

Étude protéomique de la microhétérogénéité des caséines [alpha]s1 et [bêta] équines : identification des variants transcriptionnels et de phosphorylation ; identification des sites phosphorylés de la caséine [bêta]

Aurélie Mateos

▶ **To cite this version:**

Aurélie Mateos. Étude protéomique de la microhétérogénéité des caséines [alpha]s1 et [bêta] équines : identification des variants transcriptionnels et de phosphorylation ; identification des sites phosphorylés de la caséine [bêta]. Sciences agricoles. Institut National Polytechnique de Lorraine, 2008. Français. NNT : 2008INPL089N . tel-01753053

HAL Id: tel-01753053
https://hal.univ-lorraine.fr/tel-01753053

Submitted on 29 Mar 2018

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

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

AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4
Code de la Propriété Intellectuelle. articles L 335.2- L 335.10
http://www.cfcopies.com/V2/leg/leg_droi.php
http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm

INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE

Ecole Nationale Supérieure d'Agronomie et des Industries Alimentaires

Ecole doctorale Science et Ingénierie des Ressources, Procédés, Produits, Environnement

THESE

Présentée pour l'obtention du grade de

Docteur de l'Institut National Polytechnique de Lorraine

Discipline : Sciences Agronomiques

Soutenue publiquement le 21 novembre 2008 par

Aurélie MATEOS

Etude protéomique de la microhétérogénéité des caséines α_{s1} et β équines : identification des variants post-transcriptionnels et de phosphorylation ; identification des sites phosphorylés de la caséine β.

Membres du jury :

Rapporteurs : M. Saïd BOUHALLAB, Directeur de recherche, INRA, Rennes

M. Jean-Marc CHOBERT, Directeur de Recherche, INRA, Nantes

Examinateurs : M. Jean-Luc GAILLARD, Professeur, Université de Caen (Directeur de thèse)

M. Patrice MARTIN, Directeur de recherche, INRA, Jouy-en-Josas

M. Laurent MICLO, Maître de conférences, UHP, Nancy 1

Membre invité : M. Jean-Michel GIRARDET, Ingénieur d'études, UHP, Nancy 1

Unité de Recherche Animal et Fonctionnalités des Produits Animaux
Equipe Protéolyse et Biofonctionnalité des Protéines et Peptides
Faculté des Sciences et Techniques, Université Henri Poincaré- 54500 Vandœuvre-lès-Nancy

Avant-propos

Ce travail a été réalisé au sein de l'équipe Protéolyse et Biofonctionnalités de Protéines et Peptides de l'Unité de Recherche sur l'Animal et la Fonctionnalité des Produits Animaux. Je tiens à remercier le Pr Guido Rychen, directeur d'unité, pour m'avoir accueillie au sein de son laboratoire. Je remercie également chaleureusement le Pr Annie Dary pour ces 3 années passées au sein de son équipe.

Je tiens à remercier le Pr Jean-Luc Gaillard, mon directeur de thèse, pour m'avoir donné la chance de pouvoir réaliser ce travail de thèse, moi qui n'étais pas initialement biochimiste de formation…

Un grand merci au Dr Jean-Michel Girardet pour m'avoir suivie et supportée pendant ces 3 années. C'est Jean-Michel qui m'a accueillie le premier jour et qui m'a initiée à quasiment toutes les techniques que j'ai été amenée à utiliser. Il m'a également été de bon conseil au cours de la rédaction de mes publications et de ce manuscrit, j'ai beaucoup appris grâce à lui. Il a toujours été là quand j'avais besoin de lui pour un conseil, une manip ou juste pour discuter autour d'une tasse de thé ! Merci Jean-Mi.

Merci également au Dr Laurent Miclo, mon co-directeur (pour moi en tout cas tu l'es !). Cela fut un grand plaisir et un honneur de travailler avec quelqu'un comme Laurent. Toujours disponible même quand son agenda était surbooké (surtout ces derniers temps…), ses conseils m'ont été précieux tout au long de mon travail.

Merci au Dr Daniel Mollé de l'INRA de Rennes pour avoir contribué à la réalisation de ce travail.

Je remercie aussi Delphine, Nawara, Maira, Frédérique, Faïza, Wessam et Ounki, mes collègues thésards, pour leur soutien au quotidien et leur bonne humeur. Cela fut un plaisir de travailler à vos côtés pendant ces quelques années.

Merci également à Emeline pour ton amitié et tes conseils ainsi que pour avoir été ma partenaire de club pendant tout ce temps…

Une pensée particulièrement pour Leila, une stagiaire devenue une bonne amie, une pensée également aux autres stagiaires qui ont participé de près ou de plus loin à ces travaux.

Merci également aux membres de l'équipe PB2P, Clarisse, Céline, Magali, Anne, Catherine, Chantal, Emilie, Katy, Alain, Gérard mais aussi Frank pour leur enthousiasme, leur sympathie et leur soutien.

Je souhaite également remercier les membres de l'équipe DAC, Jean, Pascal, Marielle et Jean-Noël pour m'avoir fait découvrir pour la première fois le monde de la recherche. Merci aussi à Magali avec qui se fut un plaisir de travailler. Merci Jean de m'avoir poussée à postuler pour cette thèse qui m'a fait passer de l'étude de la perche commune au lait de jument !

Un grand merci à Arnaud qui m'a soutenue et supportée pendant ces 3 années même quand la distance nous a séparé…Tu m'as épaulée dans mes moments de doutes et de désarroi, tu as toujours été là pour moi. Merci.

Merci également à la famille Griot-Campeas pour leur soutien et leur intérêt pour mon travail. Une pensée pour Albert, un homme d'un âge vénérable que j'admire beaucoup.

Le plus important, ma famille. Je remercie de tout mon coeur mes parents d'avoir toujours cru en moi et d'avoir contribué à ce que je suis aujourd'hui. Je vous remercie pour votre soutien indéfectible pendant toute ma scolarité et d'avoir toujours été là en cas de besoin. Une pensée pour Marie-José qui a réussi à trouver sa voie et finalement bien qu'étant le numéro 3 est diplômée avant moi ! L'année 2008, finalement, est celle des palmes et des lauriers ! Merci soeurette pour tous les moments partagés ensemble, les bons comme les mauvais. Un gros bisous à ma Lolo, ma petite sœur, qui est née alors que je débutais mon parcours universitaire, et nous voilà quasiment 10 ans plus tard…Que le temps passe vite !

Liste des abréviations

ACN : acétonitrile

AMP : adénosine monophosphate

Api III+ : spectromètre de masse tandem à triple quadrupôle

BAEE : N-benzoyl-L-arginine éthyl ester

β-LG : β lactoglobuline

C : rapport (en %) de la masse de bisacrylamide sur la masse d'acrylamide + bisacrylamide

CAPS : acide 3-(cyclohexylamino)-1-propanesulfonique

CHAPS : sulfonate de 3-[(3-cholamidopropyl)-diméthylammonio]-1-propane

CKG : caséine kinase golgienne

CK1 : caséine kinase I

CK2 : caséine kinase II

CN : caséine

CPP : caséinophosphopeptide

d : déphosphorylé

/d : désamidé

2-DE : électrophorèse bidimensionnelle

DEAE : diéthylaminoéthyle

DHB : acide 2,5-dihydroxybenzoïque

ΔH : enthalpie

ΔG : énergie libre

ΔQ : délétion d'un résidu Gln

ΔS : entropie

ΔX : délétion de la région codée par l'exon X

ECA : enzyme de conversion de l'angiotensine

EDTA : éthylènediaminetétracétate

ESE : *exonic splicing enhancer*

ESI : ionisation à électropsray

ESS : *exonic splicing silencer*

FPLC : *fast protein liquid chromatography* (chromatographie liquide rapide des protéines)

F : fraction chromatographique

GABA : acide gamma-aminobutyrique

Gal : galactose

GalNAc : N-acétylgalactosamine

HIC : *hydrophobic interaction chromatography* (chromatographie d'interactions hydrophobe)

hnRNP : *heterogeneous nuclear ribonucleoprotein*

HPLC : *high pefrormance liquid chromatography* (chromatographie liquide haute performance)

IEC : *ion-exchange chromatography* (chromatographie d'échange d'ions)

IEF : isoélectrofocalisation

Ig : immunoglobuline

IMAC : *immobilized metal affinity chromaography* (chromatographie d'affinité aux métaux immobilisés)

IPG : *immobilized pH gradient* (gradient de pH immobilisé)

ISE : *intron splicing enhancer*

ISS : *intron splicing silencer*

ITC: microcalorimétrie à titration isotherme

K_A : constante d'affinité

LC/ESI-MS: chromatographie liquide couplée à la spectrométrie de masse à ionisation électrospray

LSm : protéine « Sm-like »

MOAC : *metal oxide affinity chromatography* (chromatographie d'affinité aux oxydes métalliques)

m/v : masse par volume

m/z : masse sur charge

NeuAc : acide N-acétylneuraminique

XP : X groupement(s) phosphate

PAGE : *polyacrylamide gel electrophoresis* (électrophorèse sur gel de polyacrylamide)

pHi : pH isoélectrique

PIMT : L-isoaspartyl méthyltransférase

PMSF : fluorure de phénylméthylsulfonyle

PP : protéose-peptone

PVDF : difluorure de polyvinylidène

Q/Star XL : spectromètre de masse hybride quadrupôle/temps de vol

R_f : distance de migration relative

SerP : résidu de sérine phosphorylé

SDS : dodécylsulfate de sodium

SF : sous-fraction chromatographique

Sm : protéine Sm appelée ainsi car elle réagit avec les anticorps auto-immuns de sérotype Sm de patients atteints de lupus érythémateux systémique

SMN : *survival of motor neuron*

SnRNA : *small nuclear ribonucleic acid*

SnRNP : *small nuclear ribonucleoprotein*

sucWGA : *succinylated wheat germ agglutinin* (lectine de germe de blé succinylée)

T : rapport (en %) de la masse d'acrylamide + bisacrylamide sur le volume de gel

TFA : acide trifluoroacétique

TPCK : L-1-tosylamido-2-phényléthyl chlorométhyl cétone

Tris : trihydroxy-méthyl-amino-méthane

UV : ultraviolet

v/v : volume par volume

WGA : *wheat germ agglutinin* (lectine de germe de blé)

Sommaire

Avant-propos *2*

Liste des abréviations *3*

INTRODUCTION BIBLIOGRAPHIQUE *9*

1. Les protéines du lait **12**
1.1. Les protéines lactosériques 12
1.2. Les caséines 14
1.2.1. La caséine β 15
1.2.2. La caséine α_{s1} 18
1.2.3. La caséine α_{s2} 22
1.2.4. La caséine κ 23
1.2.5. La micelle de caséines 27

2. Modification post-transcriptionnelle : l'épissage conventionnel et l'épissage alternatif d'exons **31**
2.1. Le splicéosome 31
2.2. Le processus d'épissage 34
..... 37
2.3. L'épissage alternatif 38
2.3.1. Facteurs influençant la reconnaissance d'un exon 38
2.3.2. Les sites d'épissage cryptiques 43
2.3.3. Les différents types d'épissage alternatif pouvant être rencontrés 43
2.4. L'épissage des caséines 45
2.4.1. Structure des gènes des caséines bovines 45
2.4.2. Epissage alternatif des caséines 47

3. Modifications post-traductionnelles : phosphorylation et glycosylation des caséines **49**
3.1. Phosphorylation des caséines 49
3.2. Glycosylation des caséines 51

4. Les lactoprotéines, source de peptides à activités biologiques **51**
4.1. Peptides anti-microbiens 54
4.2. Peptides anti-hypertenseurs 55
4.3. Peptides antithrombotiques 56
4.4. Peptides à activité opiacée 56
4.5. Peptide anxiolytique 57
4.6. Cas des caséinophosphopeptides 58
4.6.1. Activité anti-oxydante 59
4.6.2. Activité anti-cariogène 60
4.6.3. Effet mitogène 61
4.6.4. Chélation de minéraux d'intérêt nutritionnel : exemple du calcium et du fer 61

CONTEXTE ET OBJECTIFS DU TRAVAIL *66*

MATERIELS ET METHODES **70**

1. Préparation du caséinate de sodium **71**

2. Méthodes électrophorétiques et techniques attenantes **72**
2.1. Electrophorèse en gel de polyacrylamide en présence de dodécylsulfate de sodium (PAGE-SDS) 72
2.2. Electrophorèse en gel de polyacrylamide en présence d'urée (PAGE-urée) 72
2.3. Electrophorèse bidimensionnelle couplant une isoélectrofocalisation (IEF) et une PAGE-SDS . 73
2.4. Electrotransfert de protéines 74
2.4.1. En cuve verticale 74
2.4.2. Transfert semi-sec 74
2.5. Révélation des protéines au bleu de Coomassie R-250 75
2.6. Révélation des protéines sur la membrane de PVDF 75

2.7. Révélation des glycoprotéines sur la membrane de PVDF 75

3. Hydrolyses enzymatiques **76**
3.1. Hydrolyse par la chymosine 76
3.2. Hydrolyse par la trypsine immobilisée sur bille d'agarose 76
3.3. Hydrolyse par la phosphatase alcaline 77
3.4. Hydrolyse par la phosphatase acide 77
3.5. Hydrolyse par la neuraminidase 78

4. Chromatographie liquide **78**
4.1. Chromatographie FPLC (*fast protein liquid chromatography*) d'interactions hydrophobes 78
4.2. Chromatographie FPLC d'échange d'anions 79
4.3. Chromatographie d'affinité à la lectine de germe de blé (*wheat germ agglutinin* ou WGA) 79
4.4. Chromatographie MOAC (*Metal Oxide Affinity Chromatography*) au dioxyde de titane immobilisé sur microcolonne 80
4.5. Chromatographie liquide haute performance en phase inversée (RP-HPLC) sur colonne C_{18} semi-préparative 80
4.6. Chromatographie RP-HPLC sur colonne C_{18} analytique 81
4.7. Chromatographie RP-HPLC sur colonne C_4 81

5. Analyse des protéines et des peptides par spectrométrie de masse **81**
5.1 Détermination des masses moléculaires 82
5.2 Séquençage peptidique 82

PARTIE 1 ..... ***84***

Caractérisation structurale de la caséine α_{s1} équine : variants post-transcriptionnels et post-traductionnels ***84***

1. Fractionnement du caséinate de sodium équin **86**
1.1. Chromatographie HPLC en phase inversée 86
1.2. Chromatographie FPLC d'interactions hydrophobes 88

2. Séparation de la CN-κ et de la CN-α_{s1} par chromatographie d'affinité à la lectine de germe de blé (WGA) immobilisée **90**

3. Déphosphorylation de la CN-α_{s1} et caractérisation des variants d'épissage **95**
3.1.Déphosphorylation de la CN-α_{s1} 96
3.1.1. Utilisation de la phosphatase alcaline ou de la phosphatase acide 96
3.1.2. Utilisation de la phosphatase alcaline suivie de la phosphatase acide 97
3.2. Caractérisation des variants d'épissage de la CN-α_{s1} 100

4. Caractérisation des variants de phosphorylation de la CN-α_{s1} équine **104**
4.1. Fractionnement de la CN-α_{s1} par chromatographie FPLC d'échange d'ions 104
4.2. Caractérisation du degré de phosphorylation 107

5. Conclusions et perspectives **113**

PARTIE 2 ..... ***115***

Cartographie bidimensionnelle de la caséine β équine établie après séparation de ses variants de phosphorylation - Étude de sa désamidation spontanée ***115***

1. Purification de la CN-β équine par chromatographie HIC **117**

2. Purification chromatographique des variants de phosphorylation de la CN-β équine **119**
2.1. Essai de purification par chromatographie d'affinité aux métaux immobilisés (IMAC) 119
2.2. Fractionnement de la CN-β par chromatograhie IEC 120
2.3. Caractérisation des isoformes de CN-β par LC/ESI-MS 122

3. Etude de la désamidation non-enzymatique de la CN-β équine **125**

4. Etablissement de la cartographie bidimensionnelle de la CN-β équine **130**

5. Conclusions et perspectives **137**

PARTIE 3 ..... *139*

Préparation des caséinophosphopeptides de la CN-β équine et détermination des sites de phosphorylation en fonction du degré de phosphorylation *139*

1. Hydrolyse des différents variants de phosphorylation de la CN-β équine isolés par FPLC d'échange d'ions **140**

2. Préparation d'une fraction enrichie en CPP par chromatographie d'affinité au dioxyde de titane et HPLC en phase inversée **143**

2.1. Chromatographie d'affinité au dioxyde de titane ou MOAC (Metal Oxide Affinity Chromatography) 143

2.1.1. Choix du métal immobilisé 143

2.1.2. Analyse des fractions liées à TiO_2 par spectrométrie de masse 146

2.2. Mise au point d'une technique de purification des CPP par HPLC en phase inversée 149

2.2.1. Essai avec une colonne C_{18} 149

2.2.2. Purification des CPP par HPLC en phase inversée sur colonne C_4 154

3. Localisation des sites de phosphorylation de la CN-β équine par spectrométrie de masse **158**

4. Conclusions et perspectives **164**

CONCLUSION GENERALE ET PERSPECTIVES *164*

REFERENCES BIBLIOGRAPHIQUES *164*

Liste des publications *164*

INTRODUCTION BIBLIOGRAPHIQUE

Le lait maternel est l'aliment exclusif puis principal des nouveaux-nés chez les mammifères jusqu'à la période de sevrage. Celui-ci contient les éléments indispensables à leur survie puis à leur croissance. Pendant les premiers mois de la vie, le lait a pour fonctions naturelles d'assurer la nutrition du jeune mammifère et de transmettre, de la mère vers son petit, une certaine immunité vis-à-vis des agressions microbiennes extérieures (Lestradet, 1988).

Les laits des différents mammifères renferment des constituants de même nature. Ils forment des mélanges complexes contenant une forte proportion d'eau. Les constituants du lait se répartissent entre deux phases : une phase aqueuse et une phase lipidique. La phase aqueuse contient principalement du lactose, des minéraux et des protéines. La phase lipidique (matière grasse laitière) existe sous forme de globules gras. Par ailleurs, le lait renferme des molécules à activité biologique : enzymes, vitamines, hormones, immunoglobulines. Toutefois, leurs proportions peuvent varier de façon notable d'une espèce animale à l'autre, et à un degré moindre, d'une race à l'autre au sein d'une même espèce. Par exemple, le lait provenant d'animaux de la famille des Félidés (serval, panthère, lionne, chatte), espèces carnivores, est plus riche en protéines totales (en caséines et protéines lactosériques), lipides que celui d'espèces herbivores (ânesse, jument, vache, éléphante) ou omnivore (femme, Tableau 1). Dans tous les cas, la composition nutritionnelle du lait de chaque espèce est remarquablement bien adaptée aux besoins du jeune. De ce fait, depuis quelques années, des études sur la composition du lait de certains mammifères sauvages (Félidés, Éléphantidés) se développent puisque parfois, en captivité, les jeunes ne peuvent être nourris par leur mère (De Waal *et al.*, 2004 ; Osthoff *et al.,* 2005, 2006, 2007). En effet, le lait de substitution utilisé pour l'élevage d'espèces de la famille Félidés est en général formulé à partir des connaissances acquises sur du lait d'autres espèces (Baines, 1981 ; De Waal, 1990 ; McManamon, 1993). L'utilisation d'un tel lait inadapté entraîne des complications comme des cataractes, un faible gain de poids et des intolérances (Remillard *et al.,* 1993) montrant bien que chaque lait est spécifique d'une espèce .

Tableau 1 : Comparaison de la composition globale de laits de différentes espèces

	Femme [1]	Anesse [2]	Jument [1]	Vache [3]	Eléphante [4]	Serval [5]	Panthère [6]	Lionne [7]	Chienne [8]
Protéines (g/kg lait)	9-17	17,2	15-28	31-38	21,8-25	113,9-202,7	99,6	60,2-84,6	68,4-143
Lipide (g/kg lait)	35-40	3,8	5-20	35-39	56-76	90,3-214,9	64,8	113,6-136,9	112,5-137,2
Lactose (g/kg lait)	63-70	68,8	58-70	44-49	11,8-52,5	37,3-100,1	40,2	26,5	16,6-40,2
Caséines (CN) (g/kg lait)	3,7	2,9	10,7	25,2	10-14	72,8-162,6	34,2	n.d.	n.d.
Protéines lactosériques (LS) (g/kg lait)	7,6	6,8	8,3	6,1	n.d.	33,9-47,3	65,3	n.d.	n.d.
Ratio CN/LS	1/2	1/2	1/1	1/4	n.d.	3/1	1/2	n.d.	7/3

(1) d'après Malacarne *et al.* (2002)
(2) d'après Salimei *et al.* (2004)
(3) d'après Ng-Kwai-Hang (2003)
(4) d'après Osthoff *et al.* (2005)
(5) d'après Osthoff *et al.* (2007)
(6) d'après Osthoff *et al.* (2006)
(7) d'après De Waal *et al.* (2004)
(8) d'après Adkins *et al.* (2001)

1. Les protéines du lait

Le lait est constitué de deux fractions protéiques principales, une fraction soluble et une fraction colloïdale. La fraction soluble se compose des protéines lactosériques et la fraction colloïdale est formée par les caséines, ces fractions représentant 20% et 80% des protéines du lait des ruminants, respectivement.

1.1. Les protéines lactosériques

Les protéines lactosériques ou protéines du lactosérum représentent la fraction protéique restant soluble lors d'une précipitation à pH 4,6 et à 20°C des protéines du lait écrémé d'origine bovine (Thompson *et al.,* 1965) et à pH 4,2 et 20°C pour le lait équin (Egito *et al.,* 2001). Les protéines lactosériques se composent généralement de β-lactoglobuline, d'α-lactalbumine, d'immunoglobulines, de sérum albumine, de lactoferrine et de lysozyme. La concentration des différentes protéines lactosériques peut varier fortement entre les laits équin, bovin et humain (Tableau 2). Notamment, le lysozyme et la lactoferrine ne sont trouvés en quantité appréciable que dans les laits équin et humain. Le lysozyme, très stable à la température (Jauregui-Adell, 1975), est un puissant agent antimicrobien surtout en présence de lactoferrine. En effet, le lysozyme peut détruire des bactéries $Gram^-$ en synergie avec la lactoferrine.

La β-lactoglobuline (β-LG), protéine allergénisante du lait bovin, est absente du lait humain. Dans le lait équin, il existe deux formes de la β-LG, la β-LG I (Conti *et al.,* 1984) et la β-LG II (Godovac-Zimmermann *et al*., 1985). Dans le lait bovin, les variants génétiques A et B de la β-LG sont les variants majoritaires bien qu'il existe 11 variants génétiques (A, B, C, D, E, F, G, H, I, J et H ; Farrel Jr *et al.,* 2004) de la β-LG bovine. La β-LG équine est présente en quantité plus faible dans le lait que chez l'espèce bovine mais pourrait être allergisante et être responsable d'une sensibilisation aux protéines équines chez les jeunes enfants. Cependant, ce phénomène serait moindre comparé à celui engendré par le lait bovin. Le lait équin serait mieux toléré chez les enfants sévèrement allergiques au lait bovin (Gall *et al.,* 1996 ; Businco *et al.,* 2000). Ainsi, le lait équin peut servir à la formulation de laits adaptés ou de substituts nutritionnels pour des jeunes enfants allergiques au lait bovin (Stoianova *et al.,* 1988). La β-LG peut lier des composés aromatiques (Pelletier *et al.,* 1998), des rétinoïdes (Dufour et Haertlé, 1991 ; Wang *et al.,* 1997a), de la vitamine D et du cholestérol (Wang *et*

al., 1997b). Son rôle dans le lait pourrait être de transporter des petites molécules hydrophobes à travers la paroi intestinale (Wu *et al.,* 1999).

Tableau 2 : Comparaison des distributions protéiques des laits équin, humain (Malacarne *et al.,* 2002) et bovin (Ng-Kwai-Hang, 2003).

	Lait équin	Lait humain	Lait bovin
Protéines lactosériques (%)			
β-lactoglobuline	25,2-36,3	absente	52,2
α-lactalbumine	27,5-29,7	30,3-45,4	19,7
Immunoglobulines	18,7-20,9	15,1-19,7	13,1
Sérum albumine	4,4-4,5	4,5-9,1	6,6
Lactoferrine	9,9	30,3	1,6
Lysozyme	6,6	1,7	Traces
Total (g/kg lait)	8,3	7,6	6,1
Caséines (%)			
Caséines α_s	40,2-59	11,1-12,5	50
Caséine β	40,1-51,4	62,5-66,7	36,9
Caséine κ	7,7	23,5	13,1
Total (g/kg lait)	10,7	3,7	25,2
Taille des micelles[1](nm)	100-300	200-300	500

[1] d'après Jasinska et Jaworska (1988)

L'α-lactalbumine constitue l'une des deux sous-unités du complexe enzymatique lactose synthétase. Elle permet la synthèse du lactose dans la glande mammaire en modifiant la spécificité de substrat de l'UDP-galactosyltranférase et son affinité pour le glucose (Ebner *et al.*, 1966). L'α-lactalbumine possède un site de liaison de l'ion Ca^{2+} chez toutes les espèces. L'α-lactalbumine humaine possède, quant à elle, un site secondaire de liaison au Ca^{2+} de plus faible affinité (Chandra *et al.*, 1998). La fixation du calcium joue un rôle important dans la stabilisation moléculaire de la structure spatiale de l'α-lactalbumine. L'α-lactalbumine peut aussi chélater d'autres cations tels que Zn^{2+}, Mg^{2+}, Mn^{2+}, Na^{+} et K^{+} en plus de Ca^{2+} (Permyakov et Berliner, 2000). Trois variants de l'α-lactalbumine (variants A, B et C) ont été identifiés dans le lait de jument (Kaminogawa *et al.*, 1984 ; Godovac-Zimmermann *et al.*, 1987) alors qu'il n'en existe que deux pour l'α-lactalbumine bovine (variants A et B). Le variant A équin présente 72% d'identité avec le variant B bovin. Il existerait également une forme glycosylée mineure dans le lait équin (Girardet *et al.,* 2004).

1.2. Les caséines

Les caséines (CN) sont les protéines majeures du lait des ruminants. Elles se définissent comme étant la fraction protéique précipitant à pH 4,6 et à 20°C pour le lait écrémé bovin (Thompson *et al.,* 1965) et à pH 4,2 et 20°C pour le lait équin (Egito *et al.,* 2001). Les caséines sont au nombre de quatre : caséine α_{s1}, caséine $\alpha_{s2,}$ caséine β et caséine κ. Ce sont des protéines multiphosphorylées, la CN-κ pouvant être également glycosylée. Elles sont synthétisées par l'épithélium sécrétoire mammaire sous forme de précurseurs contenant un peptide signal qui permet l'adressage et qui sera clivé au cours du passage de ces protéines à travers la membrane du réticulum endoplasmique (Gaye *et al.,* 1977 ; Mercier et Gaye, 1980). Dans le lait, les caséines sont organisées sous la forme de micelles et s'associent à un complexe minéral de phosphate de calcium. Les micelles se forment à l'intérieur des vésicules du Golgi où il a été observé en microscopie électronique l'agrégation de petites particules submicellaires en plus grandes particules micellaires (Farrell Jr., 1988 ; Patton, 2004). Un des rôles de la micelle est de transporter le phosphate et le calcium dans le lait, permettant de maintenir dans le lait une concentration élevée de calcium fortement assimilable (forme soluble).

1.2.1. La caséine β

La CN-β bovine contient 209 résidus d'acides aminés et comporte cinq résidus de sérine phosphorylés (Bandura *et al.,* 2002 ; Figure 1). Elle ne contient pas de résidu de cystéine. Cette protéine possède neuf variants génétiques (A^1, A^2, A^3, B, C, D, E, F, G ; Farrell Jr *et al.,* 2004). Cependant l'existence du variant génétique G est controversée (Visser., 2000). Le lait bovin contient 9,3 $g.L^{-1}$ de CN-β (Ribadeau-Dumas et Grappin, 1989). Elle peut être hydrolysée par la plasmine au niveau des liaisons Lys^{28}-Lys^{29}, Lys^{105}-His^{106} et Lys^{107}-Glu^{108} ce qui génère les fragments 1-28 (PP8f), 1-105/107 (PP5) et 29-105/107 (PP8s) également appelés protéoses-peptones (PP) et retrouvés dans le lactosérum. Les fragments complémentaires sont les CN-γ1 [29-209], CN-γ2 [106–209] et CN-γ3 [108–209] et sont retrouvés dans la fraction caséinique car insolubles à pH 4,6 (Gordon *et al.,* 1972 ; Kaminogawa et Yamauchi, 1972 ; Farrell Jr *et al.,* 2004).

```
  1MKVLILACLV ALALARELEE LNVPGEIVES LSSSEESITR INKKIEKFQS EEQQQTEDEL
 61 QDKIHPFAQT QSLVYPFPGP IHNSLPQNIP PLTQTPVVVP PFLQPEVMGV SKVKEAMAPK
121HKEMPFPKYP VEPFTESQSL TLTDVENLHL PLPLLQSWMH QPHQPLPPTV MFPPQSVLSL
181SQSKVLPVPQ KAVPYPQRDM PIQAFLLYQE PVLGPVRGPF PIIV224
```

Equin

```
  1MKILILACLV ALALAREKEE LNVSSETVES LSSNEPDSSS EESITHINKE KLQKFKHEGQ
 61QQREVERQDK ISRFVQPQPV VYPYAEPVPY AVVPQSILPL AQPPILPFLQ PEIMEVSQAK
121ETILPKRKVM PFLKSPIVPF SERQILNPTN GENLRLPVHL IQPFMHQVPQ SLLQTLMLPS
181QPVLSPPQSK VAPFPQPVVP YPQRDTPVQA FLLYQDPRLG PTGELDPATQ PIVAVHNPVI V241
```

Humain

```
  1MKVLILACLV ALALARETIE SLSSSEESIT EYKQKVEKVK HEDQQQGEDE HQDKIYPSFQ
 61PQPLIYPFVE PIPYGFLPQN ILPLAQPAVV LPVPQPEIME VPKAKDTVYT KGRVMPVLKS
121PTIPFFDPQI PKLTDLENLH LPLPLLQPLM QQVPQPIPQT LALPPQPLWS VPQPKVLPIP
181QQVVPYPQRA VPVQALLLNQ ELLLNPTHQI YPVTQPLAPV HNPISV226
```

Figure 1 : Comparaison des séquences des caséines β bovine, humaine et équine. Les numéros d'accession dans la banque de données SwissProt sont les suivants : NP851351 pour la séquence bovine (*Bos taurus*), PO5814 pour la séquence humaine (*Homo sapiens*) et Q9GKK3 pour la séquence équine (*Equus caballus*). Les résidus phosphorylés sont indiqués en gras et italique. Les résidus potentiellement phosphorylés sont soulignés. Le peptide signal est indiqué en bleu.

Les études par diffraction des rayons X (Kajiwara *et al.*, 1988) et par spectroscopie Raman (Byler *et al.*, 1988) montrent que la CN-β a majoritairement une structure désordonnée (Figure 2). Celle-ci résulte des contraintes de torsion de la chaîne polypeptidique imposées par les nombreux résidus de proline (au nombre de 35), et 9 à 12% d'hélices α ainsi que 11 à 25% de feuillets β. Une modélisation hypothétique de la structure tridimensionnelle réalisée par Kumosinski *et al.* (1993) montre la CN-β sous la forme d'une molécule asymétrique ressemblant à un "crabe", dont la partie hydrophile serait présente au niveau des "pinces". Les zones hydrophobes, véritables zones d'ancrage pour les enzymes, seraient probablement à l'origine de la forte capacité des monomères de caséines β à s'associer entre-eux pour former des "néomicelles" (Mercier et Ribadeau Dumas, 1972 ; Takase *et al.*, 1980 ; Kajiwara *et al.*, 1988). Les deux "pinces" contiennent les sites d'hydrolyse par la plasmine endogène du lait. La CN-β est la plus hydrophobe des caséines. Sa séquence N-terminale contient de nombreux résidus d'acides aminés chargés et le cluster phosphorylé. Le calcul de la charge nette de la protéine à pH 6,6 montre que les 21 premiers résidus d'acides aminés auraient une charge nette d'environ -11,5 et que les 21 derniers résidus d'acides aminés [190-209] ne sont pas chargés. (Farrell Jr. *et al.,* 2004).

a) b)

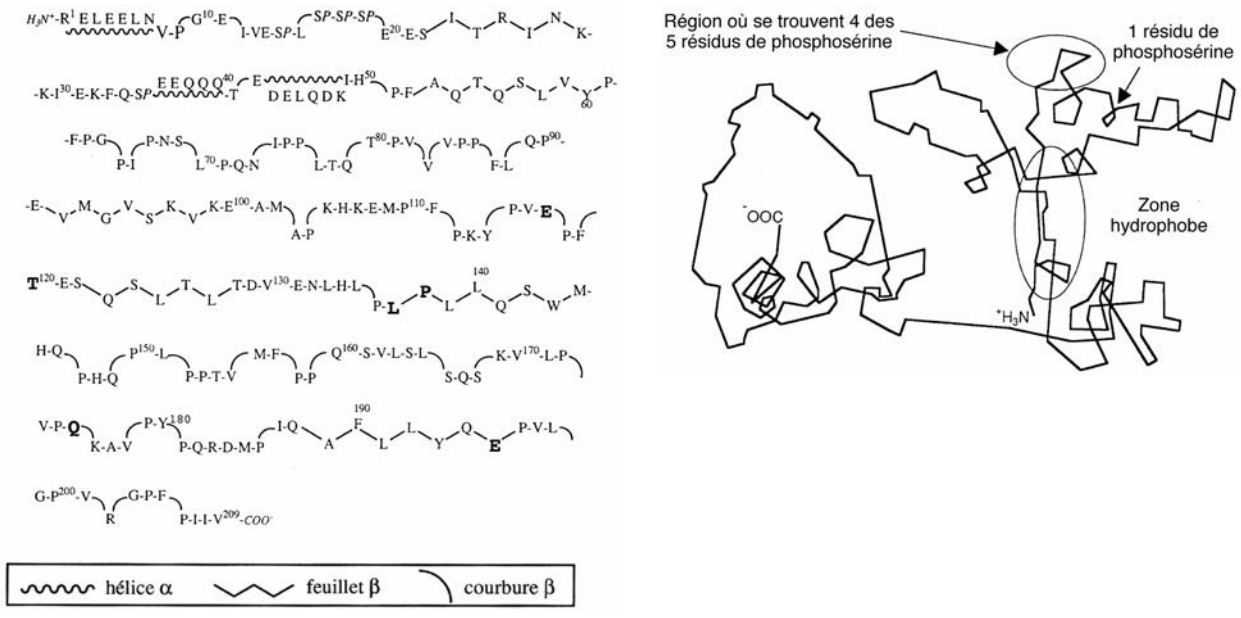


Figure 2 : Structures primaire et secondaire prédites de la caséine β bovine, variant A2 (a : d'après Ribadeau-Dumas *et al.,* 1972 ; Farrell Jr. *et al.,* 1993) et modèle hypothétique de la structure tridimensionnelle de la caséine β bovine (b : d'après Kumosinski *et al.,* 1993). Figures reproduites d'après Cayot et Lorient (1998), revues et corrigées d'après la banque de données Swissprot (numéro d'accession : P02666).

