Since SUMO and ubiquitin target the same lysine residues, SUMOylation was also predicted to stabilize proteins by competing with ubiquitin (Desterro et al. 1998; Lin et al. 2003). The dialogue between SUMO and other modifications is emerging as a common mechanism that allows control of the transcriptional activity of transcription factors (Verger et al. 2003). SUMOylation usually competes with ubiquitination, phosphorylation, and acetylation. Ubiquitination/SUMOylation

and SUMOylation/acetylation are mutually exclusive, whereas SUMOylation/phosphorylation can be agonistic or antagonistic depending on the substrates.

ΔLf also possesses a putative acetylation site at its N-terminus with Lys³ as the target for acetyltransferases. Acetylation is also a dynamic process that mainly contributes to transcriptional activation (Glozak et al. 2005; Sadoul et al. 2008; Yang and Seto 2008). Interestingly, among the 4 SUMO sites, 2 are located in areas that are targeted by other modifications (Fig. 3). At the N-terminus, the putative acetylation site is adjacent to the *O*-GlcNAcylation/phosphorylation site and to the SUMO-1 site (Fig. 3). Therefore, all 3 regulatory sites are in close proximity. The crosstalk between these sites could constitute the ΔLf code responsible for the control of the transactivation of ΔLf target genes. We know from our results and from the literature that acetylation and phosphorylation both lead to transcriptional activation whereas *O*-GlcNAcylation and SUMOylation repress it (Fig. 4). So we hypothesize that this region might be part of the ΔLf transactivation domain. On the other hand, the C-terminal SUMO-4 site is in the heart of the PEST sequence. SUMO is known to compete with ubiquitination; therefore, Lys³⁹¹ is either ubiquitinated or SUMOylated (Fig. 3) and SUMOylation would positively regulate stability. Therefore, at the PEST motif, a crosstalk between phosphorylation/SUMOylation/ubiquitination could constitute a ΔLf code responsible for the control of ΔLf stability. These agonistic and (or) antagonistic modifications, illustrated in Fig. 4, may contribute to the establishment of finely regulated mechanisms governing ΔLf transcriptional activity and stability depending on the type of target genes and cellular homeostasis.

In summary, various aspects regulate ΔLf expression and activity including transcriptional, posttranscriptional, translational, and posttranslational events.

Acknowledgment

This investigation was supported in part by the CNRS UMR 8576 (Unité de Glycobiologie Structurale et Fonctionnelle), the Institut Fédératif de Recherche No. 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, Université de Lille Nord de France) for reviewing this manuscript. Molecular graphic images were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR001081). Solvent-excluded molecular surfaces were created with the help of the MSMS package (Sanner et al. 1996).

References

- Aisen, P., and Leibman, A. 1972. Lactoferrin and transferrin: a comparative study. *Biochim. Biophys. Acta*, **257**(2): 314–323. PMID:4336719.
- Ambros, V. 2004. The functions of animal microRNAs. *Nature*, **431**(7006): 350–355. doi:10.1038/nature02871. PMID:15372042.
- Anderson, B.F., Baker, H.M., Dodson, E.J., Norris, G.E., Rumball, S.V., Waters, J.M., and Baker, E.N. 1987. Structure of human lactoferrin at 3.2-Å resolution. *Proc. Natl. Acad. Sci. U.S.A.* **84**(7): 1769–1773. doi:10.1073/pnas.84.7.1769. PMID:3470756.
- Anderson, B.F., Baker, H.M., Norris, G.E., Rice, D.W., and Baker, E.N. 1989. Structure of human lactoferrin: crystallographic structure analysis and refinement at 2.8 Å resolution. *J. Mol. Biol.* **209**(4): 711–734. doi:10.1016/0022-2836(89)90602-5. PMID:2585506.
- Bai, C., Sen, P., Hofmann, K., Ma, L., Goebel, M., Harper, J.W., and Elledge, S.J. 1996. SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. *Cell*, **86**(2): 263–274. doi:10.1016/S0092-8674(00)80098-7. PMID:8706131.
- Baker, E.N. 2005. Lactoferrin: a multi-tasking protein par excellence. *Cell. Mol. Life Sci.* **62**(22): 2529–2530. doi:10.1007/s00018-005-5366-y. PMID:16261258.
- Baker, E.N., and Baker, H.M. 2009. A structural framework for understanding the multifunctional character of lactoferrin. *Biochimie*, **91**(1): 3–10. doi:10.1016/j.biochi.2008.05.006. PMID:18541155.
- Baker, E.N., Baker, H.M., and Kidd, R.D. 2002. Lactoferrin and transferrin: functional variations on a common structural framework. *Biochem. Cell Biol.* **80**(1): 27–34. doi:10.1139/o01-153. PMID:11908640.
- Baldwin, G.S. 1993. Comparison of transferrin sequences from different species. *Comp. Biochem. Physiol. B*, **106**(1): 203–218. doi:10.1016/0305-0491(93)90028-4. PMID:8403849.
- Baltimore, D., Boldin, M.P., O'Connell, R.M., Rao, D.S., and Taganov, K.D. 2008. MicroRNAs: new regulators of immune cell development and function. *Nat. Immunol.* **9**(8): 839–845. doi:10.1038/ni.f.209. PMID:18645592.
- Benaïssa, M., Peyrat, J.-P., Hornez, L., Mariller, C., Mazurier, J., and Pierce, A. 2005. Expression and prognostic value of lactoferrin mRNA isoforms in human breast cancer. *Int. J. Cancer*, **114**(2): 299–306. doi:10.1002/ijc.20728. PMID:15543612.
- Benayoun, B.A., and Veitia, R.A. 2009. A post-translational modification code for transcription factors: sorting through a sea of signals. *Trends Cell Biol.* **19**(5): 189–197. doi:10.1016/j.tcb.2009.02.003. PMID:19328693.
- Breton, M., Mariller, C., Benaïssa, M., Caillaux, K., Browaeys, E., Masson, M., et al. 2004. Expression of delta-lactoferrin induces cell cycle arrest. *Biometals*, **17**(3): 325–329. doi:10.1023/B:BIOM.0000027712.81056.13. PMID:15222485.
- Calvo, S.E., Pagliarini, D.J., and Mootha, V.K. 2009. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc. Natl. Acad. Sci. U.S.A.* **106**(18): 7507–7512. doi:10.1073/pnas.0810916106. PMID:19372376.
- Carrano, A.C., Eytan, E., Hershko, A., and Pagano, M. 1999. SKP2 is required for ubiquitin-mediated degradation of the CDK inhibitor p27. *Nat. Cell Biol.* **1**(4): 193–199. doi:10.1038/12013. PMID:10559916.
- Carthew, R.W., and Sontheimer, E.J. 2009. Origins and mechanisms of miRNAs and siRNAs. *Cell*, **136**(4): 642–655. doi:10.1016/j.cell.2009.01.035. PMID:19239886.
- Cheng, X., and Hart, G.W. 2000. Glycosylation of the murine estrogen receptor-alpha. *J. Steroid Biochem. Mol. Biol.* **75**(2–3): 147–158. doi:10.1016/S0960-0760(00)00167-9. PMID:11226831.
- Comer, F.I., and Hart, G.W. 2001. Reciprocity between *O*-GlcNAc and *O*-phosphate on the carboxyl terminal domain of RNA polymerase II. *Biochemistry*, **40**(26): 7845–7852. doi:10.1021/bi0027480. PMID:11425311.
- Desterro, J.M., Rodriguez, M.S., and Hay, R.T. 1998. SUMO-1 modification of IκBα inhibits NF-κB activation. *Mol. Cell*,

- 2(2): 233–239. doi:10.1016/S1097-2765(00)80133-1. PMID: 9734360.
- Efthymiadis, A., Shao, H., Hübner, S., and 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**(35): 22134–22139. doi:10.1074/jbc.272.35.22134. PMID: 9268357.
- Frischmeyer, P.A., van Hoof, A., O'Donnell, K., Guerrero, A.L., Parker, R., and Dietz, H.C. 2002. An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science*, **295**(5563): 2258–2261. doi:10.1126/science.1067338. PMID:11910109.
- Gareau, J.R., and Lima, C.D. 2010. The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nat. Rev. Mol. Cell Biol.* **11**(12): 861–871. doi:10.1038/nrm3011. PMID:21102611.
- Garofalo, M., Condorelli, G.L., Croce, C.M., and Condorelli, G. 2010. MicroRNAs as regulators of death receptors signaling. *Cell Death Differ.* **17**(2): 200–208. doi:10.1038/cdd.2009.105. PMID: 19644509.
- Garré, C., Bianchi-Scarrá, G., Sirito, M., Musso, M., and Ravazzolo, R. 1992. Lactoferrin binding sites and nuclear localization in K562 (S) cells. *J. Cell. Physiol.* **153**(3): 477–482. doi:10.1002/jcp. 1041530306. PMID:1447310.
- Gill, G. 2005. Something about SUMO inhibits transcription. *Curr. Opin. Genet. Dev.* **15**(5): 536–541. doi:10.1016/j.gde.2005.07. 004. PMID:16095902.
- Glozak, M.A., Sengupta, N., Zhang, X., and Seto, E. 2005. Acetylation and deacetylation of non-histone proteins. *Gene*, **363**: 15–23. doi:10.1016/j.gene.2005.09.010. PMID:16289629.
- Goldberg, G.S., Kunimoto, T., Alexander, D.B., Suenaga, K., Ishidate, F., Miyamoto, K., et al. 2005. 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**(1): 57–64. doi:10.2174/ 1573406053402532. PMID:16789886.
- Gupta, R., and Brunak, S. 2002. Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac. Symp. Biocomput.* 310–322. PMID:11928486.
- Hames, R.S., Wattam, S.L., Yamano, H., Bacchieri, R., and 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**(24): 7117–7127. doi:10.1093/emboj/ 20.24.7117. PMID:11742988.
- Han, I., and Kudlow, J.E. 1997. Reduced O glycosylation of Sp1 is associated with increased proteasome susceptibility. *Mol. Cell. Biol.* **17**(5): 2550–2558. PMID:9111324.
- Hannoun, Z., Greenhough, S., Jaffray, E., Hay, R.T., and Hay, D.C. 2010. Post-translational modification by SUMO. *Toxicology*, **278**(3): 288–293. doi:10.1016/j.tox.2010.07.013. PMID:20674646.
- Hardivillé, S., Hoedt, E., Mariller, C., Benäissa, M., and Pierce, A. 2010. O-GlcNAcylation/phosphorylation cycling at Ser10 controls both transcriptional activity and stability of delta-lactoferrin. *J. Biol. Chem.* **285**(25): 19205–19218. doi:10.1074/jbc.M109. 080572. PMID:20404350.
- Hart, G.W., Haltiwanger, R.S., Holt, G.D., and Kelly, W.G. 1989. Glycosylation in the nucleus and cytoplasm. *Annu. Rev. Biochem.* **58**(1): 841–874. doi:10.1146/annurev.bi.58.070189.004205. PMID:2673024.
- He, J., and Furmanski, P. 1995. Sequence specificity and transcriptional activation in the binding of lactoferrin to DNA. *Nature*, **373**(6516): 721–724. doi:10.1038/373721a0. PMID:7854459.
- Herman, J.G., and Baylin, S.B. 2003. Gene silencing in cancer in association with promoter hypermethylation. *N. Engl. J. Med.* **349**(21): 2042–2054. doi:10.1056/NEJMra023075. PMID:14627790.
- Hoedt, E., Hardivillé, S., Mariller, C., Elass, E., Perraudin, J.-P., and 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*, **23**(3): 441–452. doi:10.1007/s10534-010-9305-5. PMID:20155437.
- Jans, D.A., Briggs, L.J., Gustin, S.E., Jans, P., Ford, S., and Young, I.G. 1997. A functional bipartite nuclear localisation signal in the cytokine interleukin-5. *FEBS Lett.* **406**(3): 315–320. doi:10. 1016/S0014-5793(97)00293-7. PMID:9136909.
- Johnson, E.S. 2004. Protein modification by SUMO. *Annu. Rev. Biochem.* **73**(1): 355–382. doi:10.1146/annurev.biochem.73. 011303.074118. PMID:15189146.
- Kamemura, K., and 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**: 107–136. doi:10.1016/S0079-6603(03)01004-3. PMID:12882516.
- Kim, S.J., Yu, D.Y., Pak, K.W., Jeong, S., Kim, S.W., and Lee, K.K. 1998. Structure of the human lactoferrin gene and its chromosomal localization. *Mol. Cells*, **8**(6): 663–668. PMID:9895117.
- Kirkin, V., and Dikic, I. 2007. Role of ubiquitin- and Ubl-binding proteins in cell signaling. *Curr. Opin. Cell Biol.* **19**(2): 199–205. doi:10.1016/j.cel.2007.02.002. PMID:17303403.
- Kozak, M. 2002. Pushing the limits of the scanning mechanism for initiation of translation. *Gene*, **299**(1–2): 1–34. doi:10.1016/ S0378-1119(02)01056-9. PMID:12459250.
- Lahortiga, I., Vázquez, I., Agirre, X., Larrayoz, M.J., Vizmanos, J.L., Gozzetti, A., et al. 2004. Molecular heterogeneity in AML/MDS patients with 3q21q26 rearrangements. *Genes Chromosomes Cancer*, **40**(3): 179–189. doi:10.1002/gcc.20033. PMID: 15138998.
- Lefebvre, T., Dehennaut, V., Guinez, C., Olivier, S., Drougas, L., Mir, A.-M., et al. 2010. Dysregulation of the nutrient/stress sensor O-GlcNAcylation is involved in the etiology of cardiovascular disorders, type-2 diabetes and Alzheimer's disease. *Biochim. Biophys. Acta*, **1800**(2): 67–79. doi:10.1016/j.bbagen.2009.08. 008. PMID:19732809.
- Legrand, D., Vigié, K., Said, E.A., Elass, E., Masson, M., Slomianny, M.-C., et al. 2004. Surface nucleolin participates in both the binding and endocytosis of lactoferrin in target cells. *Eur. J. Biochem.* **271**(2): 303–317. doi:10.1046/j.1432-1033.2003.03929. x. PMID:14717698.
- Legrand, D., Pierce, A., Elass, E., Carpentier, M., Mariller, C., and Mazurier, J. 2008. Lactoferrin structure and functions. *Adv. Exp. Med. Biol.* **606**: 163–194. doi:10.1007/978-0-387-74087-4_6. PMID:18183929.
- Liao, Y., Du, X., and Lönnadal, B. 2010. miR-214 regulates lactoferrin expression and pro-apoptotic function in mammary epithelial cells. *J. Nutr.* **140**(9): 1552–1556. doi:10.3945/jn.110. 124289. PMID:20610637.
- Lin, X., Liang, M., Liang, Y.-Y., Brunicardi, F.C., and Feng, X.-H. 2003. SUMO-1/Ubc9 promotes nuclear accumulation and metabolic stability of tumor suppressor Smad4. *J. Biol. Chem.* **278**(33): 31043–31048. doi:10.1074/jbc.C300112200. PMID:12813045.
- Linder, P., Lasko, P.F., Ashburner, M., Leroy, P., Nielsen, P.J., Nishi, K., et al. 1989. Birth of the D-E-A-D box. *Nature*, **337**(6203): 121–122. doi:10.1038/337121a0. PMID:2563148.
- Liu, D., Wang, X., Zhang, Z., and Teng, C.T. 2003. An intronic alternative promoter of the human lactoferrin gene is activated by Ets. *Biochem. Biophys. Res. Commun.* **301**(2): 472–479. doi:10. 1016/S0006-291X(02)03077-2. PMID:12565886.
- Liu, S.-W., Jiao, X., Liu, H., Gu, M., Lima, C.D., and Kiledjian, M.