La CN-β humaine est constituée de 211 résidus d'acides aminés. Contrairement à la CN-β bovine, la phosphorylation de la CN-β humaine est variable. Ainsi, cette protéine peut contenir de 0 à 5 groupements phosphate (0P à 5P), les variants 2P et 4P étant majoritaires (Greenberg *et al.,* 1984 ; Figure 1). Alors que le gène de la CN-β bovine possède huit exons, dans le cas de la CN-β humaine l'exon 3 est constitutivement délété (Figure 3 ; Hansson *et al.,* 1994). Lors de la biosynthèse de la CN-β, la phosphorylation se déroule de manière séquentielle et n'est pas due au hasard. Ainsi le variant 1P est phosphorylé dans 50% des cas sur le résidu de Ser^{9} et dans 50% des cas sur le résidu de Ser^{10}. Le variant 2P est phosphorylé sur les résidus de Ser^{9} et Ser^{10}, puis s'ajoutent successivement pour les autres degrés de phosphorylation le résidu de Ser^{8} puis le résidu de Ser^{6} et enfin le résidu de Thr^{3} (Greenberg *et al.,* 1984).

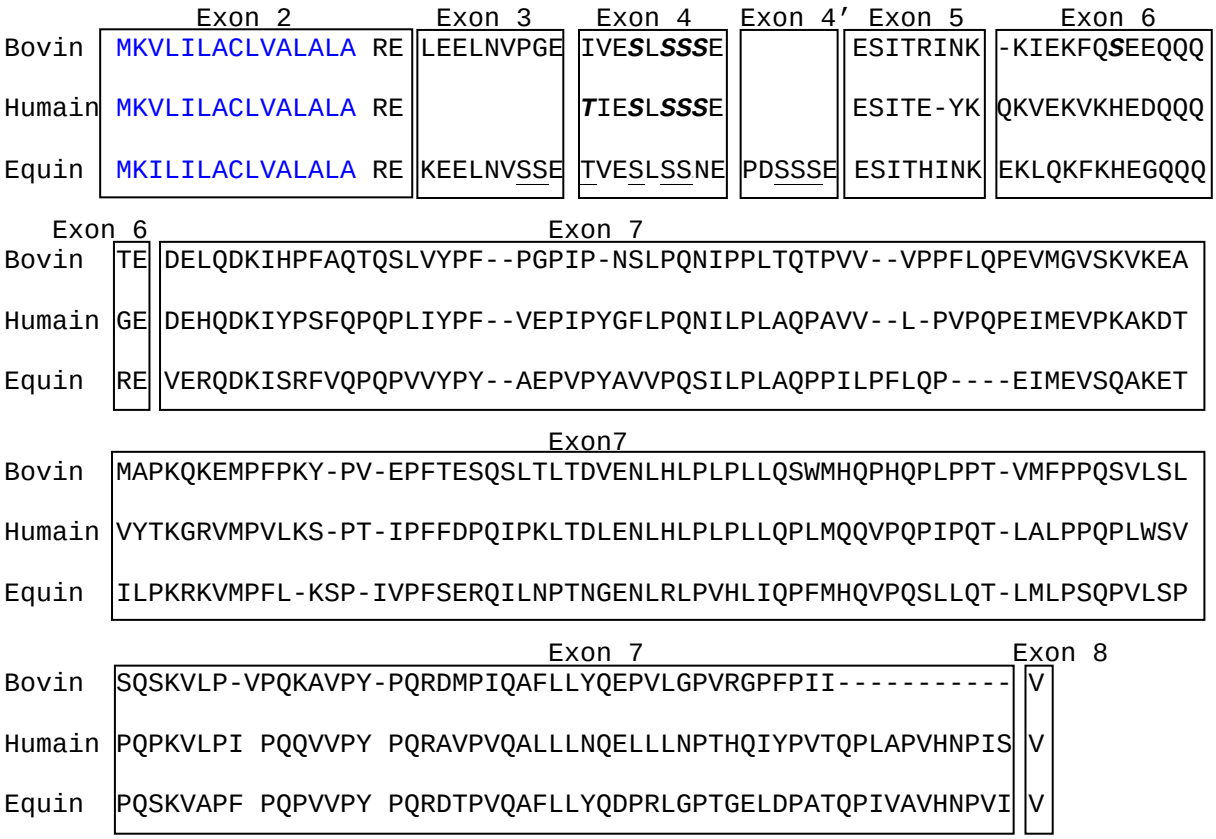


Figure 3 : Comparaison de la structure exonique des séquences des caséines β bovine, humaine et équine. Les numéros d'accession dans la banque de données SwissProt sont les suivants : NP851351 pour la séquence bovine (*Bos taurus*), PO5814 pour la séquence humaine (*Homo sapiens*) et Q9GKK3 pour la séquence équine (*Equus caballus*). Les résidus phosphorylés sont indiqués en gras et italique. Les résidus potentiellement phosphorylés sont soulignés. Le peptide signal est indiqué en bleu.

La CN-β équine dans sa forme complète (c'est-à-dire dont tous les exons sont traduits) contient 226 résidus d'acides aminés et comporte neuf sites potentiels de phosphorylation (Girardet *et al.,* 2006 ; Figure 1). Il a été montré que la CN-β équine présente un degré de phosphorylation variable, comme la CN-β humaine, allant de 3P à 7P (Girardet *et al.,* 2006). Il existe des variants d'épissage alternatif de cette protéine. En effet, l'exon 5 peut être excisé et conduire à une CN-β délétée de huit résidus d'acides aminés (Figure 3 ; Miranda *et al.,* 2004 ; Lenasi *et al.,* 2006) que l'on nommera CN-$\beta^{\Delta 5}$. Un autre variant d'épissage pour lequel 75% de l'exon 7 a été délété, dû à l'utilisation d'un site cryptique d'épissage, ce qui représente 132 résidus d'acides aminés, a été mis en évidence par Miclo *et al.* (2007) ; ce variant est appelé « low-M_r CN-β". Un ADNc d'un variant résultant de l'épissage de l'exon 5 et de l'exon 8 a été obtenu (Lenasi *et al.,* 2006) mais le produit de traduction de l'ARNm correspondant n'a encore jamais été trouvé. La CN-β équine présente une microhétérogénéité encore accrue, car elle possède un site de désamidation non-enzymatique au niveau du résidu Asn^{135} dans une séquence de type Asn-Gly (Girardet *et al.,* 2006). La désamidation de ce résidu conduit à une protéine présentant une augmentation de masse de 1 Da et l'apparition d'une charge négative supplémentaire. Les caséines β bovine et humaine ne possèdent pas de séquence Asn-Gly cible d'une désamidation potentielle. La CN-β équine, comme son homologue bovin, présente un site sensible à la chymosine au niveau de la liaison Leu^{190}-Tyr^{191} (Egito *et al.,* 2001).

1.2.2. La caséine α_{s1}

La CN-α_{s1} bovine est la protéine majeure du lait de vache : elle représente plus de 40% des caséines totales. Elle présente huit variants génétiques (A, B, C, D, E, F, G et H ; Ng-Kwai-Hang, 2003) parmi lesquels le variant B est le plus fréquent (87 à 99% chez la vache de race Prim'Holstein ; Remeuf *et al.,* 1991). Le variant B est constitué de 199 résidus d'acides aminés (Figure 4) : le variant majoritaire possèdant huit résidus toujours phosphorylés (Ser^{46}, Ser^{48}, Ser^{64}, Ser^{66}, Ser^{67}, Ser^{68}, Ser^{75} et Ser^{115}) est appelé CN-α_{s1} B-8P (Eigel *et al.,* 1984). Il existe un variant B minoritaire possédant un résidu phosphorylé supplémentaire (Ser^{41}) et qui est nommé CN-α_{s0} B-9P (Eigel *et al.,* 1984). La CN-α_{s1} aurait un rôle de charpente de la micelle.

Bovin

[1]MKLLILTCLV AVALARPKHP IKHQGLPQEV LNENLLRFFV APFPEVFGKE KVNEL***S***KDIG
[61]***S***E***S***TEDQAME DIKQMEAE***S***I ***SSS***EEIVPN***S*** VEQKHIQKDD VPSERYLGYL EQLLRLKKYK
[121]VPQLEIVPN***S*** AEERLHSMKE GIHAQQKEPM IGVNQELAYF YPELFRQFYQ LDAYPSGAWY
[181]YVPLGTQYTD APSFSDIPNP IGSENSGKTT MPLW[214]

Equin

[1]MKLLILTCLV AVALARPKLP HRQPEIIQNE QDSREKVLKE RKFPSFALEY INELNRQREL
[61]LKEKQKDEHK DTSNESTEEY LIEDPEQQES SSTSSSEEVV PINTEQKRIP REDMLYQHTL
[121]EQLRRLSKYN QLQLQAIHAQ EQLIRMKENS QRKPMRVVNQ EQAYFYLEPF QPSYQLDVYP
[181]YAAWFHPAQI MQHVAYSPFH DTAKLIASEN SEKTDIIPEW[220]

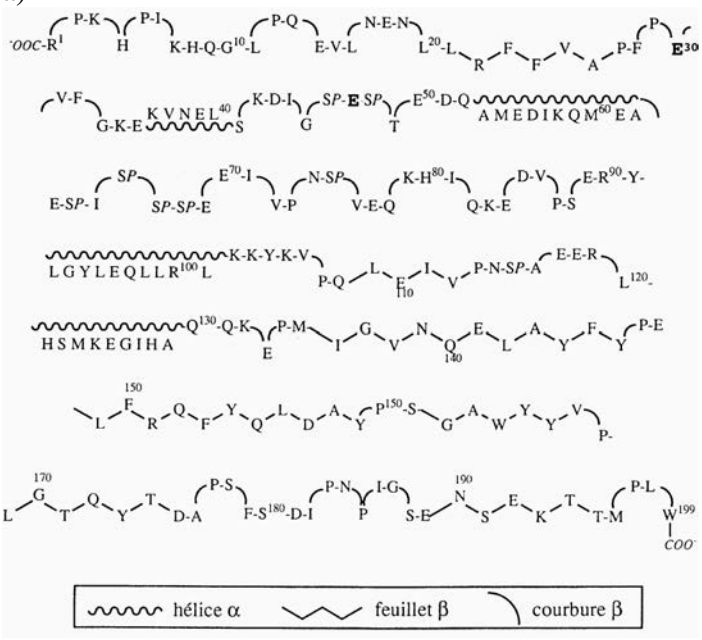
Humain

[1]MRLLILTCLV AVALARPKLP LRYPERLQNP SE***S***EPIPLE***S*** REEYMNGMNR QRNILREKQT
[61]DEIKDTRNES TQNCVVAEPE KMESSI***S***S***S*** EEMSLSKCAE QFCRLNEYNQ LQLQAAHAQE
[121]QIRRMNENSH VQVPFQQLNQ LAAYPYAVWY YPPQIMQYVP FPPFSDISNP TAHENYEKNN
[181]VMLQW[185]

Figure 4 : Comparaison des séquences des caséines α_{s1} bovine, humaine et équine. Les numéros d'accession dans la banque de données SwissProt sont les suivants : ACG63494 pour la séquence bovine (*Bos taurus*), P47710 pour la séquence humaine (*Homo sapiens*), NP001075352 et AAU01570 pour la séquence équine (*Equus caballus*). Les résidus phosphorylés sont indiqués en gras et italique. Les résidus potentiellement phosphorylés sont soulignés. Le peptide signal est indiqué en bleu.

La CN-α_{s1} bovine ne possède pas de résidus de cystéine. Elle possède peu de structures ordonnées avec un faible pourcentage d'hélices α (8 à 13%), mais avec 20-23% de feuillets β et 29-35% de coudes β (Byler *et al.*, 1988). Les hélices α seraient localisées au niveau des fragments [37-40], [50-69], [75-79], [95-100] et [116-128] (Swaisgood, 1988). Le modèle hypothétique tridimensionnel de la CN-α_{s1} établi par Farrell Jr. *et al.* (1993) la fait apparaître sous la forme d'un "chien assis" avec les résidus de phosphosérine présents dans un coude β, qui formeraient une région hydrophile séparée du reste de la protéine (Figure 5).

a)



b)

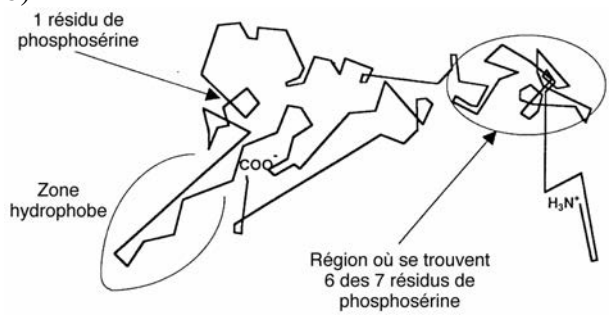


Figure 5 : Structures primaire et secondaire prédites de la caséine α_{s1} bovine, variant B (a : d'après Mercier *et al.,* 1971 ; Farrell Jr. *et al.,* 1993) et modèle hypothétique de la structure tridimensionnelle de la caséine α_{s1} bovine (b : d'après Farrel Jr. *et al.,* 1993). Figures reproduites d'après Cayot et Lorient (1998), revues et corrigées d'après la banque de données Swissprot (numéro d'accession : P02662).

La CN-α_{s1} humaine est très peu présente dans le lait humain. Trois types de transcrits d'ARNm différents ont été identifiés : le type 1 correspondant à la protéine mature complète (170 résidus d'acides aminés ; Figures 4 et 6), le type 2 correspondant à la délétion du résidu Gln^{52} et le type 3 correspondant à la délétion de l'exon 7 (Johnsen *et al.,* 1995). Le type 1 et le type 2 sont les plus abondants. Bien que la structure primaire présente neuf sites potentiels de phosphorylation (sur huit résidus de sérine et un résidu de thréonine), la CN-α_{s1} humaine a un degré de phosphorylation variable allant de 0P à 4P, les résidus phosphorylés étant Ser^{18}, Ser^{26}, Ser^{73} et Ser^{75} (Kjeldsen *et al.,* 2007). La CN-α_{s1} humaine, qui contient trois résidus Cys en position 75, 99 et 104 (Rasmussen *et al.,* 1995), a la particularité de pouvoir former des multimères avec la CN-κ *via* des liaisons par ponts disulfure.

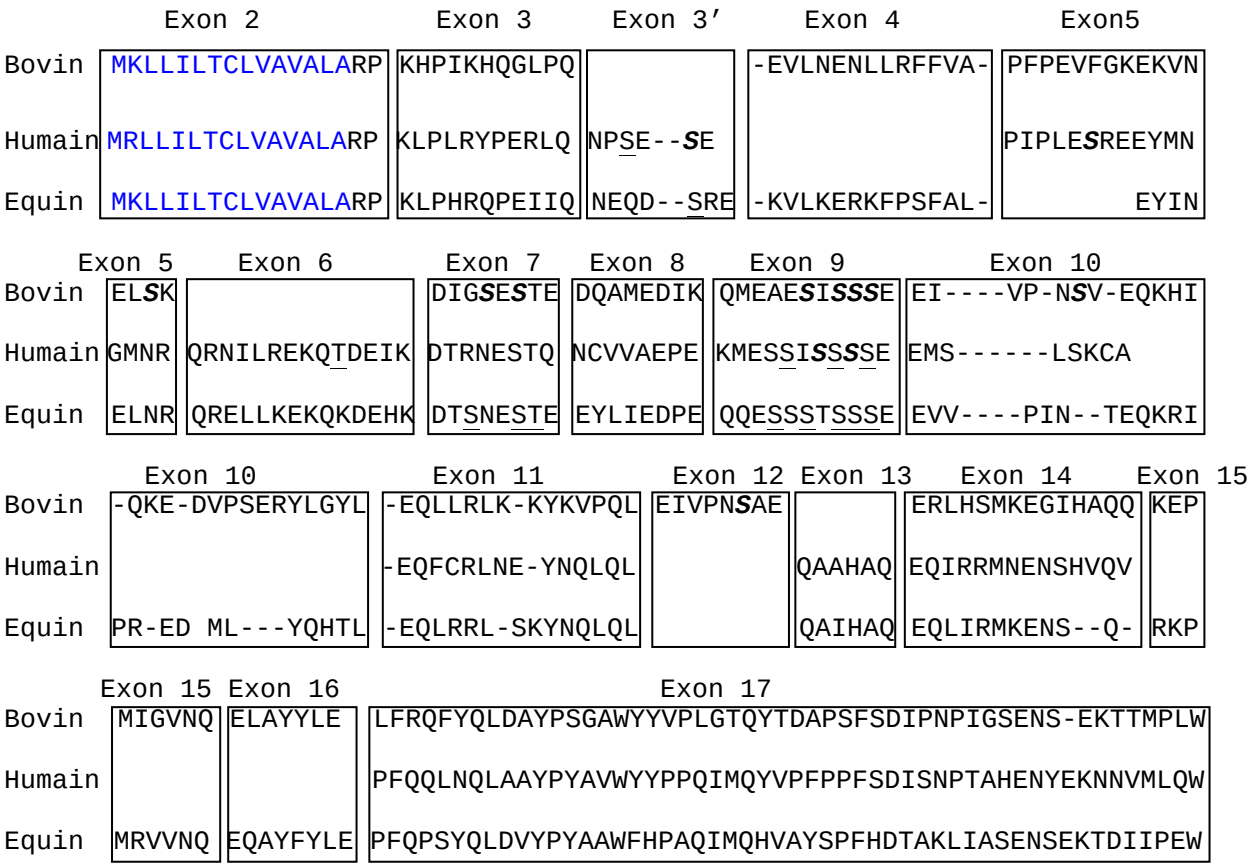


Figure 6 : Comparaison de la structure exonique des séquences des caséines α_{s1} bovine, humaine et équine. Les numéros d'accession dans la banque de données SwissProt sont les suivants : BAA00313 pour la séquence bovine (*Bos taurus*), P47710 pour la séquence humaine (*Homo sapiens*), NP001075352 et AAU01570 pour la séquence équine (*Equus caballus*). Les résidus phosphorylés sont indiqués en gras et italique. Les résidus potentiellement phosphorylés sont soulignés. Le peptide signal est indiqué en bleu.

La CN-α_{s1} équine (Figures 4 et 6) est, comme son homologue bovin, la protéine majeure du lait de jument (40 à 59% ; Malacarne *et al.,* 2002). Plusieurs ARNm ont été isolés à partir de cellules de glande mammaire, suggérant l'existence d'un phénomène d'épissage alternatif concernant les exons 7 et 14. En effet, un ARNm correspondant à la protéine CN-$\alpha_{s1}^{\Delta 7}$ a été isolé chez une jument de race Hanovre (Lenasi *et al.,* 2003), et un ARNm correspondant à CN-$\alpha_{s1}^{\Delta 14}$ a été isolé chez le poney gallois (Milenkovic *et al.,* 2002). La présence de CN-$\alpha_{s1}^{\Delta 7}$ et de CN-$\alpha_{s1}^{\Delta 7\text{-}14}$ a été fortement suggérée au niveau protéique (précision par spectrométrie de masse de 0,2%) chez le poney gallois (Miranda *et al.,* 2004). Lenasi *et al.* (2003) ont montré que l'exon 5 est constitutivement délété. La structure primaire de la protéine complète (c'est-à-dire comportant les régions codées par les exons 7 et 14) est constituée de 205 résidus d'acides aminés et contient dix sites potentiels de phosphorylation

(sur huit résidus de sérine et deux résidus de thréonine). Cette protéine ne contient pas de résidus Cys. Les profils électrophorétiques obtenus en gel de polyacrylamide en présence d'urée montrant la présence de plusieurs bandes de CN-α_{s1}, laissent supposer que, comme dans le cas de la CN-α_{s1} humaine, la CN-α_{s1} équine aurait un degré variable de phosphorylation, et donc un nombre de charges négatives variable, responsable de la complexité du profil électrophorétique (Egito *et al.,* 2002).

1.2.3. La caséine α_{s2}

La CN-α_{s2} bovine est la caséine la moins abondante (10% des caséines totales) et la plus sensible au calcium. Elle contient 207 résidus d'acides aminés et possède deux clusters de résidus Ser phosphorylés (Figure 7 ; FitzGerald, 1998). La CN-α_{s2} est le composant caséinique le plus phosphorylé avec un degré de phosphorylation variable allant de 10P à 13P (Swaisgood, 1992). Certains sites phosphorylés ont été localisés : Ser^{8}, Ser^{9}, Ser^{10}, Ser^{16}, Ser^{31}, Ser^{56}, Ser^{57}, Ser^{58}, Ser^{61}, Ser^{129}, Ser^{131} et Ser^{143}. Les sites potentiels restant sont les résidus Thr^{3}, Thr^{66}, Thr^{130} et Thr^{154} (Farrell Jr. *et al.,* 2004). Il existe quatre variants génétiques (A, B, C et D ; Eigel *et al.,* 1984) et l'isoforme CN-α_{s2} A-11P est la plus fréquente (Whitney *et al.,* 1976).

Bovin

^{1}MKFFIFTCLL AVALAKNTME HV***SSS***EESII ***S***QETYKQEKN MAINP***S***KENL CSTFCKEVVR
^{61}NANEEEYSIG ***SSS***EE***S***AEVA TEEVKITVDD KHYQKALNEI NQFYQKFPQY LQYLYQGPIV
^{121}LNPWDQVKRN AVPITPTLNR EQLST***S***EENS KKTVDME***S***TE VFTKKTKLTE EEKNRLNFLK
^{181}KISQRYQKFA LPQYLKTVYQ HQKAMKPWIQ PKTKVIPYVR YL^{222}

Figure 7 : Séquence de la caséines α_{s2} bovine (numéro d'accession dans la banque de données SwissProt : P02663). Les résidus phosphorylés sont indiqués en gras et italique. Les résidus potentiellement phosphorylés sont soulignés. Le peptide signal est indiqué en bleu.

Cette protéine a un rôle structural très important dans la micelle *via* la formation de ponts phosphocalciques. La présence de deux résidus Cys permet la formation de ponts disulfures intra- et inter-moléculaires (Rasmussen *et al.,* 1992 ; 1994). La CN-α_{s2} existe sous forme monomérique (>85%) avec un pont disulfure entre les résidus Cys^{36} et Cys^{40} (Rasmussen *et al.,* 1994) ou sous forme dimérique avec deux ponts disulfure soit en orientation parallèle

(ponts disulfure entre deux chaînes polypeptidiques 36-36 et 40-40) soit en orientation anti-parallèle (ponts disulfure 36-40 et 40-36 ; Rasmussen *et al.*, 1992). Il n'existe pas de trimère de CN-α_{s2} (Ramussen *et al.*, 1992). La CN-α_{s2} peut aussi former des ponts disulfures avec la CN-κ. La CN-α_{s2} a été étudiée par dichroïsme circulaire (Haga *et al*., 1983 ; Hoagland *et al.*, 2001) ainsi que par spectroscopie infra-rouge à transformée de Fourier (Hoagland *et al.*, 2001). Sa structure secondaire contient 20% (Haga *et al*., 1983) à 32% (Hoagland *et al.*, 2001) d'hélices α et des proportions importantes de feuillets β (environ 30%) et de coudes β (24 à 31% ; Hoagland *et al.*, 2001). La CN-α_{s2} est la plus hydrophile des caséines. D'une part, les 37 résidus d'acides aminés C-terminaux possèdent une charge nette globale positive (environ + 9,5) au pH du lait bien que cette région soit relativement hydrophobe (Swaisgood, 1992). D'autre part, les 38 résidus d'acides aminés N-terminaux possèdent une charge nette négative (-21) au pH du lait. La structure primaire de la CN-α_{s2} peut être donc divisée en quatre domaines : un domaine N-terminal hydrophile avec des clusters anioniques, un domaine central hydrophobe suivi par un autre domaine hydrophile avec des clusters anioniques et enfin un domaine C-terminal hydrophobe chargé positivement (Swaisgood, 1992). Cette structure est cohérente avec le comportement d'association de la CN-α_{s2} qui est très dépendant de la force ionique (Snoeren *et al.,* 1980). L'association est la plus forte pour une force ionique de 0,2 M ; avec une dissociation se déroulant à de faibles concentrations en sels provoquée par des répulsions électrostatiques, mais aussi à de fortes concentrations en sels par suppression de l'attraction électrostatique. Cela reflète bien les contributions des interactions hydrophobes et de l'attraction électrostatique.

La CN-α_{s2} équine n'a pas encore été séquencée, ni même purifiée, à l'heure actuelle. Des fragments peptidiques pouvant appartenir à une protéine de type CN-α_{s2} (cette hypothèse est formulée après comparaison avec les séquences d'autres espèces) ont été cependant trouvés (Ochirkhuyag *et al.,* 2000 ; Egito *et al.,* 2001 ; Miranda *et al.,* 2004). Néanmoins, aucun ARNm n'a encore été isolé.

1.2.4. La caséine κ

La CN-κ bovine contient 169 résidus d'acides aminés et comporte 1P, 2P ou 3P (Figure 8). Les trois sites phosphorylés sont les résidus Ser^{127}, Ser^{149} et Thr^{135} ; le résidu Ser^{149} est toujours phosphorylé tandis que le résidu Ser^{127} ne l'est que partiellement. La forme 3P est

quant à elle minoritaire (Mercier, 1981 ;Vreeman *et al.,* 1986 ; Mollé et Léonil, 1995 ; Holland *et al.,* 2006).

Bovin

[1]MMKSFFLVVT ILALTLPFLG AQEQNQEQPI RCEKDERFFS DKIAKYIPIQ YVLSRYPSYG
[61]LNYYQQKPVA LINNQFLPYP YYAKPAAVRS PAQILQWQVL SNTVPAKSCQ AQPTTMARHP
[121]HPHLSFMAIP PKKNQDKTEI PTINTIA***S***GE P**T**S**T**PTTEAV ES**T**VA***T***LED***S*** PEVIESPPEI
[181]NTVQVTSTAV[190]

Equin

[1]MKSFFLVVNI LALTLPFLGA EVQNQEQPTC HKNDERFFDL KTVKYIPIYY VLNSSPRYEP
[61]IYYQHRLALL INNQHMPYQY YARPAAVRPH VQIPQWQVLP NIYPSTVVRH PCPHPSFIAI
[121]PPKKLQEITV IPKIN***T***IATV EP***T***PIP***T***PEP ***T***VNNAVIPDA SSEFIIAS***T***P E***TTT***VPV***T***SP
[181]VVQKL[185]

Humain

[1]MKSFLLVVNA LALTLPFLAV EVQNQKQPAC HENDERPFYQ KTAPYVPMYY VPNSYPYYGT
[61]NLYQRRPAIA INNPYVPRTY YANPAVVRPH AQIPQRQYLP NSHPPTVVRR PNLHPSFIAI
[121]PPKKIQDKII IP***T***INTIATV EP***T***PAPA***T***EP ***T***VDSVV***T***PEA FSESII***T***S***T***P ETTTVAVTPP
[181]TA[182]

Figure 8 : Comparaison des séquences des caséines κ bovine, humaine et équine. Les numéros d'accession dans la banque de données SwissProt sont les suivants : NP776719 pour la séquence bovine (*Bos taurus*), P07498 pour la séquence humaine (*Homo sapiens*), P82187 pour la séquence équine (*Equus caballus*). Les résidus phosphorylés sont indiqués en gras et italique. Les résidus potentiellement phosphorylés sont soulignés. Les résidus glycosylés sont indiqués en rouge. Les résidus potentiellement glycosylés sont en rouge et italique. Le peptide signal est indiqué en bleu.

Cette caséine ne précipite pas en présence de calcium et permet donc de stabiliser la structure micellaire. La CN-κ a la particularité de posséder des formes glycosylées car elle présente six sites de O-glycosylation (résidus de Thr^{121}, Thr^{131}, Thr^{133}, Thr^{136}, Thr^{142} et Thr^{165}) situés dans la partie C-terminale de la protéine (Pisano *et al.,* 1993 ; Pisano *et al.,* 1994). Les glycannes sont constitués de résidus de N-acétylgalactosamine (GalNAc), de galactose (Gal) et d'acide N-acétyl neuraminique (NeuAc). La chaîne O-glycannique majoritaire est une structure tétrasaccharidique ramifiée (NeuAc(α2-3)Gal(β1-3)[NeuAc(α2-6)]GalNAc). Il a été

également montré l'existence de structures trisaccharidiques (NeuAc(α2-3)Gal(β1-3)GalNAc) ou (Gal(β1-3)[NeuAc(α2-6)]GalNAc), disaccharidiques (Gal(β1-3)GalNAc) ou encore monosaccharidique (GalNAc ; Figure 9 ; Saito et Itoh, 1992 ; Holland *et al.,* 2006).

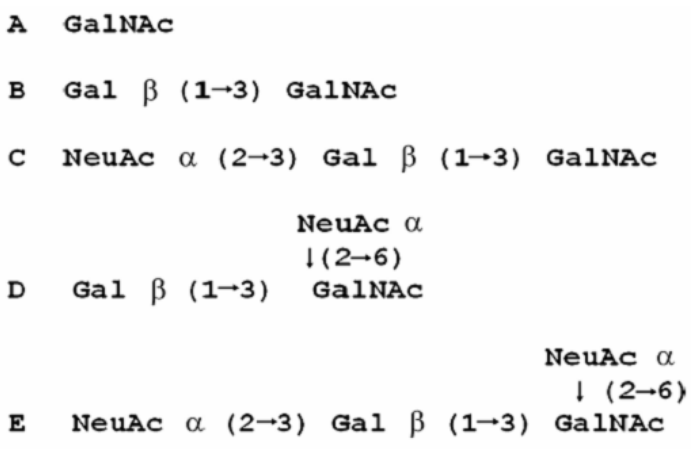


Figure 9 : Structure des différentes chaînes glycanniques retrouvées au niveau de la caséine κ bovine (d'après Saito et Itoh 1992). GalNAc, N-acétylgalactosamine ; Gal, galactose ; NeuAc, N-acétyl neuraminique.

Les quantités relatives de chaque type de chaînes glycanniques ont été déterminées par HPLC ; ainsi il y a 56% de chaînes tétrasaccharidiques, 36,9% de chaînes trisaccharidiques, 6,3% de chaînes disaccharidiques et 0,8% de résidus GalNAc seuls (Saito et Itoh 1992). Une forme non glycosylée de la CN-κ bovine (37% de la CN-κ totale) contenant deux résidus de phosphosérine (Ser^{127} et Ser^{149}), appelée CN-κ1, co-existe avec les formes glycosylées qui, elles, possèdent majoritairement un seul résidu de phosphosérine (Ser^{149}), dans de rares cas deux (Ser^{127} et Ser^{149}) mais jamais trois (Holland *et al.,* 2006). Des formes non glycosylées à 1P (Ser^{149}) et 3P (Ser^{127}, Ser^{149} et Thr^{166}) ont également été mises en évidence (Holland *et al.,* 2004, 2006). Il existe 11 variants génétiques de la CN-κ bovine cependant les variants A (75 à 80% des cas) et B (15 à 25% des cas) sont les variants majoritaires (Remeuf *et al.,* 1991). L'hydrolyse de la CN-κ bovine par la chymosine au niveau de la liaison Phe^{105}-Met^{106} génère deux fragments : le fragment 1-105 ou paracaséine-κ (hydrophobe) et le fragment 106-169 ou glycomacropeptide (hydrophile) qui contient tous les résidus glycosylés. Cette hydrolyse est à l'origine de la formation du caillé (Swaisgood, 1982). Les études de dichroïsme circulaire ont montré que la caséine κ était la plus organisée des caséines avec 14% d'hélices α et 31% de feuillets β (Loucheux-Lefebvre *et al.*, 1978). Il existerait une structure spécifique de type coude β-feuillet β-coude β aux environs de la liaison Phe^{105}-Met^{106} sensible à la chymosine (Holt, 1992). La modélisation prédictive de Kumosinski *et al.* (1991) montre que la CN-κ est

sous la forme d'un "cavalier sur son cheval", que la paracaséine-κ correspond au "cheval" et le glycomacropeptide au "cavalier". Les régions hydrophobes de la CN-κ seraient regroupées en feuillets β antiparallèles au niveau des régions représentant les "pattes du cheval" (Figure 10). Celles-ci seraient un point d'association avec d'autres molécules de CN-κ par un processus auto-associatif mais également avec les caséines α_{S1}, α_{S2} et β.

a) b)

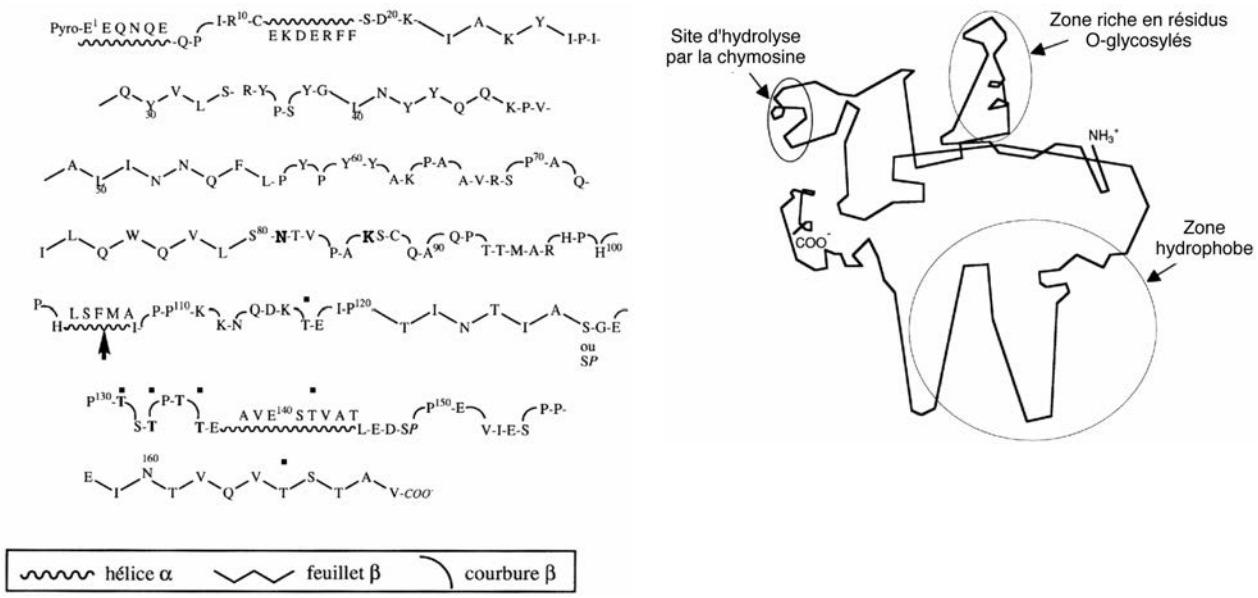


Figure 10 : Structures primaire et secondaire prédites de la caséine κ bovine, variant A (a : d'après Mercier *et al.,* 1973 ; Farrell Jr. *et al.,* 1993) et modèle hypothétique de la structure tridimensionnelle de la caséine κ bovine (b : d'après Kumosinski *et al.,* 1991). Figures reproduites d'après Cayot et Lorient (1998), revues et corrigées d'après la banque de données Swissprot (numéro d'accession : P02668). La flèche en a) indique le site d'hydrolyse par la chymosine, les résidus Thr marqués d'un point peuvent être glycosylés (Pisano *et al.,* 1994). Si en position 127 se trouve un résidu Ser phosphorylé (SP), la caséine κ n'est pas glycosylée.

La CN-κ humaine, constituée de 162 résidus d'acides aminés (Bergström *et al.,* 1992), représente 10% des caséines humaines totales. Elle possède un résidu Cys (Cys^{10}). Cette protéine possède sept sites de O-glycosylation qui sont les résidus Thr^{113}, Thr^{123}, Thr^{128}, Thr^{131}, Thr^{137}, Thr^{147} et Thr^{149} (Figure 8 ; Fiat *et al.,* 1980) et les structures glucidiques représentent de 50 à 60% de sa masse moléculaire (Kunz et Lönnerdal, 1989). La CN-κ humaine contient 5% de résidus de fucose, 17% de résidus de galactose, 18% de résidus de N-acétylglucosamine, 8% de résidus de N-acétylgalactosamine et 7% de résidus d'acide sialique

(Dev *et al.,* 1993). D'après sa structure primaire, la CN-κ humaine possède trois sites potentiels de phosphorylation : Thr^{119}, Thr^{137} et Thr^{149}. Une étude de la structure secondaire par résonance magnétique nucléaire du fragment [84-101] montre la présence d'un coude β au niveau de la région Pro^{87}-His^{90} (Plowman *et al.,* 1999).

La CN-κ équine est une protéine de 165 résidus d'acides aminés possèdant quatre sites potentiels de phosphorylation (Thr^{119}, Thr^{127}, Thr^{149} et Ser^{141} ; Iametti *et al.,* 2001). Elle possède deux résidus Cys (Cys^{10} et Cys^{92}). Elle comporte également neuf sites potentiels de O-glycosylation présentant des motifs de glycosylation typique (résidus de Thr proches de résidus Arg/Lys, Thr ou Pro ; Lenasi *et al.,* 2003) au niveau des résidus de Thr^{116}, Thr^{123}, Thr^{127}, Thr^{131}, Thr^{149}, Thr^{152}, Thr^{153}, Thr^{154} et Thr^{158} (Figure 8 ; Egito *et al.,* 2001). Elle contiendrait des motifs de type Gal(β1,3)GalNAc (Iametti *et al.,* 2001). Le site de clivage par la chymosine est différent de celui de la CN-κ bovine, l'hydrolyse ayant lieu au niveau de la liaison Phe^{97}-Ile^{98} (Egito *et al.,* 2001). Nakhasi *et al.* (1984) ont proposé une classification en deux groupes des caséines κ des différentes espèces étudiées. Le groupe I (brebis, bufflonne, chèvre et vache) diffère principalement du groupe II (femme, ratte, souris et truie) au niveau du site de coupure par la chymosine : liaison Phe-Met pour le groupe I et Phe-Ile ou Phe-Leu pour le groupe II. Cette différence reflète probablement les différences dans le mécanisme de coagulation du lait chez les ruminants et les non-ruminants (Herskovitis, 1966). La CN-κ équine appartient donc au groupe II comme la CN-κ humaine.

1.2.5. La micelle de caséines

Malgré les variations dans la composition en caséine du lait des différentes espèces, les caséines se présentent sous forme de micelles. La couleur plus ou moins blanche caractéristique du lait est principalement due à la diffusion de la lumière par les micelles. L'ultra-structure des micelles de caséines de la majorité des espèces est similaire. La présence de micelles dans le lait suggère qu'elles doivent avoir un rôle physiologique ou nutritionnel qui va au-delà de la seule valeur nutritionnelle des caséines. Le calcium et le phosphate sont nécessaires au développement des dents et des os et le taux de croissance est positivement corrélé avec les concentrations de calcium et de phosphate inorganique. Cependant, le phosphate de calcium est très peu soluble au pH du lait. Il devrait donc précipiter dans la glande mammaire provoquant la formation de concrétions bouchant les canaux galactophores

et pouvant entraîner la mort de cet organe. En formant des micelles, les caséines maintiennent l'excès de phosphate de calcium dans un état colloïdal stable et permettent la sécrétion de lait contenant une forte concentration en phosphate de calcium soluble. La coagulation des micelles dans l'estomac du nouveau-né (chez certaines espèces) par la chymosine pourrait retarder l'entrée des constituants du lait dans l'intestin grêle améliorant ainsi leur digestibilité (Fox et Brodkorb, 2008).

Les micelles sont dispersées par l'urée, le SDS, un pH élevé ou encore l'éthanol (concentration supérieure à 35%, 70°C) ce qui montre que des interactions hydrophobes et électrostatiques ainsi que des liaisons hydrogène sont impliquées dans l'intégrité de la micelle. Plus de 50% de la CN-β, la plus hydrophobe des caséines, se dissocie de la micelle de manière réversible au froid ce qui indique l'importance des interactions hydrophobes et suggère que la micelle est suffisamment poreuse pour permettre à la CN-β de diffuser en dehors. Si la température est remontée après un passage au froid, la CN-β forme une couche diffuse à la surface de la micelle (Fox et Brodkorb, 2008).

Les caséines β, α_{s1} et α_{s2} sont dites « sensibles » au calcium puisqu'elles sont précipitées par le calcium qui se lie à leurs résidus Ser phoshorylés. La CN-κ est dite « insensible » au calcium car elle est soluble en présence de calcium. La CN-κ interagit avec les autres caséines afin de les stabiliser et d'initier la formation d'un état colloïdal stable (Farrell Jr *et al.,* 2006). La localisation de la CN-κ dans la micelle doit lui permettre de stabiliser les caséines « sensibles » au calcium, d'être rapidement hydrolysée par la chymosine et d'autres protéases et de rendre possible l'association avec les protéines lactosériques lors du chauffage du lait. La localisation la plus probable de la CN-κ dans la micelle est donc à sa surface (Fox, 2003). La couche « chevelue » de CN-κ à la surface de la micelle assurerait sa stabilité à travers un mécanisme de stabilisation stérique (Holt et Horne, 1996).

Plusieurs modèles de structures de la micelle de caséine ont été proposés et parmi ces modèles, deux à l'heure actuelle sont très discutés.

L'un des modèles (modèle de Schmidt) repose sur l'existence de submicelles de caséines qui s'associent par pontage phosphocalcique pour former la micelle. La micelle (diamètre moyen de 180 nm chez la vache) serait constituée, en son cœur, de submicelles (10 nm) de caséines α_{s1} et β et, à sa périphérie, de submicelles (6 nm) de caséines α_{s1}, α_{s2} et κ exposées au solvant (Ono et Obata, 1989). Les submicelles seraient reliées entre-elles par des ponts phosphocalciques. La CN-κ serait principalement localisée à la surface de la micelle,

accessible à la chymosine. Les caséines κ les plus phosphorylées engageraient des liaisons phosphocalciques avec les autres caséines et seraient orientées vers l'intérieur de la micelle, tandis que les caséines κ O-glycosylées (les moins phosphorylées) seraient exposées au solvant et formeraient une "chevelure" dont l'épaisseur (en moyenne 12 nm) varierait selon les caractéristiques du milieu (Figure 11 ; Holt, 1992 ; Tchorbanov et Iliev, 1993). De la CN-α_{s1} pourrait également se trouver en surface avec la CN-κ (Dalgleish *et al.*, 1989). Les groupements phosphorylés de cette dernière caséine associés aux groupements glycosylés de la CN-κ joueraient un rôle stabilisateur. Walstra (1997) propose une variante aux modèles classiques de la micelle : le phosphate de calcium, sous la forme de granules, serait situé au cœur des submicelles et ces submicelles seraient liées entre-elles par des interactions hydrophobes et un "ciment" protéique constitué de CN-κ.

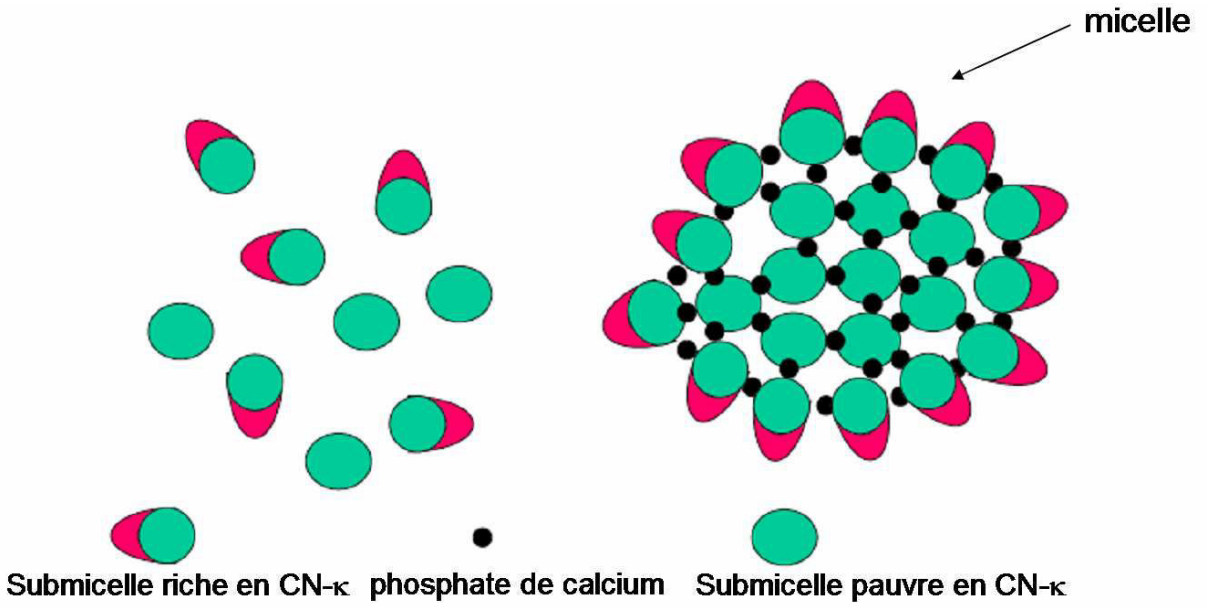


Figure 11 : Schéma (d'après Horne, 2006) représentant le modèle submicelle de formation des micelles de caséines selon le modèle proposé par Schmidt (1982).

L'autre modèle réfute la théorie des submicelles. La partie phosphorylée de la CN-β peut lier et stabiliser des agrégats phosphocalciques ce qui permet la formation de nanoclusters. Ces nanoclusters, dans ce modèle (De Kruif et Holt, 2003), seraient les pièces centrales de la structure de la micelle de caséines. La formation de nanoclusters conduirait à la formation de la micelle par liaison au hasard des protéines phosphorylées créant une micelle inversée. Ainsi, plus de protéines pourraient recouvrir cette surface hydrophobe permettant la liaison de plus de phosphate de calcium jusqu'à ce soit atteinte la taille limite de formation d'un colloïde. Les protéines non structurées ont pour fonction de lier le phosphate de calcium,

ce qui les structure. Par conséquent, dans ce modèle, il n'y a pas nécessité de structure secondaire spécifique ou d'interactions protéine-protéine sauf pour la CN-κ à la surface. Le modèle de Holt est présenté dans la figure 12a. Horne (1998) a proposé un autre modèle pouvant être considéré comme une extension du modèle de Holt. Dans son modèle, Horne prend en considération les interactions protéine-protéine. La nature amphiphile des caséines fait qu'elles agissent comme des blocs de copolymères alternant une partie chargée et une partie hydrophobe. Ainsi, la CN-β peut être considérée comme un bloc comportant une boucle chargée (phosphorylée) et une région hydrophobe (agencement inverse pour la CN-κ), les caséines α_{s1} et α_{s2} sont considérées comme un bloc contenant une région hydrophobe, une boucle chargée et une autre région hydrophobe. Ce modèle tient compte des interactions hydrophobes mais pas de la structure secondaire des protéines. La croissance des nanoclusters de phosphate de calcium débute le processus de la formation de la micelle mais la croissance des nanoclusters est limitée par la liaison aux phosphopeptides des boucles chargées. Une fois lié à la matrice inorganique amorphe, des interactions protéine-protéine lient les blocs hydrophobes entres eux (Figure 12b). La formation de la micelle conduit à une structure interne ressemblant à un gel composé de nanoclusters de phosphate de calcium. La réaction avec la CN-κ ne contenant qu'un résidu phosphorylé limite la croissance micellaire.

a) b)

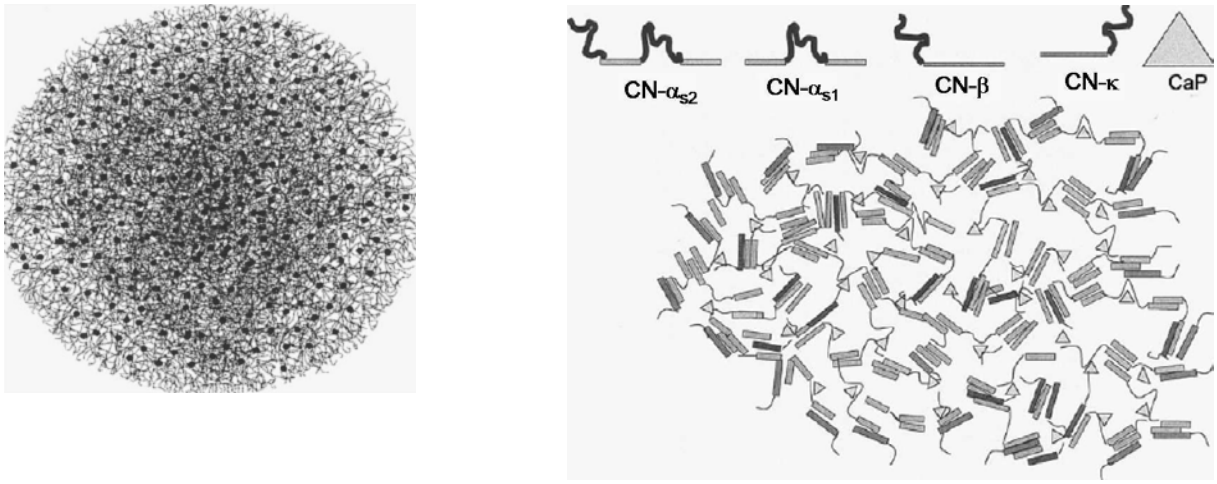


Figure 12 : Modèles de micelles de caséines bovines prenant en compte l'importance du phosphate de calcium dans la structure de la micelle et localisant la CN-κ à la surface de la micelle (d'après Farrel Jr. *et al.,* 2006). a) modèle de Holt, les monomères de caséine sont représentés sous la forme de filament et les cercles noirs représentent les nanoclusters de phosphate de calcium. b) modèle de Horn, des interactions protéine-protéine ont lieu entre les régions hydrophobes (rectangle) tandis que les régions hydrophiles (boucles) lient les clusters de phosphate de calcium (triangle). La CN-κ est monomérique et en surface.

2. Modification post-transcriptionnelle : l'épissage conventionnel et l'épissage alternatif d'exons

Chez les eucaryotes, les gènes codant les protéines sont organisés en une suite d'introns et d'exons alternés. Les introns représentent la partie non-codante de l'ARNm et les exons la partie codante. Lors de la transcription, un pré-ARNm contenant les introns et les exons est synthétisé dans le noyau. Celui-ci va subir une maturation pendant laquelle les introns vont être épissés, c'est-à-dire éliminés, et les exons vont être conservés afin de donner un ARNm dit mature qui sera ensuite traduit en protéine. Ce phénomène d'épissage se déroule dans le noyau de la cellule puis l'ARNm mature gagne le cytoplasme afin d'être traduit en protéine au niveau du réticulum endoplasmique.

2.1. Le splicéosome

L'épissage est assuré par un complexe ribonucléoprotéique appelé splicéosome (du terme anglais « splicing » signifiant épissage). Le splicéosome est constitué de plusieurs complexes, les snRNP (small nuclear ribonucleoprotein) contenant un ARN (appelé snRNA) et plusieurs protéines. Le splicéosome est capable de reconnaître des séquences spécifiques au niveau des jonctions intron/exon et à l'intérieur de l'intron. La jonction exon/intron en 5' de l'intron est appelée site donneur ou site 5', la jonction intron/exon en 3' de l'intron est appelée site accepteur ou site 3' et la séquence interne de l'intron est appelée site de branchement.

Dans la majorité des cas, l'intron est dit de type U2, où la séquence AG-GU constitue le site donneur, la séquence AG-G précédée d'une région poly-pyrimidique le site accepteur (Breathnach *et al.,* 1978) et le site de branchement contient une adénine centrale au sein de la séquence CUA/G**A**CU qui joue un rôle important dans le mécanisme d'épissage. Il existe cependant un autre type d'introns appelé type U12 ayant une faible fréquence (Figure 13) et parmi les génomes étudiés de différentes espèces, le génome humain en contient le plus grand nombre, environ 700 (Sheth *et al.,* 2006 ; Alioto, 2007). Le site accepteur 3' de l'intron de type U12 a classiquement pour séquence YAC ou YAG (où Y est une pyrimidine). Cependant, les introns de type U12 ne requièrent pas spécifiquement une séquence précise et de nombreuses autres séquences dinucléotidiques peuvent être utilisées comme site accepteur (Dietrich *et al.,* 2001 ; Levine et Durbin, 2001 ; Hastings *et al.,* 2005). La distance moyenne entre le site accepteur 3' et le site de branchement (soit. 12 à 15 nucléotides) est plus faible

que celle observée dans les introns de type U2 (soit 18 à 40 nucléotides). En outre, le site accepteur 3' ne possède pas de région poly-pyrimidique en amont de sa séquence contrairement aux introns de type U2 (Hall et Padgett, 1994 ; Dietrich *et al.,* 2001 ; Levine et Durbin, 2001). Le site donneur en 5' de l'intron de type U12 a pour séquence A/GUAUCCUUU et le site de branchement est de type UCCUUAACU, ce qui diffère de l'intron de type U2 (Figure 13).

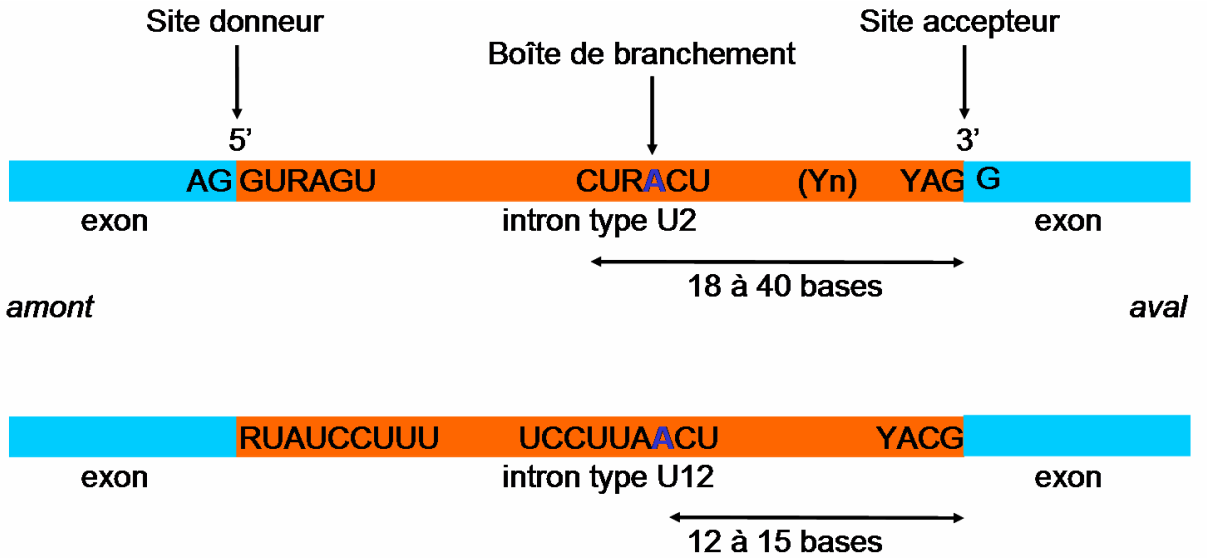


Figure 13 : Comparaison de l'architecture des jonctions exons/introns dans le cas des introns de type U2 et de type U12, d'après Will et Lührmann, 2005. L'adénosine centrale de la boîte de branchement est indiquée en bleu. Y : pyrimidine ; R : purine ; Yn : région poly-pyrimidique.

Bien que les introns de type U12 ne soient présents que dans moins de 1% des gènes humains (Levine et Durbin, 2001), ils sont retrouvés dans des gènes essentiels pour le fonctionnement de la cellule comme la réplication et la réparation de l'ADN, la transcription, la traduction de l'ARN. L'épissage des introns de type U2 et de type U12 est réalisé par des splicéosomes différents : le splicéosome de type U2 et le splicéosome de type U12, respectivement (Patel et Steitz, 2003 ; Will et Lührmann, 2005). Les splicéosomes de type U2 et de type U12 co-existent au sein des cellules eucaryotes et diffèrent de part leur composition en snRNP, entres autres U2 et U12, d'où leurs noms. Le mécanisme d'épissage utilisé par ces splicéosomes serait par contre très similaire (Shukla et Padgett, 2001).

Le splicéosome de type U2, plus fréquent que celui de type U12, est retrouvé chez tous les eucaryotes (Moore *et al.,* 1993 ; Burge *et al.,* 1999). Il est constitué de cinq snRNP

(U1, U2, U4, U5 et U6). Chaque snRNP (sauf U6) est constitué d'une molécule de snRNA, de sept protéines Sm communes (appelées ainsi parce qu'elles réagissent avec des anticorps auto-immuns de sérotype Sm de patients atteints de lupus érythémateux systémique) et d'autres protéines spécifiques. Les protéines Sm, nommées B/B', D1, D2, D3, E, F et G, s'assemblent pour former un anneau heptamérique entourant le snRNA (Will et Lührmann, 1997). Dans le cas du snRNP U6, l'anneau heptamérique est formé par des protéines Sm-like nommées LSm2, LSm3, LSm4, LSm5, LSm6, LSm7 et LSm8 (Achsel *et al.,* 1999 ; Vidal *et al.,* 1999). Au contraire des protéines Sm, en absence du snRNP U6 les protéines LSm peuvent être isolées sous la forme d'un complexe hétéromérique (Achsel *et al.,* 1999). En ce qui concerne les protéines spécifiques, elles sont en nombre variable selon le snRNP considéré. Par exemple, le complexe U1 contient trois protéines spécifiques appelées U1-70K, U1A et U1C (Will et Lührmann, 1997 ; Figure 14).

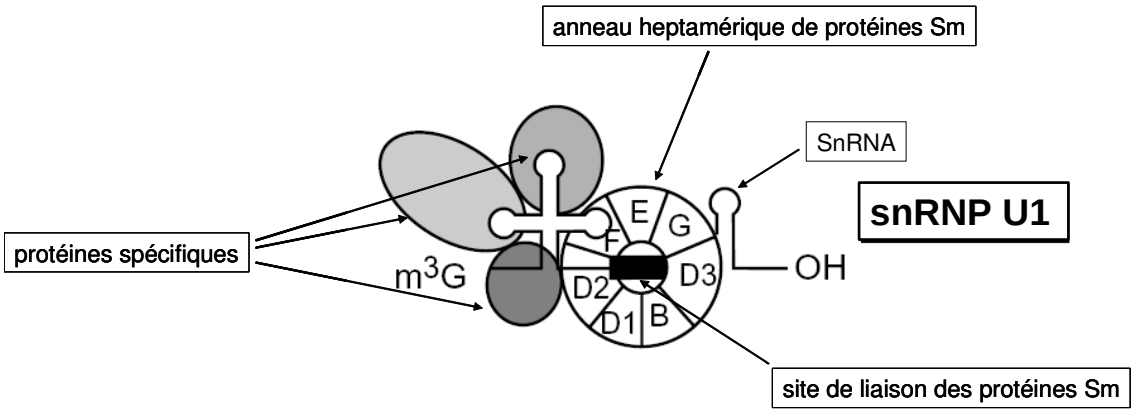


Figure 14 : Schéma représentant la structure du SnRNP U1, un des complexes protéiques constituant le splicéosome, d'après Will *et al.* (2001). m^3G, coiffe méthylée ; SnRNP, small nuclear ribonucleoprotein.

Le splicéosome type U12 assure l'épissage des introns de type U12 (Jackson, 1991 ; Hall et Padgett, 1994). Ce type de splicéosome n'est pas retrouvé chez les eucaryotes inférieurs tels que les nématodes et les levures et est en faible fréquence chez les autres eucaryotes (Burge *et al.,* 1998). Les splicéosomes de type U12 sont constitués, comme ceux de type U2, de snRNA, de protéines Sm communes et de protéines spécifiques. Seul le snRNP U5 est commun aux splicéosomes de type U12 et de type U2. En effet, les autres snRNP (U1, U2, U4 et U6) sont remplacés par des analogues fonctionnels appelés U11, U12, U4atac et U6atac (Hall et Padgett, 1996 ; Tarn et Steitz, 1996a ; Kolossova et Padgett, 1997 ; Yu et Steitz, 1997 ; Incorvaia et Padgett, 1998). Les snRNP U11 et U12 forment un complexe contenant huit protéines spécifiques ainsi que des protéines Sm (identiques à celle des SnRNP

du splicéosome de type U2) mais également sept sous-unités de facteurs d'épissage 3b (SF3b155, SF3b145, SF3b130, SF3b49, SF3b14a/p14, SF3b14b et SF3b10) (Will *et al.*, 1999 ; 2001 ; 2004).

2.2. Le processus d'épissage

D'une manière générale, l'épissage se déroule en deux temps. Tout d'abord, la liaison ester entre le phosphore en position 5' de l'intron et l'oxygène en position 3'de l'exon situé en amont est échangé contre une liaison ester avec l'oxygène en position 2' de l'adénosine du site de branchement. Ensuite, l'exon en amont de l'intron a de ce fait son extrémité 3'-OH libre qui va pouvoir réagir avec le groupement phosphate de la jonction intron/exon en aval et ainsi échanger la liaison ester entre le phosphore en position 5' de l'exon aval et l'oxygène en position 3' de l'intron contre une liaison ester avec l'oxygène en position 3' de l'exon en amont. Ainsi, l'intron est libéré sous forme cyclisée par l'intermédiaire de l'adénosine du site de branchement (appelée couramment lasso) et les exons sont reliés l'un à l'autre (Figure 15). Afin d'être recyclé, le lasso est par la suite ouvert par des enzymes de débranchement.

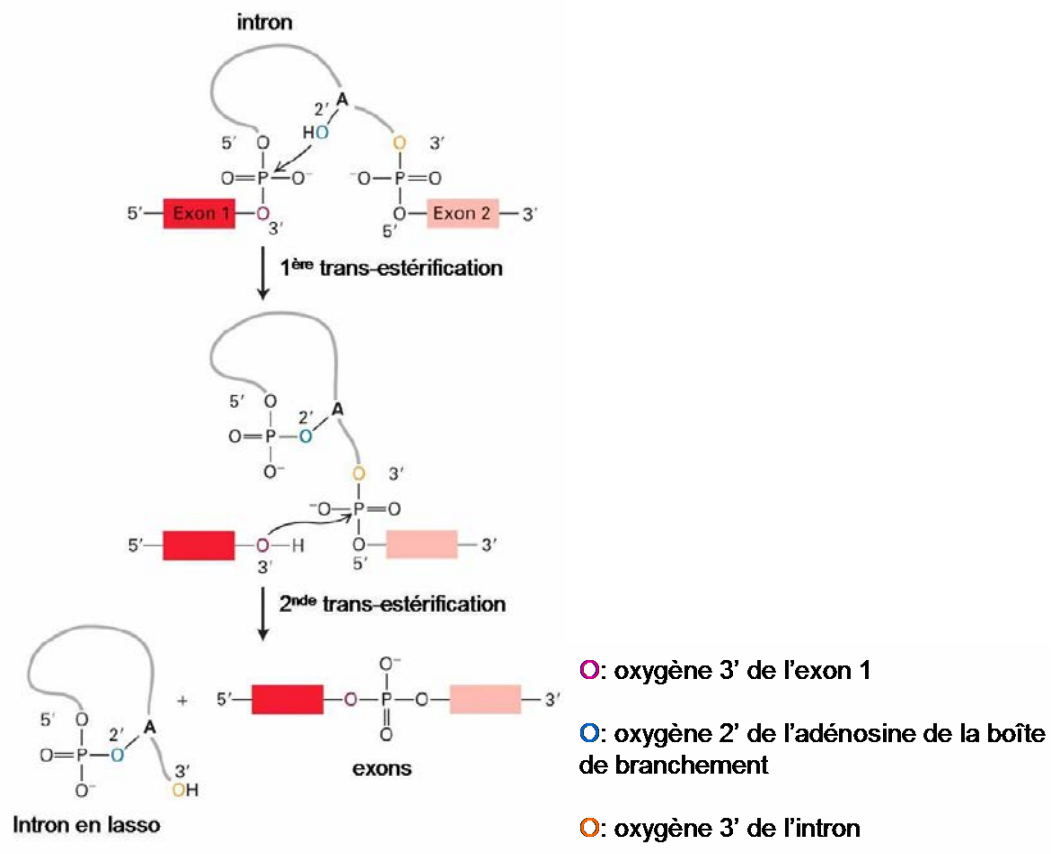


Figure 15 : Mécanisme en 2 étapes de l'épissage d'un pré-ARNm.

L'épissage est un processus dynamique suivant une succession précise d'événements où les différents snRNP vont s'associer et se dissocier du splicéosome.