2004. Functional analysis of mRNA scavenger decapping enzymes. *RNA*, **10**(9): 1412–1422. doi:10.1261/rna.7660804. PMID:15273322.
- Mann, D.M., Romm, E., and Migliorini, M. 1994. Delineation of the glycosaminoglycan-binding site in the human inflammatory response protein lactoferrin. *J. Biol. Chem.* **269**(38): 23661–23667. PMID:8089135.
- Mariller, C., Benaïssa, M., Hardivillé, S., Breton, M., Pradelle, G., Mazurier, J., and Pierce, A. 2007. Human delta-lactoferrin is a transcription factor that enhances Skp1 (S-phase kinase-associated protein) gene expression. *FEBS J.* **274**(8): 2038–2053. doi:10.1111/j.1742-4658.2007.05747.x. PMID:17371504.
- Mariller, C., Hardivillé, S., Hoedt, E., Benaïssa, M., Mazurier, J., and Pierce, A. 2009. Proteomic approach to the identification of novel delta-lactoferrin target genes: characterization of DcpS, an mRNA scavenger decapping enzyme. *Biochimie*, **91**(1): 109–122. doi:10.1016/j.biochi.2008.07.009. PMID:18725266.
- McCombs, J.L., Teng, C.T., Pentecost, B.T., Magnuson, V.L., Moore, C.M., and McGill, J.R. 1988. Chromosomal localization of human lactotransferrin gene (LTF) by in situ hybridization. *Cytogenet. Cell Genet.* **47**(1-2): 16–17. doi:10.1159/000132496. PMID:3356163.
- McGrattan, P., Logan, A., Humphreys, M., and Bowers, M. 2010. Jumping translocation in acute monocytic leukemia (M5b) with alternative breakpoint sites in the long arm of donor chromosome 3. *Med. Oncol.* **27**(3): 667–672. doi:10.1007/s12032-009-9266-0. PMID:19629764.
- Melino, G., and Knight, R.A. 2010. MicroRNAs meet cell death. *Cell Death Differ.* **17**(2): 189–190. doi:10.1038/cdd.2009.122. PMID:20062066.
- Morris, D.R., and Geballe, A.P. 2000. Upstream open reading frames as regulators of mRNA translation. *Mol. Cell. Biol.* **20**(23): 8635–8642. doi:10.1128/MCB.20.23.8635-8642.2000. PMID:11073965.
- Novoselov, S.V., Kryukov, G.V., Xu, X.-M., Carlson, B.A., Hatfield, D.L., and Gladyshev, V.N. 2007. Selenoprotein H is a nucleolar thioredoxin-like protein with a unique expression pattern. *J. Biol. Chem.* **282**(16): 11960–11968. doi:10.1074/jbc.M701605200. PMID:17337453.
- Pan, Y., Tsai, C.-J., Ma, B., and Nussinov, R. 2010. Mechanisms of transcription factor selectivity. *Trends Genet.* **26**(2): 75–83. doi:10.1016/j.tig.2009.12.003. PMID:20074831.
- Penco, S., Pastorino, S., Bianchi-Scarrà, G., and Garrè, C. 1995. Lactoferrin down-modulates the activity of the granulocyte macrophage colony-stimulating factor promoter in interleukin-1 β -stimulated cells. *J. Biol. Chem.* **270**(20): 12263–12268. doi:10.1074/jbc.270.20.12263. PMID:7744878.
- Penco, S., Scarfi, S., Giovine, M., Damonte, G., Millo, E., Villaggio, B., et al. 2001. Identification of an import signal for, and the nuclear localization of, human lactoferrin. *Biotechnol. Appl. Biochem.* **34**(3): 151–159. doi:10.1042/BA20010038. PMID:11730482.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**(13): 1605–1612. doi:10.1002/jcc.20084. PMID:15264254.
- Rechsteiner, M., and Rogers, S.W. 1996. PEST sequences and regulation by proteolysis. *Trends Biochem. Sci.* **21**(7): 267–271. PMID:8755249.
- Rocchia, W., Sridharan, S., Nicholls, A., Alexov, E., Chiabrera, A., and Honig, B. 2002. Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: applications to the molecular systems and geometric objects. *J. Comput. Chem.* **23**(1): 128–137. doi:10.1002/jcc.1161. PMID:11913378.
- Rogers, S., Wells, R., and Rechsteiner, M. 1986. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science*, **234**(4774): 364–368. doi:10.1126/science.2876518. PMID:2876518.
- Sadoul, K., Boyault, C., Pabion, M., and Khochbin, S. 2008. Regulation of protein turnover by acetyltransferases and deacetylases. *Biochimie*, **90**(2): 306–312. doi:10.1016/j.biochi.2007.06.009. PMID:17681659.
- Sanner, M.F., Olson, A.J., and Spehner, J.C. 1996. Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers*, **38**(3): 305–320. doi:10.1002/(SICI)1097-0282(199603)38:3<305::AID-BIP4>3.0.CO;2-Y. PMID:8906967.
- Shen, V., Liu, H., Liu, S.-W., Jiao, X., and Kiledjian, M. 2008. DcpS scavenger decapping enzyme can modulate pre-mRNA splicing. *RNA*, **14**(6): 1132–1142. doi:10.1261/rna.1008208. PMID:18426921.
- Siebert, P.D., and 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**(6): 2198–2203. doi:10.1073/pnas.94.6.2198. PMID:9122171.
- Son, K.-N., Park, J., Chung, C.-K., Chung, D.K., Yu, D.-Y., Lee, K.-K., and Kim, J. 2002. Human lactoferrin activates transcription of IL-1 β gene in mammalian cells. *Biochem. Biophys. Res. Commun.* **290**(1): 236–241. doi:10.1006/bbrc.2001.6181. PMID:11779159.
- Teng, C.T., Beard, C., and Gladwell, W. 2002. Differential expression and estrogen response of lactoferrin gene in the female reproductive tract of mouse, rat, and hamster. *Biol. Reprod.* **67**(5): 1439–1449. doi:10.1095/biolreprod.101.002089. PMID:12390874.
- Teng, C., Gladwell, W., Raphiou, I., and Liu, E. 2004. Methylation and expression of the lactoferrin gene in human tissues and cancer cells. *Biometals*, **17**(3): 317–323. doi:10.1023/B:BIOM.0000027711.13818.8a. PMID:15222484.
- Testoni, N., Borsaru, G., Martinelli, G., Carboni, C., Ruggeri, D., Ottaviani, E., et al. 1999. 3q21 and 3q26 cytogenetic abnormalities in acute myeloblastic leukemia: biological and clinical features. *Haematologica*, **84**(8): 690–694. PMID:10457403.
- Tsvetkov, L.M., Yeh, K.-H., Lee, S.-J., Sun, H., and Zhang, H. 1999. p27(Kip1) ubiquitination and degradation is regulated by the SCF (Skp2) complex through phosphorylated Thr187 in p27. *Curr. Biol.* **9**(12): 661–664. doi:10.1016/S0960-9822(99)80290-5. PMID:10375532.
- van Berkel, P.H., Geerts, M.E., van Veen, H.A., Mericksay, M., de Boer, H.A., and 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**(1): 145–151. PMID:9359845.
- Verger, A., Perdomo, J., and Crossley, M. 2003. Modification with SUMO. A role in transcriptional regulation. *EMBO Rep.* **4**(2): 137–142. doi:10.1038/sj.emboj.738. PMID:12612601.
- Ward, P.P., Paz, E., and Conneely, O.M. 2005. Multifunctional roles of lactoferrin: a critical overview. *Cell. Mol. Life Sci.* **62**(22): 2540–2548. doi:10.1007/s00018-005-5369-8. PMID:16261256.
- Yang, X.-J., and Seto, E. 2008. Lysine acetylation: codified crosstalk with other posttranslational modifications. *Mol. Cell.* **31**(4): 449–461. doi:10.1016/j.molcel.2008.07.002. PMID:18722172.
- Yang, X., Su, K., Roos, M.D., Chang, Q., Paterson, A.J., and Kudlow, J.E. 2001. O-linkage of N-acetylglucosamine to Sp1 activation domain inhibits its transcriptional capability. *Proc. Natl. Published by NRC Research Press*

Pagination not final/Pagination non finale

Mariller et al.

13

- Acad. Sci. U.S.A. **98**(12): 6611–6616. doi:10.1073/pnas.111099998. PMID:11371615.
- Yang, W.H., Kim, J.E., Nam, H.W., Ju, J.W., Kim, H.S., Kim, Y.S., and Cho, J.W. 2006. Modification of p53 with O-linked N-acetylglucosamine regulates p53 activity and stability. *Nat. Cell Biol.* **8**(10): 1074–1083. doi:10.1038/ncb1470. PMID:16964247.
- Zhang, H., Kobayashi, R., Galaktionov, K., and Beach, D. 1995. p19Skp1 and p45Skp2 are essential elements of the cyclin A-CDK2 S phase kinase. *Cell*, **82**(6): 915–925. doi:10.1016/0092-8674(95)90271-6. PMID:7553852.

Abbréviations

A

A	Adénine
AAA ATPase	ATP hydrolase AAA
AACT	Amino Acid Coded mass Tagging
ADN	Acide désoxyribonucléique
ADNc	ADN complémentaire
AKT/PKB	Protein Kinase B (PKB)
ALDH18A1	Aldhéhyde deshydrogénase 18A1
alpha-cyano	Acide α -cyano-4-hydroxycinnamique
AML-1a	Acute Myeloid Leukemia-1
APC/C	Anaphase-promoting complex/cyclosome
APEX	Absolute Protein EXpression measurement
apoLf	Lactoferrine non saturée en fer
AQUA	Absolute QUAntification
Arg	Arginine
ARN	Acide ribonucléique
ARNm	ARN messager

B

Bax	Bcl-2-associated X protein
BCG	Bacille de Calmette et Guérin
Bcl-2	B-cell lymphoma 2

C

C	Carbone
C	Cytosine
C/EBP (α et ϵ),	CCAAT Enhancer Binding Protein
CAT	Chloramphénicol acétyl transférase
CBC	Cap-binding complex
CCR5	C-C chemokine receptor type 5
CCT2	Chaperonin-containing T-complex protein 2
CD14	Cluster of differentiation 14
CD4	Cluster of differentiation 4
Cdc	Cell division cycle
Cdk	Cyclin dependent kinase
cdki	Cdk inhibitor
CDP/cut	CCAAT displacement protein
ChIP	Chromatin ImmunoPrecipitation
ChIP-Seq	Chromatin ImmunoPrecipitation Sequencing
CLHP	Chromatographie liquide haute pression
CMV	Cytomégavirus

c-Myc	Oncogène cellulaire myelocytose
COUP	Chicken ovalbumin upstream promoter
CRE	cAMP response element
CREB	cAMP responsive element-binding protein
Ct	C-terminal
CTD	Carboxy-terminal domain
Cul	Culline
CXCR4	C-X-C chemokine receptor type 4

D

ΔLf	Delta-lactoferrin
ΔLfRE	Delta-lactoferrin response element
d.i	diamètre interne
Da	Dalton
DBD	DNA binding domain
D-Box	Destruction box
Dcp2	mRNA-decapping enzyme 2
DcpS	scavenger mRNA decapping enzyme
DC-SIGN	Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin
DHB	Acide 2,5-dihydroxybenzoïque
DIGE	Differential In Gel Electrophoresis
DMEM	Dubelco Modified Eagle's medium
Dox	Doxycycline
DTT	Dithiothréitol
DUB	Desubiquitinase

E

E.U.	Endotoxin Unit
E2F	Elongation factor
ECD	Capture d'électron
EGF	Epidermal growth factor
eIF4E	Eukaryotic translation initiation factor 4E
ELISA	Enzyme-Linked Immunosorbent Assay
EMSA	Electrophoretic Mobility Shift Assay
ERE	Estrogen response element
ESI	Electrospray
ET-1	Endothéline-1
ETD	Dissociation par transfert d'électron
Ets	E-twenty six

F

FBP	F-box protein
FITC	Fluorescein isothiocyanate
FT-ICR	Analyseur à résonance ionique cyclotronique par transformée de Fourier

G

G	Guanine
GAPDH	Glycéraldéhyde deshydrogénase
G-CSF	Granulocyte-Colony Stimulating factor
GFAT	Glutamine:fructose-6-phosphate aminotransférase
GFP	Green fluorescent protein
GGH	Gamma glutamyl hydrolase
GlcNAc	N-acétylglucosamine
GlcNAcylation	O-N-acétylglucosaminylation
GlcNH ₂	Glucosamine
Glu-C	<i>Staphylococcus aureus Protease V8</i>
GRE	Glucocorticoid response element

H

h	Heure
H2B	Histone 2B
HAT	Histone acétyltransférase
HIC-1	Hypermethylated In Cancer 1
HIF-1	Hypoxia Inducible Factor-1
HNPL	Heterogeneous Nuclear RibonucleoProtein L
holoLf	Lactoferrine saturée en fer
HPRT	Hypoxanthine PhosphoRibosylTransferase
HRE	Hormone Response Element
HRE	Hypoxia Response Element
Hsc	Heat shock cognate protein
Hsp	Heat shock protein
HSV	Herpes simplex virus
HTLV-1	Human T-cell leukemia virus type 1

I

ICAM	Intercellular Adhesion Molecule 1
ICAT	Isotope Coded Affinity Tag
ICR	Ion Cyclotron Resonance
IEF	Isoélectrofocalisation
IFN γ	Interferon γ
Ig	Immunoglobuline
IL-6	Interleukin-6
Ion trap	Trappe ionique 3-D
IPG	Immobilized pH Gradient
IRF	Interferon Regulatory Factor
iTRAQTM	Isobaric Tags for Relative and Absolute Quantification

J

JNK c-Jun N-terminal kinase

K

kDa Kilo Dalton

KLF5 Krueppel-like factor 5

L

LBP LPS-binding protein

LC Liquid Chromatography

LC-MS Liquid chromatography-mass spectrometry

LC-MS Chromatographie liquide couplée avec la spectrométrie de masse

Lf Lactoferrine

LIT® Linear Ion Trap

LPS Lipopolysaccharide

LRP Lipoprotein Receptor-related Protein

LTQ® Linear Quadripole ion Trap

Lys Lysine

M

µM Micromolaire

MALDI Matrix-Assisted Laser Desorption/Ionisation

MALDI-TOF Matrix assisted laser desorption ionization-time of flight

MaN₂B₁ Mannosidase

mCD14 CD14 membranaire

Mcl 1 Myeloid cell leukemia sequence 1

MEC Matrice extracellulaire

mER-β Murine estrogen receptor-β

MFPaQ Mascot file parsing and quantification

min Minute

miR214 micro-ARN214

mM Millimolaire

MMP1 Métalloprotéinase

MRM Multiple Reaction Monitoring

MS/MS Spectrométrie de masse en tandem

MudPIT Multidimensional Protein Identification Technology

N

N Azote

NaCl Chlorure de sodium

NFATc1 Nuclear Factor Activated T cells

NF-κB Nuclear factor-kappa B

NK Natural killer

nL Nanolitre

NLS Nuclear localization signal

NO Nitric oxyde

NSF N-ethylmaleimide Sensitive Factor

Nt N-terminal

P

p/v	poids/volume
P1	Promoteur 1 du gène <i>Lf</i>
P2	Promoteur 2 du gène <i>Lf</i>
PAI	Protein Abundance Indices
PARP	Poly (ADP-ribose) polymérase
pb	paire de base
PBS	Phosphate buffer saline
PCBP 2	Poly(rC)-binding protein 2
pcDNA- Δ Lf	Vecteur pcDNA contenant le cDNA de la Δ Lf
pCMV-10-3XFLAG- Δ Lf	Vecteur pCMV-10-3XFLAG contenant le cDNA de la Δ Lf
pCMV-10-3XFLAG-Lf	Vecteur pCMV-10-3XFLAG contenant le cDNA de la Lf
PCR	Polymerase chain reaction
PDIA3	Protein disulfide isomerase A3
PEST	Peptide sequence rich in proline, glutamic acid, serine and threonine
PFF	Peptide Fragment Fingerprinting
Pi	Point isoélectrique
PMF	Peptide Mass Fingerprinting
pORF	Primary Open Reading Frame
PPAR	Peroxisome proliferator-activated receptors
ppm	partie par million
PRP19	Pre-mRNA processing factor 19
PSAQ	Protein Standard Absolute Quantification
PSB3	Proteasome subunit beta type 3
pTRE	Vecteur pTRE

Q

QconCAT	Quantification concatamer
qPCR	Quantitative polymerase chain reaction

R

RARE	Retinoic acid response element
Rb	Retinoblastoma susceptibility protein
RE	Response element
ROS	Reactive oxygen species
RP	Reverse Phase
RPS9	Protéine ribosomale S9

S

s	Seconde
SAB	Sérum albumine bovine
sCD14	CD14 soluble
SCF	Skp1-Cul1-F-box-protein
SDS	Sodium dodécylsulfate
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SelH	Sélénoprotéine H

Ser	Serine
SILAC	Stable Isotope Labelling by Amino acids in Cell culture
siRNA	Short interfering RNA
SISCAPA	Stable Isotope Standards and Capture by Anti-Peptide Antibodies
Skp1	S-phase-kinase-associated protein-1
SNARE	Soluble N-ethylmaleimide-sensitive fusion attachment protein receptor
Sp1	Specific protein 1
SRM	Selected Reaction Monitoring
Stat3	Signal transducer and activator of transcription 3
Stat5A	Signal transducer and activator of transcription 5A
SUMO	Small ubiquitin related modifier
SVF	Sérum de veau foetal

T

T	Thymine
TAD	TransActivation Domain
TCPB	Protéine T complex subunit β
Tet-on	Tetracycline-Controlled Transcriptional Activation
TGF	Transforming growth factor
Th1	Cellules T auxiliaires de type 1
THP1	Human acute monocytic leukemia cell line
Thr	Thrénanine
TLR	Toll-like receptor
TNF	Tumor necrosis factor
TOF	Time Of Flight
tr.min	tours.minutes

U

Ub	Ubiquitine
Ubc	Ubiquitin-conjugating enzyme
Ubl	Ubiquitin like modifier
UBR	Ubiquitin-Protein Ligase E3 Component N-Recognin
uORF	Upstream Open Reading Frame
UTR	Untranslated region

V

VCAM	Vascular cell adhesion protein
VEGF	Vascular Endothelial Growth Factor
VIH	Virus de l'immunodéficience humaine

X

XIC	eXtracted Ion Current ou eXtracted Ion Chromatogram
XRE	Xenobiotic response element
Xrn1	5'-3' exoribonuclease 1

Y

YY1	YinYang-1
-----	-----------

Références

- Actor, J. K., Hwang, S., and Kruzel, M. L. (2009). Lactoferrin as a natural immune modulator. *Curr. Pharm. Des.* *15*, 1956–1973.
- Adourian, A., Jennings, E., Balasubramanian, R., Hines, W. M., Damian, D., Plasterer, T. N., Clish, C. B., Stroobant, P., McBurney, R., Verheij, E. R., et al. (2008). Correlation network analysis for data integration and biomarker selection. *Mol Biosyst* *4*, 249.
- Aebersold, R., and Goodlett, D. R. (2001). Mass spectrometry in proteomics. *Chem Rev* *101*, 269–295.
- Agrawal, P., Yu, K., Salomon, A. R., and Sedivy, J. M. (2010). Proteomic profiling of Myc-associated proteins. *Cell Cycle* *9*, 4908–4921.
- Alban, A., David, S. O., Bjorkesten, L., Andersson, C., Sloge, E., Lewis, S., and Currie, I. (2003). A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. *Proteomics* *3*, 36–44.
- Ammendolia, M. G., Pietrantoni, A., Tinari, A., Valenti, P., and Superti, F. (2007). Bovine lactoferrin inhibits echovirus endocytic pathway by interacting with viral structural polypeptides. *Antiviral Research* *73*, 151–160.
- Andersen, J. H., Jenssen, H., Sandvik, K., and Gutteberg, T. J. (2004). Anti-HSV activity of lactoferrin and lactoferricin is dependent on the presence of heparan sulphate at the cell surface. *J. Med. Virol.* *74*, 262–271.
- Anderson, B. F., Baker, H. M., Norris, G. E., Rumball, S. V., and Baker, E. N. (1990). Apolactoferrin structure demonstrates ligand-induced conformational change in transferrins. *Nature* *344*, 784–787.
- Anderson, B. F., Baker, H. M., Dodson, E. J., Norris, G. E., Rumball, S. V., Waters, J. M., and Baker, E. N. (1987). Structure of human lactoferrin at 3.2-A resolution. *Proc Natl Acad Sci USA* *84*, 1769–1773.
- Anderson, B. F., Baker, H. M., Norris, G. E., Rice, D. W., and Baker, E. N. (1989). Structure of human lactoferrin: crystallographic structure analysis and refinement at 2.8 Å resolution. *J. Mol. Biol.* *209*, 711–734.
- Anderson, D. C., Li, W., Payan, D. G., and Noble, W. S. (2003). A New Algorithm for the Evaluation of Shotgun Peptide Sequencing in Proteomics: Support Vector Machine Classification of Peptide MS/MS Spectra and SEQUEST Scores. *J Proteome Res* *2*, 137–146.
- Anderson, L., and Anderson, N. G. (1977). High resolution two-dimensional electrophoresis of human plasma proteins. *Proc Natl Acad Sci USA* *74*, 5421–5425.
- Anderson, L., and Hunter, C. L. (2006). Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Mol Cell Proteomics* *5*, 573–588.
- Anderson, N. L., Anderson, N. G., Haines, L. R., Hardie, D. B., Olafson, R. W., and Pearson, T. W. (2004). Mass spectrometric quantitation of peptides and proteins using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA). *J Proteome Res* *3*, 235–244.
- Appelmelk, B. J., An, Y. Q., Geerts, M., Thijs, B. G., de Boer, H. A., MacLaren, D. M., de

- Graaff, J., and Nuijens, J. H. (1994). Lactoferrin is a lipid A-binding protein. *Infect Immun* *62*, 2628–2632.
- Arai, K., Takano, S., Teratani, T., Ito, Y., Yamada, T., and Nozawa, R. (2008). S100A8 and S100A9 overexpression is associated with poor pathological parameters in invasive ductal carcinoma of the breast. *Curr Cancer Drug Targets* *8*, 243–252.
- Arif, M., Senapati, P., Shandilya, J., and Kundu, T. K. (2010). Protein lysine acetylation in cellular function and its role in cancer manifestation. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* *1799*, 702–716.
- Arnold, D., Di Biase, A. M., Marchetti, M., Pietrantoni, A., Valenti, P., Seganti, L., and Superti, F. (2002). Antiadenovirus activity of milk proteins: lactoferrin prevents viral infection. *Antiviral Research* *53*, 153–158.
- Arnold, R. R., Cole, M. F., and McGhee, J. R. (1977). A bactericidal effect for human lactoferrin. *Science* *197*, 263–265.
- Arnold, R. R., Brewer, M., and Gauthier, J. J. (1980). Bactericidal activity of human lactoferrin: sensitivity of a variety of microorganisms. *Infect Immun* *28*, 893–898.
- Artym, J., Zimecki, M., and 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.
- Ashida, K., Sasaki, H., Suzuki, Y. A., and Lönnardal, B. (2004). Cellular internalization of lactoferrin in intestinal epithelial cells. *Biometals* *17*, 311–315.
- Baggiolini, M., De Duve, C., Masson, P. L., and Heremans, J. F. (1970). Association of lactoferrin with specific granules in rabbit heterophil leukocytes. *J Exp Med* *131*, 559–570.
- Bail, S., and Kiledjian, M. (2008). DcpS, a general modulator of cap-binding protein-dependent processes? *RNA Biol* *5*, 216–219.
- Baker, E., and Baker, H. (2009). A structural framework for understanding the multifunctional character of lactoferrin. *Biochimie* *91*, 3–10.
- Baldwin, A. S. (2001). Control of oncogenesis and cancer therapy resistance by the transcription factor NF-kappaB. *J Clin Invest* *107*, 241–246.
- Baldwin, D. A., Jenny, E. R., and Aisen, P. (1984). The effect of human serum transferrin and milk lactoferrin on hydroxyl radical formation from superoxide and hydrogen peroxide. *J Biol Chem* *259*, 13391–13394.
- Baldwin, G. S. (1993). Comparison of transferrin sequences from different species. *Comp. Biochem. Physiol., B* *106*, 203–218.
- Baldwin, M. A. (2003). Protein Identification by Mass Spectrometry: Issues to be Considered. *Molecular & Cellular Proteomics* *3*, 1–9.
- Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. (2007). Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem* *389*, 1017–1031.
- Barton, J. C., Huster, W. J., and Parmley, R. T. (1988). Iron-binding reactivity in mature neutrophils: relative cell content quantification by cytochemical scoring. *Journal of*

Histochemistry & Cytochemistry 36, 649–658.

- Baumrucker, C., Schanbacher, F., Shang, Y., and Green, M. (2006). Lactoferrin interaction with retinoid signaling: Cell growth and apoptosis in mammary cells. Domestic Animal Endocrinology 30, 289–303.
- Baveye, S., Elass, E., Fernig, D. G., Blanquart, C., Mazurier, J., and Legrand, D. (2000). Human lactoferrin interacts with soluble CD14 and inhibits expression of endothelial adhesion molecules, E-selectin and ICAM-1, induced by the CD14-lipopolysaccharide complex. Infect Immun 68, 6519–6525.
- Belew, A. T., Hepler, N. L., Jacobs, J. L., and Dinman, J. D. (2008). PRFdb: A database of computationally predicted eukaryotic programmed -1 ribosomal frameshift signals. BMC Genomics 9, 339.
- Beljaars, L., van der Strate, B. W. A., Bakker, H. I., Reker-Smit, C., van Loenen-Weemaes, A., Wiegmans, F. C., Harmsen, M. C., Molema, G., and Meijer, D. K. F. (2004). Inhibition of cytomegalovirus infection by lactoferrin in vitro and in vivo. Antiviral Research 63, 197–208.
- Bellamy, W. R., Wakabayashi, H., TAKASE, M., Kawase, K., Shimamura, S., and Tomita, M. (1993). Role of cell-binding in the antibacterial mechanism of lactoferricin B. J. Appl. Bacteriol. 75, 478–484.
- Benaïssa, M., Peyrat, J., Hornez, L., Mariller, C., Mazurier, J., and Pierce, A. (2005). Expression and prognostic value of lactoferrin mRNA isoforms in human breast cancer. Int J Cancer 114, 299–306.
- Benayoun, B. A., and Veitia, R. A. (2009). A post-translational modification code for transcription factors: sorting through a sea of signals. Trends Cell Biol. 19, 189–197.
- Bergeron, J. J. M., and Hallett, M. (2007). Peptides you can count on. Nat Biotechnol 25, 61–62.
- Berkhout, B., Floris, R., Recio, I., and Visser, S. (2004). The antiviral activity of the milk protein lactoferrin against the human immunodeficiency virus type 1. Biometals 17, 291–294.
- Berlutti, F., Ajello, M., Bosso, P., Morea, C., Petrucca, A., Antonini, G., and Valenti, P. (2004). Both lactoferrin and iron influence aggregation and biofilm formation in *Streptococcus mutans*. Biometals 17, 271–278.
- Bernay, B., Gaillard, M., Guryca, V., Emadali, A., Kuhn, L., Bertrand, A., Detraz, I., Carcenac, C., Savasta, M., Brouillet, E., et al. (2009). Discovering new bioactive neuropeptides in the striatum secretome using in vivo microdialysis and versatile proteomics. Molecular & Cellular Proteomics 8, 946–958.
- Bessler, H. C., de Oliveira, I. R., and Giugliano, L. G. (2006). Human milk glycoproteins inhibit the adherence of *Salmonella typhimurium* to HeLa cells. Microbiol. Immunol. 50, 877–882.
- Beynon, R. J., and Pratt, J. M. (2005). Metabolic labeling of proteins for proteomics. Mol Cell Proteomics 4, 857–872.
- Bезault, J., Bhimani, R., Wiprovnick, J., and Furmanski, P. (1994). Human lactoferrin inhibits growth of solid tumors and development of experimental metastases in mice.

- Bi, B. Y., Leveugle, B., Liu, J. L., Collard, A., Coppe, P., Roche, A. C., Nillesse, N., Capron, M., Spik, G., and Mazurier, J. (1994). Immunolocalization of the lactotransferrin receptor on the human T lymphoblastic cell line Jurkat. European Journal of Cell Biology 65, 164–171.
- Bianchi, M. E. (2006). DAMPs, PAMPs and alarmins: all we need to know about danger. J. Leukoc. Biol. 81, 1–5.
- Biemann, K. (1990). Appendix 5. Nomenclature for peptide fragment ions (positive ions). Meth. Enzymol. 193, 886–887.
- Biserte, G., Havez, R., and Cuvelier, R. (1963). [the glycoproteins of bronchial secretions]. Expos Annu Biochim Med 24, 85–120.
- Blagoev, B., Kratchmarova, I., Ong, S., Nielsen, M., Foster, L. J., and Mann, M. (2003). A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. Nat Biotechnol 21, 315–318.
- Blais, A., Malet, A., Mikogami, T., Martin-Rouas, C., and Tomé, D. (2009). Oral bovine lactoferrin improves bone status of ovariectomized mice. Am. J. Physiol. Endocrinol. Metab. 296, E1281–8.
- Blondeau, F., Ritter, B., Allaire, P. D., Wasiak, S., Girard, M., Hussain, N. K., Angers, A., Legendre-Guillemin, V., Roy, L., Boismenu, D., et al. (2004). Tandem MS analysis of brain clathrin-coated vesicles reveals their critical involvement in synaptic vesicle recycling. Proc Natl Acad Sci USA 101, 3833–3838.
- Blueggel, M., Chamrad, D., and Meyer, H. E. (2004). Bioinformatics in proteomics. Curr Pharm Biotechnol 5, 79–88.
- Bode, A. M., and Dong, Z. (2004). Post-translational modification of p53 in tumorigenesis. Nat Rev Cancer 4, 793–805.
- Boeckmann, B. (2003). The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res 31, 365–370.
- Bouyssié, D., Gonzalez de Peredo, A., Mouton, E., Albigot, R., Roussel, L., Ortega, N., Cayrol, C., Burlet-Schiltz, O., Girard, J., and Monsarrat, B. (2007). Mascot file parsing and quantification (MFPaQ), a new software to parse, validate, and quantify proteomics data generated by ICAT and SILAC mass spectrometric analyses: application to the proteomics study of membrane proteins from primary human endothelial cells. Mol Cell Proteomics 6, 1621–1637.
- Bowers, G. N., Fassett, J. D., and White, E. (1993). Isotope dilution mass spectrometry and the National Reference System. Anal Chem 65, 475R–479R.
- Bowman, B. H., Yang, F. M., and Adrian, G. S. (1988). Transferrin: evolution and genetic regulation of expression. Adv. Genet. 25, 1–38.
- Brandenburg, K., Jürgens, G., Müller, M., Fukuoka, S., and Koch, M. H. (2001). Biophysical characterization of lipopolysaccharide and lipid A inactivation by lactoferrin. Biol. Chem. 382, 1215–1225.
- Breton, M., Mariller, C., Benaïssa, M., Caillaux, K., Browaeys, E., Masson, M., Vilain, J.,

- Mazurier, J., and Pierce, A. (2004). Expression of delta-lactoferrin induces cell cycle arrest. *Biometals* *17*, 325–329.
- Brewis, I. A., and Brennan, P. (2010). Proteomics technologies for the global identification and quantification of proteins. *Adv Protein Chem Struct Biol* *80*, 1–44.
- Broekhuyse, R. M. (1974). Tear lactoferrin: a bacteriostatic and complexing protein. *Invest Ophthalmol* *13*, 550–554.
- Broxmeyer, H. E., Williams, D. E., Hangoc, G., Cooper, S., Gillis, S., Shadduck, R. K., and Bicknell, D. C. (1987). Synergistic myelopoietic actions in vivo after administration to mice of combinations of purified natural murine colony-stimulating factor 1, recombinant murine interleukin 3, and recombinant murine granulocyte/macrophage colony-stimulating factor. *Proc Natl Acad Sci USA* *84*, 3871–3875.
- Brun, V., Masselon, C., Garin, J., and Dupuis, A. (2009). Isotope dilution strategies for absolute quantitative proteomics. *J Proteomics* *72*, 740–749.
- Brun, V., Dupuis, A., Adrait, A., Marcellin, M., Thomas, D., Court, M., Vandenesch, F., and Garin, J. (2007). Isotope-labeled Protein Standards: Toward Absolute Quantitative Proteomics. *Molecular & Cellular Proteomics* *6*, 2139–2149.
- Calvio, C., Neubauer, G., Mann, M., and Lamond, A. I. (1995). Identification of hnRNP P2 as TLS/FUS using electrospray mass spectrometry. *RNA* *1*, 724–733.
- Carr, S. A., and Anderson, L. (2008). Protein quantitation through targeted mass spectrometry: the way out of biomarker purgatory? *Clin Chem* *54*, 1749–1752.
- Carthagena, L., Becquart, P., Hocini, H., Kazatchkine, M. D., Bouhla, H., and Belec, L. (2011). Modulation of HIV Binding to Epithelial Cells and HIV Transfer from Immature Dendritic Cells to CD4 T Lymphocytes by Human Lactoferrin and its Major Exposed LF-33 Peptide. *Open Virol J* *5*, 27–34.
- Chakravarti, B., Fathy, P., Sindicich, M., Mallik, B., and Chakravarti, D. N. (2010). Comparison of SYPRO Ruby and Flamingo fluorescent stains for application in proteomic research. *Anal Biochem* *398*, 1–6.
- Chan, S., and Cheng, S. (2005). The Prp19-associated complex is required for specifying interactions of U5 and U6 with pre-mRNA during spliceosome activation. *J Biol Chem* *280*, 31190–31199.
- Chandra Mohan, K. V. P., Devaraj, H., Prathiba, D., Hara, Y., and 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.
- Chelius, D., and Bondarenko, P. V. (2002). Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. *J Proteome Res* *1*, 317–323.
- Choe, L., D'ascenzo, M., Relkin, N. R., Pappin, D., Ross, P., Williamson, B., Guertin, S., Pribil, P., and Lee, K. H. (2007). 8-Plex quantitation of changes in cerebrospinal fluid protein expression in subjects undergoing intravenous immunoglobulin treatment for Alzheimer's disease. *Proteomics* *7*, 3651–3660.
- Choi, J., and Husain, M. (2006). Calmodulin-mediated cell cycle regulation: new

mechanisms for old observations. *Cell Cycle* 5, 2183–2186.

- Choksi, S., Lin, Y., Pobezinskaya, Y., Chen, L., Park, C., Morgan, M., Li, T., Jitkaew, S., Cao, X., Kim, Y., et al. (2011). A HIF-1 Target, AT1A, Protects Cells from Apoptosis by Modulating the Mitochondrial Thioredoxin, TRX2. *Molecular Cell* 42, 597–609.
- Cicciaro, E., and Blair, I. A. (2010). Stable-isotope dilution LC–MS for quantitative biomarker analysis. *Bioanalysis* 2, 311–341.
- Clouser, K. R., Baker, P., and Burlingame, A. L. (1999). Role of Accurate Mass Measurement (± 10 ppm) in Protein Identification Strategies Employing MS or MS/MS and Database Searching. *Anal Chem* 71, 2871–2882.
- Coddeville, B., Strecker, G., Wieruszseski, J. M., Vliegenthart, J. F., van Halbeek, H., Peter-Katalinić, J., Egge, H., and Spik, G. (1992). Heterogeneity of bovine lactotransferrin glycans. Characterization of alpha-D-Galp-(1-->3)-beta-D-Gal- and alpha-NeuAc-(2-->6)-beta-D-GalpNAc-(1-->4)- beta-D-GlcNAc-substituted N-linked glycans. *Carbohydr. Res.* 236, 145–164.
- Coghlin, C., Carpenter, B., Dundas, S. R., Lawrie, L. C., Telfer, C., and Murray, G. I. (2006). Characterization and over-expression of chaperonin t-complex proteins in colorectal cancer. *J. Pathol.* 210, 351–357.
- Cohen, I., Pappo, O., Elkin, M., San, T., Bar-Shavit, R., Hazan, R., Peretz, T., Vlodavsky, I., and Abramovitch, R. (2006). Heparanase promotes growth, angiogenesis and survival of primary breast tumors. *Int J Cancer* 118, 1609–1617.
- Cole, R. B. (2010). Electrospray and MALDI Mass Spectrometry: Fundamentals, Instrumentation, Practicalities, and Biological Applications. 1–897.
- Coller, J., and Parker, R. (2004). Eukaryotic mRNA decapping. *Annu. Rev. Biochem.* 73, 861–890.
- Comi, G. P., Fortunato, F., Lucchiari, S., Bordoni, A., Prelle, A., Jann, S., Keller, A., Ciscato, P., Galbiati, S., Chiveri, L., et al. (2001). Beta-enolase deficiency, a new metabolic myopathy of distal glycolysis. *Ann. Neurol.* 50, 202–207.
- Cornish, J., Palmano, K., Callon, K. E., Watson, M., Lin, J. M., Valenti, P., Naot, D., Grey, A. B., and Reid, I. R. (2006). Lactoferrin and bone; structure-activity relationships. *Biochem Cell Biol* 84, 297–302.
- Cornish, J., Callon, K. E., Naot, D., Palmano, K. P., Banovic, T., Bava, U., Watson, M., Lin, J., Tong, P. C., Chen, Q., et al. (2004). Lactoferrin is a potent regulator of bone cell activity and increases bone formation in vivo. *Endocrinology* 145, 4366–4374.
- Cornish, T. J., and Cotter, R. J. (1993). A curved-field reflectron for improved energy focusing of product ions in time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* 7, 1037–1040.
- Corthals, G. L., Wasinger, V. C., Hochstrasser, D. F., and Sanchez, J. C. (2000). The dynamic range of protein expression: a challenge for proteomic research. *Electrophoresis* 21, 1104–1115.
- Cox, J., Matic, I., Hilger, M., Nagaraj, N., Selbach, M., Olsen, J. V., and Mann, M. (2009). A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat Protoc* 4, 698–705.

- Cravatt, B. F., Simon, G. M., and Yates, J. R., III (2007). The biological impact of mass-spectrometry-based proteomics. *Nature* *450*, 991–1000.
- Cross, S. S., Hamdy, F. C., Deloulme, J. C., and Rehman, I. (2005). Expression of S100 proteins in normal human tissues and common cancers using tissue microarrays: S100A6, S100A8, S100A9 and S100A11 are all overexpressed in common cancers. *Histopathology* *46*, 256–269.
- Crouch, S. P., and Fletcher, J. (1992). Effect of ingested pentoxifylline on neutrophil superoxide anion production. *Infect Immun* *60*, 4504–4509.
- Curran, C. S., Demick, K. P., and Mansfield, J. M. (2006). Lactoferrin activates macrophages via TLR4-dependent and -independent signaling pathways. *Cellular Immunology* *242*, 23–30.
- Daher, al, S., de Gasperi, R., Daniel, P., Hall, N., Warren, C. D., and Winchester, B. (1991). The substrate-specificity of human lysosomal alpha-D-mannosidase in relation to genetic alpha-mannosidosis. *Biochem. J.* *277* (*Pt 3*), 743–751.
- Damiens, E., Mazurier, J., Yazidi, el, I., Masson, M., Duthille, I., Spik, G., and Boilly-Marer, Y. (1998). Effects of human lactoferrin on NK cell cytotoxicity against haematopoietic and epithelial tumour cells. *Biochim Biophys Acta* *1402*, 277–287.
- Damiens, E., Yazidi, el, I., Mazurier, J., Duthille, I., Spik, G., and Boilly-Marer, Y. (1999). Lactoferrin inhibits G1 cyclin-dependent kinases during growth arrest of human breast carcinoma cells. *J. Cell. Biochem.* *74*, 486–498.
- Darula, Z., Chalkley, R., Baker, P., Burlingame, A., and Medzihradszky, K. (2010). Mass spectrometric analysis, automated identification and complete annotation of O-linked glycopeptides. *Eur. J. Mass Spectrom.* *16*, 421.
- Davis, M. T., Stahl, D. C., Hefta, S. A., and Lee, T. D. (1995). A microscale electrospray interface for on-line, capillary liquid chromatography/tandem mass spectrometry of complex peptide mixtures. *Anal Chem* *67*, 4549–4556.
- Deligny, A., Denys, A., Marcant, A., Melchior, A., Mazurier, J., van Kuppevelt, T. H., and Allain, F. (2010). Synthesis of heparan sulfate with cyclophilin B-binding properties is determined by cell type-specific expression of sulfotransferases. *Journal of Biological Chemistry* *285*, 1701–1715.
- Desterro, J. M., Rodriguez, M. S., and Hay, R. T. (1998). SUMO-1 modification of IkappaBalphalpha inhibits NF-kappaB activation. *Molecular Cell* *2*, 233–239.
- Dhennin-Duthille, I., Masson, M., Damiens, E., Fillebeen, C., Spik, G., and Mazurier, J. (2000). Lactoferrin upregulates the expression of CD4 antigen through the stimulation of the mitogen-activated protein kinase in the human lymphoblastic T Jurkat cell line. *J. Cell. Biochem.* *79*, 583–593.
- Di Biase, A. M., Pietrantoni, A., Tinari, A., Siciliano, R., Valenti, P., Antonini, G., Seganti, L., and Superti, F. (2003). Heparin-interacting sites of bovine lactoferrin are involved in anti-adenovirus activity. *J. Med. Virol.* *69*, 495–502.
- Dial, E. J., and Lichtenberger, L. M. (2002). Effect of lactoferrin on Helicobacter felis induced gastritis. *Biochem Cell Biol* *80*, 113–117.
- Dole, M. (1968). Molecular Beams of Macroions. *J. Chem. Phys.* *49*, 2240.