L'assemblage du splicéosome de type U2 (Figure 16) est initié par l'interaction du snRNP U1 avec le site donneur 5' formant ainsi le complexe E. Cette interaction implique des associations de nucléotides entre le snRNA U1 et le site donneur 5' qui sont facilitées par les protéines U1-70K et U1-C (Reed, 1996 ; Will et Lührmann, 1997). Le snRNP U1 lié au site donneur facilite l'association du facteur d'épissage U2AF avec la région poly-pyrimidique (Brow, 2002). Ce facteur U2AF est nécessaire au recrutement ultérieur du snRNP U2. Dans l'étape suivante, le snRNP U2 interagit de manière stable avec le site de branchement pour former le complexe A ou pré-splicéosome. La stabilité de l'interaction entre le snRNP U2 et le pré-ARNm est due à la présence de protéines, notamment de facteurs d'épissage (U2AF, SF3a, SF3b…) au niveau du site de branchement ou de la région poly-pyrimidique (Reed, 1996 ; Will et Lührmann, 1997). Dans ce complexe le snRNA U2 s'apparie avec le site de branchement et les protéines SF3a et SF3b jouent un rôle critique dans l'attachement du snRNP U2 au pré-ARNm en s'appariant avec une région de 20 nucléotides juste en amont du site de branchement (Champion-Arnaud et Reed, 1994 ; Gozani *et al.,* 1996). Ensuite, le complexe snRNP U4/U6 s'associe au snRNP U5 et ces trois snRNP vont s'assembler avec les snRNP U1 et U2 pour former le complexe B (Moore *et al.,* 1993). Des réarrangements structuraux majeurs vont mener à la formation du complexe B* (Figure 16). L'interaction entre les snRNP U4 et U6 est rompue et le snRNP U6 va s'apparier avec le site donneur déplaçant ainsi le snRNA U1. Le snRNP U6 va également s'apparier avec le snRNA U2 formant plusieurs complexes snRNA-snRNA appelés hélices I, II, III. Une région du snRNP U6 va former une boucle interne qui est impliquée dans la réaction de catalyse de l'épissage. Le snRNA U5, quant à lui, va aller interagir avec des nucléotides de l'exon du site donneur mais aussi avec des nucléotides de l'exon proche du site accepteur (Madhani et Guthrie, 1994 ; Nilsen, 1994). Tous ces réarrangements provoquent la déstabilisation et la dissociation des snRNP U1 et U4 du splicéosome. Chez les mammifères, les protéines SR (pour riche en dipeptides sérine/arginine) sont aussi impliquées dans l'étape d'assemblage du splicéosome car elles faciliteraient l'incorporation du complexe U4/U5/U6 dans les splicéosomes (Roscigno et Garcia-Blanco, 1995) et stimuleraient l'appariement entre les bases nucléotidiques de snRNA U2 et de snRNA U6 (Tarn et Steitz, 1995). Le complexe B* contenant les snRNP U2/U5/U6 constitue l'entité catalytique du splicéosome qui va réaliser les deux étapes permettant l'excission des introns. Les snRNA U2 et U6 interagissent avec le

site accepteur et le site de branchement leur permettant de se rapprocher et ainsi de réaliser la première réaction de transestérification conduisant à la formation du complexe C (Madhani et Guthrie, 1994 ; Yu *et al.,* 1998 ; Nilsen, 1994 ; 1998). Le snRNA U5 est impliqué dans le positionnement des exons pour la deuxième réaction de transestérification conduisant à la ligation des exons et l'excision des introns (Newman, 1997). Finalement, suite à la deuxième réaction de transestérification, le splicéosome se dissocie, relâchant les exons ligaturés ainsi que les introns excisés et les snRNP, ces deux dernières catégories seront recyclées pour prendre part à un nouveau processus d'épissage. Le splicéosome possède une activité métallo-enzymatique, car il requiert la présence de cations divalents qui se fixent au niveau du snRNA U6 pour catalyser les réactions de transestérification (Sontheimer *et al .,* 1997 ; Gordon *et al.,* 2000 ; Yean *et al.,* 2000).

L'assemblage du splicéosome de type U12 est similaire à celui de type U2 à l'exception de la première étape (Figure 16). En effet, avant de s'associer au pré-ARNm, les snRNP U11 et U12 forment un complexe très stable qui se lie au site accepteur et au site de branchement lors de la première étape de l'assemblage du splicéosome conduisant directement à la formation du complexe A (Wassarman et Steitz, 1992 ; Frilander et Steitz, 1999). Dans l'étape suivante, le complexe snRNP U4atac/U6atac/U5 interagit avec le pré-ARNm pour former le complexe B (Tarn et Steitz, 1996a ; 1996b). Comme pour le splicéosome U2, l'activation du splicéosome U12 implique la déstabilisation des snRNP U11 et U4atac et ceux-ci ne sont donc pas impliqués directement dans les réactions de catalyse. En effet, les snRNA U4 et U4atac seraient des molécules chaperonnes séquestrant les snRNA U6 et U6atac, respectivement, via des appariements de base avant l'activation catalytique du splicéosome (Guthrie et Patterson, 1988 ; Tarn et Steitz, 1996a). Les étapes suivantes sont similaires à celles du splicéosome U2.

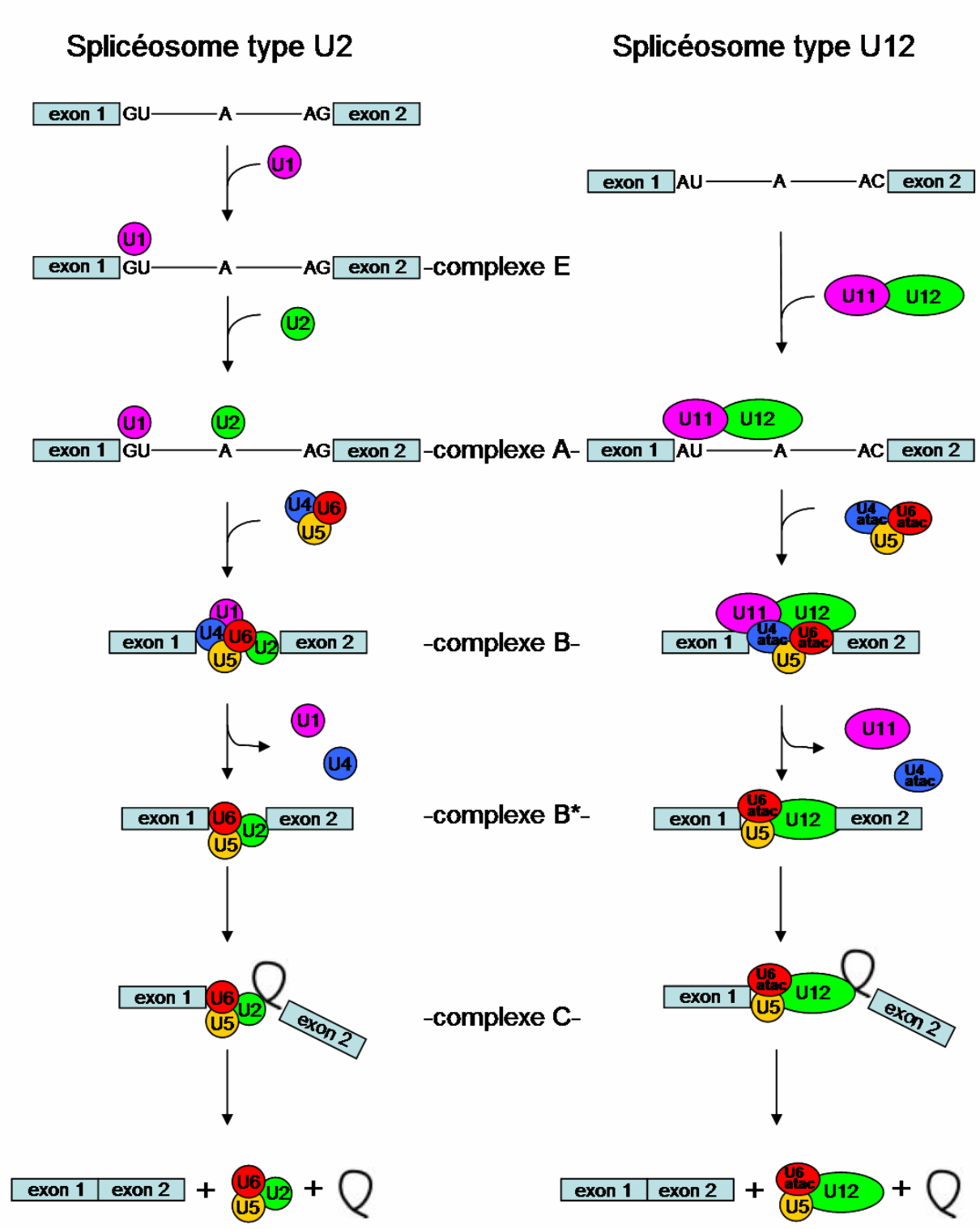


Figure 16 : Comparaison des différentes étapes permettant l'assemblage des splicéosomes de type U2 et U12 (d'après Will et Lührmann, 2005).

2.3. L'épissage alternatif

L'épissage alternatif d'un seul pré-ARNm permet la génération de plusieurs ARNm matures codant pour des isoformes protéiques ce qui enrichit considérablement la diversité du protéome des organismes eucaryotes (Maniatis et Tasic, 2002 ; Johnson *et al.,* 2003). Il existe deux types d'exons : les exons dits constitutifs qui seront toujours inclus dans l'ARNm et les exons dits alternatifs qui pourront se retrouver épissés avec les introns. Les exons constitutifs possèdent des sites d'épissage dits forts qui seront toujours reconnus par le splicéosome. Au contraire, les exons alternatifs possèdent des sites d'épissage dits faibles pour lesquels les séquences consensus de reconnaissance du site d'épissage sont moins bien conservées et donc moins bien reconnues par le splicéosome.

2.3.1. Facteurs influençant la reconnaissance d'un exon

Une des étapes importantes de l'épissage est la reconnaissance des sites d'épissage par le splicéosome. Le site donneur 5' est caractérisé par une seule séquence nucléotidique. Le site accepteur 3' est quant à lui caractérisé par trois séquences nucléotidiques qui se situent à une distance d'environ 40 nucléotides en amont de la jonction 3' intron/exon : le site de branchement, la région poly-pyrimidique et la jonction 3' intron/exon (Reed, 1996). La reconnaissance initiale des jonctions exon/intron est basée sur l'interaction de séquences spécifiques entre le snRNP U1 et le site donneur 5', le facteur d'épissage U2AF et la région poly-pyrimidique et le snRNP U2 avec le site de branchement. Les interactions mises en jeu étant de type pré-ARNm/snRNA, les sites d'épissage sont classés selon la complémentarité du site donneur 5' avec le snRNA U1 et la longueur de la région poly-pyrimidique. Ainsi, plus la région poly-pyrimidique sera longue et plus la complémentarité site donneur 5'/snRNA U1 sera forte, plus l'affinité des sites d'épissage pour les composants splicéosomaux sera grande et donc mieux l'exon sera retenu (Reed, 1996).

Il existe des séquences d'ARN *cis*-régulatrices présentes au sein de l'exon favorisant sa conservation dans l'ARNm mature. Ces séquences servent de sites de liaison pour l'assemblage de complexes activateurs d'épissage et sont appelées des ESE (*exonic splicing enhancer*). Les ESE sont généralement reconnus par au moins un des membres de la famille des protéines SR qui recrute le splicéosome au niveau de l'intron adjacent (Graveley, 2000 ; Black, 2003). Les protéines SR permettent le recrutement de facteurs d'épissage et de snRNP

par interaction entre protéines. Cependant, ces ESEs ne sont pas les seuls éléments *cis*-régulateurs à agir dans la régulation de l'épissage du pré-ARNm. En effet, d'autres facteurs activateurs ou répresseurs de l'épissage au niveau exonique (ESS : *exonic splicing silencer*) ou intronique (ISE : *intronic splicing enhancer* ; ISS : *intronic splicing silencer*) peuvent influencer la sélection du site d'épissage. Les ESS sont généralement reconnus par des membres de la famille des hnRNP (*heterogeneous nuclear* RNP) tandis que les ISS se lient le plus souvent à une protéine liée à la région poly-pyrimidique (Black, 2003). Plusieurs mécanismes ont été proposés pour expliquer la répression de l'épissage exercée par les ESS et les ISS. Les hnRNP liant les répresseurs peuvent réprimer l'assemblage du splicéosome en se multimérisant le long des exons et en empêchant ainsi l'accès au ESE (Zhu *et al.,* 2001), en bloquant le recrutement des snRNP (Tange *et al.,* 2001) ou en créant une boucle dans un intron pour rapprocher deux exons adjacents et provoquer l'inclusion d'un exon à l'intérieur de cette boucle (Figure 17 ; Martinez-Contreras *et al.,* 2006).

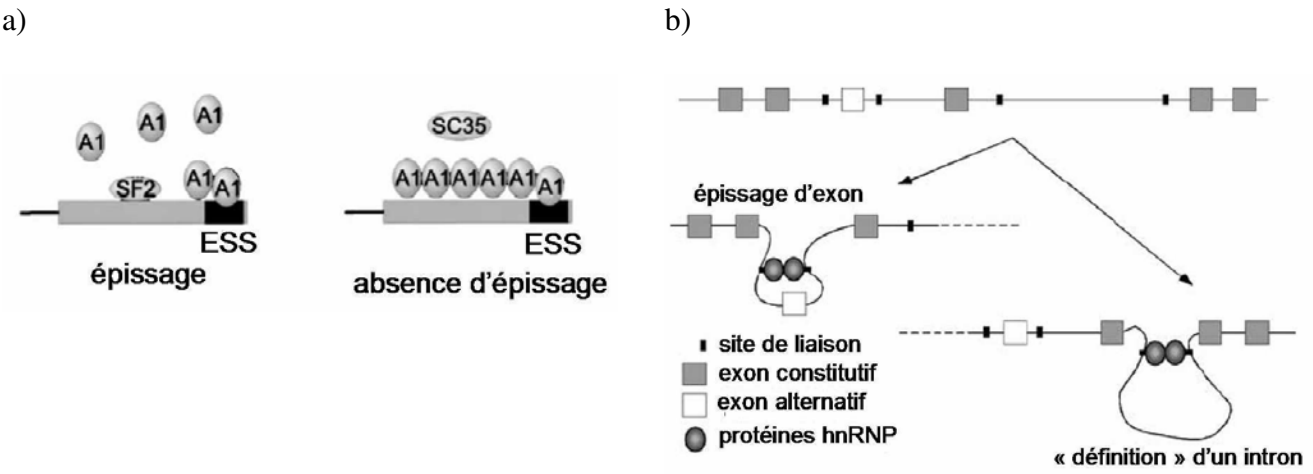


Figure 17 : Exemples de mécanismes réprimant la régulation de l'épissage exercé par les ESS et les ISS. a) A gauche, le hnRNP A1 est lié au répresseur ESS, cependant le facteur d'épissage SF2 est quant à lui lié à un motif activateur ESE ce qui bloque la liaison d'autres hnRNP A1 et leur polymérisation le long de l'exon vers l'extrémité 5' ; l'épissage de l'exon a lieu. A droite, le facteur d'épissage SC35 ne peut pas empêcher la polymérisation des hnRNP A1 le long de l'exon ; l'exon n'est pas épissé (d'après Zhu *et al.,* 2001). b) Les protéines hnRNP se lient aux sites de liaison présents sur l'intron provoquant une boucle au niveau de celui-ci, et « définit » l'intron. Une telle liaison des hnRNP peut se produire sur des sites appartenant à deux introns différents et inclure un exon dans la boucle formée provoquant son épissage avec les introns (d'après Martinez-Contreraz *et al.,* 2006).

Classiquement, les activateurs et les répresseurs sont présents à proximité des jonctions exon/intron, ce qui suggère que l'équilibre entre activation et répression par des éléments *cis*-régulateurs module la probabilité d'inclusion ou non d'un exon dans l'ARNm mature. Par exemple, des études menées sur l'épissage du pré-ARNm du SMN (*survival of motor neuron*) ont montré l'existence d'un certain nombre d'éléments activateurs et répresseurs sur l'exon 7 et ses introns adjacents. L'épissage alternatif de l'exon 7 est mis en cause dans le cas d'amyotrophie spinale (Lorson et Androphy, 2000 ; Cartegni et Krainer, 2002 ; Kashima et Manley, 2003 ; Singh *et al.,* 2004). Ces observations suggèrent que la reconnaissance de chaque exon est influencée par de nombreux éléments *cis*-régulateurs.

L'architecture exon/intron influence également la reconnaissance des sites d'épissage (Talerico et Berget, 1994 ; Berget, 1995 ; Sterner *et al.,* 1996). Par exemple, l'augmentation de la taille des exons chez les mammifères entraîne l'épissage alternatif d'exons. Néanmoins, ces mêmes exons seront conservés si les introns adjacents en 5' et en 3' sont petits (Sterner *et al.,* 1996). Par conséquent, la reconnaissance du site d'épissage est plus efficace quand les introns et les exons sont de petite taille. Les introns sont généralement courts chez les eucaryotes inférieurs, mais chez les mammifères ils peuvent être de très grande taille par rapport à celle relativement réduite des exons. Devant l'hétérogénéité de taille des introns et des exons, il a été supposé que le splicéosome puisse reconnaître les sites d'épissage selon deux mécanismes appelés la définition de l'exon et la définition de l'intron (Berget, 1995). Dans le cas de longs introns (> 20 kb ; Fox-Walsh *et al.,* 2005), le mécanisme de reconnaissance des sites d'épissage est celle de la définition de l'exon où ce sont les extrémités des exons qui sont reconnues plutôt que celles des introns. Ainsi, dans ce cas, la machinerie d'épissage va rechercher une paire de sites d'épissage 5' et 3' rapprochée et qui va appartenir à un même exon. Quand cette paire est trouvée, l'exon est défini par la liaison du snRNP U1, du snRNP U2 et des facteurs associés (Green, 1991 ; Eperon *et al.,* 1993 ; Fu et Maniatis, 1992 ; Kohtz *et al.,* 1994 ; Zuo et Manley, 1994). Ensuite, les exons voisins se rapprochent les uns des autres, probablement par des interactions entre des facteurs reconnaissant les exons (Figure 18), et l'épissage se déroule comme expliqué précédemment. Néanmoins, la reconnaissance des sites d'épissage sur les introns est significativement plus efficace que celle des sites d'épissage sur les exons et favorise la rétention au sein de l'ARNm d'exons contenant des sites d'épissage faibles (exons alternatifs). Ainsi, la taille des introns influence grandement la probabilité qu'un exon soit épissé alternativement. Dans le génome humain, l'épissage d'exons a une plus forte chance de se produire quand les introns adjacents

sont de grandes tailles (87% des introns sont supérieurs à 25 kb). Par ailleurs, des analyses expérimentales et informatiques ont montré que la longueur de l'intron en amont de l'exon est plus importante que celle de l'intron en aval dans l'induction d'un épissage alternatif. Donc, l'architecture exon/intron est un autre facteur important dans les mécanismes de reconnaissance des sites d'épissage et influence la fréquence de l'épissage alternatif d'un pré-ARNm (Fox-Walsh, 2005).

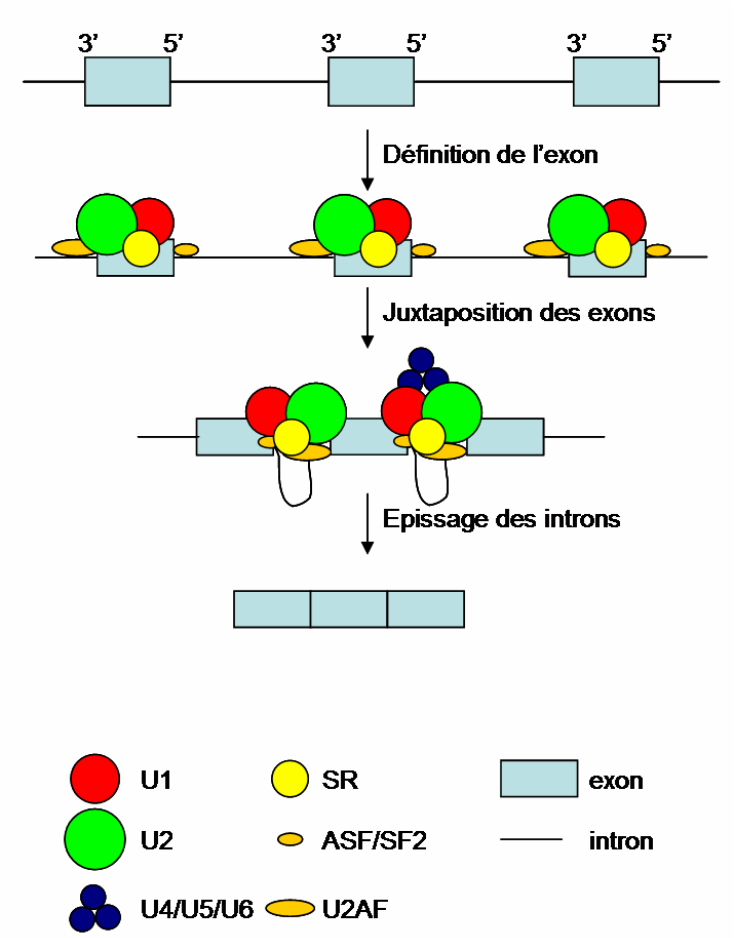


Figure 18 : « Définition » des exons dans le cas de pré-ARNm contenant des exons de petites tailles et des introns de type U2 de grandes tailles. Les protéines SnRNP U1 et U2 et les protéines SR interagissent avec les exons isolés. Le complexe U4/U5/U6 permet la liaison entre deux exons juxtaposés ce qui permet ainsi l'épissage des introns (d'après Berget, 1995).

L'ARN simple brin peut parfois adopter des structures secondaires locales et des interactions tertiaires qui peuvent impliquer jusqu'à plus de 100 nucléotides. Bien que l'ARNm soit généralement linéaire, il existe cependant des portions de l'ARNm qui adoptent une conformation en double brin et forment des structures secondaires. Selon les cas, ces structures peuvent être stable thermodynamiquement et interférer avec la reconnaissance des

sites d'épissage. Lors du processus de reconnaissance des sites d'épissage des facteurs d'épissage interagissent avec des activateurs ou des répresseurs au niveau du pré-ARNm simple brin. La présence de structures secondaires locales au niveau de l'ARNm peut interférer avec l'assemblage du splicéosome en occultant soit les sites d'épissage, soit les sites de liaison des activateurs. De même, ces structures secondaires peuvent au contraire favoriser l'assemblage du splicéosome si elles masquent les sites de liaisons de répresseurs. Ainsi, la présence de structures secondaires locales joue un rôle significatif dans la modulation de la reconnaissance des sites d'épissage (Hertel, 2008).

Des études récentes ont montré que le processus de reconnaissance des sites d'épissage peut être concomitant à la transcription. Cela signifie que les sites d'épissage d'un exon peuvent être reconnus par le splicéosome alors que les exons en aval ne sont pas encore synthétisés par l'ARN polymérase II (Görnemann *et al.,* 2005 ; Kornblihtt, 2005 ; Tardiff *et al.,* 2006). Ainsi, l'épissage des introns est lié de manière physique et temporelle à la transcription de l'ARN. Une étude a montré que l'épissage alternatif d'un gène rapporteur peut varier selon la structure du promoteur de l'ARN polymérase II ayant donné naissance au transcrit. Cela suggère un modèle dans lequel les facteurs d'épissage s'associent avec l'ARN polymérase II proche du promoteur (Cramer *et al.,* 1997). Par conséquent, des différences dans la structure du promoteur pourraient mener à des différences dans les facteurs d'épissage recrutés au niveau de la machinerie transcriptionnelle (de la Mata et Kornblihtt, 2006). Un modèle complémentaire est basé sur l'hypothèse que la cinétique de transcription par l'ARN polymérase II pourrait influencer l'épissage alternatif. L'ARN polymérase II fonctionne dans le sens 5'-3', les exons alternatifs peuvent alors être transcrits avant ou après la synthèse d'exons voisins compétiteurs. La vitesse relative de production de ces exons compétiteurs peut produire des changements dans le motif d'épissage (de la Mata *et al.,* 2003 ; Hicks *et al.,* 2006).

La sélection d'un exon est donc influencée par une combinaison de facteurs. Tous les exons comportent leurs propres éléments permettant leur reconnaissance par le splicéosome. En effet, chaque exon est bordé par une paire de sites d'épissage unique et contient un groupe de répresseurs et d'activateurs et des structures secondaires qui lui sont propres. Une échelle sur laquelle la somme des contributions de chacun des éléments est rapportée (variation des sites d'épissage, architecture exon/intron, nombre de répresseurs et d'activateurs, structures secondaires) permet de définir le potentiel global de reconnaissance d'un exon ou l'affinité de

liaison du splicéosome. Les extrémités de cette échelle sont représentées d'une part par les exons qui sont constitutivement épissés et d'autre part par ceux qui sont toujours inclus dans l'ARNm mature. Ces deux types d'exons seront insensibles aux changements possibles dans leur environnement lors de l'épissage. Le centre de l'échelle représente les exons qui ne sont pas toujours inclus dans l'ARNm mature, c'est-à-dire les exons alternatifs, qui eux par contre seront sensibles même à des changements mineurs dans l'efficacité de l'épissage (Hertel, 2008).

2.3.2. Les sites d'épissage cryptiques

Les sites d'épissage se situent en règle générale au niveau des jonctions exon/intron. Néanmoins, des séquences consensus correspondant à des sites d'épissage peuvent être retrouvées à l'intérieur d'un exon ou d'un intron. Ces sites d'épissage, qui ne sont pas les sites classiques, sont appelés des sites d'épissage cryptiques.

Une mutation peut modifier un site d'épissage classique au niveau des jonctions exon/intron et celui-ci n'est alors plus reconnu par la machinerie d'épissage. Dans ce contexte, le splicéosome peut reconnaître le site d'épissage de l'exon suivant et l'exon muté sera délété ou alors le splicéosome peut reconnaître et activer un site d'épissage cryptique qui n'est généralement pas très éloigné du site classique muté (Krawczak *et al.,* 2007). L'activation de ce site cryptique entraîne la production d'un ARNm mature modifié qui peut contenir un morceau d'intron, un exon avec une partie manquante ce qui, dans la plupart des cas, modifie le cadre de lecture et peut provoquer l'apparition rapide d'un codon stop.

Cependant, une étude récente a montré que les sites d'épissage cryptiques peuvent aussi être utilisés dans des conditions normales, c'est-à-dire sans mutation du site d'épissage classique, contrairement à ce qui était pensé jusqu'alors (Haj Khelil *et al.,* 2008). L'utilisation par le splicéosome d'un site d'épissage classique ou d'un site d'épissage cryptique n'est pas due au hasard, mais due à des modulations intrinsèques (éléments *cis*-régulateurs), et cellulaires (facteurs *trans*-régulateurs, facteurs cellulaires) au sein de la cellule (Haj Khelil *et al.,* 2008).

2.3.3. Les différents types d'épissage alternatif pouvant être rencontrés

Au regard des différents facteurs pouvant entraîner un épissage alternatif, plusieurs types d'ARNm matures peuvent être engendrés (Figure 19). Ainsi, lorsqu'un exon alternatif

se situe entre deux exons constitutifs et est soit inclu, soit exclu de l'ARNm, il s'agit du type cassette. Dans le cas où deux exons alternatifs se situent entre deux exons constitutifs, la machinerie d'épissage détermine lequel des deux exons sera épissé, c'est le type multi-cassette ou mutuellement exclusif. Il peut également exister comme évoqué précédemment des sites cryptiques provoquant la délétion d'une partie d'un exon ou la conservation d'une partie d'un intron. Enfin, il peut y avoir absence d'épissage, ce qui provoque l'inclusion d'un intron dans l'ARNm mature (Graveley, 2001).

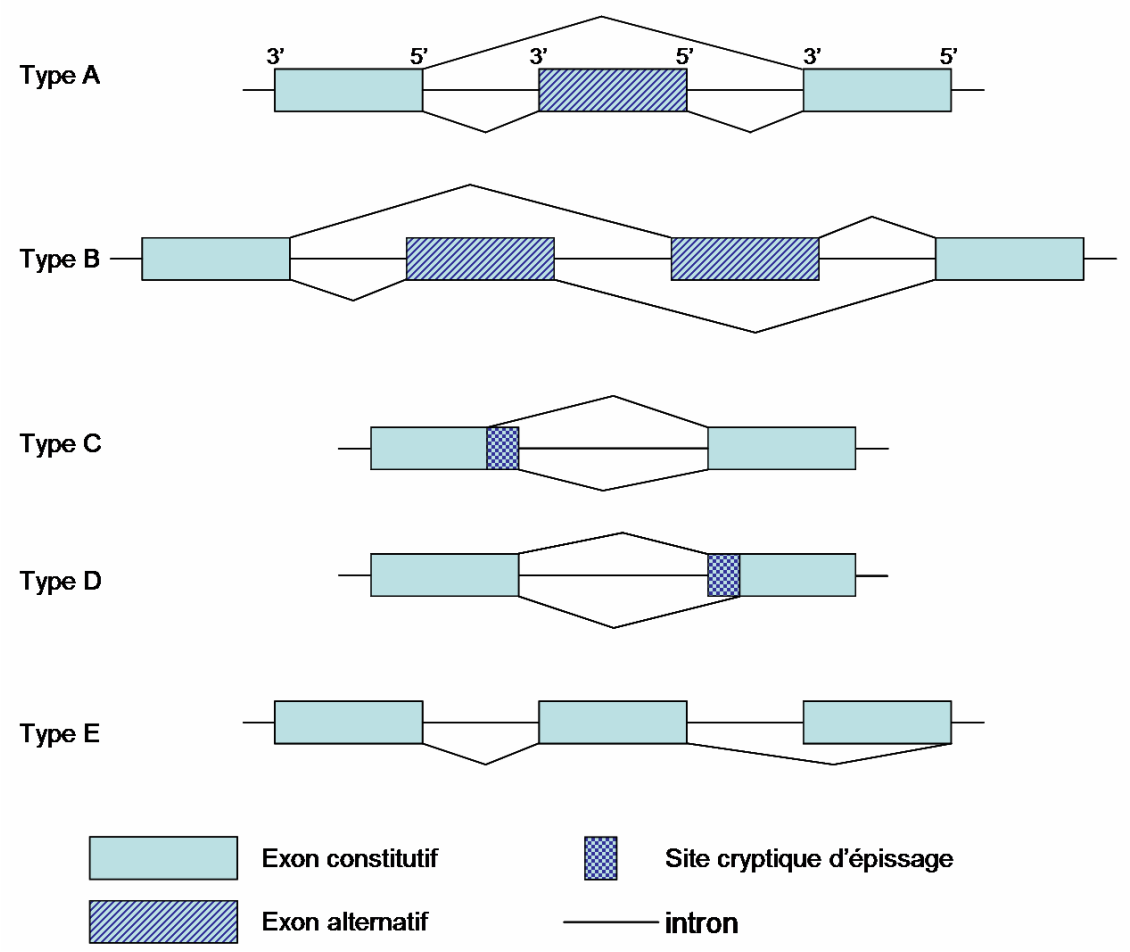


Figure 19 : Les différents types d'épissage alternatif pouvant être rencontrés (d'après Graveley, 2001). Légende :
Type A : épissage de type « cassette »
Type B : épissage de type « multi-cassette » ou mutuellement exclusif
Type C et D : épissage de site crytique (5' ou 3')
Type E : épissage avec intron retenu

2.4. L'épissage des caséines

2.4.1. Structure des gènes des caséines bovines

Les caséines β, α_{s1}, α_{s2} et κ sont chacune généralement codées par un seul gène (Figure 20) et plusieurs variants génétiques peuvent exister.

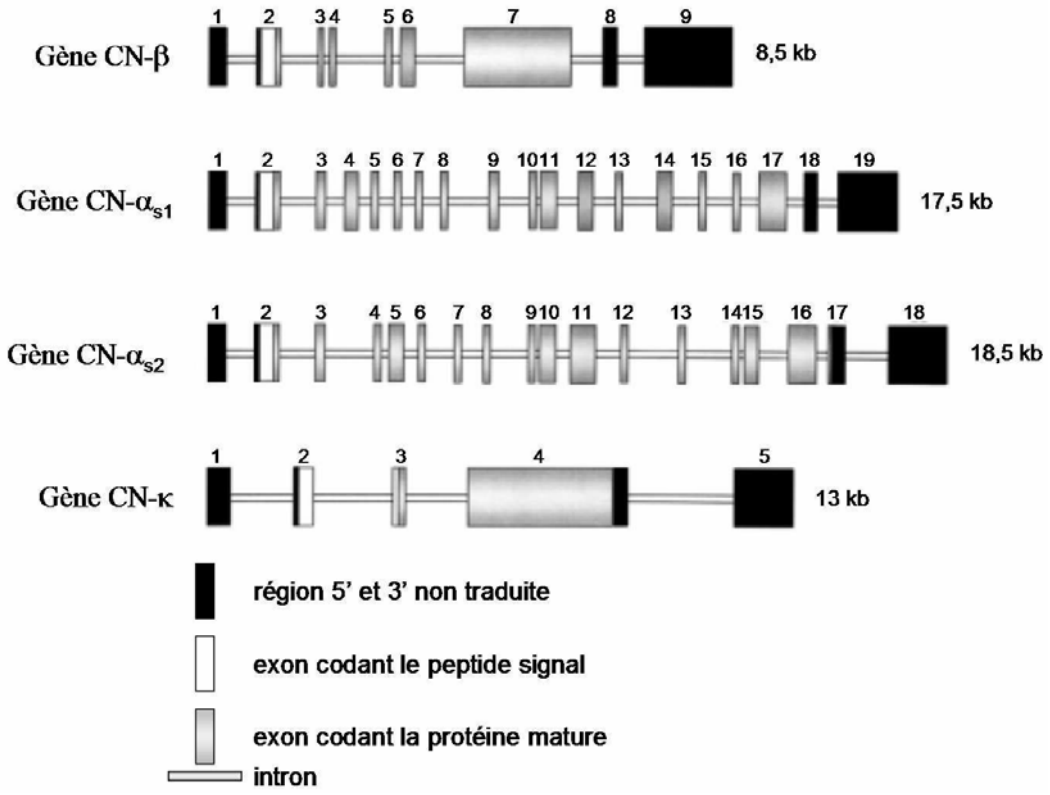


Figure 20 : Organisation structurale des gènes des quatre caséines bovines (d'après Martin *et al.,* 2002). La taille de chaque gène est indiquée en kilobases. Chaque exon est numéroté.