- Domon, B., and Aebersold, R. (2006a). Challenges and opportunities in proteomics data analysis. *Mol Cell Proteomics* *5*, 1921–1926.
- Domon, B., and Aebersold, R. (2006b). Mass spectrometry and protein analysis. *Science* *312*, 212–217.
- Dzitko, K., Dziadek, B., Dziadek, J., and Długońska, H. (2007). Toxoplasma gondii: inhibition of the intracellular growth by human lactoferrin. *Pol. J. Microbiol.* *56*, 25–32.
- Efthymiadis, A., Shao, H., Hübner, S., and 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*, 22134–22139.
- Elass, E., Masson, M., Mazurier, J., and Legrand, D. (2002). Lactoferrin inhibits the lipopolysaccharide-induced expression and proteoglycan-binding ability of interleukin-8 in human endothelial cells. *Infect Immun* *70*, 1860–1866.
- Elass-Rochard, E., Legrand, D., Salmon, V., Roseanu, A., Trif, M., Tobias, P. S., Mazurier, J., and Spik, G. (1998). Lactoferrin inhibits the endotoxin interaction with CD14 by competition with the lipopolysaccharide-binding protein. *Infect Immun* *66*, 486–491.
- Elass-Rochard, E., Roseanu, A., Legrand, D., Trif, M., Salmon, V., Motas, C., Montreuil, J., and Spik, G. (1995). Lactoferrin-lipopolysaccharide interaction: involvement of the 28–34 loop region of human lactoferrin in the high-affinity binding to Escherichia coli O55B5 lipopolysaccharide. *Biochem. J.* *312* (Pt 3), 839–845.
- Ellison, R. T., Giehl, T. J., and LaForce, F. M. (1988). Damage of the outer membrane of enteric gram-negative bacteria by lactoferrin and transferrin. *Infect Immun* *56*, 2774–2781.
- Ellison, R. T., LaForce, F. M., Giehl, T. J., Boose, D. S., and Dunn, B. E. (1990). Lactoferrin and transferrin damage of the gram-negative outer membrane is modulated by Ca²⁺ and Mg²⁺. *J. Gen. Microbiol.* *136*, 1437–1446.
- Ellison, R. T., and Giehl, T. J. (1991). Killing of gram-negative bacteria by lactoferrin and lysozyme. *J Clin Invest* *88*, 1080–1091.
- Emmott, E., Rodgers, M. A., Macdonald, A., McCrory, S., Ajuh, P., and Hiscox, J. A. (2010). Quantitative proteomics using stable isotope labeling with amino acids in cell culture reveals changes in the cytoplasmic, nuclear, and nucleolar proteomes in Vero cells infected with the coronavirus infectious bronchitis virus. *Molecular & Cellular Proteomics* *9*, 1920–1936.
- Ewing, R. M., Chu, P., Elisma, F., Li, H., Taylor, P., Climie, S., McBroom-Cerajewski, L., Robinson, M. D., O'Connor, L., Li, M., et al. (2007). Large-scale mapping of human protein–protein interactions by mass spectrometry. *Mol Syst Biol* *3*.
- Eyers, C., Simpson, D., Wong, S., Beynon, R., and Gaskell, S. (2008). QCAL—a Novel Standard for Assessing Instrument Conditions for Proteome Analysis. *J. Am. Soc. Mass Spectrom.* *19*, 1275–1280.
- Feng, X., and Siegel, M. M. (2007). FTICR-MS applications for the structure determination of natural products. *Anal Bioanal Chem* *389*, 1341–1363.
- Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989). Electrospray

ionization for mass spectrometry of large biomolecules. *Science* **246**, 64–71.

- Fillebeen, C., Ruchoux, M. M., Mitchell, V., Vincent, S., Benaïssa, M., and 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.
- Fillebeen, C., Descamps, L., Dehouck, M. P., Fenart, L., Benaïssa, M., Spik, G., Cecchelli, R., and Pierce, A. (1999). Receptor-mediated transcytosis of lactoferrin through the blood-brain barrier. *J Biol Chem* **274**, 7011–7017.
- Fischer, R., Debbabi, H., Dubarry, M., Boyaka, P., and Tomé, D. (2006). Regulation of physiological and pathological Th1 and Th2 responses by lactoferrin. *Biochem Cell Biol* **84**, 303–311.
- Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., and Merrick, J. M. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496–512.
- Frank, A., Tanner, S., Bafna, V., and Pevzner, P. (2005). Peptide sequence tags for fast database search in mass-spectrometry. *J Proteome Res* **4**, 1287–1295.
- Fujita, K., Matsuda, E., Sekine, K., Iigo, M., and 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., and Tsuda, H. (2004b). Lactoferrin modifies apoptosis-related gene expression in the colon of the azoxymethane-treated rat. *Cancer Lett.* **213**, 21–29.
- Fukami, M., Stierna, P., Veress, B., and Carlsöö, B. (1993). Lysozyme and lactoferrin in human maxillary sinus mucosa during chronic sinusitis. An immunohistochemical study. *Eur Arch Otorhinolaryngol* **250**, 133–139.
- Furmanski, P., Li, Z. P., Fortuna, M. B., Swamy, C. V., and Das, M. R. (1989). Multiple molecular forms of human lactoferrin. Identification of a class of lactoferrins that possess ribonuclease activity and lack iron-binding capacity. *J Exp Med* **170**, 415–429.
- Gareau, J. R., and Lima, C. D. (2010). The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nat Rev Mol Cell Biol* **11**, 861–871.
- Garrè, C., Bianchi-Scarrà, G., Sirito, M., Musso, M., and Ravazzolo, R. (1992). Lactoferrin binding sites and nuclear localization in K562(S) cells. *J. Cell. Physiol.* **153**, 477–482.
- Garrido, C., Schmitt, E., Candé, C., Vahsen, N., Parcellier, A., and Kroemer, G. (2003). HSP27 and HSP70: potentially oncogenic apoptosis inhibitors. *Cell Cycle* **2**, 579–584.
- Gartner, C. A., Elias, J. E., Bakalarski, C. E., and Gygi, S. P. (2007). Catch-and-release reagents for broadscale quantitative proteomics analyses. *J Proteome Res* **6**, 1482–1491.
- Geiger, T., Cox, J., and Mann, M. (2010a). Proteomics on an Orbitrap Benchtop Mass Spectrometer Using All-ion Fragmentation. *Molecular & Cellular Proteomics* **9**, 2252–2261.
- Geiger, T., Cox, J., Ostasiewicz, P., Wisniewski, J. R., and Mann, M. (2010b). Super-SILAC

- mix for quantitative proteomics of human tumor tissue. *Nat Meth* 7, 383–385.
- Geng, K., LI, Y., Bezault, J., and Furmanski, P. (1998). Induction of lactoferrin expression in murine ES cells by retinoic acid and estrogen. *Exp Cell Res* 245, 214–220.
- Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., and Gygi, S. P. (2003). Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc Natl Acad Sci USA* 100, 6940–6945.
- Gery, S., Park, D. J., Vuong, P. T., Chih, D. Y., Lemp, N., and Koeffler, H. P. (2004). Retinoic acid regulates C/EBP homologous protein expression (CHOP), which negatively regulates myeloid target genes. *Blood* 104, 3911–3917.
- Gevaert, K., Impens, F., Ghesquière, B., Van Damme, P., Lambrechts, A., and Vandekerckhove, J. (2008). Stable isotopic labeling in proteomics. *Proteomics* 8, 4873–4885.
- Gill, G. (2005). Something about SUMO inhibits transcription. *Curr. Opin. Genet. Dev.* 15, 536–541.
- Gingras, A. C., Raught, B., and Sonenberg, N. (1999). eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu. Rev. Biochem.* 68, 913–963.
- Glozak, M. A., Sengupta, N., Zhang, X., and Seto, E. (2005). Acetylation and deacetylation of non-histone proteins. *Gene* 363, 15–23.
- Goldberg, G. S., Kunimoto, T., Alexander, D. B., Suenaga, K., Ishidate, F., Miyamoto, K., Ushijima, T., Teng, C. T., Yokota, J., Ohta, T., et al. (2005). 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, 57–64.
- Gomez, H. F., Ochoa, T. J., Carlin, L. G., and Cleary, T. G. (2003). Human lactoferrin impairs virulence of *Shigella flexneri*. *J Infect Dis* 187, 87–95.
- Görg, A., Postel, W., Weser, J., Patutschnick, W., and Cleve, H. (1985). Improved resolution of PI (alpha 1-antitrypsin) phenotypes by a large-scale immobilized pH gradient. *Am. J. Hum. Genet.* 37, 922–930.
- Görg, A., Postel, W., and Günther, S. (1988). The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 9, 531–546.
- Graumann, J., Hubner, N. C., Kim, J. B., Ko, K., Moser, M., Kumar, C., Cox, J., Scholer, H., and Mann, M. (2007). Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC) and Proteome Quantitation of Mouse Embryonic Stem Cells to a Depth of 5,111 Proteins. *Molecular & Cellular Proteomics* 7, 672–683.
- Grey, A., Banovic, T., Zhu, Q., Watson, M., Callon, K., Palmano, K., Ross, J., Naot, D., Reid, I. R., and Cornish, J. (2004). The low-density lipoprotein receptor-related protein 1 is a mitogenic receptor for lactoferrin in osteoblastic cells. *Mol. Endocrinol.* 18, 2268–2278.
- Griesbeck-Zilch, B., Meyer, H. H. D., Kühn, C. H., Schwerin, M., and 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.

- Griffin, J. L., Atherton, H., Shockcor, J., and Atzori, L. (2011). Metabolomics as a tool for cardiac research. *Nat Rev Cardiol* *8*, 630–643.
- Groot, F., Geijtenbeek, T. B. H., Sanders, R. W., Baldwin, C. E., Sanchez-Hernandez, M., Floris, R., van Kooyk, Y., de Jong, E. C., and Berkhout, B. (2005). Lactoferrin prevents dendritic cell-mediated human immunodeficiency virus type 1 transmission by blocking the DC-SIGN--gp120 interaction. *J. Virol.* *79*, 3009–3015.
- Grove, M. L. (1960). The isolation of a red protein from milk. *J. Am. Chem. Soc.* *82*, 3345–3350.
- Gstaiger, M., and Aebersold, R. (2009). Applying mass spectrometry-based proteomics to genetics, genomics and network biology. *Nat Rev Genet* *10*, 617–627.
- Guo, H. Y., Jiang, L., Ibrahim, S. A., Zhang, L., Zhang, H., Zhang, M., and Ren, F. Z. (2009). Orally administered lactoferrin preserves bone mass and microarchitecture in ovariectomized rats. *Journal of Nutrition* *139*, 958–964.
- Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999). Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* *17*, 994–999.
- Ha, N. H., Nair, V. S., Reddy, S. D. N., Mudvari, P., Ohshiro, K., Ghanta, K. S., Pakala, S. B., Li, D. Q., Costa, L., Lipton, A., et al. (2011). Lactoferrin-Endothelin-1 Axis Contributes to the Development and Invasiveness of Triple Negative Breast Cancer Phenotypes. *Cancer Res.*
- Hames, R. S., Wattam, S. L., Yamano, H., Bacchieri, R., and 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*, 7117–7127.
- Hanke, S., and Mann, M. (2009). The Phosphotyrosine Interactome of the Insulin Receptor Family and Its Substrates IRS-1 and IRS-2. *Molecular & Cellular Proteomics* *8*, 519–534.
- Hannoun, Z., Greenhough, S., Jaffray, E., Hay, R. T., and Hay, D. C. (2010). Post-translational modification by SUMO. *Toxicology* *278*, 288–293.
- Hardivillé, S., Hoedt, E., Mariller, C., Benaïssa, M., and Pierce, A. (2010). O-GlcNAcylation/phosphorylation cycling at Ser10 controls both transcriptional activity and stability of delta-lactoferrin. *Journal of Biological Chemistry* *285*, 19205–19218.
- Harmsen, M. C., Swart, P. J., de Béthune, M. P., Pauwels, R., de Clercq, E., The, T. H., and Meijer, D. K. (1995). Antiviral effects of plasma and milk proteins: lactoferrin shows potent activity against both human immunodeficiency virus and human cytomegalovirus replication in vitro. *J Infect Dis* *172*, 380–388.
- Hart, G. W., Haltiwanger, R. S., Holt, G. D., and Kelly, W. G. (1989). Nucleoplasmic and cytoplasmic glycoproteins. *Ciba Found. Symp.* *145*, 102–12, discussion 112–8.
- Hartmer, R., Kaplan, D. A., Gebhardt, C. R., Ledertheil, T., and Brekenfeld, A. (2008). Multiple ion/ion reactions in the 3D ion trap: Selective reagent anion production for ETD and PTR from a single compound. *International Journal of Mass Spectrometry* *276*, 82–90.
- Hasegawa, K., Motsuchi, W., Tanaka, S., and Dosako, S. (1994). Inhibition with lactoferrin of

- in vitro infection with human herpes virus. *Jpn. J. Med. Sci. Biol.* *47*, 73–85.
- Havlis, J., Thomas, H., Sebela, M., and Shevchenko, A. (2003). Fast-response proteomics by accelerated in-gel digestion of proteins. *Anal Chem* *75*, 1300–1306.
- Havlis, J., and Shevchenko, A. (2004). Absolute Quantification of Proteins in Solutions and in Polyacrylamide Gels by Mass Spectrometry. *Anal Chem* *76*, 3029–3036.
- Håversen, L. (2002). Lactoferrin down-regulates the LPS-induced cytokine production in monocytic cells via NF-κB. *Cellular Immunology* *220*, 83–95.
- He, J., and Furmanski, P. (1995). Sequence specificity and transcriptional activation in the binding of lactoferrin to DNA. *Nature* *373*, 721–724.
- Hendrixson, D. R., Qiu, J., Shewry, S. C., Fink, D. L., Petty, S., Baker, E. N., Plaut, A. G., and St Geme, J. W. (2003). Human milk lactoferrin is a serine protease that cleaves Haemophilus surface proteins at arginine-rich sites. *Mol. Microbiol.* *47*, 607–617.
- Henry, J., Hsu, C. Y., Haftek, M., Nachat, R., de Koning, H. D., Gardinal-Galera, I., Hitomi, K., Balica, S., Jean-Decoster, C., Schmitt, A. M., et al. (2011). Hornerin is a component of the epidermal cornified cell envelopes. *The FASEB Journal* *25*, 1567–1576.
- Henzel, W. J., Billeci, T. M., Stults, J. T., Wong, S. C., Grimley, C., and Watanabe, C. (1993). Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proc Natl Acad Sci USA* *90*, 5011–5015.
- Hidaka, S., Yasutake, T., Takeshita, H., Kondo, M., Tsuji, T., Nanashima, A., Sawai, T., Yamaguchi, H., Nakagoe, T., Ayabe, H., et al. (2000). Differences in 20q13.2 copy number between colorectal cancers with and without liver metastasis. *Clin. Cancer Res.* *6*, 2712–2717.
- Hoedt, E., Hardivillé, S., Mariller, C., Elass, E., Perraudin, J., and 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* *23*, 441–452.
- Horn, D. M., Peters, E. C., Klock, H., Meyers, A., and Brock, A. (2004). Improved protein identification using automated high mass measurement accuracy MALDI FT-ICR MS peptide mass fingerprinting. *International Journal of Mass Spectrometry* *238*, 189–196.
- Hubner, N. C., and Mann, M. (2011). Extracting gene function from protein–protein interactions using Quantitative BAC Interactomics (QUBIC). *Methods* *53*, 453–459.
- Huillet, C., Adrait, A., Lebert, D., Picard, G., Trauchessec, M., Louwagie, M., Dupuis, A., Hittinger, L., Ghaleh, B., Le Corvoisier, P., et al. (2011). Accurate quantification of cardiovascular biomarkers in serum using protein standard absolute quantification (PSAQTM) and selected reaction monitoring. *Molecular & Cellular Proteomics*.
- Hutchens, T. W., Magnuson, J. S., and Yip, T. T. (1989). Interaction of human lactoferrin with DNA: one-step purification by affinity chromatography on single-stranded DNA-agarose. *Pediatr. Res.* *26*, 618–622.
- Hwang, S., Kruzel, M., and Actor, J. (2005). Lactoferrin augments BCG vaccine efficacy to generate T helper response and subsequent protection against challenge with virulent. *International Immunopharmacology* *5*, 591–599.

- Iigo, M., Alexander, D., Long, N., Xu, J., Fukamachi, K., Futakuchi, M., Takase, M., and 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., and Tsuda, H. (1999). Inhibitory effects of bovine lactoferrin on colon carcinoma 26 lung metastasis in mice. *Clin. Exp. Metastasis* *17*, 35–40.
- Iijima, H., Tomizawa, Y., Iwasaki, Y., Sato, K., Sunaga, N., Dobashi, K., Saito, R., Nakajima, T., Minna, J. D., and Mori, M. (2006). Genetic and epigenetic inactivation of LTF gene at 3p21.3 in lung cancers. *Int J Cancer* *118*, 797–801.
- Inoue, M., Yamada, J., Kitamura, N., Shimazaki, K., Andrén, A., and Yamashita, T. (1993). Immunohistochemical localization of lactoferrin in bovine exocrine glands. *Tissue Cell* *25*, 791–797.
- Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappaport, J., and Mann, M. (2005). Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics* *4*, 1265–1272.
- Ishii, T., Uto, T., Mori, K., and Fujikawa, R. (2009). Integrin-linked kinase is involved in lactoferrin-induced anchorage-independent cell growth and survival in PC12 cells. *Life Sciences* *84*, 530–536.
- Ishikado, A., Imanaka, H., Takeuchi, T., Harada, E., and Makino, T. (2005). Liposomalization of lactoferrin enhanced its anti-inflammatory effects via oral administration. *Biol. Pharm. Bull.* *28*, 1717–1721.
- Jameson, G. B., Anderson, B. F., Norris, G. E., Thomas, D. H., and Baker, E. N. (1998). Structure of human apolactoferrin at 2.0 Å resolution. Refinement and analysis of ligand-induced conformational change. *Acta Crystallogr. D Biol. Crystallogr.* *54*, 1319–1335.
- Jans, D. A., Briggs, L. J., Gustin, S. E., Jans, P., Ford, S., and Young, I. G. (1997). A functional bipartite nuclear localisation signal in the cytokine interleukin-5. *FEBS Lett* *406*, 315–320.
- Japelj, B. (2005). Structural Origin of Endotoxin Neutralization and Antimicrobial Activity of a Lactoferrin-based Peptide. *Journal of Biological Chemistry* *280*, 16955–16961.
- Jenssen, H., and Hancock, R. (2009). Antimicrobial properties of lactoferrin. *Biochimie* *91*, 19–29.
- Jesaitis, A. J., Franklin, M. J., Berglund, D., Sasaki, M., Lord, C. I., Bleazard, J. B., Duffy, J. E., Beyenal, H., and Lewandowski, Z. (2003). Compromised host defense on *Pseudomonas aeruginosa* biofilms: characterization of neutrophil and biofilm interactions. *J Immunol* *171*, 4329–4339.
- Johansson, B. Isolation of an iron-containing red protein from human milk. *Acta Chemica Scandinavica* *14*, 510–512.
- Johnson, E. S. (2004). Protein modification by SUMO. *Annu. Rev. Biochem.* *73*, 355–382.
- Johnson, R. S., and Taylor, J. A. (2002). Searching sequence databases via de novo peptide sequencing by tandem mass spectrometry. *Mol. Biotechnol.* *22*, 301–315.

- Kai, K., Komine, K., Komine, Y., Kuroishi, T., Kozutsumi, T., Kobayashi, J., Ohta, M., Kitamura, H., and Kumagai, K. (2002). Lactoferrin stimulates A *Staphylococcus aureus* killing activity of bovine phagocytes in the mammary gland. *Microbiol. Immunol.* *46*, 187–194.
- Kamiya, H., Ehara, T., and Matsumoto, T. (2011). Inhibitory effects of lactoferrin on biofilm formation in clinical isolates of *Pseudomonas aeruginosa*. *J Infect Chemother.*
- Kanwar, J. R., Palmano, K. P., Sun, X., Kanwar, R. K., Gupta, R., Haggarty, N., Rowan, A., Ram, S., and Krissansen, G. W. (2008). “Iron-saturated” lactoferrin is a potent natural adjuvant for augmenting cancer chemotherapy. *Immunol Cell Biol* *86*, 277–288.
- Karas, M., Glückmann, M., and Schäfer, J. (2000). Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors. *J Mass Spectrom* *35*, 1–12.
- Karas, M., and Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* *60*, 2299–2301.
- Katunuma, N., Le, Q., Murata, E., Matsui, A., Majima, E., Ishimaru, N., Hayashi, Y., and Ohashi, A. (2006). A novel apoptosis cascade mediated by lysosomal lactoferrin and its participation in hepatocyte apoptosis induced by d-galactosamine. *FEBS Lett* *580*, 3699–3705.
- Kelly, T. (2005). Expression of Heparanase by Primary Breast Tumors Promotes Bone Resorption in the Absence of Detectable Bone Metastases. *Cancer Res* *65*, 5778–5784.
- Keshishian, H., Addona, T., Burgess, M., Kuhn, E., and Carr, S. A. (2007). Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. *Mol Cell Proteomics* *6*, 2212–2229.
- Khanna-Gupta, A., Zibello, T., Kolla, S., Neufeld, E. J., and Berliner, N. (1997). CCAAT displacement protein (CDP/cut) recognizes a silencer element within the lactoferrin gene promoter. *Blood* *90*, 2784–2795.
- Khanna-Gupta, A., Zibello, T., Sun, H., Gaines, P., and Berliner, N. (2003). Chromatin immunoprecipitation (ChIP) studies indicate a role for CCAAT enhancer binding proteins alpha and epsilon (C/EBP alpha and C/EBP epsilon) and CDP/cut in myeloid maturation-induced lactoferrin gene expression. *Blood* *101*, 3460–3468.
- Khanna-Gupta, A., Zibello, T., Simkevich, C., Rosmarin, A. G., and Berliner, N. (2000). Sp1 and C/EBP are necessary to activate the lactoferrin gene promoter during myeloid differentiation. *Blood* *95*, 3734–3741.
- Kim, C., Son, K., Choi, S., and Kim, J. (2006). Human lactoferrin upregulates expression of KDR/Flk-1 and stimulates VEGF-A-mediated endothelial cell proliferation and migration. *FEBS Lett* *580*, 4332–4336.
- Kim, S. J., Yu, D. Y., Pak, K. W., Jeong, S., Kim, S. W., and Lee, K. K. (1998). Structure of the human lactoferrin gene and its chromosomal localization. *Mol. Cells* *8*, 663–668.
- Kim, S. Y., Lee, M., Cho, K. C., Choi, Y., Choi, J., Sung, K., Kwon, O., Kim, H., Kim, I., and Jeong, S. (2003). Alterations in mRNA expression of ribosomal protein S9 in hydrogen peroxide-treated neurotumor cells and in rat hippocampus after transient ischemia. *Neurochem. Res.* *28*, 925–931.

- Kirkin, V., and Dikic, I. (2007). Role of ubiquitin- and Ubl-binding proteins in cell signaling. *Curr Opin Cell Biol* 19, 199–205.
- Kirkpatrick, D. S., Gerber, S. A., and Gygi, S. P. (2005). The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. *Methods* 35, 265–273.
- Kitagawa, K., and Hieter, P. (2001). Evolutionary conservation between budding yeast and human kinetochores. *Nat Rev Mol Cell Biol* 2, 678–687.
- Klimek, J., Eddes, J. S., Hohmann, L., Jackson, J., Peterson, A., Letarte, S., Gafken, P. R., Katz, J. E., Mallick, P., Lee, H., et al. (2008). The Standard Protein Mix Database: A Diverse Data Set To Assist in the Production of Improved Peptide and Protein Identification Software Tools. *J Proteome Res* 7, 96–103.
- Klintworth, G. K., Sommer, J. R., Obrian, G., Han, L., Ahmed, M. N., Qumsiyeh, M. B., Lin, P. Y., Basti, S., Reddy, M. K., Kanai, A., et al. (1998). Familial subepithelial corneal amyloidosis (gelatinous drop-like corneal dystrophy): exclusion of linkage to lactoferrin gene. *Mol. Vis.* 4, 31.
- Kozak, M. (2001). Constraints on reinitiation of translation in mammals. *Nucleic Acids Res* 29, 5226–5232.
- Krishnan, N., Dickman, M. B., and Becker, D. F. (2008). Proline modulates the intracellular redox environment and protects mammalian cells against oxidative stress. *Free Radic. Biol. Med.* 44, 671–681.
- Kruzel, M. L., Harari, Y., Mailman, D., Actor, J. K., and Zimecki, M. (2002). Differential effects of prophylactic, concurrent and therapeutic lactoferrin treatment on LPS-induced inflammatory responses in mice. *Clin Exp Immunol* 130, 25–31.
- Kruzel, M. L., Actor, J. K., Radak, Z., Bacsi, A., Saavedra-Molina, A., and Boldogh, I. (2010). Lactoferrin decreases LPS-induced mitochondrial dysfunction in cultured cells and in animal endotoxemia model. *Innate Immunity* 16, 67–79.
- Kruzel, M. L., Bacsi, A., Choudhury, B., Sur, S., and Boldogh, I. (2006). Lactoferrin decreases pollen antigen-induced allergic airway inflammation in a murine model of asthma. *Immunology* 119, 159–166.
- Kurose, I., Yamada, T., Wolf, R., and Granger, D. N. (1994). P-selectin-dependent leukocyte recruitment and intestinal mucosal injury induced by lactoferrin. *J. Leukoc. Biol.* 55, 771–777.
- la Rosa, de, G., Yang, D., Tewary, P., Varadhachary, A., and Oppenheim, J. J. (2008). Lactoferrin acts as an alarmin to promote the recruitment and activation of APCs and antigen-specific immune responses. *J Immunol* 180, 6868–6876.
- Lange, V., Picotti, P., Domon, B., and Aebersold, R. (2008). Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol Syst Biol* 4, 1–14.
- Lanucara, F., and Eyers, C. E. (2011). Methods in Enzymology (Elsevier).
- Lata, S., Schoehn, G., Solomons, J., Pires, R., Göttlinger, H. G., and Weissenhorn, W. (2009). Structure and function of ESCRT-III. *Biochem. Soc. Trans.* 37, 156.
- Lebert, D., Dupuis, A., Garin, J., Bruley, C., and Brun, V. (2011). Methods in Molecular

Biology. 753, 93–115.

- Lee, H., Park, C. J., and Lee, G. (2010a). Measurement of progesterone in human serum by isotope dilution liquid chromatography–tandem mass spectrometry and comparison with the commercial chemiluminescence immunoassay. *Anal Bioanal Chem* 396, 1713–1719.
- Lee, H., Park, J., Seok, S., Baek, M., Kim, D., Lee, B., Kang, P., Kim, Y., and Park, J. (2005). Potential antimicrobial effects of human lactoferrin against oral infection with *Listeria monocytogenes* in mice. *J. Med. Microbiol.* 54, 1049–1054.
- Lee, M. O., Liu, Y., and 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, 4194–4207.
- Lee, S., Park, S., Pyo, C., Yoo, N., Kim, J., and Choi, S. (2009). Requirement of the JNK-associated Bcl-2 pathway for human lactoferrin-induced apoptosis in the Jurkat leukemia T cell line. *Biochimie* 91, 102–108.
- Lee, S., Hwang, H., Pyo, C., Hahm, D. H., and Choi, S. (2010b). E2F1-directed activation of Bcl-2 is correlated with lactoferrin-induced apoptosis in Jurkat leukemia T lymphocytes. *Biometals* 23, 507–514.
- Legrand, D., Mazurier, J., Aubert, J. P., Loucheux-Lefebvre, M. H., Montreuil, J., and Spik, G. (1986). Evidence for interactions between the 30 kDa N- and 50 kDa C-terminal tryptic fragments of human lactotransferrin. *Biochem. J.* 236, 839–844.
- Legrand, D., Elass, E., Carpentier, M., and Mazurier, J. (2006). Interactions of lactoferrin with cells involved in immune function. *Biochem Cell Biol* 84, 282–290.
- Legrand, D., Pierce, A., Elass, E., Carpentier, M., Mariller, C., and Mazurier, J. (2008). Lactoferrin structure and functions. *Adv Exp Med Biol* 606, 163–194.
- Legrand, D., Elass, E., Carpentier, M., and Mazurier, J. (2005). Lactoferrin. *Cell Mol Life Sci* 62, 2549–2559.
- Legrand, D., Mazurier, J., Elass, A., Rochard, E., Vergoten, G., Maes, P., Montreuil, J., and Spik, G. (1992). Molecular interactions between human lactotransferrin and the phytohemagglutinin-activated human lymphocyte lactotransferrin receptor lie in two loop-containing regions of the N-terminal domain I of human lactotransferrin. *Biochemistry* 31, 9243–9251.
- Legrand, D., Mazurier, J., Colavizza, D., Montreuil, J., and Spik, G. (1990). Properties of the iron-binding site of the N-terminal lobe of human and bovine lactotransferrins. Importance of the glycan moiety and of the non-covalent interactions between the N- and C-terminal lobes in the stability of the iron-binding site. *Biochem. J.* 266, 575–581.
- Legrand, D., Vigie, K., Said, E. A., Elass, E., Masson, M., Slomianny, M., Carpentier, M., Briand, J., Mazurier, J., and Hovanessian, A. G. (2004). Surface nucleolin participates in both the binding and endocytosis of lactoferrin in target cells. *Eur J Biochem* 271, 303–317.
- Leitch, E. C., and Willcox, M. D. (1999). Elucidation of the antistaphylococcal action of lactoferrin and lysozyme. *J. Med. Microbiol.* 48, 867–871.
- Leonard, W. J., and O'Shea, J. J. (1998). Jak and STATs: biological implications. *Annu Rev*

Immunol 16, 293–322.

- León-Sicairos, N., López-Soto, F., Reyes-López, M., Godínez-Vargas, D., Ordaz-Pichardo, C., and la Garza, de, M. (2006). Amoebicidal activity of milk, apo-lactoferrin, sIgA and lysozyme. Clin Med Res 4, 106–113.
- Leśniak, W., Słomnicki, Ł. P., and Filipek, A. (2009). S100A6 – New facts and features. Biochem Biophys Res Commun 390, 1087–1092.
- Lewis, J. D., and Izaurrealde, E. (1997). The role of the cap structure in RNA processing and nuclear export. Eur J Biochem 247, 461–469.
- Li, W. Y., Li, Q. W., Han, Z. S., Jiang, Z. L., Yang, H., Li, J., and Zhang, X. B. (2011). Growth suppression effects of recombinant adenovirus expressing human lactoferrin on cervical cancer in vitro and in vivo. Cancer Biother. Radiopharm. 26, 477–483.
- Li, Y., Limmon, G., Imani, F., and Teng, C. (2009). Induction of lactoferrin gene expression by innate immune stimuli in mouse mammary epithelial HC-11 cells. Biochimie 91, 58–67.
- Liao, Y., Du, X., and Lönnadal, B. (2010). miR-214 Regulates Lactoferrin Expression and Pro-Apoptotic Function in Mammary Epithelial Cells. Journal of Nutrition 140, 1552–1556.
- Lin, J., Chen, Q., Tian, H., Gao, X., Yu, M., Xu, G., and Zhao, F. (2008). Stain efficiency and MALDI-TOF MS compatibility of seven visible staining procedures. Anal Bioanal Chem 390, 1765–1773.
- Lin, T., Chiou, S., Chen, M., and Kuo, C. (2005). Human lactoferrin exerts bi-directional actions on PC12 cell survival via ERK1/2 pathway. Biochem Biophys Res Commun 337, 330–336.
- Lin, T., Chu, C., and Chiu, C. (2002). Lactoferrin inhibits enterovirus 71 infection of human embryonal rhabdomyosarcoma cells in vitro. J Infect Dis 186, 1161–1164.
- Lin, X., Liang, M., Liang, Y., Brunicardi, F. C., and Feng, X. (2003). SUMO-1/Ubc9 promotes nuclear accumulation and metabolic stability of tumor suppressor Smad4. J Biol Chem 278, 31043–31048.
- Lindström, M. S., and Nistér, M. (2010). Silencing of Ribosomal Protein S9 Elicits a Multitude of Cellular Responses Inhibiting the Growth of Cancer Cells Subsequent to p53 Activation. PLoS ONE 5, e9578.
- Liu, D., Wang, X., Zhang, Z., and Teng, C. T. (2003). An intronic alternative promoter of the human lactoferrin gene is activated by Ets. Biochem Biophys Res Commun 301, 472–479.
- Liu, H., Sadygov, R. G., and Yates, J. R. (2004). A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem 76, 4193–4201.
- Liu, H., and Kiledjian, M. (2006). Decapping the message: a beginning or an end. Biochem. Soc. Trans 34, 35–38.
- Liu, H., and Kiledjian, M. (2005). Scavenger decapping activity facilitates 5“ to 3” mRNA decay. Mol Cell Biol 25, 9764–9772.

- Lorget, F., Clough, J., Oliveira, M., Daury, M., Sabokbar, A., and Offord, E. (2002). Lactoferrin reduces in vitro osteoclast differentiation and resorbing activity. *Biochem Biophys Res Commun* **296**, 261–266.
- Loser, K., Vogl, T., Voskort, M., Lueken, A., Kupas, V., Nacken, W., Klenner, L., Kuhn, A., Foell, D., Sorokin, L., et al. (2010). The Toll-like receptor 4 ligands MRP8 and MRP14 are crucial in the development of autoreactive CD8+ T cells. *Nat. Med.* **16**, 713–717.
- Lu, P., Vogel, C., Wang, R., Yao, X., and Marcotte, E. M. (2007). Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat Biotechnol* **25**, 117–124.
- Lu, X., and Legerski, R. J. (2007). The Prp19/Pso4 core complex undergoes ubiquitylation and structural alterations in response to DNA damage. *Biochem Biophys Res Commun* **354**, 968–974.
- Ma, B., Zhang, K., Hendrie, C., Liang, C., Li, M., Doherty-Kirby, A., and Lajoie, G. (2003). PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **17**, 2337–2342.
- Makarov, A. (2000). Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Anal Chem* **72**, 1156–1162.
- Makarov, A., Denisov, E., Lange, O., and Horning, S. (2006). Dynamic range of mass accuracy in LTQ orbitrap hybrid mass spectrometer. *J. Am. Soc. Mass Spectrom.* **17**, 977–982.
- Mann, D. L., Chase, G. W., and Eitenmiller, R. R. (2001). Liquid chromatographic analysis of vitamin B6 in soy-based infant formula. *J AOAC Int* **84**, 1593–1599.
- Mann, D. M., Romm, E., and Migliorini, M. (1994). Delineation of the glycosaminoglycan-binding site in the human inflammatory response protein lactoferrin. *J Biol Chem* **269**, 23661–23667.
- Mann, M., Højrup, P., and Roepstorff, P. (1993). Use of mass spectrometric molecular weight information to identify proteins in sequence databases. *Biol. Mass Spectrom.* **22**, 338–345.
- Marchetti, M., Trybala, E., Superti, F., Johansson, M., and Bergström, T. (2004). Inhibition of herpes simplex virus infection by lactoferrin is dependent on interference with the virus binding to glycosaminoglycans. *Virology* **318**, 405–413.
- Marchetti, M., Superti, F., Ammendolia, M. G., Rossi, P., Valenti, P., and Seganti, L. (1999). Inhibition of poliovirus type 1 infection by iron-, manganese- and zinc-saturated lactoferrin. *Med. Microbiol. Immunol.* **187**, 199–204.
- Marchetti, M., Longhi, C., Conte, M. P., Pisani, S., Valenti, P., and Seganti, L. (1996). Lactoferrin inhibits herpes simplex virus type 1 adsorption to Vero cells. *Antiviral Research* **29**, 221–231.
- Marchetti, M., Pisani, S., Antonini, G., Valenti, P., Seganti, L., and Orsi, N. (1998). Metal complexes of bovine lactoferrin inhibit in vitro replication of herpes simplex virus type 1 and 2. *Biometals* **11**, 89–94.
- Mariller, C., Benaïssa, M., Hardivillé, S., Breton, M., Pradelle, G., Mazurier, J., and 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., and 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.
- Mariller, C., Hardivillé, S., Hoedt, E., Huvent, I., Pina-Canseco, S., and Pierce, A. (2012). Delta-lactoferrin, an intracellular lactoferrin isoform that acts as a transcription factor. This article is part of a Special Issue entitled Lactoferrin and has undergone the Journal's usual peer review process. *Biochem Cell Biol.*
- Mason, D. Y., and Taylor, C. R. (1978). Distribution of transferrin, ferritin, and lactoferrin in human tissues. *J. Clin. Pathol.* 31, 316–327.
- Masson, P., Heremans, J. F., and Prignot, J. (1965). Immunohistochemical localization of the iron-binding protein lactoferrin in human bronchial glands. *Experientia* 21, 604–605.
- Masson, P. L., Heremans, J. F., and Schonne, E. (1969). Lactoferrin, an iron-binding protein in neutrophilic leukocytes. *J Exp Med* 130, 643–658.
- Masson, P. L., Heremans, J. F., and 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, 679–689.
- Masuda, C., Wanibuchi, H., Sekine, K., Yano, Y., Otani, S., Kishimoto, T., Tsuda, H., and Fukushima, S. (2000). Chemopreventive effects of bovine lactoferrin on N-butyl-N-(4-hydroxybutyl)nitrosamine-induced rat bladder carcinogenesis. *Jpn. J. Cancer Res.* 91, 582–588.
- Mattsby-Baltzer, I., Roseanu, A., Motas, C., Elverfors, J., Engberg, I., and Hanson, L. A. (1996). Lactoferrin or a fragment thereof inhibits the endotoxin-induced interleukin-6 response in human monocytic cells. *Pediatr. Res.* 40, 257–262.
- Mayya, V., Rezual, K., Wu, L., Fong, M. B., and Han, D. K. (2006). Absolute quantification of multisite phosphorylation by selective reaction monitoring mass spectrometry: determination of inhibitory phosphorylation status of cyclin-dependent kinases. *Mol Cell Proteomics* 5, 1146–1157.
- McAbee, D. D., Bennett, D. J., and Ling, Y. Y. (1998). Identification and analysis of a CA(2+)-dependent lactoferrin receptor in rat liver. Lactoferrin binds to the asialoglycoprotein receptor in a galactose-independent manner. *Adv Exp Med Biol* 443, 113–121.
- McCombs, J. L., Teng, C. T., Pentecost, B. T., Magnuson, V. L., Moore, C. M., and McGill, J. R. (1988). Chromosomal localization of human lactotransferrin gene (LTF) by in situ hybridization. *Cytogenet Cell Genet* 47, 16–17.
- McCormack, A. L., Schieltz, D. M., Goode, B., Yang, S., Barnes, G., Drubin, D., and Yates, J. R. (1997). Direct analysis and identification of proteins in mixtures by LC/MS/MS and database searching at the low-femtomole level. *Anal Chem* 69, 767–776.
- Mead, J. A., Shadforth, I. P., and Bessant, C. (2007). Public proteomic MS repositories and pipelines: available tools and biological applications. *Proteomics* 7, 2769–2786.
- Medzihradszky, K. F., Campbell, J. M., Baldwin, M. A., Falick, A. M., Juhasz, P., Vestal, M. L., and Burlingame, A. L. (2000). The Characteristics of Peptide Collision-Induced

Dissociation Using a High-Performance MALDI-TOF/TOF Tandem Mass Spectrometer. *Anal Chem* 72, 552–558.