Les gènes des caséines bovines se situent sur un même locus sur le chromosome 6 (Figure 21 ; Ferretti *et al.,* 1990 ; Threadgill et Womack, 1990 ; Gallagher *et al.,* 1994).

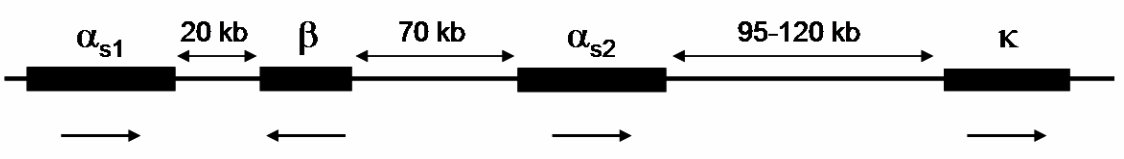


Figure 21 : Organisation génomique du locus des gènes de caséines (CN) bovines α_{s1}, β, α_{s2} et κ localisé sur le chromosome 6. La flèche simple représente le sens de transcription de chaque gène. La flèche double indique la distance intergénique en kilobases (d'après Rinjkels *et al.,* 1997).

L'organisation générale de ces gènes est conservée entre les espèces (Mercier et Vilotte, 1993). Les gènes des caséines équines se situent sur le chromosome 3 tandis que ceux des caséines humaines sont sur le chromosome 4. Les gènes sont organisés de manière ordonnée sur le locus, l'ordre étant le suivant : α_{s1}-β-α_{s2}-κ (Ferretti *et al.,* 1990 ; Threadgill et Womack, 1990). Les trois gènes codant pour les caséines β, α_{s1} et α_{s2} dérivent d'un gène ancestral commun et auraient évolué par duplication intra- et intergénique et par brassage d'exons (Groenen *et al.,* 1993). Ces gènes partagent également les mêmes motifs régulateurs situés en position 5' (Groenen *et al.,* 1992). Le gène de la CN-κ n'a pas d'évolution commune avec les trois autres gènes même si son expression et la protéine qu'il code sont essentielles pour la formation et la stabilité de la micelle (Alexander *et al.,* 1988). La région 5' est différente de celle des gènes des autres caséines. Ce gène pourrait néanmoins dériver d'un gène ancestral ayant donné naissance également à ceux de la famille des fibrinogènes (Jollès *et al.,* 1986 ; Alexander *et al.,* 1988) Les quatre gènes des caséines s'expriment néanmoins de manière coordonnée et l'expression est tissu-spécifique. La présence d'éléments *cis*-régulateurs dominants, permettant l'expression coordonnée à haut niveau des quatre gènes, est suspectée dans la région α_{s1}-β (Rijnkels *et al.,* 1997). L'induction des gènes des protéines du lait est sous un contrôle multihormonal complexe impliquant essentiellement la prolactine, les glucocorticoïdes, l'insuline, l'hormone de croissance et d'autres facteurs lactogéniques (Topper et Freeman, 1980 ; Vonderhaar et Ziska, 1989 ; Zwierzchowski, 1996 ; Rosen *et al.,* 1999).

Les gènes des caséines β et α_{s1} ont un sens de transcription opposé avec des extrémités 3' convergentes, le gène de la caséine α_{s1} étant en aval du gène de la caséine β. La distance entre ces deux gènes est d'environ 20 kb. Les gènes de la caséine α_{s2} et de la caséine κ sont transcrits dans le même sens que le gène de la caséine α_{s1}. La distance intergénique β-α_{s2} est de 70 kb et la distance intergénique α_{s2}-κ est de 95 kb. Ainsi, la taille du locus des caséines bovines est d'environ 250 kb du début du site de transcription du gène de la caséine α_{s1} jusqu'à la fin du gène de la caséine κ (Rijnkels *et al.,* 1997 ; Figure 21). Il existe d'une espèce de mammifère à l'autre des différences de distance entre les gènes des caséines et un nombre variable de gènes (notamment la présence d'un gène de la caséine α_{s2} supplémentaire chez la souris ; Rijnkels *et al.,* 1997).

2.4.2. Epissage alternatif des caséines

Les caséines, notamment les CN-β et CN-α_{s1}, peuvent subir les deux types d'épissage alternatif : épissage dû à la présence d'un site cryptique et épissage d'exon(s).

La délétion du résidu Gln^{78} de CN-α_{s1} fut d'abord mis en évidence chez la chèvre (Leroux *et al.,* 1992 ; Ferranti *et al.,* 1997). Cette délétion est cependant fréquente puisqu'elle est retrouvée chez la brebis (Ferranti *et al.,* 1995), la vache et la bufflonne d'eau (Ferranti *et al.,* 1999). Le résidu Gln^{78} correspond au premier codon de l'exon 11 qui est codé par le triplet CAG. Ce codon se situe donc au niveau de la jonction intron 10 / exon 11. La séquence de reconnaissance du site accepteur en position 3' de l'intron par le splicéosome est de type AG. Le splicéosome reconnaît en général la première séquence AG qui se situe juste en aval du site de branchement et de la région poly-pyrimidique. La présence du triplet CAG au début de l'exon fait apparaître une séquence AG qui peut être choisie alternativement par le splicéosome à la place de celle de l'intron (Smith *et al.,* 1993 ; Figure 22).

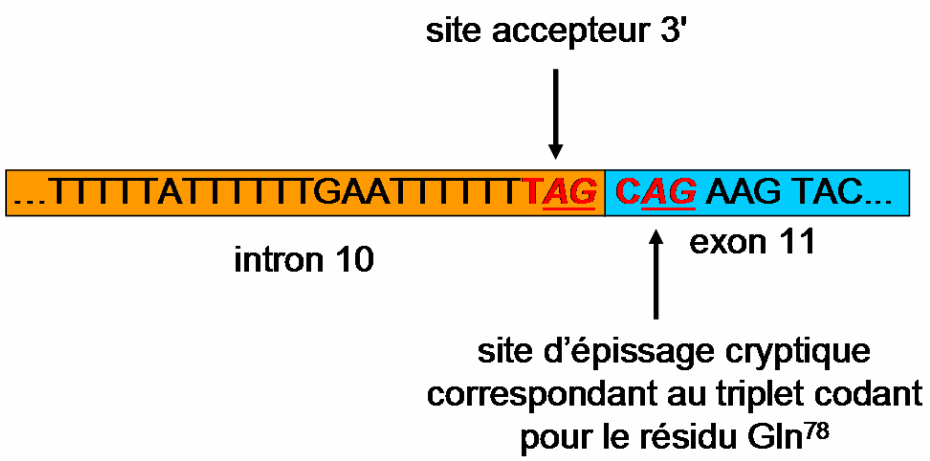


Figure 22 : Séquence nucléotidique à la jonction intron 10/exon 11 de la caséine α_{s1} chez les ruminants pouvant induire l'utilisation d'un site cryptique d'épissage par le splicéosome. Les doublets AG correspondant à un site d'épissage reconnu par le splicéosome sont indiqués en rouge et soulignés (d'après Martin *et al.,* 2002).

Ce phénomène ne semble pas être spécifique aux ruminants, car il a été également montré que le résidu Gln^{37} de la CN-α_{s1} humaine correspondant au premier codon de l'exon 6 peut lui aussi être délété (Johnsen *et al.,* 1995 ; Martin *et al.,* 1996). La taille relativement petite de l'intron et de la région poly-pyrimidique pourrait favoriser le choix du site cryptique pour l'épissage de cet intron entraînant ainsi la délétion du premier codon de l'exon aval. En

effet, l'intron 10 chez les ruminants a une longueur de 100 pb (Koczan *et al.,* 1991 ; Mao *et al.,* 1992) et l'intron 6 chez l'espèce humaine de 150 pb (Martin *et al.,* 1996). Un autre type d'épissage dû à l'utilisation d'un site cryptique a été observé pour la CN-β équine. En effet, l'utilisation par le splicéosome d'un site cryptique accepteur qui se situe dans la partie terminale de l'exon 7 implique la délétion de 75% de cet exon (Lenasi *et al.,* 2006 ; Miclo *et al.,* 2007).

Chez la chèvre plusieurs isoformes de CN-α_{s1} co-existent. Ainsi, il existe une forme contenant 199 résidus d'acides aminés correspondant à la traduction d'un ARNm complet, une forme délétée de la région 110-117 et une forme délétée de la région 141-148 (Ferranti *et al.,* 1997). Ces deux régions délétées correspondent respectivement à l'épissage alternatif des exons 13 et 16. Ainsi, la CN-α_{s1} caprine peut subir un épissage alternatif d'exons. Chez la brebis, l'exon 16 est aussi épissé alternativement (Passey *et al.,* 1996). Selon Passey *et al.* (1996), une substitution au niveau du site donneur en position 5' pourrait être responsable de cet épissage alternatif, qui touche environ 20% de l'ARNm produit. En effet, cette substitution provoquerait l'apparition d'une structure secondaire au niveau de l'ARNm qui serait inhibitrice, car le site donneur se retrouverait enfoui dans cette structure et deviendrait alors non accessible au splicéosome. La CN-α_{s1} humaine présente également une isoforme résultant de l'épissage alternatif de l'exon 7 (Johnsen *et al.,* 1995). Chez la CN-α_{s1} porcine, les exons 12 et 13' pourraient subir un épissage alternatif (Alexander *et al.,* 1992). Chez la jument, la CN-α_{s1} subirait des phénomènes d'épissage alternatif impliquant les exons 7 et 14 (Milenkovic *et al.,* 2002 ; Lenasi *et al.,* 2003 ; Miranda *et al.,* 2004). L'épissage de l'exon 14 pourrait être dû à l'apparition d'une structure secondaire au niveau de l'ARNm (Lenasi *et al.,* 2003). Chez la vache, le variant A de la CN-α_{s1} correspond à une isoforme dont la région [14-26] est délétée, or cette région correspond à celle codée par l'exon 4, ce variant est donc issu d'un épissage alternatif d'exon (Mohr *et al.,* 1994).

La CN-α_{s2} ovine présente une délétion de la région 34-42 correspondant à neuf résidus d'acides aminés, ce qui affecterait 30 à 40% des ARNm produits. Dans ce cas, un épissage alternatif partiel d'exon serait également possible (Boisnard *et al.,* 1991).

En ce qui concerne la CN-β, il n'y a pas de phénomène d'épissage alternatif mis en évidence sauf chez la jument. En effet, la CN-β équine subit un épissage alternatif de l'exon 5 (Miranda *et al.,* 2004 ; Lenasi *et al.,* 2006) et peut être aussi de l'exon 8 mais la protéine correspondante n'a pas encore été mise en évidence (Lenasi *et al.,* 2006). Il existe dans l'intron 1 du pré-ARNm de la CN-β équine un puissant élément *cis*-régulateur, appelé ISE1,

qui favorise l'inclusion des exons 5 et 8 dans l'ARNm. Cet élément *cis*-régulateur a la particularité d'avoir un effet global (agit sur tous les exons dits alternatifs), d'agir à distance et d'être promoteur-indépendant (Lenasi *et al.,* 2006).

3. Modifications post-traductionnelles : phosphorylation et glycosylation des caséines

3.1. Phosphorylation des caséines

La phosphorylation des caséines permet la liaison aux ions Ca^{2+} et la formation des micelles de caséines qui sont stables dans le lait (West, 1986). Les caséines contiennent plusieurs sites de phosphorylation qui, pour la plupart d'entre-eux sont regroupés en clusters caractéristiques du type -Ser-Ser-Ser-Glu-Glu-.

La phosphorylation des caséines est un phénomène post-traductionnel qui se déroule dans l'appareil de Golgi des lactocytes (Bingham *et al.,* 1972). Les caséines sont d'abord synthétisées au niveau des polyribosomes du réticulum endoplasmique des cellules épithéliales de la glande mammaire (Beitz *et al,* 1969 ; Heald et Saacke, 1972 ; Gaye *et al.,* 1973) puis le peptide signal permet l'adressage des caséines nouvellement synthétisées vers la face *cis* de l'appareil de Golgi. Dans l'appareil de Golgi, les caséines vont alors être phosphorylées puis seront secrétées dans des vésicules golgiennes où se formeront les micelles (Carroll *et al.,* 1970 ; Beeryl *et al.,* 1971 ; Heald et Saacke, 1972).

Le phosphore des caséines provient du phosphate inorganique sanguin qui est prélevé par la glande mammaire pour la synthèse des caséines ; l'incorporation du phosphate dans les caséines est rapide (environ 2 h ; Bingham et Farrell Jr., 1977).

Les premières études concernant la phosphorylation des caséines sont anciennes. Ainsi, dès 1954, Burnett et Kennedy ont montré qu'une protéine kinase (EC 2.7.1.37) mitochondriale de foie de rat pouvait phosphoryler des caséines selon la réaction :

$$\text{Protéine} + \text{ATP} \xrightarrow{Mg^{2+}} \text{Protéine-P} + \text{ADP}$$

Les auteurs ont montré que l'ATP provient du phosphate inorganique par phosphorylation oxydative dans la mitochondrie et que l'ion Mg^{2+} est indispensable à la réaction. Parmi tous les substrats qu'ils ont testés, seules les caséines sont phosphorylées par cette enzyme.

Ensuite, une caséine kinase a été mise en évidence (Bingham et Farrell Jr., 1974) et purifiée (Duncan *et al.,* 2000) au niveau de l'appareil de Golgi de la glande mammaire de rat (CKG : caséine kinase golgienne). Celle-ci possède un pH optimum de 7,6 et utilise l'ATP comme donneur de phosphate. Cette enzyme reconnaît la séquence -Ser/Thr-Xxx-Glu/SerP- où Xxx représente n'importe quel acide aminé et SerP un résidu Ser phosphorylé (Bingham, 1976). Elle peut aussi reconnaître la séquence -Ser-Xxx-Gln-Xxx-Xxx-Asp-Glu-Glu- (Brunati *et al.,* 2000). La réaction nécessite la présence de cations divalents tels que Mg^{2+} ou Ca^{2+} et n'est pas AMP cyclique dépendante. Une faible activité CKG a été mise en évidence dans le lait suggérant que cette kinase peut être sécrétée en même temps que les autres protéines du lait (Duncan *et al.,* 2000). L'activité détectée pourrait être due à la présence d'un fragment actif de l'enzyme résultant d'une protéolyse limitée de la CKG qui est transmembranaire. Il est possible que la CKG ait un rôle physiologique dans le lait, par exemple dans la phosphorylation post-sécrétoire de caséines non-phosphorylées qui auraient été sécrétées et stockées dans la lumière de l'alvéole mammaire (Duncan *et al.,* 2000).

D'autres kinases ont été mises en évidence aux niveaux cytosolique et nucléaire de toutes les cellules eucaryotes étudiées, dont la caséine kinase I (CK1) et la caséine kinase II (CK2 ; Allende et Allende, 1995). Ces enzymes, cependant au niveau physiologique ne sont pas en contact avec les caséines. La CK1 reconnaît la séquence -SerP-Xxx-Xxx-Ser- (Flotow *et al.,* 1990). La CK2 reconnaît la séquence -Ser/Thr-Xxx-Xxx-Glu/SerP-. La CK2 a un pH optimum de 8,0 (Chardot *et al.,* 1995). Cette enzyme, *in vitro*, a une affinité plus marquée pour la CN-β puis la CN-α_{s1} et enfin la CN-κ, que ces caséines soient sous forme isolée ou sous forme micellaire. Pour que l'enzyme fonctionne, il faudrait que la micelle se déstructure légèrement afin que la CK2 puisse y pénétrer pour phosphoryler la CN-β qui est généralement considérée comme étant moins accessible que les autres caséines (Pitois *et al.,* 1999).

Une activité caséine kinase a été détectée au sein du lait et dans différentes fractions micellaires lors d'une étude sur l'action de la CK2 sur la phosphorylation des caséines dans le lait. Cette activité ne serait due ni à CK1, ni à CK2 car la réaction est activée par l'ion Ca^{2+} et son substrat préférentiel est la CN-α_{s1} (Lovisi *et al.,* 2003).

D'après des travaux menés sur la CN-β humaine qui possède un degré de phosphorylation variable allant de 0 à 5 groupements phosphate, la phosphorylation semblerait se dérouler de manière séquentielle (Greenberg *et al.,* 1976). Le motif -Ser-Xxx-

Glu- serait phosphorylé en premier générant un résidu SerP qui entraînerait la phosphorylation des résidus suivants.

3.2. Glycosylation des caséines

Il existe deux types de glycosylation, la N-glycosylation (la plus fréquente) et la O-glycosylation. Chez toutes les espèces étudiées, parmi les caséines, seule la CN-κ est glycosylée. La CN-κ est une protéine O-glycosylée. Généralement, dans le cas d'une O-glycosylation, l'addition de glucides a lieu au niveau des chaînes latérales des résidus Ser et Thr et la réaction se déroule dans la lumière de l'appareil de Golgi. Le premier résidu glucidique ajouté est un résidu de N-acétyl-galactosamine (Gal NAc). Ce type de réaction est catalysé par des enzymes de type glycosyl-transférases. La O-glycosylation de la CN-κ se déroule bien dans l'appareil de Golgi. En effet, dans le cas de la CN-κ ovine, les enzymes impliquées dans la formation de l'unité disaccharidique Gal(β1-3)GalNAc sont localisées au niveau des membranes golgiennes de la glande mammaire de brebis (Soulier et Gaye, 1981).

La CN-κ bovine n'est glycosylée que sur des résidus Thr, les résidus Ser qui pourraient être potentiellement glycosylés ou phosphorylés sont préférentiellement phosphorylés. Cependant, le résidu Thr^{145} a été identifié comme étant phosphorylé dans une isoforme CN-κ 3P bovine, il y aurait alors une compétition possible entre la glycosylation et la phosphorylation de ce résidu (Holland *et al.,* 2006).

La chaîne glycannique majoritaire chez la CN-κ bovine est une structure tétrasaccharidique (Saito et Itoh, 1992 ; Holland *et al.,* 2006). A l'heure actuelle, le mécanisme de formation des formes mineures tri-, di-, et monosaccharidiques est encore méconnu. Elles pourraient provenir soit d'une synthèse incomplète du tétrasaccharide dans la glande mammaire, soit d'une dégradation du tétrasaccharide après synthèse, ou après sécrétion de la CN-κ dans la lumière alvéolaire de la glande mammaire (Holland *et al.,* 2006).

4. Les lactoprotéines, source de peptides à activités biologiques

Depuis plusieurs années, beaucoup de peptides physiologiquement actifs issus de protéines alimentaires ont été identifiés. Il faut bien faire la distinction entre les protéines biologiquement actives présentes naturellement dans l'aliment (facteurs de croissance,

immunoglobulines dans le lait) et les peptides provenant soit d'un processus de digestion de protéines ingérées intactes ou soit de l'utilisation d'hydrolysats protéiques entrant dans la formulation d'aliments (Kitts et Weiler, 2003). Un peptide bioactif est défini comme étant le produit issu de la dégradation d'une protéine par une ou plusieurs protéases, par exemple du tractus gastro-intestinal, et dont l'activité biologique n'apparaît qu'après sa génération à partir de la protéine-mère. Ces peptides bioactifs sont appelés des cryptéines (Autelitano *et al.,* 2006). De tels peptides peuvent être générés à partir de protéines alimentaires par trois voies : (1) pendant la digestion par les enzymes du tractus gastro-intestinal, (2) pendant le processus de fabrication de l'aliment comme la maturation du fromage, la fermentation ou encore la digestion par des enzymes microbiennes, (3) par protéolyse *in vitro* avec des enzymes choisies (Korhonen et Pihlanto, 2006). Les peptides bioactifs sont donc des fragments spécifiques de protéines qui ont un impact positif sur une fonction ou un état physiologique et par voie de conséquence, peuvent avoir une influence sur la santé (Tranter et Board, 1982). Beaucoup d'aliments sont sources de peptides bioactifs comme le lait, le poisson, les céréales ou encore l'œuf. Les propriétés souvent étudiées sont les activités anti-hypertensives, anti-microbiennes, anti-oxydantes, immunomodulatrices, anxiolytiques, opiacées (Kitts et Weiler, 2003). Des aliments enrichis en peptides bioactifs pourraient être utilisés dans la prévention de pathologies ou en complément d'un traitement médicamenteux. Des produits laitiers contenant des peptides bioactifs pour lesquels des effets bénéfiques pour la santé sont revendiqués sont commercialisés dans de nombreux pays (Tableau 3). Dans cette partie, nous allons nous focaliser sur les peptides bioactifs issus du lait (Figure 23).

Tableau 3 : Ingrédients et produits laitiers commercialisés ayant des effets revendiqués sur la santé basés sur des peptides bioactifs issus du lait (d'après Korhonen et Pihlanto, 2006).

Nom commercial	Type de produit	Peptides bioactifs	Effet revendiqué sur la santé	Fabricant
Calpis	Lait caillé	Val-Pro-Pro issus de CN-β Ile-Pro-Pro issus de CN-β et CN-κ	Diminution de la pression sanguine	Calpis Co., Japon
Evolus	Boisson au lait fermenté enrichie en calcium	Val-Pro-Pro issus de CN-β Ile-Pro-Pro issus de CN-β et CN-κ	Diminution de la pression sanguine	Valio Oy, Finlande
BioZate	Hydrolysat de lactosérum	Fragments de β-lactoglobuline	Diminution de la pression sanguine	Davisco, Etats-Unis
BioPURE-GMP	Lactosérum isolé	Glycomacropeptide	Prévention des caries dentaires, influence sur la coagulation du sang, protection contre les virus et les bactéries	Davisco, Etats-Unis
PRODIET F200/ Lactium™	Boisson lactée aromatisée, gélules	CN-α_{s1} [91-100]	Diminution des effets liés au stress	Ingredia, France
Festivo	Fromage fermenté allégé	CN-α_{s1} [1-9], CN-α_{s1} [1-7], CN-α_{s1} [1-6]	Aucun effet encore revendiqué	MTT Agrifood research, Finlande
Cysteine peptide	Hydrolysat	Peptide issu de protéines du lait	Aide à élever le niveau énergétique et aide au sommeil	DMV International, Pays-Bas
C12	Hydrolysat	Peptide issu de caséine	Diminution de la pression sanguine	DMV International, Pays-Bas
Capolac	Hydrolysat	Caséinophosphopeptide	Aide à l'absorption des minéraux	Arla Foods Ingrédients, Suède
PeptoPro	Hydrolysat	Peptide issu de caséine	Améliore les performances athlétiques et la récupération musculaire	DSM Food Specialties, Pays-Bas
Vivinal Alpha	Hydrolysat	Peptide issu de protéines du lactosérum	Aide à la relaxation et au sommeil	Borculo Domo Ingredients Pays-Bas

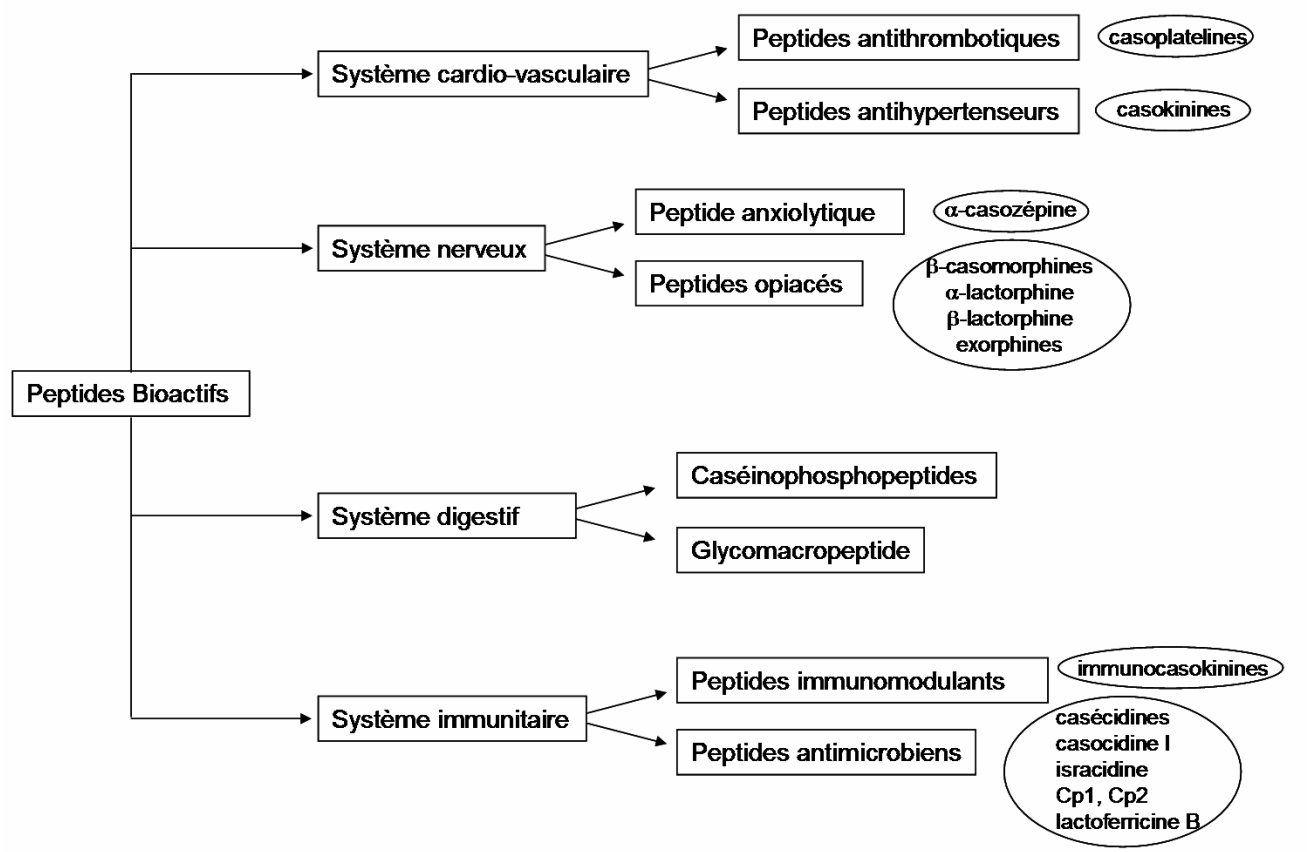


Figure 23 : Fonctions de certains peptides bioactifs issus du lait (adapté à partir de Silva et Malcata, 2005).

4.1. Peptides anti-microbiens

Les casécidines sont issues de l'hydrolyse par la chymosine des caséines du lait bovin après chauffage. Il s'agit de peptides polycationiques de faible masse moléculaire (Lahov et Regelson, 1996). Elles font partie des peptides anti-microbiens les plus purifiés actuellement et elles montrent une activité *in vitro* contre *Staphylococcus*, *Sarcina*, *Bacillus subtilis*, *Diplococcus pneumoniae* et *Streptococcus pyogenes* (Lahov et Regelson, 1996).

La casocidine-I est un peptide cationique issu de la CN- α_{s2} bovine (CN-α_{s2}-[165-203]). Ce peptide inhibe la croissance d'*Escherichia coli* et de *Staphylococcus carnosus* (Zucht *et al.,* 1995). Un autre peptide anti-bactérien (bactériostatique) issu de l'hydrolyse de la CN-α_{s2} bovine (CN-α_{s2}-[f183-207]), appelé Cp2, est actif contre *Escherichia coli* (Recio et Visser, 1999 ; McCann *et al.,* 2006).

L'isracidine est un peptide (CN-α_{s1}-[1-23]) issu de l'hydrolyse de la CN-α_{s1} bovine par la chymosine et la chymotrypsine. Ce peptide protège les souris contre *Staphylococcus aureus* après injection par voie sous-cutanée et contre *Candida albicans* après injection par

voie intraveineuse mais protège également les brebis et les vaches contre les mammites quand il est injecté dans la mamelle à des niveaux comparables à ceux utilisés lors d'un traitement antibiotique standard (Lahov et Regelson, 1996). Un autre peptide issu de l'hydrolyse de la CN-α_{s1} (CN-α_{s1}-[99-109]) a été identifié après hydrolyse de caséinate de sodium bovin par des enzymes du tractus gastro-intestinal (pepsine, trypsine, α-chymotrypsine ; McCann *et al.,* 2006). Ce peptide, appelé Cp1, a une action bactériostatique contre les bactéries $Gram^+$ et les bactéries $Gram^-$.

Un hydrolysat de lactoferrine bovine a une action bactéricide *in vitro* contre des bactéries $Gram^+$ (*Bacillus, Listeria* et *Streptococcus)* et des bactéries $Gram^-$ (*Escherichia coli*, *Klebsiella*, *Salmonella*, *Proteus* et *Pseudomonas*). Une bactérie intestinale pathogène, *E. coli* O111, y serait également sensible (Tomita *et al.,* 1991). Un peptide issu de l'hydrolyse de la lactoferrine bovine par la pepsine, la lactoferricine B, présente un puissant pouvoir bactéricide contre les bactéries $Gram^+$ et $Gram^-$ (Tomita *et al.,* 1991 ; Jones *et al.,* 1994). Ainsi, la lactoferricine B a une activité bactéricide contre *E coli* O157H:7, une bactérie entérohémorragique isolée chez des patients, à des concentrations significativement inférieures à celles de la lactoferrine ou de son hydrolysat (Shin *et al.,* 1998). Cette propriété est due à la charge nette positive de la lactoferricine B qui pourrait permettre l'augmentation de la perméabilité membranaire cellulaire *via* une attraction électrostatique entre le peptide chargé positivement et la membrane bactérienne chargée négativement. Une étude récente a montré que la lactoferricine B bovine provoque l'apparition de pores au niveau de la bicouche phospholipidique de la membrane bactérienne rendant ainsi la membrane perméable (Umeyama *et al.,* 2006) ce qui permet de tuer les microorganismes sensibles. Une production de lactoferricine a été mise en évidence dans l'estomac humain après ingestion de lactoferrine (Kuwata *et al.,* 1998).

4.2. Peptides anti-hypertenseurs

Les peptides inhibant l'enzyme de conversion de l'angiotensine I (ECA) pourraient présenter une activité anti-hypertensive (Yamamoto *et al.,* 1994 ; Nakamura *et al.,* 1995a, 1995b). L'ECA est une peptidyl-dipeptidase qui convertit l'angiotensine I en angiotensine II, un puissant vasoconstricteur, et qui inactive la bradykinine qui a une action vasodilatatrice. Tout ceci a pour conséquence une augmentation de la pression artérielle. Des peptides inhibiteurs de l'ECA issus des caséines, les casokinines, sont générés par hydrolyse de la CN-

α_{s1}, de la CN-β et de la CN-α_{s2} bovines (Maruyama et Suzuki, 1982 ; Maruyama *et al.,* 1987 ; Schlimme et Meisel, 1993 ; Meisel et Schlimme, 1994 ; Maeno *et al.,* 1996 ; Tauzin *et al.,* 2002). D'autres séquences ont été mises en évidence dans les structures primaires des CN-α_{s1} et CN-κ humaines et pourraient présenter une activité inhibitrice de l'ECA une fois les peptides correspondants libérés (Kohmura *et al.,* 1989 ; Kohmura *et al.,* 1990 ; Shah, 2000). L'hydrolyse des caséines humaines par la pepsine et la pancréatine a permis d'identifier un peptide de la CN-β, HLPLP ([125-129]), qui a une forte activité inhibitrice de l'ECA (IC_{50}= 21 μM ; Hernandez-Ledesma *et al.,* 2007). Des peptides inhibiteurs de l'ECA ont été mis en évidence dans certains fromages comme le parmesan (Addeo *et al.,* 1992), le gouda (Meisel *et al.,* 1997) ou le gorgonzola (Smacchi et Gobbeti, 1998) entre autres.

4.3. Peptides antithrombotiques

Les mécanismes impliqués dans la coagulation du lait par l'action de la chymosine sur la CN-κ et dans la coagulation du sang par l'action de la thrombine sur le fibrinogène présentent des similitudes (Jollès, 1975 ; Jollès et Henschen, 1982 ; Fiat *et al.,* 1993). Ainsi, plusieurs peptides issus de la CN-κ bovine ([106-116], [106-112], [113-116]) et appelés casoplatélines affectent la fonction plaquettaire. Le décapeptide [106-116] possède trois résidus d'acides aminés particuliers (Ile^{108}, Lys^{112} et Asp^{115}) qui ont une position homologue à ceux présents dans la chaine γ du fibrinogène humain (Fiat et Jollès, 1989). En effet, les casoplatélines inhibent l'aggrégation des plaquettes humaines en empêchant généralement la liaison de la chaine γ du fibrinogène sur son récepteur spécifique au niveau des plaquettes activées par l'ADP (Jollès *et al.,* 1986 ; Fiat et Jollès, 1989 ; Schlimme et Meisel, 1995). Néanmoins, le fragment [103-111] de la CN-κ bovine peut inhiber l'aggrégation plaquettaire sans affecter la liaison du fibrinogène à son récepteur plaquettaire (Fiat *et al.,* 1993).