Meier, U. T., and Blobel, G. (1990). A nuclear localization signal binding protein in the nucleolus. *J Cell Biol* 111, 2235–2245.

Meilinger, M., Haumer, M., Szakmary, K. A., Steinböck, F., Scheiber, B., Goldenberg, H., and Huettlinger, M. (1995). Removal of lactoferrin from plasma is mediated by binding to low density lipoprotein receptor-related protein/alpha 2-macroglobulin receptor and transport to endosomes. *FEBS Lett* 360, 70–74.

Mendelev, N., Witherspoon, S., and Li, P. A. (2009). Overexpression of human selenoprotein H in neuronal cells ameliorates ultraviolet irradiation-induced damage by modulating cell signaling pathways. *Experimental Neurology* 220, 328–334.

Mendelev, N., Mehta, S. L., Witherspoon, S., He, Q., Sexton, J. Z., and Li, P. A. (2011). Upregulation of human selenoprotein H in murine hippocampal neuronal cells promotes mitochondrial biogenesis and functional performance. *Mitochondrion* 11, 76–82.

Meng, C. K., and Fenn, J. B. (1990). Analyzing organic molecules with electrospray mass spectrometry. *Am. Biotechnol. Lab.* 8, 54–60.

Metz-Boutigue, M. H., Jollès, J., Mazurier, J., Schoentgen, F., Legrand, D., Spik, G., Montreuil, J., and Jollès, P. (1984). Human lactotransferrin: amino acid sequence and structural comparisons with other transferrins. *Eur J Biochem* 145, 659–676.

Miyazaki, S. (2003). DNA Data Bank of Japan (DDBJ) in XML. *Nucleic Acids Res* 31, 13–16.

Mizutani, K., Toyoda, M., and Mikami, B. (2011). X-ray structures of transferrins and related proteins. *Biochimica et Biophysica Acta (BBA) - General Subjects*.

Molina, H., Yang, Y., Ruch, T., Kim, J., Mortensen, P., Otto, T., Nalli, A., Tang, Q., Lane, M. D., Chaerkady, R., et al. (2009). Temporal Profiling of the Adipocyte Proteome during Differentiation Using a Five-Plex SILAC Based Strategy. *J Proteome Res* 8, 48–58.

Monteoliva, L., and Albar, J. P. (2004). Differential proteomics: an overview of gel and non-gel based approaches. *Brief Funct Genomic Proteomic* 3, 220–239.

Montreuil, J., Tonnelat, J., and MULLET, S. (1960). Préparation et propriétés de la lactosidérophiline (lactotransferrine) du lait de femme. *Biochim Biophys Acta* 45, 413–421.

Moore, S. A., Anderson, B. F., Groom, C. R., Haridas, M., and Baker, E. N. (1997). Three-dimensional structure of diferric bovine lactoferrin at 2.8 Å resolution. *J. Mol. Biol.* 274, 222–236.

Moriuchi, M., and Moriuchi, H. (2006). Induction of Lactoferrin Gene Expression in Myeloid or Mammary Gland Cells by Human T-Cell Leukemia Virus Type 1 (HTLV-1) Tax: Implications for Milk-Borne Transmission of HTLV-1. *J. Virol.* 80, 7118–7126.

Mortensen, P., Gouw, J. W., Olsen, J. V., Ong, S., Rigbolt, K. T. G., Bunkenborg, J., Cox, J., Foster, L. J., Heck, A. J. R., Blagoev, B., et al. (2010). MSQuant, an open source platform for mass spectrometry-based quantitative proteomics. *J Proteome Res* 9, 393–403.

Mosquito, S., Ochoa, T. J., Cok, J., and Cleary, T. G. (2010). Effect of bovine lactoferrin in *Salmonella* ser. *Typhimurium* infection in mice. *Biometals* 23, 515–521.

- Mujtaba, S., He, Y., Zeng, L., Yan, S., Plotnikova, O., Sachchidanand, Sanchez, R., Zeleznik-Le, N. J., Ronai, Z., and Zhou, M. (2004). Structural mechanism of the bromodomain of the coactivator CBP in p53 transcriptional activation. *Molecular Cell* *13*, 251–263.
- Mulligan, P., White, N. R. J., Monteleone, G., Wang, P., Wilson, J. W., Ohtsuka, Y., and Sanderson, I. R. (2006). Breast milk lactoferrin regulates gene expression by binding bacterial DNA CpG motifs but not genomic DNA promoters in model intestinal cells. *Pediatr. Res.* *59*, 656–661.
- Nagaraj, N., D’Souza, R. C. J., Cox, J., Olsen, J. V., and Mann, M. (2010). Feasibility of Large-Scale Phosphoproteomics with Higher Energy Collisional Dissociation Fragmentation. *J Proteome Res* *9*, 6786–6794.
- Naik, M. U., Naik, T. U., Suckow, A. T., Duncan, M. K., and Naik, U. P. (2008). Attenuation of Junctional Adhesion Molecule-A Is a Contributing Factor for Breast Cancer Cell Invasion. *Cancer Res* *68*, 2194–2203.
- Nakajima, K., Nakamura, M., Gao, X., and Kozakai, T. (2008). Possible involvement of prolactin in the synthesis of lactoferrin in bovine mammary epithelial cells. *Biosci. Biotechnol. Biochem.* *72*, 1103–1106.
- Naot, D., Chhana, A., Matthews, B. G., Callon, K. E., Tong, P. C., Lin, J., Costa, J. L., Watson, M., Grey, A. B., and Cornish, J. (2011). Molecular mechanisms involved in the mitogenic effect of lactoferrin in osteoblasts. *Bone* *49*, 217–224.
- Nesvizhskii, A. I., Vitek, O., and Aebersold, R. (2007). Analysis and validation of proteomic data generated by tandem mass spectrometry. *Nat Meth* *4*, 787–797.
- Nguyen, S., and Fenn, J. B. (2007). Gas-phase ions of solute species from charged droplets of solutions. *Proc Natl Acad Sci USA* *104*, 1111–1117.
- Norrby, K. (2004). Human apo-lactoferrin enhances angiogenesis mediated by vascular endothelial growth factor A in vivo. *J. Vasc. Res.* *41*, 293–304.
- Norrby, K., Mattsby-Baltzer, I., Innocenti, M., and Tuneberg, S. (2001). Orally administered bovine lactoferrin systemically inhibits VEGF(165)-mediated angiogenesis in the rat. *Int J Cancer* *91*, 236–240.
- Novoselov, S. V., Kryukov, G. V., Xu, X., Carlson, B. A., Hatfield, D. L., and Gladyshev, V. N. (2007). Selenoprotein H is a nucleolar thioredoxin-like protein with a unique expression pattern. *J Biol Chem* *282*, 11960–11968.
- O’Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* *250*, 4007–4021.
- Obermajer, N., Premzl, A., Zavašnik Bergant, T., Turk, B., and Kos, J. (2006). Carboxypeptidase cathepsin X mediates β 2-integrin-dependent adhesion of differentiated U-937 cells. *Exp Cell Res* *312*, 2515–2527.
- Ochoa, T. J., Chea-Woo, E., Campos, M., Pecho, I., Prada, A., McMahon, R., and Clearly, T. G. (200). Impact of lactoferrin supplementation on growth and prevalence of Giardia colonization children. *Clin Infect Dis* *12*, 1881–1886.
- Ochoa, T. J., and Clearly, T. G. (2004). Lactoferrin disruption of bacterial type III secretion systems. *Biometals* *17*, 257–260.

- Oda, Y., Huang, K., Cross, F. R., Cowburn, D., and Chait, B. T. (1999). Accurate quantitation of protein expression and site-specific phosphorylation. *Proc Natl Acad Sci USA* *96*, 6591–6596.
- Oh, S., Pyo, C., Kim, Y., and Choi, S. (2004). Neutrophil lactoferrin upregulates the human p53 gene through induction of NF-κB activation cascade. *Oncogene* *23*, 8282–8291.
- Oh, S. M. (2001). Human Neutrophil Lactoferrin trans-Activates the Matrix Metalloproteinase 1 Gene through Stress-activated MAPK Signaling Modules. *Journal of Biological Chemistry* *276*, 42575–42579.
- Old, W. M., Meyer-Arendt, K., Aveline-Wolf, L., Pierce, K. G., Mendoza, A., Sevinsky, J. R., Resing, K. A., and Ahn, N. G. (2005). Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol Cell Proteomics* *4*, 1487–1502.
- Ong, S., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* *1*, 376–386.
- Ong, S., and Mann, M. (2005). Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.* *1*, 252–262.
- Paesano, R., Pietropaoli, M., Gessani, S., and Valenti, P. (2009). The influence of lactoferrin, orally administered, on systemic iron homeostasis in pregnant women suffering of iron deficiency and iron deficiency anaemia. *Biochimie* *91*, 44–51.
- Palumbo, A. M., Smith, S. A., Kalcic, C. L., Dantus, M., Stemmer, P. M., and Reid, G. E. (2011). Tandem mass spectrometry strategies for phosphoproteome analysis. *Mass Spectrom. Rev.* *30*, 600–625.
- Pan, C., Gnad, F., Olsen, J. V., and Mann, M. (2008). Quantitative phosphoproteome analysis of a mouse liver cell line reveals specificity of phosphatase inhibitors. *Proteomics* *8*, 4534–4546.
- Pan, S., Aebersold, R., Chen, R., Rush, J., Goodlett, D. R., McIntosh, M. W., Zhang, J., and Brentnall, T. A. (2009). Mass Spectrometry Based Targeted Protein Quantification: Methods and Applications. *J Proteome Res* *8*, 787–797.
- Pan, Y., Tsai, C., Ma, B., and Nussinov, R. (2010). Mechanisms of transcription factor selectivity. *Trends in Genetics* *26*, 75–83.
- Park, I., Schaeffer, E., Sidoli, A., Baralle, F. E., Cohen, G. N., and Zakin, M. M. (1985). Organization of the human transferrin gene: direct evidence that it originated by gene duplication. *Proc Natl Acad Sci USA* *82*, 3149–3153.
- Paša-Tolić, L., Jensen, P. K., Anderson, G. A., Lipton, M. S., Peden, K. K., Martinović, S., Tolić, N., Bruce, J. E., and Smith, R. D. (1999). High Throughput Proteome-Wide Precision Measurements of Protein Expression Using Mass Spectrometry. *J. Am. Chem. Soc.* *121*, 7949–7950.
- Penco, S., Scarfi, S., Giovine, M., Damonte, G., Millo, E., Villaggio, B., Passalacqua, M., Pozzolini, M., Garrè, C., and Benatti, U. (2001). Identification of an import signal for, and the nuclear localization of, human lactoferrin. *Biotechnol. Appl. Biochem.* *34*, 151–159.
- Pentecost, B. T., and Teng, C. T. (1987). Lactotransferrin is the major estrogen inducible

protein of mouse uterine secretions. *J Biol Chem* **262**, 10134–10139.

Perkins, D. N., Pappin, D. J., Creasy, D. M., and Cottrell, J. S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567.

Pflieger, D., Jünger, M. A., Müller, M., Rinner, O., Lee, H., Gehrig, P. M., Gstaiger, M., and Aebersold, R. (2008). Quantitative proteomic analysis of protein complexes: concurrent identification of interactors and their state of phosphorylation. *Molecular & Cellular Proteomics* **7**, 326–346.

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

Pierce, A., Colavizza, D., Benaïssa, M., Maes, P., Tartar, A., Montreuil, J., and Spik, G. (1991). Molecular cloning and sequence analysis of bovine lactotransferrin. *Eur J Biochem* **196**, 177–184.

Pietrantoni, A., Di Biase, A. M., Tinari, A., Marchetti, M., Valenti, P., Seganti, L., and Superti, F. (2003). Bovine lactoferrin inhibits adenovirus infection by interacting with viral structural polypeptides. *Antimicrob. Agents Chemother.* **47**, 2688–2691.

Plaut, A. G., Qiu, J., and St Geme, J. W. (2000). Human lactoferrin proteolytic activity: analysis of the cleaved region in the IgA protease of *Haemophilus influenzae*. *Vaccine* **19 Suppl 1**, S148–52.

Powell, M. J., and Ogden, J. E. (1990). Nucleotide sequence of human lactoferrin cDNA. *Nucleic Acids Res* **18**, 4013.

Pownall, H. J., and Ehnholm, C. (2006). The unique role of apolipoprotein A-I in HDL remodeling and metabolism. *Curr. Opin. Lipidol.* **17**, 209–213.

Pratt, J. M., Simpson, D. M., Doherty, M. K., Rivers, J., Gaskell, S. J., and Beynon, R. J. (2006). Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. *Nat Protoc* **1**, 1029–1043.

Prestage, J., Williams, A., Maleki, L., Djomehri, M., and Harabetian, E. (1991). Dynamics of charged particles in a Paul radio-frequency quadrupole trap. *Phys. Rev. Lett.* **66**, 2964–2967.

Pruitt, K. D., Tatusova, T., Klimke, W., and Maglott, D. R. (2009). NCBI Reference Sequences: current status, policy and new initiatives. *Nucleic Acids Res* **37**, D32–D36.

Puddu, P., Borghi, P., GESSIONI, S., Valenti, P., Belardelli, F., and Seganti, L. (1998). Antiviral effect of bovine lactoferrin saturated with metal ions on early steps of human immunodeficiency virus type 1 infection. *Int J Biochem Cell Biol* **30**, 1055–1062.

Puddu, P., Latorre, D., Carollo, M., Catizone, A., Ricci, G., Valenti, P., and Gessani, S. (2011). Bovine Lactoferrin Counteracts Toll-Like Receptor Mediated Activation Signals in Antigen Presenting Cells. *PLoS ONE* **6**, e22504.

Qiu, J., Hendrixson, D. R., Baker, E. N., Murphy, T. F., St Geme, J. W., and Plaut, A. G. (1998). Human milk lactoferrin inactivates two putative colonization factors expressed by *Haemophilus influenzae*. *Proc Natl Acad Sci USA* **95**, 12641–12646.

Rabilloud, T. (2002). Two-dimensional gel electrophoresis in proteomics: old, old fashioned,

but it still climbs up the mountains. *Proteomics* **2**, 3–10.

- Rabilloud, T., Chevallet, M., Luche, S., and Lelong, C. (2008). Fully denaturing two-dimensional electrophoresis of membrane proteins: A critical update. *Proteomics* **8**, 3965–3973.
- Rado, T. A., Wei, X. P., and Benz, E. J. (1987). Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA during normal and leukemic myelopoiesis. *Blood* **70**, 989–993.
- Rappsilber, J., Ryder, U., Lamond, A. I., and Mann, M. (2002). Large-scale proteomic analysis of the human spliceosome. *Genome Res* **12**, 1231–1245.
- Ray, P. D., Huang, B., and Tsuji, Y. (2012). Reactive oxygen species (ROS) homeostasis and redox regulation in cellular signaling. *Cellular Signalling*.
- Rechsteiner, M., and Rogers, S. W. (1996). PEST sequences and regulation by proteolysis. *Trends Biochem. Sci.* **21**, 267–271.
- Redeker, V. (2010). *Methods in Cell Biology* (Elsevier).
- Rey, M. W., Woloshuk, S. L., deBoer, H. A., and Pieper, F. R. (1990). Complete nucleotide sequence of human mammary gland lactoferrin. *Nucleic Acids Res* **18**, 5288.
- Ricquier, D. (2005). Maladies héréditaires du métabolisme et apports de la métabolomique. *Med Sci (Paris)* **21**, 512–516.
- Riter, L. S., Gooding, K. M., Hodge, B. D., and Julian, R. K. (2006). Comparison of the Paul ion trap to the linear ion trap for use in global proteomics. *Proteomics* **6**, 1735–1740.
- Rivers, J., Simpson, D. M., Robertson, D. H. L., Gaskell, S. J., and Beynon, R. J. (2007). Absolute multiplexed quantitative analysis of protein expression during muscle development using QconCAT. *Mol Cell Proteomics* **6**, 1416–1427.
- Robbins, J., Dilworth, S. M., Laskey, R. A., and 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.
- Rochard, E., Legrand, D., Lecocq, M., Hamelin, R., Crepin, M., Montreuil, J., and Spik, G. (1992). Characterization of lactotransferrin receptor in epithelial cell lines from non-malignant human breast, benign mastopathies and breast carcinomas. *Anticancer Res.* **12**, 2047–2051.
- Ross, P. L. (2004). Multiplexed Protein Quantitation in *Saccharomyces cerevisiae* Using Amine-reactive Isobaric Tagging Reagents. *Molecular & Cellular Proteomics* **3**, 1154–1169.
- Rothrock, C. R., House, A. E., and Lynch, K. W. (2005). HnRNP L represses exon splicing via a regulated exonic splicing silencer. *EMBO J.* **24**, 2792–2802.
- Roychoudhury, P., Paul, R. R., Chowdhury, R., and Chaudhuri, K. (2007). HnRNP E2 is downregulated in human oral cancer cells and the overexpression of hnRNP E2 induces apoptosis. *Mol. Carcinog.* **46**, 198–207.
- Sachs, M. S., and Geballe, A. P. (2006). Downstream control of upstream open reading frames. *Genes Dev.* **20**, 915–921.

- Saidi, H., Eslahpazir, J., Eslaphazir, J., Carboneil, C., Carthagena, L., Requena, M., Nassreddine, N., and Belec, L. (2006). Differential modulation of human lactoferrin activity against both R5 and X4-HIV-1 adsorption on epithelial cells and dendritic cells by natural antibodies. *J Immunol* 177, 5540–5549.
- Sallmann, F. R., Baveye-Descamps, S., Pattus, F., Salmon, V., Branza, N., Spik, G., and Legrand, D. (1999). Porins OmpC and PhoE of Escherichia coli as specific cell-surface targets of human lactoferrin. Binding characteristics and biological effects. *J Biol Chem* 274, 16107–16114.
- Sawatzki, G., and Rich, I. N. (1989). Lactoferrin stimulates colony stimulating factor production in vitro and in vivo. *Blood Cells* 15, 371–385.
- Schevzov, G., Vrhovski, B., Bryce, N. S., Elmira, S., Qiu, M. R., O'neill, G. M., Yang, N., Verrills, N. M., Kavallaris, M., and Gunning, P. W. (2005). Tissue-specific tropomyosin isoform composition. *Journal of Histochemistry & Cytochemistry* 53, 557–570.
- Schmidt, A., Kellermann, J., and Lottspeich, F. (2005). A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* 5, 4–15.
- Schneider, E., and Ryan, T. J. (2006). Gamma-glutamyl hydrolase and drug resistance. *Clin Chim Acta* 374, 25–32.
- Schulze, W. X., and Mann, M. (2004). A novel proteomic screen for peptide-protein interactions. *J Biol Chem* 279, 10756–10764.
- Schwartz, J. C., Senko, M. W., and Syka, J. E. P. (2002). A two-dimensional quadrupole ion trap mass spectrometer. *J. Am. Soc. Mass Spectrom.* 13, 659–669.
- Schwerin, M., Solinas Toldo, S., Eggen, A., Brunner, R., Seyfert, H. M., and Fries, R. (1994). The bovine lactoferrin gene (LTF) maps to chromosome 22 and syntenic group U12. *Mamm. Genome* 5, 486–489.
- Scigelova, M., and Makarov, A. (2006). Orbitrap mass analyzer--overview and applications in proteomics. *Proteomics* 6 Suppl 2, 16–21.
- Seyfried, N. T., Gozal, Y. M., Dammer, E. B., Xia, Q., Duong, D. M., Cheng, D., Lah, J. J., Levey, A. I., and Peng, J. (2010). Multiplex SILAC Analysis of a Cellular TDP-43 Proteinopathy Model Reveals Protein Inclusions Associated with SUMOylation and Diverse Polyubiquitin Chains. *Molecular & Cellular Proteomics* 9, 705–718.
- Shaheduzzaman, S., Vishwanath, A., Furusato, B., Cullen, J., Chen, Y., Bañez, L., Nau, M., Ravindranath, L., Kim, K., Mohammed, A., et al. (2007). Silencing of Lactotransferrin expression by methylation in prostate cancer progression. *Cancer Biol. Ther.* 6, 1088–1095.
- Shen, Y., Zhao, R., Berger, S. J., Anderson, G. A., Rodriguez, N., and Smith, R. D. (2002). High-Efficiency Nanoscale Liquid Chromatography Coupled On-Line with Mass Spectrometry Using Nanoelectrospray Ionization for Proteomics. *Anal Chem* 74, 4235–4249.
- Shevchenko, A., Sunyaev, S., Loboda, A., Shevchenko, A., Bork, P., Ens, W., and Standing, K. G. (2001). Charting the Proteomes of Organisms with Unsequenced Genomes by MALDI-Quadrupole Time-of-Flight Mass Spectrometry and BLAST Homology Searching. *Anal Chem* 73, 1917–1926.

- Shimamura, M., Yamamoto, Y., Ashino, H., Oikawa, T., Hazato, T., Tsuda, H., and Iigo, M. (2004). Bovine lactoferrin inhibits tumor-induced angiogenesis. *Int J Cancer* *111*, 111–116.
- Shimazaki, K., and Kushida, T. (2010). A preliminary approach to creating an overview of lactoferrin multi-functionality utilizing a text mining method. *Biometals* *23*, 453–463.
- Shimizu, K., Matsuzawa, H., Okada, K., Tazume, S., Dosako, S., Kawasaki, Y., Hashimoto, K., and Koga, Y. (1996). Lactoferrin-mediated protection of the host from murine cytomegalovirus infection by a T-cell-dependent augmentation of natural killer cell activity. *Arch. Virol.* *141*, 1875–1889.
- Shinoda, I., TAKASE, M., Fukuwatari, Y., Shimamura, S., Köller, M., and König, W. (1996). Effects of lactoferrin and lactoferricin on the release of interleukin 8 from human polymorphonuclear leukocytes. *Biosci. Biotechnol. Biochem.* *60*, 521–523.
- Shirsat, N. V., Bittenbender, S., Kreider, B. L., and Rovera, G. (1992). Structure of the murine lactotransferrin gene is similar to the structure of other transferrin-encoding genes and shares a putative regulatory region with the murine myeloperoxidase gene. *Gene* *110*, 229–233.
- Shubbar, E., Vegfors, J., Carlström, M., Petersson, S., and Enerbäck, C. (2011). Psoriasin (S100A7) increases the expression of ROS and VEGF and acts through RAGE to promote endothelial cell proliferation. *Breast Cancer Res Treat.*
- Siciliano, R., Rega, B., Marchetti, M., Seganti, L., Antonini, G., and Valenti, P. (1999). Bovine lactoferrin peptidic fragments involved in inhibition of herpes simplex virus type 1 infection. *Biochem Biophys Res Commun* *264*, 19–23.
- Siebert, P. D., and 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 USA* *94*, 2198–2203.
- Silva, J. C. (2005). Absolute Quantification of Proteins by LCMS: A Virtue of Parallel ms Acquisition. *Molecular & Cellular Proteomics* *5*, 144–156.
- Singh, P. K., Parsek, M. R., Greenberg, E. P., and Welsh, M. J. (2002). A component of innate immunity prevents bacterial biofilm development. *Nature* *417*, 552–555.
- Son, H., Lee, S., and Choi, S. (2006). Human lactoferrin controls the level of retinoblastoma protein and its activity. *Biochem Cell Biol* *84*, 345–350.
- Son, K., Park, J., Chung, C., Chung, D. K., Yu, D., Lee, K., and Kim, J. (2002). Human lactoferrin activates transcription of IL-1beta gene in mammalian cells. *Biochem Biophys Res Commun* *290*, 236–241.
- Sorimachi, K., Akimoto, K., Hattori, Y., Ieiri, T., and Niwa, A. (1997). Activation of macrophages by lactoferrin: secretion of TNF-alpha, IL-8 and NO. *Biochem. Mol. Biol. Int.* *43*, 79–87.
- Spik, G., Coddeville, B., and Montreuil, J. (1988). Comparative study of the primary structures of sero-, lacto- and ovotransferrin glycans from different species. *Biochimie* *70*, 1459–1469.
- Spik, G., Strecker, G., Fournet, B., Bouquelet, S., Montreuil, J., Dorland, L., van Halbeek, H., and Vliegenthart, J. F. (1982). Primary structure of the glycans from human lactotransferrin. *Eur J Biochem* *121*, 413–419.

- Srivastava, C. H., Rado, T. A., Bauerle, D., and Broxmeyer, H. E. (1991). Regulation of human bone marrow lactoferrin and myeloperoxidase gene expression by tumor necrosis factor-alpha. *J Immunol* *146*, 1014–1019.
- Stafford, G. (2002). Ion trap mass spectrometry: a personal perspective. *J. Am. Soc. Mass Spectrom.* *13*, 589–596.
- Stahl-Zeng, J., Lange, V., Ossola, R., Eckhardt, K., Krek, W., Aebersold, R., and Domon, B. (2007). High Sensitivity Detection of Plasma Proteins by Multiple Reaction Monitoring of N-Glycosites. *Molecular & Cellular Proteomics* *6*, 1809–1817.
- Stewart, I. I., Thomson, T., and Figgeys, D. (2001). ¹⁸O Labeling: a tool for proteomics. *Rapid Commun. Mass Spectrom.* *15*, 2456–2465.
- Stine, M. J., Wang, C. J., Moriarty, W. F., Ryu, B., Cheong, R., Westra, W. H., Levchenko, A., and Alani, R. M. (2011). Integration of Genotypic and Phenotypic Screening Reveals Molecular Mediators of Melanoma-Stromal Interaction. *Cancer Res* *71*, 2433–2444.
- Stoesser, G. (2003). The EMBL Nucleotide Sequence Database: major new developments. *Nucleic Acids Res* *31*, 17–22.
- Stokes, K. (2004). Estrogen response element and the promoter context of the human and mouse lactoferrin genes influence estrogen receptor -mediated transactivation activity in mammary gland cells. *Journal of Molecular Endocrinology* *33*, 315–334.
- Strande, V., Canelle, L., Tastet, C., Burlet-Schiltz, O., Monsarrat, B., and Hondermarck, H. (2009). The proteome of the human breast cancer cell line MDA-MB-231: Analysis by LTQ-Orbitrap mass spectrometry. *Prot. Clin. Appl.* *3*, 41–50.
- Stroobant, E. D. H. A. V. (2007). Mass Spectrometry Principles and Applications. 1–502.
- Superti, F., Ammendolia, M. G., Valenti, P., and Seganti, L. (1997). Antirotaviral activity of milk proteins: lactoferrin prevents rotavirus infection in the enterocyte-like cell line HT-29. *Med. Microbiol. Immunol.* *186*, 83–91.
- Suzuki, Y. A., Lopez, V., and Lönnnerdal, B. (2005). Mammalian lactoferrin receptors: structure and function. *Cell Mol Life Sci* *62*, 2560–2575.
- Suzuki, Y. A., Shin, K., and Lönnnerdal, B. (2001). Molecular cloning and functional expression of a human intestinal lactoferrin receptor. *Biochemistry* *40*, 15771–15779.
- Swart, P. J., Kuipers, M. E., Smit, C., Pauwels, R., deBéthune, M. P., de Clercq, E., Meijer, D. K., and Huisman, J. G. (1996). Antiviral effects of milk proteins: acylation results in polyanionic compounds with potent activity against human immunodeficiency virus types 1 and 2 in vitro. *AIDS Res. Hum. Retroviruses* *12*, 769–775.
- Tabb, D. L., Saraf, A., and Yates, J. R. (2003). GutenTag: High-Throughput Sequence Tagging via an Empirically Derived Fragmentation Model. *Anal Chem* *75*, 6415–6421.
- Takakura, N., Wakabayashi, H., Ishibashi, H., Teraguchi, S., Tamura, Y., Yamaguchi, H., and Abe, S. (2003). Oral lactoferrin treatment of experimental oral candidiasis in mice. *Antimicrob. Agents Chemother.* *47*, 2619–2623.
- Takayama, Y., Takahashi, H., Mizumachi, K., and Takezawa, T. (2003). Low density lipoprotein receptor-related protein (LRP) is required for lactoferrin-enhanced collagen gel contractile activity of human fibroblasts. *J Biol Chem* *278*, 22112–22118.

- Takayama, Y., and Mizumachi, K. (2009). Effect of lactoferrin-embedded collagen membrane on osteogenic differentiation of human osteoblast-like cells. *J. Biosci. Bioeng.* **107**, 191–195.
- Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., and Yoshida, T. (1988). Protein and Polymer Analyses up to m/z 100 000 by Laser Ionization Time-of flight Mass Spectrometry. *Rapid communication in mass spectrometry* **2**, 151–153.
- Tanaka, T., Kawabata, K., Kohno, H., Honjo, S., Murakami, M., Ota, T., and Tsuda, H. (2000). Chemopreventive effect of bovine lactoferrin on 4-nitroquinoline 1-oxide-induced tongue carcinogenesis in male F344 rats. *Jpn. J. Cancer Res.* **91**, 25–33.
- Tanida, T., Okamoto, T., Okamoto, A., Wang, H., Hamada, T., Ueta, E., and Osaki, T. (2003). Decreased excretion of antimicrobial proteins and peptides in saliva of patients with oral candidiasis. *J. Oral Pathol. Med.* **32**, 586–594.
- Tanner, S., Shu, H., Frank, A., Wang, L., Zandi, E., Mumby, M., Pevzner, P. A., and Bafna, V. (2005). InsPecT: Identification of Posttranslationally Modified Peptides from Tandem Mass Spectra. *Anal Chem* **77**, 4626–4639.
- Tasaki, T., Zakrzewska, A., Dudgeon, D. D., Jiang, Y., Lazo, J. S., and Kwon, Y. T. (2009). The substrate recognition domains of the N-end rule pathway. *J Biol Chem* **284**, 1884–1895.
- Teng, C., Gladwell, W., Raphiou, I., and Liu, E. (2004a). Methylation and expression of the lactoferrin gene in human tissues and cancer cells. *Biometals* **17**, 317–323.
- Teng, C., Gladwell, W., Raphiou, I., and Liu, E. (2004b). Methylation and expression of the lactoferrin gene in human tissues and cancer cells. *Biometals* **17**, 317–323.
- Teng, C. T. (2002). Lactoferrin gene expression and regulation: an overview. *Biochem Cell Biol* **80**, 7–16.
- Teng, C. T., Pentecost, B. T., Marshall, A., Solomon, A., Bowman, B. H., Lalley, P. A., and Naylor, S. L. (1987). Assignment of the lactotransferrin gene to human chromosome 3 and to mouse chromosome 9. *Somat. Cell Mol. Genet.* **13**, 689–693.
- Teng, C. T., Beard, C., and Gladwell, W. (2002). Differential expression and estrogen response of lactoferrin gene in the female reproductive tract of mouse, rat, and hamster. *Biol. Reprod.* **67**, 1439–1449.
- Teng, C. T., Liu, Y., Yang, N., Walmer, D., and Panella, T. (1992). Differential molecular mechanism of the estrogen action that regulates lactoferrin gene in human and mouse. *Mol. Endocrinol.* **6**, 1969–1981.
- Thiede, B., Höhenwarter, W., Krah, A., Mattow, J., Schmid, M., Schmidt, F., and Jungblut, P. R. (2005). Peptide mass fingerprinting. *Methods* **35**, 237–247.
- Thompson, A., Schäfer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., and Hamon, C. (2003). Tandem Mass Tags: A Novel Quantification Strategy for Comparative Analysis of Complex Protein Mixtures by MS/MS. *Anal Chem* **75**, 1895–1904.
- Togawa, J., Nagase, H., Tanaka, K., Inamori, M., Nakajima, A., Ueno, N., Saito, T., and Sekihara, H. (2002). Oral administration of lactoferrin reduces colitis in rats via modulation of the immune system and correction of cytokine imbalance. *J. Gastroenterol. Hepatol.* **17**, 1291–1298.

- Tsuda, H., Sekine, K., Fujita, K., and Iigo, M. (2002). Cancer prevention by bovine lactoferrin and underlying mechanisms--a review of experimental and clinical studies. *Biochem Cell Biol* **80**, 131–136.
- Tsuda, H., Ohshima, Y., Nomoto, H., Fujita, K., Matsuda, E., Iigo, M., Takasuka, N., and Moore, M. A. (2004). Cancer prevention by natural compounds. *Drug Metab. Pharmacokinet.* **19**, 245–263.
- Tsuda, H., Sekine, K., Nakamura, J., Ushida, Y., Kuhara, T., Takasuka, N., Kim, D. J., Asamoto, M., Baba-Toriyama, H., Moore, M. A., et al. (1998). Inhibition of azoxymethane initiated colon tumor and aberrant crypt foci development by bovine lactoferrin administration in F344 rats. *Adv Exp Med Biol* **443**, 273–284.
- Tsuda, H., Sekine, K., Takasuka, N., Toriyama-Baba, H., and Iigo, M. (2000). Prevention of colon carcinogenesis and carcinoma metastasis by orally administered bovine lactoferrin in animals. *Biofactors* **12**, 83–88.
- Tuccari, G., Villari, D., Giuffrè, G., Simone, A., Squadrito, G., Raimondo, G., and Barresi, G. (2002). Immunohistochemical evidence of lactoferrin in hepatic biopsies of patients with viral or cryptogenetic chronic liver disease. *Histol. Histopathol.* **17**, 1077–1083.
- Ushida, Y., Sekine, K., Kuhara, T., Takasuka, N., Iigo, M., and Tsuda, H. (1998). Inhibitory effects of bovine lactoferrin on intestinal polyposis in the Apc(Min) mouse. *Cancer Lett.* **134**, 141–145.
- Ushida, Y., Sekine, K., Kuhara, T., Takasuka, N., Iigo, M., Maeda, M., and Tsuda, H. (1999). Possible chemopreventive effects of bovine lactoferrin on esophagus and lung carcinogenesis in the rat. *Jpn. J. Cancer Res.* **90**, 262–267.
- van Berkel, P. H., Geerts, M. E., van Veen, H. A., Mericksay, M., de Boer, H. A., and 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)**, 145–151.
- van Dijk, E., Le Hir, H., and Séraphin, B. (2003). DcpS can act in the 5'-3' mRNA decay pathway in addition to the 3“-5” pathway. *Proc Natl Acad Sci USA* **100**, 12081–12086.
- van Veen, H. A., Geerts, M. E. J., van Berkel, P. H. C., and Nuijens, J. H. (2004). The role of N-linked glycosylation in the protection of human and bovine lactoferrin against tryptic proteolysis. *Eur J Biochem* **271**, 678–684.
- van Vugt, H., van Gool, J., Ladiges, N. C., and Boers, W. (1975). Lactoferrin in rabbit bile: its relation to iron metabolism. *Q J Exp Physiol Cogn Med Sci* **60**, 79–88.
- Varadhachary, A., Wolf, J. S., Petrak, K., O'Malley, B. W., Spadaro, M., Curcio, C., Forni, G., and Pericle, F. (2004). Oral lactoferrin inhibits growth of established tumors and potentiates conventional chemotherapy. *Int J Cancer* **111**, 398–403.
- Viseux, N., de Hoffmann, E., and Domon, B. (1998). Structural assignment of permethylated oligosaccharide subunits using sequential tandem mass spectrometry. *Anal Chem* **70**, 4951–4959.
- Viswanathan, C. T., Bansal, S., Booth, B., DeStefano, A. J., Rose, M. J., Sailstad, J., Shah, V. P., Skelly, J. P., Swann, P. G., and Weiner, R. (2007). Quantitative Bioanalytical Methods Validation and Implementation: Best Practices for Chromatographic and Ligand Binding Assays. *Pharm Res* **24**, 1962–1973.