4.4. Peptides à activité opiacée

Des peptides ayant une activité opiacée ont été identifiés dans plusieurs fractions de caséines hydrolysées par des enzymes du tractus digestif (Brantl *et al.,* 1979 ; Pihlanto-Leppälä *et al.,* 1994 ; Teschemacher, 2003). Ces peptides sont des ligands pour les récepteurs opiacés et ont une activité agoniste ou antagoniste. Ces récepteurs opiacés (types μ, δ et κ) ont généralement un rôle dans les systèmes immunitaire, nerveux et endocrinien mais également

au niveau du tractus gastro-intestinal des mammifères. Des ligands endogènes ou des ligands exogènes agonistes ou antagonistes peuvent interagir avec ces récepteurs. L'administration orale de peptides opiacés pourrait moduler le phénomène d'absorption intestinale et influencer la fonction gastro-intestinale de deux façons : en stimulant les muscles lisses ce qui réduit la durée du transit intestinal, en affectant le transport intestinal des électrolytes ce qui explique les propriétés anti-sécrétoires de ces peptides. Les effets physiologiques des peptides opiacés dérivés du lait restent encore à confirmer (Korhonen et Pihlanto, 2006). Des peptides opiacés issus de la CN-β, appelés β-casomorphines, ainsi que leurs précurseurs ont été détectés dans le chyme duodénal de cochon nain (Meisel et Frister, 1989), dans le plasma de veaux nouveaux-nés (Umbach *et al.,* 1985) et dans l'intestin grêle d'adultes humains (Svedberg *et al.,* 1985) après l'ingestion orale de caséines ou de lait. Chez les mammifères adultes, de la β-casomorphine a été retrouvée dans le plasma de femmes enceintes ou post-partum, cependant aucun peptide n'est détecté dans le plasma chez l'homme ou la femme non-enceinte (Koch *et al.,* 1988). L'hydrolyse des CN-$\alpha_{s1,}$ CN-κ et CN-γ produit également des peptides opiacés (Loukas *et al.,* 1983 ; Pihlanto-Leppälä *et al.,* 1994 ; Perpetuo *et al.,* 2003). Des peptides opiacés, appelés casoxines, exercent une activité antagoniste en supprimant l'activité agoniste de l'enképhaline (Chiba *et al.,* 1989 ; Yoshikawa *et al.,* 1994). Les casoxines A et B sont issues de CN-κ bovine et humaine (Chiba et Yoshikawa, 1986 ; Chiba *et al.,* 1989), la casoxine C dérive de la CN-κ bovine (Chiba *et al.,* 1989) tandis que la casoxine D est issue de la CN-α_{s1} humaine (Yoshikawa *et al.,* 1994). L'extrémité C-terminale de ces casoxines peut être méthylée lors de leur préparation ce qui les rend plus actif comparé à leurs homologues non méthylés (Chiba et Yoshikawa, 1986 ; Chiba *et al.,* 1989).

4.5. Peptide anxiolytique

L'hydrolysat trypsique de la CN-α_{s1} bovine présente une activité benzodiazépine-mimétique. Ainsi, une injection par voie intrapéritonéale de 3 $mg.kg^{-1}$ de cet hydrolysat diminue les symptômes épileptiques provoqués par le pentylènetétrazole chez le rat. La même dose diminue l'anxiété des rats lors du test du labyrinthe en croix surélevée et lors du test d'enfouissement défensif conditionné. Dans l'hydrolysat trypsique de CN-α_{s1} seul un peptide présente une activité benzodiazépine-mimétique : le fragment [91-100] appelé α-casozépine (Miclo *et al.,* 2001). L'α-casozépine se fixe sur le récepteur $GABA_A$ qui est un récepteur multimérique présentant un site de fixation des benzodiazépines, ces dernières étant des

molécules anxiolytiques et anti-convulsivantes. Ce peptide a fait l'objet d'un dépôt de brevet européen en 1996 (Miclo *et al.*). La société Ingredia commercialise sous le nom de Lactium™, un hydrolysat industriel de CN-α_{s1} contenant l'α-casozépine. Lactium™ est utilisé pour la préparation de compléments alimentaires et de médicaments destinés, pour ces derniers, à la médecine vétérinaire. Chez le rat, l'ingestion orale de cet hydrolysat préviendrait les troubles du sommeil induit par un stress chronique (Guesdon *et al.,* 2006). Chez le chat, il a un effet positif sur la gestion des troubles anxieux comme les phobies sociales (Beata *et al.,* 2007). Chez l'Homme, une ingestion de 1,2 g de Lactium™ en trois prises avant une période de stress diminue significativement l'augmentation de pression sanguine induite par le stress. De plus, la concentration du cortisol plasmatique diminue chez les sujets traités en comparaison avec les individus ayant pris un placebo à base de poudre de lait (Messaoudi *et al.,* 2005). Dans une autre étude, des femmes volontaires ont ingéré des capsules de Lactium™ à une dose de 150 mg par jour pendant 30 jours. Les individus recevant le Lactium™ et le placebo ont été croisés et il fut alors observé une diminution des symptômes digestifs, cardiovasculaires, intellectuels, émotionnels et sociaux liés au stress (Kim *et al.,* 2007).

4.6. Cas des caséinophosphopeptides

Le terme de caséinophosphopeptide (CPP) a été employé pour la première fois par Mellander en 1950 pour désigner un peptide phosphorylé issu de caséines qui améliorait la calcification osseuse indépendamment de la vitamine D chez des enfants rachitiques. Le terme a ensuite été repris plus largement pour désigner les peptides phosphorylés issus de l'hydrolyse enzymatique des caséines α_{s1}, α_{s2} et β *in vitro* ou formés dans le système digestif. La majorité des CPP possèdent le cluster caractéristique du type -Ser-Ser-Ser-Glu-Glu-. Les charges négatives des chaînes latérales, en particulier les groupements phosphates des résidus SerP mais aussi le groupement carboxylique des résidus Glu, constituent des sites de liaison où peuvent se fixer des cations nutritionnellement intéressants notamment le calcium, le fer ou encore le zinc. Cela permet d'assurer la stabilité et la solubilité de ces cations dans différentes conditions physico-chimiques et donc leur absorption au niveau de l'intestin (Brulé et Fauquant, 1982 ; Meisel et Schlimme, 1993 ; Meisel, 1998). La fixation de certains ions aux CPP fait d'eux des peptides pouvant avoir une activité biologique.

4.6.1. Activité anti-oxydante

L'oxydation des lipides dans les aliments peut entraîner la production de radicaux libres qui attaquent les acides gras et génèrent le développement d'odeurs et de flaveurs rances indésirables (Nawar, 1996 ; Akoh et Min, 2002). Ainsi l'élaboration d'aliments transformés stables est un véritable challenge d'autant plus que les acides gras polyinsaturés qui les composent sont très sensibles à l'oxydation.

Les métaux de transition comme le fer se comportent comme des agents pro-oxydants dans les systèmes alimentaires en accélérant le clivage des hydroperoxydes lipidiques en radicaux hydroxy $^{\bullet}OH$ très réactifs (Decker et McClements, 2001).

Les CPP ont une activité anti-oxydante primaire par piégeage de radicaux libres $^{\bullet}OH$ (*via* les résidus Lys, His, Pro et Tyr) mais également une activité anti-oxydante secondaire due à la chélation de métaux pro-oxydants (Díaz et Decker, 2004 ; Kitts et Nakamura, 2006).

Par exemple, le radical $^{\bullet}OH$ peut être dosé indirectement par le dosage « désoxyribose » (Kitts, 2005). Le désoxyribose peut être dégradé par les radicaux $^{\bullet}OH$ qui sont générés lors de la réaction de Fenton. Les CPP peuvent séquestrer l'ion ferreux, qui est un pro-oxydant dans la réaction de Fenton, ce qui entraîne une diminution de la production de radicaux $^{\bullet}OH$. La réaction de Fenton est la suivante :

$$H_2O_2 + Fe^{2+} \rightarrow Fe^{3+} + {}^{\bullet}OH + OH^-$$

Le désoxyribose peut former un complexe avec l'ion Fe^{2+} et réagir avec une molécule de H_2O_2 formant ainsi des radicaux $^{\bullet}OH$ qui endommagent immédiatement la molécule de désoxyribose (Aruoma *et al.,* 1989). Les CPP entrent en compétition avec le désoxyribose pour la chélation du fer réduisant ainsi l'interaction entre le métal pro-oxydant et son substrat. Il en résulte la protection du désoxyribose contre la dégradation causée par les radicaux $^{\bullet}OH$ issus de la réaction de Fenton. Lors de l'ajout d'EDTA, afin de réduire la liaison entre les ions ferreux et le désoxyribose, entraînant une augmentation de la production de radicaux $^{\bullet}OH$, les CPP peuvent piéger les radicaux $^{\bullet}OH$ formés, ce qui entraîne une diminution de l'oxydation du désoxyribose (Kitts, 2005). Les résidus de SerP catalysent l'oxydation du fer ferreux en fer ferrique permettant la formation d'un complexe CPP/Fe^{3+} très stable.

La présence de CPP dans le tractus gastro-intestinal mise en évidence après la digestion de produits laitiers par les enzymes digestives pourrait avoir un effet bénéfique pour la santé en contribuant à la réduction du stress oxydatif intestinal et au maintien du bon fonctionnement de l'intestin (Kitts, 2005).

Les CPP ont montré des propriétés anti-oxydantes dans des systèmes aqueux et des modèles d'émulsions (liposomes) mais également dans de la viande de bœuf (Díaz et Decker, 2004 ; Kansci *et al.,* 2004 ; Kitts et Nakamura, 2006). Les CPP pourraient être utilisés comme antioxydants dans des préparations culinaires à base de viande (Díaz *et al.,* 2003 ; Sakanaka *et al.,* 2005).

4.6.2. Activité anti-cariogène

L'émail dentaire est une substance polymérique constituée de phosphate de calcium cristallin (hydroxyapatite) entourant une matrice protéique. Un environnement acide provoque une déminéralisation de l'émail dentaire avec solubilisation du calcium et du phosphate provoquant l'apparition de caries. La déminéralisation peut être directe, c'est-à-dire juste après la consommation d'aliments. Elle peut aussi être indirecte due aux produits de fermentation issus de la dégradation de particules alimentaires résiduelles (accumulées entres les dents ou adhérentes à la plaque dentaire) par des bactéries odontopathogènes de la plaque dentaire. Les CPP peuvent prévenir la déminéralisation et favoriser la reminéralisation de l'émail dentaire. En effet, ils peuvent inhiber la croissance de bactéries odontopathogènes comme par exemple *Streptococcus mutans*. Les CPP forment également des nanoclusters avec du phosphate de calcium amorphe à la surface de la dent ce qui constitue un réservoir de calcium et de phosphate afin de maintenir un état de sursaturation au niveau de l'émail. Cela permet à la fois de tamponner le pH au niveau de la plaque dentaire et d'apporter les ions nécessaires à la reminéralisation (Aimutis, 2004). Il a été montré que la protection contre les caries par les CPP est de 60% à pH 4,2. L'augmentation de la protection anti-carie est corrélée avec l'augmentation de la concentration en CPP jusqu'à 10 mg/mL. La protection maximale est atteinte pour des concentrations de 4 à 6 mg/mL et celle-ci est constante pour des pH compris entre 2,5 et 4,5 (Kanekanian *et al.,* 2008). Il a également été observé que la fixation préalable de calcium aux CPP ne permet pas d'améliorer l'efficacité de protection contre les caries par rapport aux CPP libres. Il se pourrait que les CPP se lient très fortement à l'hydroxyapatite formant ainsi une couche protectrice au niveau de l'émail *via* un processus d'adsorption à la surface de la dent, couche qui serait résistante à des étapes de rinçage de la bouche (Kanekanian *et al.,* 2008).

Les CPP sont utilisés dans des compositions de gommes à mâcher, dentifrices, bains de bouches ainsi que dans des additifs alimentaires afin de permettre un contact prolongé avec

la plaque dentaire pour prévenir la formation de caries (Gaucheron *et al.,* 1995 ; Aimutis, 2004).

4.6.3. Effet mitogène

Les CPP bovins correspondant aux peptides CN-α_{s2} [1-32] et CN-β [1-28] augmentent la prolifération cellulaire induite par une stimulation par la concanavaline A, la phytohémoglutinine et le liposacharride, ainsi que la production d'immunoglobuline dans des cultures de cellules spléniques chez la souris (Hata *et al.,* 1999). Les résidus de SerP contenus dans la séquence de type -SerP-Xxx-SerP- du CPP bovin CN-β (f1-28) ont été identifiés comme étant des activateurs de cette fonction immunomodulatrice (Otani *et al.,* 2001). Ainsi, il a été montré que l'administration orale de CPP (CN-α_{s2} [1-32] et CN-β [1-28]) améliore les taux intestinaux et plasmatiques en IgA chez la souris (Otani *et al.,* 2003).

Kawahara *et al.* (2004) ont étudié l'effet du CPP CN-β [1-28] sur la prolifération de lignées cellulaires humaines (lymphocytes T, B et monocytes) et sur la production d'immunoglobulines et l'expression de cytokinines par ces cellules. Le CPP CN-β [1-28] améliore les réponses prolifératives de toutes ces lignées et stimule la production d'IgA par les lymphocytes B. Donc des produits alimentaires contenant le CPP CN-β [1-28] pourraient promouvoir les réponses immunitaires dans l'organisme humain.

4.6.4. Chélation de minéraux d'intérêt nutritionnel : exemple du calcium et du fer

Cas du calcium :

Les CPP peuvent former des sels de phosphate organique et servent de transporteurs pour différents minéraux comme le calcium. La liaison du calcium aux CPP empêche sa précipitation dans le tractus gastro-intestinal. Ainsi, les CPP permettent le maintien du calcium dans une forme soluble à pH alcalin dans la lumière intestinale ce qui améliorerait sa disponibilité pour une absorption paracellulaire (Kitts et Yuan, 1992). De nombreuses études ont été réalisées sur l'effet des CPP sur l'absorption intestinale du calcium et les résultats obtenus peuvent être contradictoires. Des études menées *in vivo* sur des rats nourris avec des protéines de soja supplémentées en CPP montrent une meilleure absorption du calcium pour ces rats supplémentés par rapport à des rats non supplémentés (Kitts *et al.,* 1992). La biodisponibilité du calcium est également augmentée chez de jeunes ratons nourris avec des

aliments riches en phytate enrichis en CPP (Hansen *et al.,* 1996) et chez des rattes âgées ovariectomisées nourries avec un aliment enrichi en CPP-Ca en prévenant la chute de la densité minérale (Tsuchita *et al.,* 1996). Les CPP améliorent aussi l'absorption du calcium (mesure de la quantité de ^{47}Ca quittant la lumière intestinale), mesurée par la technique *in vivo* de l'intestin ligaturé et la technique *in vitro* de l'intestin retourné, chez des poulets rachitiques mais aussi chez des poulets normaux (Mykkanen et Wasserman, 1980). Cependant, d'autres études *in vivo* sur des porcs en croissance, de jeunes rattes et des rattes adultes montrent qu'une supplémentation en CPP de l'aliment n'influence ni l'absorption du calcium, ni la minéralisation osseuse (Pointillard et Guéguen, 1989 ; Brommage *et al.,* 1991 ; Kopra *et al.,* 1992). Une étude menée chez l'homme sur l'effet de boissons enrichies en CPP ne montre pas une augmentation de l'absorption de calcium en présence de CPP (Teucher *et al.,* 2006), ce qui avait déjà été observé dans de précédents travaux (Meisel *et al.,* 2003 ; Narva *et al.,* 2003). Néanmoins, dans toutes ces études menées *in vivo* les CPP sont généralement considérés comme des agents capables de maintenir le calcium sous forme soluble dans l'intestin ce qui faciliterait l'absorption du minéral au niveau intestinal. Il n'est cependant pas tenu compte du fait que la présence ou l'absence d'autres substances dans les aliments tels que le phosphate ou le phytate peuvent influer sur la solubilité du calcium. En effet, ces composés peuvent former des complexes ou des sels de calcium insolubles. Ceci pourrait expliquer que les différentes études rapportées ont conduit à des résultats contradictoires obtenus *in vivo* (Ferraretto *et al.,* 2001). De même, ces résultats contradictoires obtenus pourraient être dus aux conditions expérimentales utilisées qui diffèrent par les modèles expérimentaux utilisés, l'âge et l'état physiologique des animaux, la durée d'administration des CPP, la composition des aliments ou encore la complexité et la variabilité des préparations de CPP utilisées (Bouhallab et Bouglé, 2004). Ainsi, des travaux ont été réalisés *in vitro* sur un modèle d'épithélium intestinal utilisé dans les études d'absorption intestinale (cellules tumorales intestinales humaine : HT-29) afin d'étudier l'influence de la seule présence de CPP sur l'absorption intestinale du calcium. Les résultats montrent une augmentation de la concentration intracellulaire de calcium indiquant qu'il y a eu une absorption de ce calcium (Ferraretto *et al.,* 2001, 2003 ; Gravaghi *et al.,* 2007). Il a été montré que les CPP n'influençaient pas les récepteurs membranaires et les canaux ioniques au calcium mais interagiraient directement avec la membrane. Une hypothèse formulée est que les CPP pourraient jouer le rôle de ionophores au calcium en s'enchâssant dans la membrane cellulaire créant ainsi leurs propres canaux (Ferraretto *et al.,* 2001) ; les CPP pouvant s'agréger en oligomères ils pourraient avoir les propriétés physico-chimiques nécessaires pour

s'insérer dans la membrane (Reynolds, 1993). Une autre hypothèse est que les CPP joueraient le rôle de transporteur de calcium à travers la membrane, par exemple, par endocytose (Ferraretto *et al.,* 2001). Il est observé dans le modèle HT-29 que le CPP CN-β [1-25] permet une meilleure absorption du calcium comparativement au CPP CN-α_{s1} [59-79] (Ferraretto *et al.,* 2003). Afin de mieux comprendre le mécanisme d'absorption du calcium, des expériences ont été menées sur le CPP CN-β [1-25]. Ainsi, un peptide synthétique CN-β [17-21] constitué du motif -SerP-SerP-SerP-Glu-Glu- est inefficace sur l'absorption du calcium. Il en est de même lorsque le CPP CN-β (f1-25) est totalement déphosphorylé. L'absence des cinq résidus d'acides aminés N-terminaux ou une modification de la séquence N-terminale du CPP CN-β [1-25], par l'inversion lors de la synthèse chimique des portions [1-4] et [8-11], provoque une perte totale d'activité. Cependant, l'absence de la partie C-terminale [21-24] n'a pas d'effet sur l'activité qui est donc conservée. Ainsi, l'absorption du calcium par les cellules HT-29 en présence du CPP CN-β [1-25] semble dépendre à la fois de la séquence phosphorylée et de la région N-terminale du peptide. L'interaction entre le CPP et la membrane plasmique requiert donc une structure et une conformation peptidique précise (Ferraretto *et al.,* 2003). Le modèle cellulaire HT-29 a permis d'établir un ratio molaire optimal Ca/CPP de 20 – 4 mmol/L Ca^{2+} / 200 µmol/L CPP CN-β [1-25] – pour obtenir un effet sur l'absorption du calcium (Gravaghi *et al.,* 2007), confirmant ainsi des résultats obtenus *in vivo* avec des animaux ou avec de l'intestin inversé (poulets et rats ; Mykkänen et Wasserman, 1980 ; Saito *et al.,* 1998 ; Tsuchita *et al.,* 2001 ; Erba *et al.,* 2002).

Cas du fer :

La carence martiale est un problème de santé publique dans les pays en voie de développement mais aussi en dans les pays industrialisés. Par exemple, 15 à 20% des nourrissons et 20 à 30% des femmes enceintes ou en âge de procréer présentent une carence biologique en fer pouvant être responsable d'anémie, de perturbations immunitaires ou de perturbations du développement psychomoteur (Hercberg et Preziosi, 2000). L'absorption du fer dépend de facteurs diététiques qui peuvent favoriser ou au contraire inhiber cette absorption. Ainsi, la présence de vitamine C ou d'hémoglobine augmente l'absorption du fer alors que la présence de calcium ou de phytate la diminue (Barton *et al.,* 1983 ; West, 1986 ; Hurrell *et al.,* 1992 ; Conrad, 1993).

Le CPP CN-β [1-25] peut chélater l'ion Fe^{2+} et cette liaison du fer semble améliorer sa biodisponibilité et le statut martial chez le jeune rat carencé (Aït-Oukhatar *et al.,* 1997). Les

CPP possèdent une capacité plus élevée pour guérir une anémie et restaurer les stocks de fer dans les tissus comparativement aux caséines non hydrolysées et aux sels inorganiques (Aït-Oukhatar *et al.,* 1999). Ce rôle positif sur l'absorption du fer est également observé avec le CPP CN-β [1-25] chez des rats non déficients en fer (Bouhallab *et al.,* 2002). En effet, la liaison de l'ion Fe^{2+} au CPP provoque son oxydation en ion Fe^{3+} (Bouhallab *et al.,* 1991), cette liaison permet de maintenir le fer soluble dans le tractus digestif ce qui prévient la formation d'hydroxyde de fer de masse moléculaire élevée qui est faiblement absorbé (Aït-Oukhatar *et al.,* 2002). La liaison du fer au CPP CN-β [1-25] empêche l'effet inhibiteur du calcium (Pérès *et al.,* 1997) et du zinc (Pérès *et al.,* 1999) sur l'absorption intestinale du fer. Le CPP CN-β [1-25] possède quatre résidus de SerP et peut fixer de quatre à cinq atomes de fer (Brulé et Fauquant, 1982, Meisel et Schlimme, 1993). De plus, les CPP ont une affinité 100 fois plus forte pour le fer que pour le calcium ou les autres cations (Brulé et Fauquant 1982 ; Emery, 1991).

Des expériences sur du duodénum de rat ligaturé ont montré que le CPP CN-β [1-25] lié ou non au fer est sensible aux enzymes digestives (protéases, phosphatases, peptidases) (Bouhallab *et al.,* 1999). En effet, plusieurs peptides dérivant du CPP CN-β [1-25] sont retrouvés dans le contenu luminal de l'intestin ligaturé mais néanmoins l'hydrolyse semble être plus lente quand le CPP CN-β [1-25] est lié au fer. Le peptide CN-β [15-24] contenant les résidus phosphorylés, issu de l'hydrolyse du CPP CN-β [1-25], est le seul peptide à ne pas être retrouvé dans le contenu luminal. Cela pourrait suggérer que, de par sa masse moléculaire faible le complexe CN-β [15-24]/Fe puisse franchir la bordure en brosse et serait donc l'entité moléculaire impliquée dans le transfert du fer à travers la membrane et non le CPP CN-β [1-25] (Bouhallab *et al.,* 1999).

L'absorption du fer est fortement dépendante de la séquence du CPP. En effet, le CPP CN-β [1-25] permet une meilleure absorption du fer que l'ensemble des CPP en mélange ou que les CPP issus des CN-α_{s1} et CN-α_{s2} dans le modèle de l'intestin de rat perfusé (duodénum ; Bouhallab *et al.,* 2002). La différence d'absorption du fer observée entre le lait bovin et le lait humain, meilleure pour ce dernier, pourrait alors être due à la différence de composition protéique de ces deux laits, le lait humain contenant peu de CN-α_{s1} au contraire du lait bovin (Bouhallab *et al.,* 2002). Plusieurs hypothèses peuvent être formulées pour expliquer la différence d'absorption du fer observé selon l'origine du CPP impliqué. En effet, la liaison du fer aux CPP implique des changements structuraux qui sont différents pour les CPP issus de la CN-β et des CN-α_s (Gaucheron *et al.,* 1997). Ainsi, pour un CPP donné il

existerait une corrélation entre la complexation avec le minéral et l'accessibilité du complexe CPP/Fe à la membrane apicale et aux systèmes de transport. La constante d'affinité entre le fer et le CPP doit être optimale pour générer un effet biologique positif. Le CPP CN-β [1-25] ayant quatre résidus SerP et sept résidus Glu possède une bonne constante d'affinité. Au contraire, les CPP issus des CN-α_s ayant un ou deux résidus SerP ne posséderaient pas assez de sites de liaison au fer pour être efficaces et ceux ayant quatre ou cinq résidus de SerP fixeraient trop fortement le fer de par leur forte charge négative (Bouhallab *et al.,* 2002).

Le fer est majoritairement absorbé (65%) par transport actif énergie-dépendante, le reste étant absorbé par voie passive (35%). Dans cette voie passive, l'endocytose représenterait 20% du fer absorbé (Pérès *et al.,* 1999 ; Bouhallab et Bouglé, 2004). Le rôle de la phosphatase alcaline de la bordure en brosse dans l'absorption du complexe CPP/Fe a été étudié (Kibangou *et al.,* 2005). La présence d'un inhibiteur de la phosphatase alcaline (Na_2WO_4) diminue la déphosphorylation du CPP CN-β [1-25] non lié au fer qui ne perd en présence de cet inhibiteur que un ou deux groupements phosphates. Des analyses de spectrométrie de masse montrent que l'hydrolyse du CPP est initiée dans sa partie N-terminale, puis la chaîne polypeptidique résiduelle est déphosphorylée et enfin hydrolysée totalement. Quand le fer est lié au CPP CN-β [1-25] celui-ci est moins sensible à la protéolyse et à la déphosphorylation enzymatique, la liaison du fer permet de diminuer l'activité de la phosphatase alcaline comme le fait Na_2WO_4. L'inhibition de la phosphatase alcaline augmente la disparition du complexe CPP/Fe dans le contenu luminal suggérant ainsi que le fer est donc partiellement absorbé sous sa forme liée au CPP (Ani-Kibangou *et al.,* 2005). Le complexe CN-α_{s1} [59-79]/Fe a une absorption beaucoup plus faible que le complexe CN-β [1-25]/Fe. Au contraire du complexe CN-β [1-25]/Fe, l'inhibition de la phosphatase alcaline de la bordure en brosse ne modifie pas l'absorption du complexe CN-α_{s1} [59-79]/Fe dans le modèle cellulaire Caco-2 et dans le modèle du duodénum de rat perfusé. Ce phénomène est probablement dû à une faible accessibilité des résidus SerP du fragment [59-79] aux enzymes intestinales (Kibangou *et al.,* 2005). En effet, la liaison du fer aux CPP CN-α_{s1} [59-79] et CN-β [1-25] induirait des conformations différentes en solution comme dans le cas du calcium (Cross *et al.,* 2001). Ainsi, la conformation acquise par le CPP CN-α_{s1} [59-79] en solution et sa forte charge négative (due à la présence de cinq résidus SerP) pourraient provoquer une trop forte liaison au fer et donc un complexe CN-α_{s1} [59-79]/Fe peu biodisponible (Kibangou *et al.,* 2005).

CONTEXTE ET OBJECTIFS DU TRAVAIL

En Europe de l'Est et en Asie Centrale, le lait de jument est traditionnellement consommé depuis des siècles par les populations pastorales. Dans ces pays, le lait de jument est réputé car il est supposé avoir des vertus thérapeutiques, cosmétiques et « vivifiantes ». Le lait de jument est surtout consommé sous forme de boisson fermentée, le koumis. Celui-ci est considéré comme un remède efficace dans le traitement des infections urinaires et uro-génitales. Le lait de jument est également utilisé depuis le XIX^{e} siècle comme aliment de substitution en pédiatrie en raison de ressemblances existant entre le lait de jument et le lait humain (Lozovich, 1995, Ochirkhuyag *et al.,* 2000). Néanmoins, les vertus supposées du lait de jument doivent encore être démontrées scientifiquement.

Dans les pays occidentaux, les élevages de chevaux sont plutôt destinés aux loisirs, au travail et à la consommation de viande. Cependant, depuis une vingtaine d'années la production de lait de jument connaît un certain succès et des jumenteries voient le jour notamment en Europe de l'Ouest. Le lait de jument est commercialisé en poudre sous forme lyophilisée, en comprimés, en gélules, en lait cru surgelé mais également comme composé intégré dans des formulations de crèmes cosmétiques, de savons et de shampoings.

Le polymorphisme des caséines, aussi bien du point de vue de la séquence que des modifications post-traductionnelles, est connu pour être impliqué dans les propriétés biologiques et fonctionnelles du lait. C'est le cas, par exemple, de l'allergénicité du lait caprin en relation avec le polymorphisme de la CN-α_{s1} (Bevilacqua *et al.,* 2001). Des modifications mineures affectant les caséines peuvent avoir des conséquences importantes comme dans le cas de la consommation du variant A1 de la CN-β bovine qui pourrait entraîner le développement de certains diabètes (Hill et Boland, 2001). La consommation de lait de jument se développant de plus en plus, il semble donc important d'étudier de manière plus approfondie les constituants protéiques de ce lait.

Des études récentes ont permis d'établir que la CN-β équine a, comme la CN-β humaine, la particularité de posséder un degré de phosphorylation variable, allant de 3P à 7P (Girardet *et al.,* 2006). Il a également été mis en évidence la présence d'un variant d'épissage alternatif (Miranda *et al.,* 2004 ; Lenasi *et al.,* 2006) et d'un variant d'épissage utilisant un site d'épissage cryptique (Lenasi *et al.,* 2006 ; Miclo *et al.,* 2007). De même, un site de désamidation non enzymatique a été mis en évidence dans la structure primaire de la CN-β, ce qui pourrait augmenter l'hétérogénéité de cette protéine, qui gagne une charge négative une fois désamidée (Girardet *et al.,* 2006). En ce qui concerne la CN-α_{s1}, les profils électrophorétiques observés (Egito *et al.,* 2002) laissent supposer que cette caséine

possèderait, comme la CN-β, un degré de phosphorylation variable. De plus, un phénomène d'épissage alternatif concernant les exons 7 et 14 a été mis en évidence grâce à l'isolement d'ARNm à partir de cellules de glandes mammaires (Milenkovic *et al.,* 2002 ; Lenasi *et al.,* 2003).

La phosphorylation variable de la CN-β équine, et celle supposée de la CN-α_{s1}, pourrait avoir des répercussions sur l'affinité des CPP générés enzymatiquement pour des minéraux d'intérêts nutritionnels et pourrait donc modifier la biodisponibilité des minéraux au niveau gastro-intestinal. Il pourrait également y avoir des conséquences sur la chélation de métaux pro-oxydants et donc les variations structurales des caséines pourraient moduler les propriétés anti-oxydantes des CPP. En effet, des études portant sur les CPP issus des CN-β et CN-α_{s1} bovines ont montré que la séquence et le degré de phosphorylation ont une influence sur l'absorption du fer (Gaucheron *et al.,* 1997 ; Bouhallab *et al.,* 2002 ; Ani-Kibangou *et al.,* 2005).

L'objectif de ce travail est de mieux définir la microhétérogénéité des caséines α_{s1} et β équines due à la présence (probable pour la CN-α_{s1}) d'une phosphorylation variable et de variants d'épissage, ceci afin de pouvoir réaliser ultérieurement l'étude de l'influence de la séquence et du degré de phosphorylation et sur le pouvoir chélateur des CPP. Les CPP équins, de par leur diversité de structure, nous permettront de mieux comprendre les mécanismes de chélation et de transport de minéraux, et par conséquent de mieux comprendre comment les CPP bovins (séquences totalement phosphorylées) et les CPP humains (séquences phosphorylées à des degrés variables) chélatent les minéraux.

Dans la première partie de notre travail, la microhétérogénéité de la CN-α_{s1} a été étudiée au niveau protéique. Cette caséine a d'abord été purifiée à partir de caséinate de sodium équin fractionné par FPLC d'interactions hydrophobes et les différents variants post-transcriptionnels (résultants d'épissage alternatif) et post-traductionnels (phosphorylation) de la CN-α_{s1} ont été totalement déphosphorylés par les actions successives de la phosphatase acide et de la phosphatase alcaline. Grâce à ce double traitement enzymatique, différents variants d'épissage ont pu être identifiés et leurs degrés de phosphorylation déterminés par spectrométrie de masse.

Dans la deuxième partie de notre travail, nous nous sommes focalisés sur l'étude de la microhétérogénéité de la CN-β équine et sur la localisation des sites phosphorylés au niveau de sa structure primaire. Contrairement au cas de la CN-α_{s1}, le degré de phosphorylation de la CN-β équine ainsi que ses variants d'épissage sont déjà connus. Cependant, les profils

électrophorétiques mono- et bidimensionnels de cette protéine restent difficiles à interpréter de part la présence de nombreuses isoformes natives et la présence potentielle d'isoformes modifiées par désamidation non enzymatique. Dans le présent travail, il a donc été nécessaire de déterminer si un phénomène de désamidation non enzymatique de la CN-β pouvait se produire dans le lait ou encore pendant la préparation de la CN-β. Les différents variants de phosphorylation de la CN-β ont été purifiés individuellement par FPLC d'échange d'anions et la génération *in vitro* de leurs formes désamidées respectives a permis d'établir une cartographie bidimensionnelle des différentes isoformes de la CN-β équine. La séparation des différents variants de phosphorylation de la CN-β a également permis de générer les CPP correspondant à chacun des degrés de phosphorylation de 4P à 7P, de purifier ces CPP par une technique récente de chromatographie d'affinité au dioxyde de titane et de localiser les sites phosphorylés de chacune des isoformes par spectrométrie de masse en tandem . Cette étude structurale de la CN−β équine est un préalable indispensable à toute étude du mécanisme de chélation de minéraux ou d'oligo-éléments d'intérêt nutritionnel par la CN-β ou de ses CPP.

MATERIELS ET METHODES

1. Préparation du caséinate de sodium

Le lait utilisé est un lait cru provenant de juments de race Haflinger de la société « le domaine de la voie lactée » basée à Volmunster en Moselle. Il ne s'agit pas d'un lait provenant d'une jument unique mais d'un lait de petit mélange. Le Haflinger est un poney ayant une taille idéale entre 144 et 149 cm (Figure 24). La robe de ce poney est toujours alezan crin lavés. C'est un poney principalement utilisé pour une équitation de loisirs (attelage, randonnée, endurance…) mais également pour la production de lait de jument. Le premier étalon de la race, Folie, a vu le jour en 1874, près de Hafling, un village autrichien. Les ascendants paternels de tous les étalons de France remontent sans discontinuer à Folie.