- Vlodavsky, I., Friedmann, Y., Elkin, M., Aingorn, H., Atzmon, R., Ishai-Michaeli, R., Bitan, M., Pappo, O., Peretz, T., Michal, I., et al. (1999). Mammalian heparanase: gene cloning, expression and function in tumor progression and metastasis. *Nat. Med.* *5*, 793–802.
- Wakabayashi, H., Abe, S., Okutomi, T., Tansho, S., Kawase, K., and Yamaguchi, H. (1996). Cooperative anti-Candida effects of lactoferrin or its peptides in combination with azole antifungal agents. *Microbiol. Immunol.* *40*, 821–825.
- Wakabayashi, H., Kurokawa, M., Shin, K., Teraguchi, S., Tamura, Y., and Shiraki, K. (2004). Oral lactoferrin prevents body weight loss and increases cytokine responses during herpes simplex virus type 1 infection of mice. *Biosci. Biotechnol. Biochem.* *68*, 537–544.
- Wakabayashi, H., Kondo, I., Kobayashi, T., Yamauchi, K., Toida, T., Iwatsuki, K., and Yoshie, H. (2010). Periodontitis, periodontopathic bacteria and lactoferrin. *Biometals* *23*, 419–424.
- Walther, T. C., and Mann, M. (2010). Mass spectrometry-based proteomics in cell biology. *J Cell Biol* *190*, 491–500.
- Wang, J., Li, Q., Ou, Y., Han, Z., Li, K., Wang, P., and Zhou, S. (2011a). Inhibition of tumor growth by recombinant adenovirus containing human lactoferrin through inducing tumor cell apoptosis in mice bearing EMT6 breast cancer. *Arch. Pharm. Res.* *34*, 987–995.
- Wang, J., Chen, L., Li, Y., and Guan, X. (2011b). Overexpression of Cathepsin Z Contributes to Tumor Metastasis by Inducing Epithelial-Mesenchymal Transition in Hepatocellular Carcinoma. *PLoS ONE* *6*, e24967.
- Ward, P. P., Paz, E., and Conneely, O. M. (2005). Lactoferrin. *Cell Mol Life Sci* *62*, 2540–2548.
- Wei, Y., Tong, J., Taylor, P., Strumpf, D., Ignatchenko, V., Pham, N., Yanagawa, N., Liu, G., Jurisica, I., Shepherd, F. A., et al. (2011). Primary tumor xenografts of human lung adenocarcinoma and squamous cell carcinoma express distinct proteomic signatures. *J Proteome Res* *10*, 161–174.
- Weinberg, E. D. (1977). Infection and iron metabolism. *Am. J. Clin. Nutr.* *30*, 1485–1490.
- Weinberg, E. D. (2004). Suppression of bacterial biofilm formation by iron limitation. *Med. Hypotheses* *63*, 863–865.
- Weinberg, E. D. (1993). The development of awareness of iron-withholding defense. *Perspect. Biol. Med.* *36*, 215–221.
- Wells, J. M., Badman, E. R., and Cooks, R. G. (1998). A quadrupole ion trap with cylindrical geometry operated in the mass-selective instability mode. *Anal Chem* *70*, 438–444.
- Weng, T., Chen, L., Shyu, H., Chen, S., Wang, J., Yu, C., Lei, H., and Yeh, T. (2005). Lactoferrin inhibits enterovirus 71 infection by binding to VP1 protein and host cells. *Antiviral Research* *67*, 31–37.
- Westbrook, J., Feng, Z., Chen, L., Yang, H., and Berman, H. M. (2003). The Protein Data Bank and structural genomics. *Nucleic Acids Res* *31*, 489–491.
- Whiteaker, J. R., Zhao, L., Zhang, H. Y., Feng, L., Piening, B. D., Anderson, L., and Paulovich, A. G. (2007). Antibody-based enrichment of peptides on magnetic beads for mass-

- spectrometry-based quantification of serum biomarkers. *Anal Biochem* 362, 44–54.
- Wichmann, L., Vaalasti, A., Vaalasti, T., and Tuohimaa, P. (1989). Localization of lactoferrin in the male reproductive tract. *Int. J. Androl.* 12, 179–186.
- Wiesner, J., Premsler, T., and Sickmann, A. (2008). Application of electron transfer dissociation (ETD) for the analysis of posttranslational modifications. *Proteomics* 8, 4466–4483.
- Wildgruber, R., Harder, A., Obermaier, C., Boguth, G., Weiss, W., Fey, S. J., Larsen, P. M., and Görg, A. (2000). Towards higher resolution: two-dimensional electrophoresis of *Saccharomyces cerevisiae* proteins using overlapping narrow immobilized pH gradients. *Electrophoresis* 21, 2610–2616.
- Wilkins, M. R., Sanchez, J. C., Gooley, A. A., Appel, R. D., Humphery-Smith, I., Hochstrasser, D. F., and Williams, K. L. (1996). Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnol. Genet. Eng. Rev.* 13, 19–50.
- Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., and Mann, M. (1996). Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* 379, 466–469.
- Wilm, M., and Mann, M. (1996). Analytical properties of the nanoelectrospray ion source. *Anal Chem* 68, 1–8.
- Wolf, J. S., Li, G., Varadhachary, A., Petrak, K., Schneyer, M., Li, D., Ongkasuwan, J., Zhang, X., Taylor, R. J., Strome, S. E., et al. (2007). Oral Lactoferrin Results in T Cell-Dependent Tumor Inhibition of Head and Neck Squamous Cell Carcinoma In vivo. *Clinical Cancer Research* 13, 1601–1610.
- Wolters, D. A., Washburn, M. P., and Yates, J. R. (2001). An automated multidimensional protein identification technology for shotgun proteomics. *Anal Chem* 73, 5683–5690.
- Wu, C. H., Yeh, L. L., Huang, H., Arminski, L., Castro-Alvear, J., Chen, Y., Hu, Z., Kourtesis, P., Ledley, R. S., Suzek, B. E., et al. (2003). The Protein Information Resource. *Nucleic Acids Res* 31, 345–347.
- Wu, W. W., Wang, G., Baek, S. J., and Shen, R. (2006). Comparative study of three proteomic quantitative methods, DIGE, cICAT, and iTRAQ, using 2D gel- or LC-MALDI TOF/TOF. *J Proteome Res* 5, 651–658.
- Wu, Y., Le, W., and Jankovic, J. (2011). Preclinical biomarkers of Parkinson disease. *Arch. Neurol.* 68, 22–30.
- Xiao, Y. (2004). Lactoferrin Down-Regulates G1 Cyclin-Dependent Kinases during Growth Arrest of Head and Neck Cancer Cells. *Clinical Cancer Research* 10, 8683–8686.
- Xu, X. X., Jiang, H. R., Li, H. B., Zhang, T. N., Zhou, Q., and Liu, N. (2010). Apoptosis of stomach cancer cell SGC-7901 and regulation of Akt signaling way induced by bovine lactoferrin. *J Dairy Sci* 93, 2344–2350.
- Xu, Y. Y., Samaranayake, Y. H., Samaranayake, L. P., and Nikawa, H. (1999). In vitro susceptibility of *Candida* species to lactoferrin. *Med. Mycol.* 37, 35–41.
- Yamaguchi, H., Abe, S., and Takakura, N. (2004). Potential usefulness of bovine lactoferrin

- for adjunctive immunotherapy for mucosal Candida infections. *Biometals* *17*, 245–248.
- Yamauchi, K., Tomita, M., Giehl, T. J., and Ellison, R. T. (1993). Antibacterial activity of lactoferrin and a pepsin-derived lactoferrin peptide fragment. *Infect Immun* *61*, 719–728.
- Yang, D., la Rosa, de, G., Tewary, P., and Oppenheim, J. J. (2009). Alarmins link neutrophils and dendritic cells. *Trends Immunol.* *30*, 531–537.
- Yang, N., Shigeta, H., Shi, H., and Teng, C. T. (1996). Estrogen-related receptor, hERR1, modulates estrogen receptor-mediated response of human lactoferrin gene promoter. *J Biol Chem* *271*, 5795–5804.
- Yang, X., and Seto, E. (2008). Lysine Acetylation: Codified Crosstalk with Other Posttranslational Modifications. *Molecular Cell* *31*, 449–461.
- Yao, M., Ma, L., Humphreys, W. G., and Zhu, M. (2008). Rapid screening and characterization of drug metabolites using a multiple ion monitoring-dependent MS/MS acquisition method on a hybrid triple quadrupole-linear ion trap mass spectrometer. *J Mass Spectrom* *43*, 1364–1375.
- Yates, J. R., Speicher, S., Griffin, P. R., and Hunkapiller, T. (1993). Peptide mass maps: a highly informative approach to protein identification. *Anal Biochem* *214*, 397–408.
- Yates, J. R., Cociorva, D., Liao, L., and Zabrouskov, V. (2006). Performance of a Linear Ion Trap-Orbitrap Hybrid for Peptide Analysis. *Anal Chem* *78*, 493–500.
- Yi, M., Kaneko, S., Yu, D. Y., and Murakami, S. (1997). Hepatitis C virus envelope proteins bind lactoferrin. *J. Virol.* *71*, 5997–6002.
- Yoshida, A., Rzhetsky, A., Hsu, L. C., and Chang, C. (1998). Human aldehyde dehydrogenase gene family. *Eur J Biochem* *251*, 549–557.
- Zamboni, A., Di Carli, M., Guzzo, F., Stocchero, M., Zenoni, S., Ferrarini, A., Tononi, P., Toffali, K., Desiderio, A., Lilley, K. S., et al. (2010). Identification of Putative Stage-Specific Grapevine Berry Biomarkers and Omics Data Integration into Networks. *PLANT PHYSIOLOGY* *154*, 1439–1459.
- Zemann, N., Klein, P., Wetzel, E., Huettinger, F., and Huettinger, M. (2010). Lactoferrin induces growth arrest and nuclear accumulation of Smad-2 in HeLa cells. *Biochimie* *92*, 880–884.
- Zhang, H., Feng, X., Liu, W., Jiang, X., Shan, W., Huang, C., Yi, H., Zhu, B., Zhou, W., Wang, L., et al. (2011). Underlying mechanisms for LTF inactivation and its functional analysis in nasopharyngeal carcinoma cell lines. *J. Cell. Biochem.* *112*, 1832–1843.
- Zhang, R., Sioma, C. S., Thompson, R. A., Xiong, L., and Regnier, F. E. (2002). Controlling deuterium isotope effects in comparative proteomics. *Anal Chem* *74*, 3662–3669.
- Zhang, W., and Chait, B. T. (2000). ProFound: An Expert System for Protein Identification Using Mass Spectrometric Peptide Mapping Information. *Anal Chem* *72*, 2482–2489.
- Zheng, J., Ather, J. L., Sonstegard, T. S., and Kerr, D. E. (2005). Characterization of the infection-responsive bovine lactoferrin promoter. *Gene* *353*, 107–117.
- Zhi, W., Wang, M., and She, J. (2011). Selected reaction monitoring (SRM) mass spectrometry without isotope labeling can be used for rapid protein quantification. *Rapid*

Commun. Mass Spectrom. 25, 1583–1588.

- Ziere, G. J., Bijsterbosch, M. K., and van Berkel, T. J. (1993). Removal of 14 N-terminal amino acids of lactoferrin enhances its affinity for parenchymal liver cells and potentiates the inhibition of beta- very low density lipoprotein binding. *J Biol Chem* 268, 27069–27075.
- Zimecki, M., Mazurier, J., Spik, G., and Kapp, J. A. (1995). Human lactoferrin induces phenotypic and functional changes in murine splenic B cells. *Immunology* 86, 122–127.
- Zimecki, M., Chodaczek, G., Kocieba, M., and Kruzel, M. L. (2004). Lethality in LPS-induced endotoxemia in C3H/HeCr mice is associated with prevalence of proinflammatory cytokines: lack of protective action of lactoferrin. *FEMS Immunol. Med. Microbiol.* 42, 167–172.
- Zuccotti, G. V., Vigano, A., Borelli, M., Saresella, M., Giacomet, V., and Clerici, M. (2007). Modulation of innate and adaptive immunity by lactoferrin in human immunodeficiency virus (HIV)-infected, antiretroviral therapy-naïve children. *International Journal of Antimicrobial Agents* 29, 353–355.

Résumé

La transcription du gène de la lactoferrine conduit à deux isoformes : une isoforme sécrétée appelée lactoferrine (Lf) et une isoforme intracellulaire appelée delta-lactoferrine (Δ Lf). Néanmoins, les deux isoformes possèdent une séquence consensus de localisation nucléaire et il a été décrit qu'elles peuvent être localisées dans le noyau. La séquence primaire de la Δ Lf correspond à celle de la Lf tronquée de ses 45 premiers résidus d'acides aminés dont la séquence signal. La Lf, découverte en 1960, est dotée de nombreuses propriétés parmi lesquelles des activités antimicrobienne, immunomodulatrice et anticancéreuse. La Δ Lf a été décrite pour la première fois en 1997 et les travaux effectués au laboratoire ont montré que la Δ Lf possède également une activité anticancéreuse par le biais de son activité transcriptionnelle. Lf et Δ Lf sont des suppresseurs de tumeur potentiels et leur expression est sous-régulée dans le cas de cancer. Elles sont considérées comme des marqueurs sains puisque leur expression est corrélée à un facteur de bon pronostic dans le cancer du sein. Toutes deux sont impliquées dans le contrôle de la prolifération cellulaire via le contrôle de la transition G1/S du cycle cellulaire et l'induction de l'apoptose.

En tant que facteur de transcription, la Δ Lf exerce son activité grâce à la transactivation de plusieurs gènes cibles tels que *Skp1* (cycle cellulaire) et *Bax* (apoptose) en se fixant sur un élément de réponse appelé Δ LfRE (Δ Lf Response Element). Par contre, la Lf active des voies de signalisation qui diffèrent selon le type cellulaire étudié. Néanmoins, certaines études montrent que la Lf pourrait également intervenir en tant que régulateur transcriptionnel.

Au cours de la thèse, j'ai tout d'abord développé une technique qui permet de distinguer les deux isoformes de Lf à l'aide de la PCR quantitative Taqman. J'ai pu étudier leur expression dans le cas de cancer et dans un contexte inflammatoire. J'ai également confirmé la sous-expression ou l'absence des transcrits des deux isoformes dans les lignées cancéreuses par rapport aux lignées saines correspondantes. Puis j'ai montré, que bien que les promoteurs des deux isoformes contiennent des éléments de réponse putatifs pour des *stimuli* inflammatoires, seuls les transcrits de la Lf sont fortement surexprimés en réponse au LPS (lipopolysaccharide bactérien).

J'ai ensuite étudié le protéome de cellules HEK-293 exprimant la Δ Lf. J'ai réussi à détecter 14 protéines surexprimées en présence de Δ Lf grâce à l'électrophorèses 2-D et identifié 8 d'entre elles par spectrométrie de masse. Parmi ces protéines, seule DcpS a montré également une surexpression au niveau génique. L'emploi d'un gène rapporteur et du ChIP a démontré que la Δ Lf interagit spécifiquement avec le promoteur de *DcpS* *in vivo*. Cette étude m'a donc permis de mettre en évidence une nouvelle cible de l'activité transcriptionnelle de la Δ Lf qui est une enzyme clé responsable de la maturation des ARNm. En parallèle, ces travaux ont aussi montré une régulation de la stabilité et de l'activité transcriptionnelle de la Δ Lf par la O-GlcNAcylation.

Puis, j'ai utilisé les cellules de glande mammaire cancéreuse MDA-MB-231 comme modèle afin de comparer les effets différentiels des isoformes de la lactoferrine sur la croissance et la survie cellulaire. Pour cela, j'ai établi une lignée cellulaire MDA-MB-231 exprimant stablement la Δ Lf en réponse à une induction par la doxycycline. J'ai également purifié la hLf exempte de toute contamination par le LPS à partir du lait maternel. Ensuite, une approche protéomique basée sur le SILAC (Stable Isotope Labelled Amino-acids in Cell culture) et la spectrométrie de masse a été réalisée en collaboration avec l'équipe du Dr. Bernard Monsarrat (Institut de Protéomique et de Biologie Structurale, Toulouse). 5030 protéines différentes ont pu être identifiées et quantifiées. Parmi elles, 3 protéines en présence de Δ Lf et 10 protéines en présence de Lf sont surexprimées avec un ratio supérieur à 1,5. De même, 12 et 11 protéines sont sous-exprimées respectivement en réponse à la Δ Lf et la Lf. L'aldéhyde déshydrogénase 18 (ALDH18A1), la protéine ribosomale S9 (RPS9) et la sélénoprotéine H (SelH) sont surexprimées en réponse aux deux isoformes de lactoferrine. Ces protéines sont principalement impliquées dans la prolifération, l'apoptose et le stress oxydatif. Plusieurs des gènes correspondants possèdent un élément de réponse à la Δ Lf mais seule la fonctionnalité du Δ LfRE de SelH a été étudiée en détail. En résumé, Cette partie de mon travail montre que le protéome des cellules MDA-MB-231 est modifié par la Lf ou la Δ Lf et elle a permis d'identifier le gène SelH en tant que cible de l'activité transcriptionnelle de la Lf et de la Δ Lf, suggérant que la Lf sécrétée possède également un rôle de facteur de transcription.

Abstract

Transcription of the *lactoferrin* gene results in the production of two isoforms: a secreted isoform known as lactoferrin (Lf) and intracellular delta-lactoferrin (Δ Lf). Nevertheless, both Lf forms contain consensus nuclear targeting signals and were reported to enter the nucleus. The primary sequence of Δ Lf corresponds to a truncated Lf lacking its first 45 amino acid residues including the signal sequence. Lf, which was discovered in 1960, has numerous functions including antimicrobial immunomodulatory and antitumoral activities. Δ Lf, which was first described in 1997, has been studied in our laboratory and we showed that it is a transcription factor, the only activity it shares with Lf being its anticancer activity. Lf and Δ Lf are potential tumor suppressors whose expression is downregulated in the case of cancer. Therefore they may be considered as healthy markers, the expression of which is correlated with a good prognosis in the case of breast cancer. Both proteins are involved in the control of cell proliferation through the control of the G1/S transition and trigger apoptosis. Δ Lf transactivates different target genes such as *Skp1* (cell cycle) and *Bax* (apoptosis) by binding to a Δ LfRE (Δ Lf Response Element) present in their promoter regions, whereas Lf acts by activating signaling pathways that differ depending on the cell type studied. However, some studies reported that Lf may also act as a transcriptional regulator. During my PhD thesis, we first developed an assay to evaluate the level of expression of the transcripts of the two Lf isoforms using a quantitative Taqman PCR method. We measured their level of expression either in the case of cancer or in an inflammatory context. We confirmed the underexpression and/or the absence of both isoform mRNAs in cancer cell lines compared to their normal counterparts. On the other hand, although the two *lactoferrin* promoter regions contain putative response elements to inflammatory stimuli, only Lf transcription was strongly upregulated in response to LPS (bacterial lipopolysaccharide).

We then studied the proteome profile of Δ Lf-expressing cells. A total of fourteen differentially expressed proteins were visualized by 2D electrophoresis and silver staining and eight proteins were identified by mass spectrometry analyses, all of which were up-regulated. Among these proteins, only DcpS was also up-regulated at the mRNA level. Reporter gene analyses and ChIP assays demonstrated that Δ Lf interacts specifically with the *DcpS* promoter *in vivo*. Our data established that DcpS, a key enzyme in mRNA decay, is a new target of Δ Lf transcriptional activity. In parallel, we demonstrated that Δ Lf stability and transcriptional activity are highly regulated by O-GlcNAcylation.

Finally, in order to compare the differential effects of the two antitumoral Lf isoforms on cell growth and survival, we chose the cancerous mammary gland MDAMB-231 cell line as a model. We produced a MDAMB-231 cell line stably expressing DLf under doxycyclin induction. We also purified hLf devoid of LPS contamination from human milk. We used a high-throughput proteomic approach (Stable Isotope Labelling with Amino acids in Cell culture, SILAC) coupled to mass spectrometry in order to carry out highly accurate quantitative and systematic proteome profiling in collaboration with the team of Dr. Bernard Monsarrat (Institute of Proteomics and Structural Biology, Toulouse). A total of 5030 proteins were identified and quantified, among which expression was increased by 1.5-fold or more for 3 proteins in Δ Lf expressing cells and for 10 proteins in Lf treated cells, compared to the control. On the other hand, 12 and 11 proteins were found to be downregulated in response to DLf and Lf, respectively. Aldehyde dehydrogenase 18 (ALDH18A1), ribosomal protein S9 (RPS9) and selenoprotein H (SelH) are upregulated in response to both lactoferrin isoforms. These proteins are mainly implicated in the regulation of processes such as proliferation, apoptosis, and oxidative stress. Some of the corresponding genes possess a DLf response element in their promoter region but only the functionality of SelH DLfRE has been investigated. Our study showed that the MDAMB-231 proteome was modified by either Lf or DLf, allowed the identification of SelH as a target of both Lf and Δ Lf transcriptional activity and suggested that secreted Lf also acts as a transcription factor.