Le lait est additionné de 0,05% (v/v) de toluène puis est écrémé par centrifugation (30 min, 32°C, 4000 x *g*). Les caséines et les protéines du lactosérum sont préparées à partir du lait écrémé selon la méthode décrite par Egito *et al.* (2002). Les caséines sont précipitées à pH 4,2 par addition d'HCl 1 M. Le coagulum est ensuite centrifugé (20 min, 20°C, 4000 x *g*). Le culot est lavé deux fois sous agitation avec de l'eau distillée additionnée de 0,05% (v/v) de toluène puis dissous par remontée du pH à 7,0 avec NaOH 1 M pendant une nuit à 4°C. Ce cycle précipitation-lavage-dissolution est répété deux fois. Le caséinate de sodium obtenu est ensuite dialysé contre de l'eau distillée en présence de thymol pendant 72 h à 4°C puis lyophilisé. Le lyophilisat est conservé à -20°C.



Figure 24 : Photographies de poneys de la race Haflinger.

2. Méthodes électrophorétiques et techniques attenantes

2.1. Electrophorèse en gel de polyacrylamide en présence de dodécylsulfate de sodium (PAGE-SDS)

La méthode en tampons discontinus de Laemmli et Favre (1973) est utilisée avec un système de migration verticale (modèle électrophorétique 2001, Amersham Pharmacia, Uppsala, Suède). Le gel est constitué d'un gel de concentration (T : 4,9% ; C : 2,7%) en tampon Tris-HCl 0,125 M, pH 6,8 contenant 0,1% (m/v) de SDS et d'un gel de séparation (T : 15,4 % ; C : 2,7 %) en tampon Tris-HCl 0,38 M, pH 8,8 contenant 0,1% (m/v) de SDS. Le gel a une épaisseur de 1,5 mm, une hauteur de 16 cm et une largeur de 18 cm. Le tampon d'électrodes (pH 8,3) est constitué de Tris 50 mM, glycine 384 mM et de 0,1% (m/v) SDS. Les échantillons (lyophilisats) sont solubilisés à 2 $mg.mL^{-1}$ dans un tampon Tris-HCl 0,125 M, pH 6,8 contenant 2% (m/v) de SDS, 5% (v/v) de 2-mercaptoéthanol, 10% (v/v) de glycérol et 0,1% (m/v) de bleu de bromophénol. Les échantillons sont chauffés à 100°C pendant 3 min. Un volume de 20 µL (40 µg de protéines) par échantillon est déposé dans les puits du gel. La migration électrophorétique est effectuée sous des conditions électriques de 60 mA, 30 W, 500 V (pour un gel) pendant environ 150 min à 4°C. Les marqueurs de masses moléculaires sont constitués d'un mélange (Bio-Rad, Hercules, CA, Etats-Unis) de myosine (200,0 kDa), β-galactosidase (116,25 kDa), phosphorylase b (97,4 kDa), sérum albumine bovine (66,2 kDa), ovalbumine (45,0 kDa), anhydrase carbonique (31,0 kDa), inhibiteur trypsique (21,5 kDa), lysozyme (14,4 kDa ; cependant sa masse moléculaire apparente est de 15,5 kDa), aprotinine (6,5 kDa).

2.2. Electrophorèse en gel de polyacrylamide en présence d'urée (PAGE-urée)

La méthode utilisée est adaptée de celle de Ng-Kwai-Hang et Kroeker (1984). Un gel de polyacrylamide (T : 8,2% ; C : 2,7%) en tampon 75 mM Tris-HCl, pH 8,9 contenant 4 M d'urée, d'épaisseur 1,5 mm, est coulé dans le système de plaques verticales (18 x 16 cm). Le tampon d'électrode contient 80 mM de glycine et son pH est ajusté à 8,3 par ajout de Tris. Les échantillons (lyophilisats) sont solubilisés à 2 $mg.mL^{-1}$ dans un tampon 75 mM Tris-HCl pH 8,9 contenant 4 M d'urée, 5% (v/v) de 2-mercaptoéthanol, 10% (v/v) de glycérol et 0,1% (m/v) de bleu de bromophénol. Les échantillons subissent alors un traitement thermique à

100°C pendant 3 min avant d'être analysés. Un dépôt de 20 µL (40 µg) par échantillon est réalisé dans chaque puits du gel. La migration électrophorétique est effectuée sous des conditions électriques de 60 mA, 30 W, 500 V (pour un gel) pendant environ 150 min à 4°C.

2.3. Electrophorèse bidimensionnelle couplant une isoélectrofocalisation (IEF) et une PAGE-SDS

La méthode utilisée est adaptée de celle de Herbert *et al.* (1998). La première dimension (IEF) consiste à faire migrer les protéines selon leur point isoélectrique dans un gradient de pH pré-établi. Les caséines (50 µg) sont solubilisées dans 200 µL d'une solution contenant 8 M d'urée, 4% (m/v) de sulfonate de 3-[(3-cholamidopropyl)-diméthylammonio]-1-propane (CHAPS), 2 mM de tributyl phosphine, 40 mM de Tris, 0,2% (m/v) d'ampholines 3-10 (Bio-Lytes 3-10 ; Bio-Rad) et 0,001% (m/v) de bleu de bromophénol. Des bandelettes (IPG-strip d'une longueur de 11 cm et d'une largeur de 3,3 mm) constituées d'un gel de polyacrylamide (T : 4% ; C : 3%), sur lesquelles des ampholines ont été fixées de manière covalente afin de réaliser un gradient linéaire de pH, sont utilisées. Le gradient de pH utilisé est compris entre 4,0 et 7,0 pour les caséines (IPG ReadyStrip pH 4-7 ; Bio-Rad). Les IPG-strip sont placés dans l'appareil Protean IEF Cell (Bio-Rad) et sont réhydratés sous des conditions actives (*i.e.*, voltage constant de 50V) avec les solutions protéiques correspondantes pendant 15 h à 10°C ou à 20°C selon le cas. La focalisation se déroule ensuite en trois étapes à 10°C ou 20°C. Tout d'abord, un faible voltage (250 V pendant 15 min) est requis pour éliminer les ions salins en excès. Ensuite, une montée en gradient de voltage de 250 V à 8000 V est appliquée sur 150 min avec un courant limité à 50 µA par IPG-strip. Enfin, le voltage est maintenu constant à 8000 V pendant au moins 5 h. A l'issue de la migration les IPG-strips peuvent être conservés à -20°C en attente de la deuxième dimension.

La seconde dimension consiste à séparer les protéines selon leur masses moléculaires par PAGE-SDS. Les IPG-strips issus de la première dimension sont équilibrés pendant 15 min dans un tampon 125 mM Tris-HCl, pH 6,8 contenant 6 M d'urée, 2% (m/v) de SDS, 20% (v/v) de glycérol et 5 mM de tributyl phosphine. Les IPG-strips sont ensuite scellés au sommet d'un gel SDS-PAGE grâce à une solution de 1% (m/v) d'agarose fondu. La migration SDS-PAGE se déroule ensuite comme décrit ci-dessus. Les pHi apparents des protéines sont déterminés à l'aide d'un mélange de proteins standards (2-D SDS-PAGE Standards ; Bio-Rad) contenant de l'inhibiteur trypsique (pHi = 4,5), de l'actine (pHi = 5,1), de l'albumine

sérique (pHi = 5,5), de l'anhydrase carbonique (pHi = 6,2), de la conalbumine (pHi = 6,3 et 6,6), de la myoglobine (pHi = 7,0) et de la glycéraldéhyde-3-phosphate déshydrogénase (pHi = 6,9 et 7,6). Les relations entre les pHi apparents et les migrations relatives (R_f) lors de la première dimension sont linéaires avec l'IPG ReadyStrip pH 4-7 de 11 cm (R_f mesurée à partir du côté le plus acide du gel correspondant à pH 4,0). L'équation utilisée est : pH = 3.34 R_f + 3.90, r = 0.98.

2.4. Electrotransfert de protéines

2.4.1. En cuve verticale

Une membrane de difluorure de polyvinylidène (PVDF) est incubée pendant 2 secondes dans 100% de méthanol puis lavée 2 min dans de l'eau ultra-pure et équilibrée 10 min dans du tampon de transfert composé de 10 mM d'acide 3-(cyclohexylamino)-1-propanesulfonique (CAPS), pH 11,0 contenant 10% (v/v) de méthanol. Après migration électrophorétique, le gel de polyacrylamide est équilibré 15 min dans le tampon de transfert puis mis en contact avec la membrane de PVDF. L'ensemble est comprimé légèrement entre deux blocs de Scotch-Brite® en évitant les bulles d'air et le tout est placé dans une cuve verticale (modèle TE-50 Transphor Cell ; Hoeffer Scientific Instrument, San Francisco, CA, Etats-Unis) remplie de tampon de transfert. Le gel est placé côté cathode et la membrane côté anode. La migration des protéines du gel vers la membrane est réalisée à 4°C avec un voltage constant de 100 V et un courant de 0,5 A pendant 4 h.

2.4.2. Transfert semi-sec

Le traitement de la membrane PVDF et du gel de polyacrylamide est identique à ce qui est décrit ci-dessus. Des feuilles de papier buvard, préalablement découpées à la taille du gel, sont trempées dans le tampon de transfert afin qu'elles en soient totalement imbibées (6 feuilles par gel). L'électrotransfert est réalisé grâce à un appareil de transfert semi-sec (système TE 77 Power, 21 x 26 cm, Amersham Biosciences, Piscataway, NJ, Etats-Unis). Trois feuilles de buvards imbibées de tampon de transfert sont placées sur l'électrode inférieure (anode) en évitant les bulles d'air. Ensuite, la membrane PVDF, le gel de polyacrylamide et les feuilles de papier buvard restantes sont empilées les unes sur les autres. Le couvercle de l'appareil contenant la cathode est refermé. Le courant à appliquer est

fonction de la surface des gels à transférer (0,8 $mA.cm^{-2}$) : la surface du gel étant de 182 cm^2, le courant appliqué est de 146 mA (durée du transfert : 90 min).

2.5. Révélation des protéines au bleu de Coomassie R-250

Après la migration électrophorétique, les protéines sont fixées dans le gel par une solution d'acide trichloroacétique à 12% (m/v) pendant 20 min (PAGE-urée) ou 60 min (PAGE-SDS). Ensuite, les gels sont révélés pendant 2 h (PAGE-urée) ou une nuit (PAGE-SDS) dans un mélange constitué de 0,1% (m/v) de bleu de Coomassie R-250, de 2% (m/v) d'acide trichloroacétique et de 50% (v/v) d'éthanol. La décoloration s'effectue par des lavages successifs dans des bains constitués d'un mélange de 30% (v/v) d'éthanol et de 7,5% (v/v) d'acide acétique. Les gels sont ensuite séchés entre 2 feuilles de cellophane dans un four à chaleur tournante (GelAir Dryer, BioRad) et conservés.

2.6. Révélation des protéines sur la membrane de PVDF

Après électrotransfert, la membrane est révélée une nuit dans un mélange contenant 0,1% (m/v) de bleu de Coomassie G-250, 1% (v/v) d'acide acétique et 40% (v/v) de méthanol. Elle est ensuite décolorée dans un mélange constitué de 50% (v/v) de méthanol dans l'eau. Enfin, la membrane est séchée à l'air libre.

2.7. Révélation des glycoprotéines sur la membrane de PVDF

Les glycoprotéines sont révélées par le réactif de Schiff après oxydation par l'acide périodique selon la méthode de Kapitany et Zebrowski (1973). La membrane est tout d'abord lavée pendant 5 min dans 12% (m/v) d'acide trichloroacétique, puis incubée pendant 15 min dans 1% (m/v) d'acide périodique. La membrane est ensuite lavée 3 fois 5 min dans 15% (v/v) d'acide acétique et révélée pendant 30 min dans 50 mL de réactif de Schiff. Elle est finalement lavée par des bains successifs dans 7,5% d'acide acétique pendant 2 h et séchée à l'air libre. Toutes les étapes se déroulent à 4°C et à l'obscurité.

3. Hydrolyses enzymatiques

3.1. Hydrolyse par la chymosine

Les caséines (10 $mg.mL^{-1}$) sont solubilisées dans un tampon d'hydrolyse constitué de 100 mM de phosphate de sodium, pH 6,4 et de 0,02% (m/v) d'azoture de sodium (NaN_3). La chymosine (E.C. 3.4.23.4) de caillette de veau (51 unités enzymatiques par mg d'enzyme ; une unité enzymatique coagule 10 mL de lait bovin par min à 30°C ; Sigma Chemical Co., Saint Louis, MO, Etats-Unis) est ajoutée (2,3 unités enzymatiques par mL) à la solution protéique et le volume réactionnel est incubé à 37°C pendant 24 h. La réaction est stoppée par l'ajout de 3 volumes de tampon Tris-HCl 125 mM, pH 6.8 contenant 2% (m/v) de SDS, 5% (v/v) de 2-mercaptoéthanol, 50% (v/v) de glycérol et 0,01% (m/v) de bleu de bromophénol. Les échantillons sont alors traités thermiquement à 100°C pendant 3 min avant d'être analysés par SDS-PAGE.

3.2. Hydrolyse par la trypsine immobilisée sur bille d'agarose

Les caséines (2 $mg.mL^{-1}$) sont solubilisées dans un tampon 0,1 M Tris-HCl, pH 8,0 contenant 10% (v/v) d'acétonitrile et 0,02% (m/v) d'azoture de sodium. L'enzyme est la trypsine (EC 3.4.21.4) isolée de pancréas bovin, immobilisée sur billes d'agarose et traitée à la L-1-tosylamido-2-phényléthyl chlorométhyl cétone (TPCK) par le fournisseur (Sigma). Les billes, stockées dans une solution à 10 mM d'acide acétique, pH 3,2, contenant 0.01% (m/v) de thimérosal, sont préalablement lavées 3 fois à l'eau ultra-pure par centrifugation à 4000 x *g* pendant 5 min. La trypsine immobilisée (0,14 unité par mg de caséines ; 1 unité hydrolyse 1.0 µmole d'hydrochlorure de Nα-benzoyl L-arginine éthyl ester ou BAEE par min à pH 8,0 et à 30°C) est ajoutée aux caséines solubilisées dans le tampon d'hydrolyse. Le volume réactionnel est mis à incuber sous agitation au bain-marie à 37°C pendant 210 min. A l'issue de l'incubation, le volume réactionnel est centrifugé à 4000 x *g* pendant 5 min ; le surnageant constitue l'hydrolysat qui est congelé à -20°C avant lyophilisation. Les billes sont récupérées, lavées 3 fois par centrifugation à 4000 x *g* pendant 5 min avec de l'eau ultrapure puis stockées à 4°C dans une solution à 10 mM d'acide acétique, pH 3,2 contenant 0.01% (m/v) de thimérosal.

3.3. Hydrolyse par la phosphatase alcaline

La méthode utilisée est dérivée de celle de Yeung *et al.* (2001) utilisée pour la déphosphorylation des caséines bovines. L'échantillon de caséine est solubilisé (3 $mg.mL^{-1}$) dans 800 µL d'un tampon d'hydrolyse 20 mM Tris-HCl, pH 8,5 contenant 1 mM de fluorure phénylméthanesulphonique (PMSF) et 0,02% (m/v) d'azoture de sodium. Un volume de 100 µL d'une solution de déphosphorylation constituée de 1 mM de $MgCl_2$ et de 1 mM de $Zn(CH_3COO)_2$ est ajouté à l'échantillon. Le mélange obtenu est incubé 10 min à 37°C. La phosphatase alcaline (E.C. 3.1.3.1), isolée à partir de muqueuse intestinale bovine (1109 $unités.mg^{-1}$; 1 unité hydrolyse 1,0 µmol de *p*-nitrophenyl phosphate par minute à pH 9,8 et à 37°C ; Sigma), est solubilisée (0,5 $mg.mL^{-1}$) dans la solution de déphosphorylation. La réaction enzymatique est déclenchée par l'ajout de 10 µL de la solution enzymatique dans l'échantillon et se déroule pendant 24 h à 37°C. Pendant la déphosphorylation, le volume réactionnel est dialysé contre le tampon d'hydrolyse pour empêcher une éventuelle rétro-inhibition par les ions phosphates libérés dans le milieu. A l'issue de l'incubation, le volume réactionnel est dialysé contre de l'eau distillé en présence de thymol et lyophilisé. Le lyophilisat est conservé à -20°C.

3.4. Hydrolyse par la phosphatase acide

La technique utilisée dérive de celle de Carles et Ribadeau-Dumas (1986) utilisée pour la caséine β bovine. L'échantillon est solubilisé (3 $mg.mL^{-1}$) dans 950 µL d'un tampon d'hydrolyse constitué de citrate de sodium 50 mM, pH 5,8 contenant 1 mM de PMSF et 0,02% (m/v) d'azoture de sodium. La phosphatase acide issue de pomme de terre (E.C. 3.1.3.2 ; 6,7 $unités.mg^{-1}$; 1 unité hydrolyse 1,0 µmol de *p*-nitrophenyl phosphate par minute à pH 4,8 et à 37°C ; Sigma) est solubilisée (0,2 $mg.mL^{-1}$) dans un tampon sulfate d'ammonium 0,24 M, pH 6,0. La réaction est déclenchée par l'ajout de 8 µL de solution enzymatique dans l'échantillon et se déroule à 30°C pendant 24 h. Le volume réactionnel est dialysé contre du tampon d'hydrolyse tout au long de la réaction pour éviter une rétro-inhibition éventuelle de l'enzyme par les ions phosphates. Après incubation, le volume réactionnel est dialysé contre de l'eau distillée en présence de thymol puis lyophilisé. Le lyophilisat est conservé à -20°C.

3.5. Hydrolyse par la neuraminidase

La méthode utilisée dérive de celle décrite par Coddeville *et al.* (1989). La neuraminidase type V (E.C. 3.2.1.18 ; 1 unité libère 1.0 µmole d'acide N-acétylneuraminique par min à pH 5,0 et à 37°C avec, comme substrat, la mucine sous-maxillaire bovine) est issue de *Clostridium perfringens* (Sigma). L'échantillon (5 mg) est solubilisé dans 500 µL de tampon d'hydrolyse constitué de phosphate de sodium 0,5 M, pH 7,0 et de 0,02% (m/v) d'azoture de sodium. Un volume de 50 µL de neuraminidase est ajouté à la solution et la réaction se déroule à 37°C pendant 24 h. Le volume réactionnel est ensuite dialysé contre de l'eau distillée en présence de thymol puis lyophilisé. Le lyophilisat est conservé à -20°C.

4. Chromatographie liquide

4.1. Chromatographie FPLC (*fast protein liquid chromatography*) d'interactions hydrophobes

Les caséines du lait de jument sont fractionnées par chromatographie d'interactions hydrophobes avec une colonne semi-préparative TSK Phenyl 5-PW (21,5 x 150 mm ; taille des particules 13 µm ; Interchim, Montluçon, France) connecté à un système FPLC (ÄKTA-FPLC, Amersham Pharmacia). La méthode utilisée est celle décrite par Tauzin *et al.* (2003) pour fractionner les caséines bovines. Les caséines sont solubilisées (10 $mg.mL^{-1}$) dans un tampon dihydrogénophosphate de sodium à 480 mM, pH 6,4 contenant 2,5 M d'urée. Un volume de 10 mL est injecté sur la colonne préalablement équilibrée avec ce même tampon. Un gradient linéaire décroissant de 480–37 mM de dihydrogénophosphate de sodium est appliqué pendant 47 min, suivi d'une étape isocratique à 37 mM de dihydrogénophosphate de sodium de 12 min. Le débit appliqué est de 6 $mL.min^{-1}$ et l'absorbance est mesurée à 280 nm. Les fractions collectées sont dialysées contre de l'eau distillée en présence de thymol pendant 48 h puis lyophilisées. Les lyophilisats sont conservés à -20°C.

4.2. Chromatographie FPLC d'échange d'anions

La méhode utilisée est celle décrite par Aoki *et al.* (1992) pour séparer les isoformes de phosphorylation de la caséine β humaine. Les caséines sont fractionnées sur une colonne TSK DEAE 5-PW semi-préparative (21,5x150 mm ; taille de particule 10 µm ; Amersham Pharmacia) connectée au système ÄKTA-FPLC. L'échantillon est solubilisé (6 $mg.mL^{-1}$) dans un tampon imidazole à 0,02 M, pH 8,0 contenant 3,3 M d'urée et 0,08 M de NaCl. Un volume de 10 mL est alors injecté sur la colonne préalablement équilibrée avec ce même tampon. Un gradient linéaire croissant de 0,08–0,28 M de NaCl à pH 8,0 est réalisé pendant 80 min, suivi d'une étape isocratique à 0,28 M de NaCl de 15 min. Le débit est de 4 $mL.min^{-1}$ et l'absorbance est mesurée à 280 nm. Les fractions collectées sont dialysées contre de l'eau distillée en présence de thymol pendant 48 h puis lyophilisées.

Lorsqu'une colonne analytique TSK DEAE 5PW (7,5 x 75 mm ; taille de particule 10 µm, Amersham Pharmacia) est utilisée, l'échantillon est solubilisé à 15 $mg.mL^{-1}$ et un volume de 200 µL est injecté sur la colonne. Un débit de 0,5 $mL.min^{-1}$ est appliqué. Un gradient linéaire croissant de 0,08–0,28 M de NaCl à pH 8,0 est réalisé pendant 70 min, suivi d'une étape isocratique à 0,28 M de NaCl de 13 min. Les fractions collectées sont dialysées contre de l'eau distillée en présence de thymol pendant 48 h puis ensuite lyophilisées. Les lyophilisats sont conservés à -20°C.

4.3. Chromatographie d'affinité à la lectine de germe de blé (*wheat germ agglutinin* ou WGA)

La méthode utilisée est celle décrite par Egito *et al.* (2001). Les caséines (20 mg) sont solubilisées dans 10 mL de tampon Tris-HCl 20 mM, pH 8,0 contenant 0,02% (m/v) d'azoture de sodium et fractionnées dans une colonne (10 x 1 cm) contenant un gel d'agarose sur lequel la lectine WGA a été immobilisée sans être modifiée chimiquement (WGA) ou immobilisée après succinylation (sucWGA ; Vector Laboratories, Burlingame, CA, Etats-Unis). Un débit de 10 $mL.h^{-1}$ est appliqué. La fraction non retenue dans la colonne (fraction WGA^- ou $sucWGA^-$) est éluée avec du tampon Tris-HCl 20 mM, pH 8,0 contenant 1 M de NaCl et 0.02% (m/v) d'azoture de sodium. La fraction retenue dans la colonne (fraction WGA^+ ou $sucWGA^+$) est éluée avec du tampon Tris-HCl 20 mM, pH 8,0 contenant 0,1 M de N-acétyl-α-D-glucosamine et 0,02% (m/v) d'azoture de sodium. Les deux fractions collectées

sont dialysées contre de l'eau distillée en présence de thymol et lyophilisées. Les lyophilisats sont conservés à -20°C.

4.4. Chromatographie MOAC (*Metal Oxide Affinity Chromatography*) au dioxyde de titane immobilisé sur microcolonne

Les microcolonnes (UptiTip ^{TM} Packed Titanium Dioxide, Interchim) sont des cônes de 200 µL adaptables sur une micropipette automatique et remplis d'une phase de silice greffée de dioxyde de titane (TiO_2) ; elles permettent un enrichissement en phosphopeptides. L'échantillon (hydrolysat trypsique de caséine) est solubilisé (6 $mg.mL^{-1}$) dans un mélange de 80% (v/v) d'acétonitrile (ACN ; Rathburn, Walkerburn, Grande-Bretagne) et de 0,1% (v/v) d'acide trifluoroacétique (TFA ; Sigma). Un volume de 3 mL est injecté sur la microcolonne préalablement équilibrée avec le mélange ACN/TFA. L'éluat est collecté et constitue la fraction « non liée TiO_2 ». Un volume de 200 µL d'une solution à 50% (v/v) d'ACN et un volume de 200 µL d'un mélange de 80% (v/v) ACN et de 0,1% (v/v) de TFA sont injectés successivement sur la microcolonne. L'éluat est collecté et constitue la fraction « lavage TiO_2 ». L'élution des phosphopeptides est réalisée grâce à l'injection de 700 µL d'un tampon hydroxyde d'ammonium à 0,25% (v/v), pH 10,5. L'éluat collecté constitue la fraction « liée TiO_2 » Les 3 fractions collectées sont lyophilisées. Les lyophilisats sont conservés à -20°C.

4.5. Chromatographie liquide haute performance en phase inversée (RP-HPLC) sur colonne C_{18} semi-préparative

Les caséines sont séparées avec une colonne semi-préparative Uptisphère C_{18} (250 x 10 mm ; taille de particules 5 µm ; porosité 30 nm ; Interchim) reliée à un système HPLC Alliance 2690 (Waters, Milford, MA, Etats-Unis). Les caséines sont solubilisées (5 $mg.mL^{-1}$) dans une solution à 30% (v/v) d'ACN contenant 0,1% (v/v) de TFA. Le débit est de 3 $mL.min^{-1}$ et le volume d'échantillon injecté sur la colonne est de 500 µL. Un gradient linéaire de 30–42% (v/v) d'ACN est appliqué en présence de 0,1% (v/v) de TFA pendant 50 min. La détection des protéines est réalisée à 215 nm avec un détecteur à barrette de diodes 996 (Waters) relié en ligne à la chaîne HPLC. Les résultats sont exploités avec le logiciel Millenium 32, version 3.05.01 (Waters) selon la méthode de Miclo *et al*. (1995) et Perrin *et*

al. (1995) en ce qui concerne l'analyse des spectres UV. Les fractions collectées sont directement lyophilisées. Les lyophilisats sont conservés à -20°C.

4.6. Chromatographie RP-HPLC sur colonne C_{18} analytique

La colonne utilisée est une colonne Lichrospher C_{18} (150 x 2 mm, taille de particules 5µm, Cluzeau Info Labo) reliée au système HPLC Alliance 2690. Les échantillons (fractions peptidiques à 4 $mg.mL^{-1}$) sont solubilisés dans une solution à 5% (v/v) d'ACN contenant 0,1% (v/v) de TFA. Le débit est de 250 $\mu L.min^{-1}$ et le volume d'échantillon injecté sur la colonne est de 500 µL. Une étape isocratique à 5% (v/v) d'ACN en présence de 0,1% (v/v) de TFA est appliquée pendant 5 min puis un gradient linéaire de 5–50% d'ACN en présence de 0,1% (v/v) de TFA est appliqué pendant 90 min. Les fractions collectées sont directement lyophilisées. Les lyophilisats sont conservés à -20°C.

4.7. Chromatographie RP-HPLC sur colonne C_4

La colonne utilisée est une colonne analytique Nucleosil C_4 (150 x 2 mm ; taille de particules 5 µm ; Cluzeau Info Labo, Sainte Foy la Grande, France) reliée au système HPLC Alliance 2690. Les échantillons (fractions peptidiques à 4 $mg.mL^{-1}$) sont solubilisés dans une solution à 4% (v/v) d'ACN contenant 0,1% (v/v) de TFA. Le débit est de 250 $\mu L.min^{-1}$ et le volume d'échantillon injecté sur la colonne est de 500 µL. Un gradient linéaire de 4–48% d'ACN en présence de 0,1% (v/v) de TFA est appliqué pendant 15 min, suivi d'une étape isocratique à 64% (v/v) d'ACN en présence de 0,1% (v/v) de TFA de 5 min. Les fractions collectées sont directement lyophilisées. Les lyophilisats sont conservés à -20°C.

5. Analyse des protéines et des peptides par spectrométrie de masse

Les analyses de spectrométrie de masse ont été réalisées au Laboratoire de Sciences et Technologie du Lait et de l'œuf (STLO, INRA de Rennes) dans le cadre d'une collaboration.

5.1 Détermination des masses moléculaires

Les masses moléculaires des protéines et peptides sont déterminées par chromatographie liquide couplée à la spectrométrie de masse à ionisation électrospray (LC/ESI-MS) grâce à un spectromètre de masse tandem à triple quadripôle équipé d'une source d'ionisation électrospray (API III$^+$, MDS SCIEX, Toronto, Canada) ou grâce à un spectromètre hybride quadrupôle/temps de vol (Q/Star XL ; MDS SCIEX).

Un couplage LC/ESI-MS est tout d'abord réalisé sur une chaîne HPLC Hewlett-Packard 1100 (Agilent Technologies, Massy, France) couplée avec le spectromètre API III$^+$. Les échantillons sont séparés puis ionisés dans la source ESI avec une tension optimisée aux alentours de 4800 V. Cette première étape permet de mesurer des masses, mais avec une précision assez faible liée au quadripôle (± 0,01%). Les fractions majoritaires sont collectées puis sont analysées avec le spectromètre Q/Star XL grâce à une infusion par nanopointe ESI (NanoES spray capillaries, Proxéon) pour obtenir avec plus de précision (± 0,001%) les masses déterminées précédemment avec l'API III$^+$. Pour mémoire, la résolution obtenue avec un analyseur quadripolaire est de 1000 au m/z 1000 alors que celle d'un analyseur temps de vol est de 10000 au même m/z.

Les masses moléculaires théoriques sont calculées à partir du logiciel ProtParam Tool disponible sur l'adresse internet http://expasy.org/tools/protparam.html (Gasteiger *et al.,* 2005).

5.2 Séquençage peptidique

La détermination des sites de phosphorylation de la caséine β a été réalisée à partir de deux manipulations successives : i) les échantillons des différents CPPs de la CN β ont été analysés par couplage LC/ESI-MS avec la chaîne HPLC Hewlett-Packard 1100 sur colonne Vydac C4 (218 TP5215) couplée au spectromètre API III+ afin d'obtenir les masses individuelles contenues dans chaque échantillon, les CPPs identifiés ont été collectés et soumis à une hydrolyse par l'endoprotéinase GluC (ou V8 ; E.C. 3.4.21.19) pour avoir des peptides de taille comprise entre 500 et 2000 Da compatible avec une bonne fragmentation MS/MS, ii) les hydrolysats V8 ont été analysés en LC/ESI-MS sur colonne symmetry C18 (5µm, 2,1 x 150 mm), les différents peptides ont été collectés puis infusés en nanopointe sur

le spectromètre Q/Star XL pour être analysés en mode nanospray par ESI-MS/MS afin de déterminer la séquence des peptides et la position des sites phosphorylés.

Pour l'hydrolyse par la V8, les fractions collectées sont séchées au moyen d'un concentrateur-évaporateur Speed-Vac ; Jouan, Saint Herblain, France) puis solubilisées dans un tampon Tris-HCl 50 mM, pH 7,8 contenant du bicarbonate d'ammonium 20 mM. L'enzyme est ajoutée (21 $\mu g.mL^{-1}$) à la solution et l'hydrolyse est réalisée pendant 20 h à 30°C. La réaction est stoppée avec 10 µL de TFA 5% (v/v).

PARTIE 1

Caractérisation structurale de la caséine α_{s1} équine : variants post-transcriptionnels et post-traductionnels

La CN-α_{s1} équine présente un profil électrophorétique complexe qui reflèterait une grande hétérogénéité de cette protéine. Cette hétérogénéité pourrait être due à la présence de variants de phosphorylation et de variants d'épissage. En effet, la CN-α_{s1} équine possède 10 sites potentiels (Figure 25) de phosphorylation et cette phosphorylation pourrait être variable comme dans le cas de la CN-β équine qui possède de 3P à 7P (Girardet *et al.,* 2006).

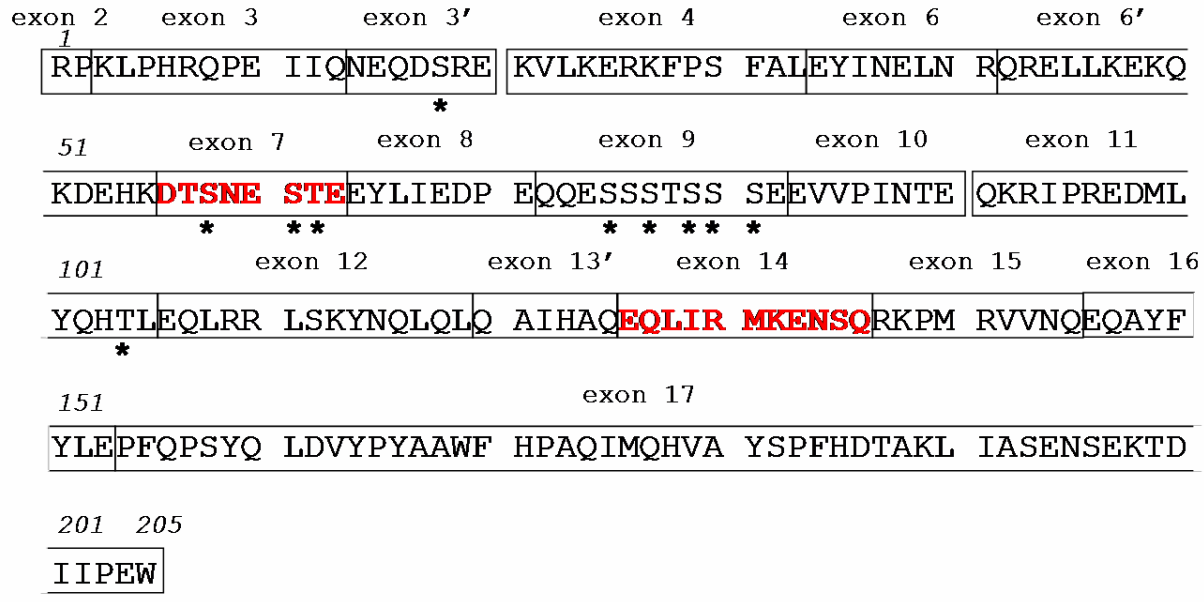


Figure 25 : Structure exonique de la CN-α_{s1} complète déduite à partir de la séquence d'ARNm isolée à partir de cellules de la glande mammaire chez le poney gallois (numéro d'accession GenBank AY049939 ; Milenkovic *et al.,* 2002). La nomenclature des exons est celle utilisée par Miranda *et al.* (2004). Les sites potentiels de phosphorylation sont indiqués à l'aide d'un astérisque. Les exons pouvant être alternativement épissés sont indiqués en rouge.

Le niveau de phosphorylation de la CN-α_{s1} équine n'a pour le moment jamais été déterminé. Au niveau transcriptionnel, il a été montré que les exons 7 et 14 étaient la cible de phénomènes d'épissage alternatif. Chez les juments hanovriennes (race de selle allemande), un ARNm exempt de l'exon 7 (correspondant alors à un variant d'épissage appelé CN-$\alpha_{s1}^{\Delta 7}$) a été isolé à partir de cellules de glande mammaire (Lenasi *et al.,* 2003). Chez le poney gallois, un ARNm sans l'exon 14 (correspondant à la CN-$\alpha_{s1}^{\Delta 14}$) a été isolé (Milenkovic *et al.,* 2002) et la présence de protéines CN-$\alpha_{s1}^{\Delta 7}$ et de CN-$\alpha_{s1}^{\Delta 7\text{-}14}$ dans le lait est fortement suggéré (Miranda *et al.,* 2004). Cependant, la protéine complète, c'est-à-dire contenant tous les exons potentiels, n'a pas été pour l'heure caractérisée.

Le but de ce travail est de déterminer les degrés de phosphorylation ainsi que d'identifier les éventuels variants d'épissage de la CN-α_{s1} qui seraient présents dans le lait de juments Haflinger. L'identification de tous les variants de la CN-α_{s1} équine est indispensable notamment pour des travaux ultérieurs concernant l'étude du mécanisme de chélation de minéraux d'intérêt nutritionnel par les CPP. En effet, pour étudier l'influence du degré et de la séquence de phosphorylation des CPP sur la chélation, la protéine d'origine doit être parfaitement caractérisée.

1. Fractionnement du caséinate de sodium équin

1.1. Chromatographie HPLC en phase inversée

Le caséinate de sodium préparé à partir de lait cru congelé est fractionné par HPLC en phase inversée sur une colonne semi-préparative C_{18} grâce à une méthode mise au point au laboratoire lors de précédents travaux (Girardet *et al.,* 2006). Le fractionnement permet l'obtention de trois fractions appelées A, B et C (Figure 26).

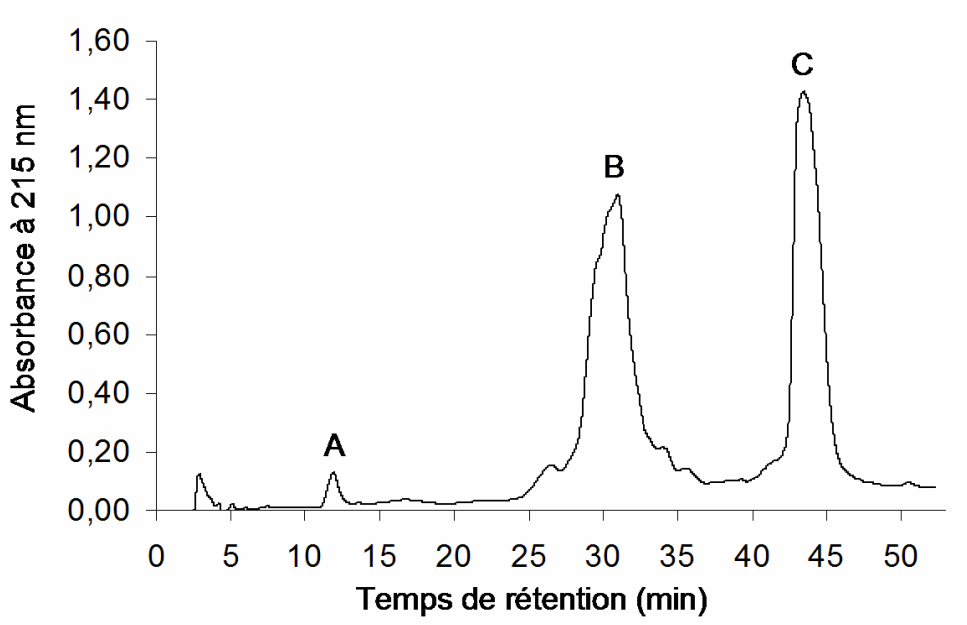


Figure 26 : Séparation du caséinate de sodium par HPLC en phase inversée sur colonne semi-préparative C_{18}. Une quantité de 2,5 mg de protéines a été déposée sur la colonne. Un gradient linéaire de 30 à 42% d'acétonitrile en présence de 0,1% de TFA a été appliqué pendant 50 min à un débit de 3 mL.min^{-1}. A, B et C : fractions collectées.

Le contenu de chacune des fractions est analysé par PAGE-urée (Figure 27). Ainsi, la fraction A contient essentiellement de la low-M_r CN-β, la fraction B est un mélange de CN-α_{s1} et de CN-κ (dont la nature glycosylée est révélée au réactif de Schiff ; Figure 27b) et la fraction C contient la CN-β, l'identification des différentes caséines étant réalisée d'après des résultats antérieurs obtenus au laboratoire (Egito *et al.,* 2002 ; Girardet *et al.,* 2006 ; Miclo *et al.,* 2007).

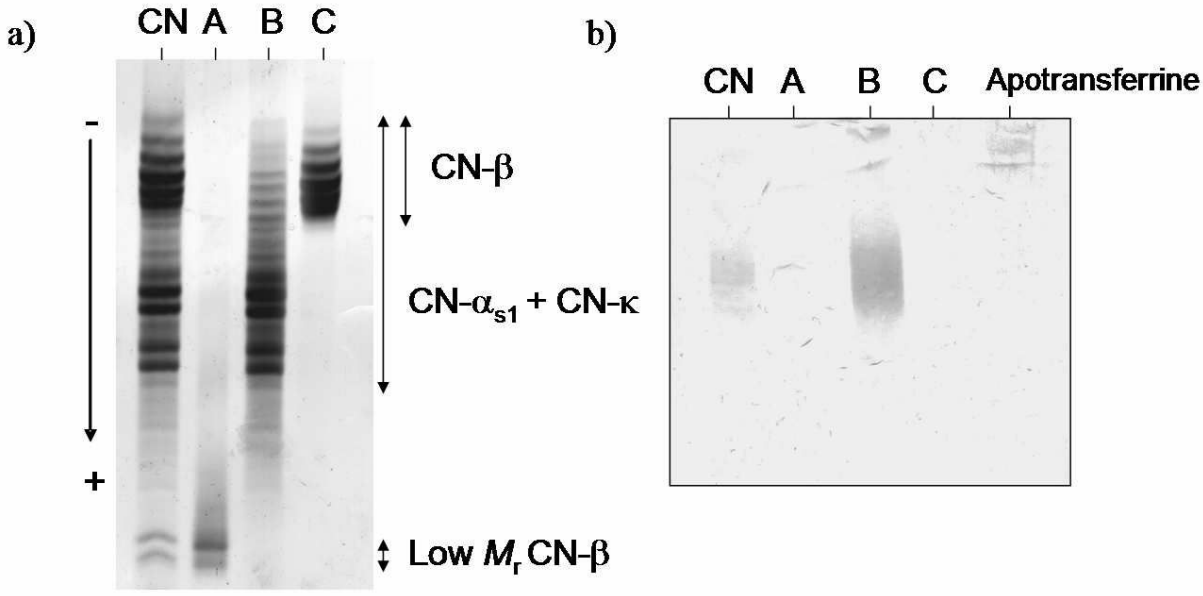


Figure 27 : Caractérisation électrophorétique (PAGE-urée) des fractions A, B et C purifiées par HPLC en phase inversée. Des volumes de 20 µL a) ou 40 µL b), correspondant à 40 µg et 80 µg de protéines respectivement, ont été déposés dans chaque puits. a) La coloration est réalisée au bleu de Coomassie. b) Les protéines sont électrotransférées sur une membrane PVDF. La membrane est ensuite révélée avec le réactif de Schiff après oxydation par 1% d'acide periodique. De l'apotransferrine a été déposée dans le gel en tant que témoin positif dans la révélation au réactif de Schiff. CN, caséinate de sodium.

Bien que la méthode utilisée soit qualifiée de semi-préparative, il est difficile d'obtenir une quantité suffisante de caséines fractionnées, notamment du mélange CN-α_{s1}/CN-κ, pour les études menées dans cette partie (2,5 mg de caséinate de sodium déposé par fractionnement HPLC). Il a donc semblé nécessaire de s'orienter vers une autre méthode de fractionnement du caséinate de sodium permettant d'obtenir de plus grandes quantités de CN-α_{s1}.

1.2. Chromatographie FPLC d'interactions hydrophobes

Les caséines équines sont séparées sur une colonne semi-préparative TSK-Phenyl 5-PW. Le profil chromatographique obtenu montre l'existence de quatre fractions nommées F1 à F4 (Figure 28). Ces fractions sont analysées par PAGE-urée (Figure 29), les bandes principales obtenues étant numérotées de 1 à 14 selon la nomenclature proposée par Egito *et al.* (2002). Ainsi, l'électrophorégramme montre que la fraction F1 contient deux bandes majeures (bandes 13 et 14), qui correspondent à des variants isoélectriques de la low-M_r CN-β (Miclo *et al.*, 2007), tandis que la fraction F2 contient les variants de phosphorylation de la CN-β dans les bandes majoritaires 1 à 5 (Egito *et al.*, 2002 ; Girardet *et al.*, 2006).

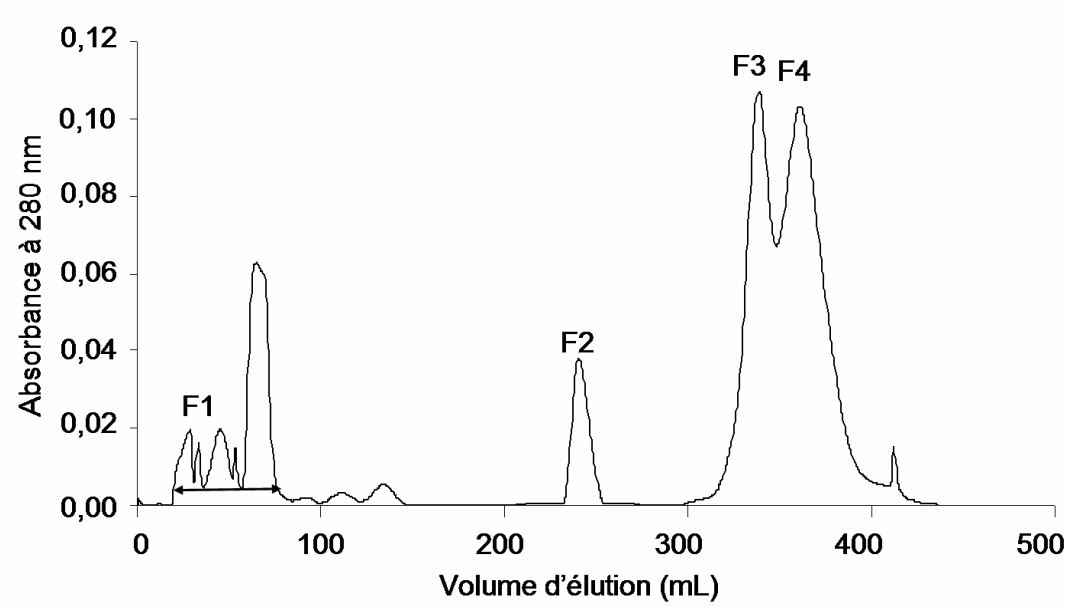


Figure 28 : Séparation du caséinate de sodium par FPLC d'interactions hydrophobes sur colonne semi-préparative TSK Phenyl 5-PW. Une quantité protéique de 100 mg de caséinate de sodium a été déposée sur la colonne. Un gradient linéaire décroissant de 480–37 mM de dihydrogénophosphate de sodium, pH 6,4, est appliqué pendant 47 min, suivi d'une étape isocratique à 37 mM de dihydrogénophosphate de sodium pH 6,4 de 12 min. Le débit appliqué est de 6 mL.min^{-1} et l'absorbance est mesurée à 280 nm.

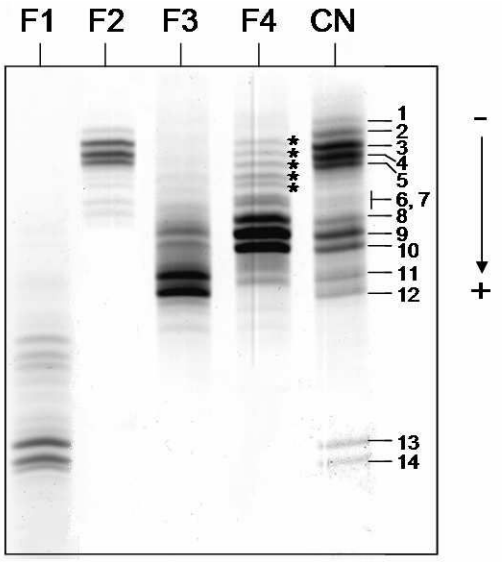


Figure 29 : Caractérisation électrophorétique (PAGE-urée) des fractions F1 à F4 purifiées par FPLC d'interactions hydrophobes. Une quantité de 40 µg de protéines est déposée dans chaque puits. Les protéines sont révélées au bleu de Coomassie. Les bandes sont numérotées de 1 à 14 selon la nomenclature établie par Egito *et al.* (2002). Les astérisques indiquent les bandes mineures de la fraction F4. CN, caséinate de sodium.

Egito *et al.* (2002) ont montré par microséquençage chimique que les bandes 6 à 12 correspondent à des variants isoélectriques de la CN-α_{s1}. Dans le présent travail, les bandes 11 et 12 sont les bandes principalement contenues dans la fraction F3, tandis que les bandes 8 à 10 sont majoritairement présentes dans la fraction F4 (Figure 29). La CN-α_{s1} est donc présente dans deux fractions, dont la séparation chromatographique est toutefois faiblement résolue. Sur le profil électrophorétique de la fraction F4 (Figure 29), cinq bandes mineures, se trouvant dans une zone délimitée par les bandes 2 et 6, migrent au même niveau que les variants isoélectriques de la CN-κ identifiés par Egito *et al.* (2002) comme s'intercalant entre les bandes de CN-β 1 à 5. Après séparation par PAGE-SDS du contenu des fractions F1 à F4 et électrotransfert sur membrane PVDF, la réaction au réactif de Schiff effectuée montre que des protéines glycosylées se concentrent uniquement dans la fraction F4 (Figure 30). La CN-κ équine est totalement glycosylée (Egito *et al.,* 2001), contrairement à son homologue bovin, et semble donc être entièrement désorbée de la colonne d'interactions hydrophobes lors de la collecte de la fraction F4, contaminant ainsi une partie des variants de la CN-α_{s1}.

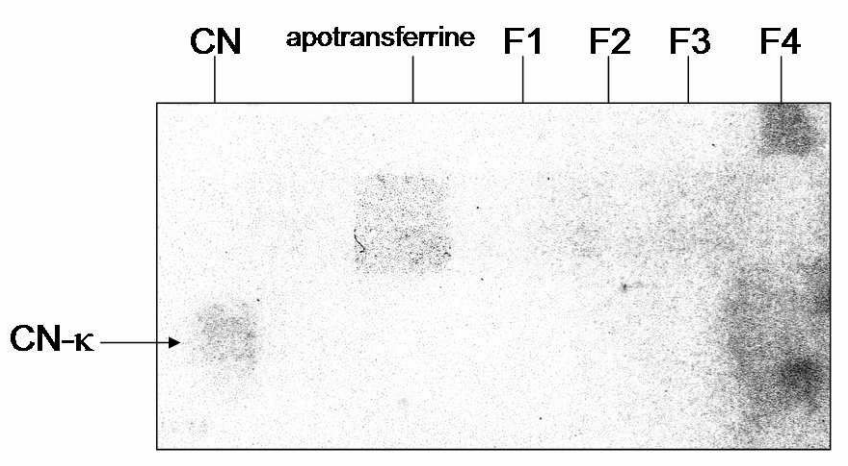


Figure 30 : Caractérisation électrophorétique (PAGE-SDS) des fractions F1 à F4 purifiées par FPLC d'interactions hydrophobes. Une quantité protéique de 80 µg a été déposé par puits. Les protéines glycosylées sont révélées dans le gel par le réactif de Schiff après oxydation par 1% d'acide periodique. CN, caséinate de sodium.

2. Séparation de la CN-κ et de la CN-α_{s1} par chromatographie d'affinité à la lectine de germe de blé (WGA) immobilisée

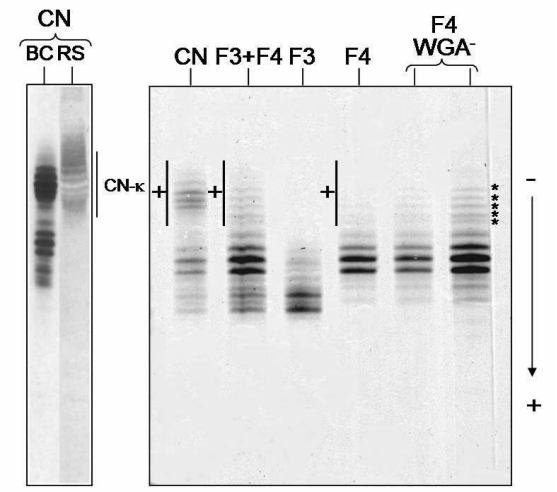
La caséine α_{s1} n'étant pas glycosylée et la caséine κ étant par contre O-glycosylée, la chromatographie d'affinité à la lectine de *Triticum vulgaris* (WGA ; *wheat germ agglutinin*) a été choisie afin de séparer la CN-κ, contenue dans la fraction F4, de la CN-α_{s1}.

Les lectines sont des (glyco)protéines d'origine végétale, animale ou bactérienne reconnaissant spécifiquement des structures glycanniques, mais elles ne sont ni des enzymes, ni des anticorps. Elles possèdent deux sites de fixation aux glycannes et sont donc capables de précipiter les glycoconjugués et d'agglutiner les cellules comme le font les immunoglobulines. Lors de la reconnaissance structurale, les chaînes glycanniques des glycoprotéines interagissent *via* des liaisons hydrogène avec les chaînes polypeptidiques des lectines (Bhattacharyva et Brewer, 1988). Certaines lectines sont souvent, et depuis longtemps, utilisées comme ligands pour purifier des glycoconjugués par chromatographie d'affinité (Dulaney, 1979).

La WGA présente une forte affinité pour les groupes structuraux GlcNAcβ(1-4)GlcNAc des chaînes glycanniques ainsi que pour les résidus NeuAc (Wu et Herp, 1985). Egito *et al.* (2001) ont utilisé la chromatographie d'affinité à la WGA immobilisée pour purifier la CN-κ équine. Cette dernière est retenue en totalité dans la colonne d'affinité,

montrant ainsi qu'elle est toujours glycosylée dans le lait de jument Haflinger (Egito *et al.*, 2001). Dans notre étude, l'utilisation de la chromatographie d'affinité à la WGA a permis de séparer, après ajout du compétiteur méthyl-α-D-Glc, la CN-α_{s1}, retrouvée dans la fraction non affine appelée WGA^-, de la CN-κ éluée dans la fraction affine (notée WGA^+). En effet, aucun composé de la fraction WGA^- ne réagit avec le réactif de Schiff après séparation par PAGE-urée et électrotransfert, alors que les variants de glycosylation de la CN-κ sont mis en évidence dans le caséinate de sodium, le mélange des fractions F3+F4 et la fraction F4 (Figure 31a). La fraction WGA^+ contient, quant à elle, les bandes glycosylées de CN-κ (résultat non montré). Afin de conforter ce résultat (la réaction au réactif de Schiff présente l'inconvénient d'avoir une faible sensibilité), les différentes fractions d'intérêt ont été hydrolysées par la chymosine, ceci afin de mettre en évidence par SDS-PAGE la présence ou l'absence de paraCN-κ (fragment 1-97) dans les hydrolysats obtenus, et, par conséquent, la présence ou l'absence de CN-κ dans les fractions avant hydrolyse. La paraCN-κ est localisée par PAGE-SDS à une masse moléculaire apparente de 16 kDa (Egito *et al.*, 2001). Elle est présente dans les hydrolysats de caséinate de sodium, de la fraction F4, du mélange de fractions F3+F4 ainsi que dans l'hydrolysat de la fraction WGA^+, alors que la CN-κ est absente dans les hydrolysats des fractions F3 et WGA^- (Figure 31b). Ceci confirme le précédent résultat obtenu par PAGE-urée, à savoir que la CN-κ est bien affine pour la WGA. Cependant, on note que les cinq bandes mineures migrant à hauteur de la CN-κ sont toujours présentes dans la fraction WGA^- alors qu'aucune réaction positive au réactif de Schiff n'a été observée (Figure 31a). Ces bandes pourraient être des variants isoélectriques de la CN-α_{s1}, voire de la CN-α_{s2}, mais ce dernier composant n'a pour l'heure pas été mis en évidence dans le lait équin (même si certaines séquences polypeptidiques rapportées semblent attester de sa présence ; Ochirkhuyag *et al.,* 2000 ; Egito *et al.,* 2001 ; Miranda *et al.,* 2004).

a)



b)

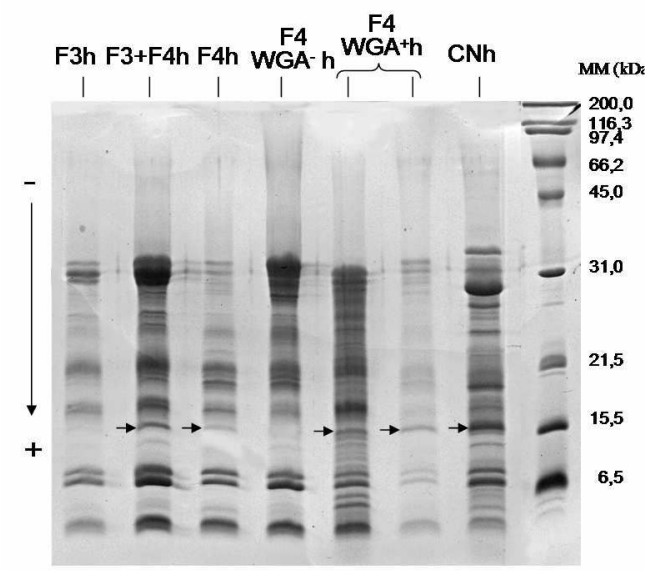
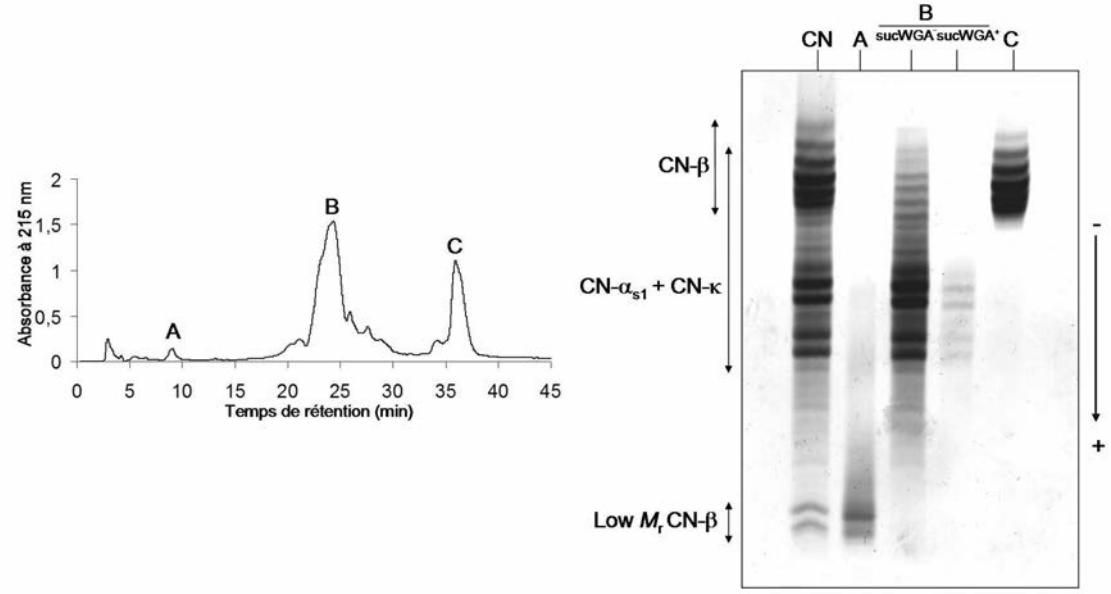


Figure 31 : Détection électrophorétique de la CN-κ dans les fractions chromatographiques collectées par FPLC d'interactions hydrophobes et par chromatographie d'affinité à la WGA.
a) PAGE-urée des fractions F3 et F4 seules ou en mélange obtenues par chromatographie d'interactions hydrophobes et de la fraction WGA$^-$ non retenue par chromatographie d'affinité à la WGA de F4. Une quantité protéique de 40 µg a été déposée dans le gel. Les protéines sont révélées dans le gel au bleu de Coomassie. Les glycoprotéines sont révélées sur membrane PVDF après électrotransfert par le réactif de Schiff après oxydation par 1% d'acide periodique et sont repérées par le symbole « + ». Les astérisques indiquent les bandes mineures de la fraction F4.
b) hydrolyse par la chymosine des fractions F3 et F4 seules ou en mélange obtenues par chromatographie d'interactions hydrophobes et des fractions WGA$^-$ et WGA$^+$ issues de la chromatographie d'affinité à la WGA de F4. Les hydrolysats sont analysés par PAGE-SDS. Une quantité protéique de 40 µg a été déposée dans le gel. Les protéines sont révélées dans le gel par le bleu de Coomassie. La flèche représente la localisation de la paraCN-κ.
BC, bleu de coomassie ; RS, réactif de Schiff ; CN, caséinate de sodium ; h, hydrolyse par la chymosine ; MM, masse moléculaire ; WGA, wheat germ agglutinin ; WGA$^-$ et WGA$^+$, fractions non retenue et retenue par la lectine de germe de blé.

L'utilisation en chromatographie d'affinité de WGA modifiée chimiquement par succinylation (sucWGA) permet de nous renseigner quant à la nature de certains résidus glycanniques. La sucWGA est beaucoup plus spécifique des résidus GlcNAc que la WGA, dans la mesure où elle n'a plus d'affinité pour les résidus NeuAc. Dans une première expérience, après séparation HPLC en phase inversée de caséinate de sodium en trois fractions A, B et C, la fraction B contenant le mélange de CN-α_{s1} et de CN-κ est ensuite séparée par chromatographie d'affinité à la sucWGA immobilisée (Figure 32a). Il est obtenu une fraction non affine (fraction sucWGA$^-$) et une fraction affine (fraction sucWGA$^+$). L'analyse par électrophorèse PAGE-urée et coloration au bleu de Coomassie montre que la fraction sucWGA$^+$ ne semble pas contenir de CN-κ, seules des bandes peu intenses sont visibles et sembleraient appartenir à la CN-α_{s1}. Dans une seconde expérience, le caséinate de sodium est fractionné dans un premier temps par chromatographie d'affinité à la sucWGA immobilisée puis la fraction sucWGA$^-$ non affine obtenue est finalement séparée par HPLC en phase inversée en deux sous-fractions, notées I et II (Figure 32b). Après migration par PAGE-urée et électrotransfert sur membrane de ces sous-fractions, la CN-κ a été mise en évidence par le réactif de Schiff dans la sous-fraction I issue de la fraction sucWGA$^-$ (Figure 32b). La fraction sucWGA$^+$, quant à elle, ne semble pas contenir de CN-κ (aucune bande détectée par le réactif de Schiff ; résultat non montré). La CN-κ ne se lie donc pas de façon significative à la lectine de blé succinylée. Ainsi, l'absence d'affinité de la CN-κ pour la sucWGA et le fait qu'elle soit affine pour la WGA montrent (i) que la CN-κ est sialylée et (ii) que la CN-κ ne possède pas de résidu GlcNAc ou alors que les résidus NeuAc empêchent l'interaction entre les éventuels résidus GlcNAc et la lectine de blé succinylée par répulsion électrostatique ou par encombrement stérique. Afin d'éclairer ce dernier point, du caséinate de sodium a été traité par la neuraminidase, l'hydrolyse des résidus NeuAc de la CN-κ permettant ainsi de rendre les éventuels résidus GlcNAc plus accessibles à la lectine succinylée. Cependant, après PAGE-urée et révélation au réactif de Schiff, la CN-κ désialylée a été retrouvée dans la fraction sucWGA$^-$ (résultat non montré), ce qui suggère que la CN-κ équine ne contiendrait pas de résidu GlcNAc. Ses structures glycanniques sembleraient donc plus proches de celles de la CN-κ bovine (qui est sialylée et qui ne contient pas de résidus GlcNAc) que de celles de la CN-κ humaine. La protéine humaine contient, entre autres, une forte proportion de résidus GlcNAc (17 résidus de GlcNAc par molécule de glycomacropeptide humain selon Fiat *et al.*, 1980 ; 18% de la masse de CN-κ selon Dev *et al.*, 1993) mais ses structures glycanniques ne sont pas encore élucidées à notre connaissance.

a)



b)

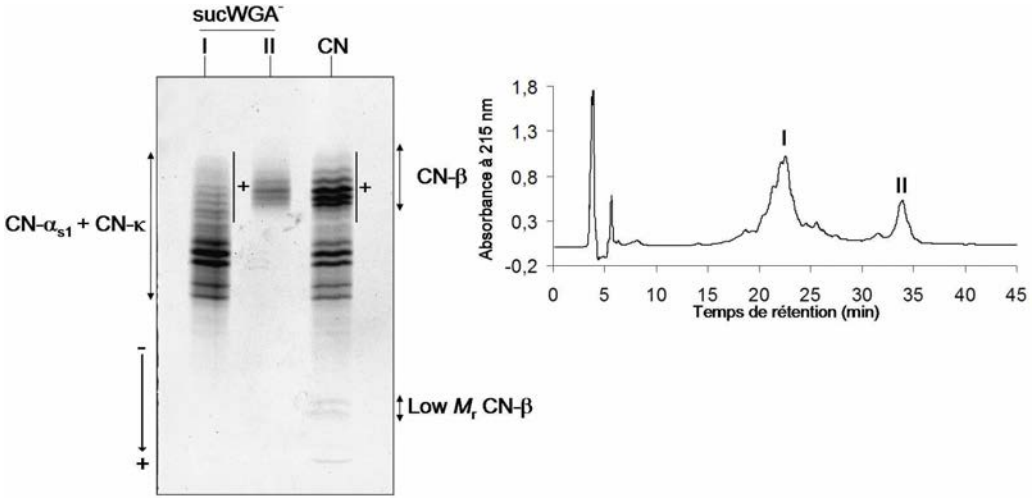


Figure 32 : Caractérisation par PAGE-urée des fractions obtenues par chromatographie d'affinité à la WGA succinylée (sucWGA).
a) le caséinate de sodium a été séparé par HPLC en phase inversée sur colonne C_{18} semi-préparative. Une quantité de 2,5 mg de protéines a été déposée sur la colonne. Un gradient linéaire de 30 à 42% d'acétonitrile en présence de 0,1% de TFA a été appliqué pendant 50 min à un débit de 3 mL.min^{-1}. La fraction B obtenue par HPLC a été séparée en deux fractions non retenue (sucWGA$^-$) et retenue (sucWGA$^+$) par chromatographie d'affinité à la sucWGA. Des quantités de 40 µg de protéines sont déposéees dans le gel. Les protéines sont révélées dans le gel par le bleu de Coomassie.
b) le caséinate de sodium a été séparé par chromatographie d'affinité à la sucWGA puis la fraction sucWGA$^-$ obtenue a été séparée par HPLC en phase inversée comme décrit ci-dessus. Des quantités de 40 µg de protéines sont déposéees dans le gel. Les protéines sont révélées dans le gel par le bleu de Coomassie. Les glycoprotéines électrotransférées sur membrane sont révélées par le réactif de Schiff après oxydation par 1% d'acide periodique et sont repérées par le symbole « + ».
CN, caséinate de sodium ; I et II sous-fractions obtenues par HPLC en phase inversée de la fraction suc WGA$^-$.

3. Déphosphorylation de la CN-α_{s1} et caractérisation des variants d'épissage

Les fractions F3 et F4 contenant principalement la CN-α_{s1} sont analysées par électrophorèse bidimensionnelle (2-DE ; Figure 33). Le profil électrophorétique de la fraction F3 montre de nombreux spots répartis en quatre groupes de masse moléculaire apparente d'environ 30, 33, 37 et 39 kDa. Pour chaque masse moléculaire apparente, il est observé au moins sept spots ayant des pHi allant de 4,85 à 6,02. Dans le cas de la fraction F4, les spots se répartissent en deux groupes de masses moléculaires apparentes de 27 et 30 kDa. Chaque groupe de masse moléculaire comporte au moins six spots avec des pHi compris entre 5,15 et 6,55. La complexité des profils électrophorétiques obtenus soutient l'hypothèse d'une présence simultanée de variants de phosphorylation et de variants d'épissage, dans la mesure où une variabilité de masse moléculaire et une variabilité isoélectrique sont observées.

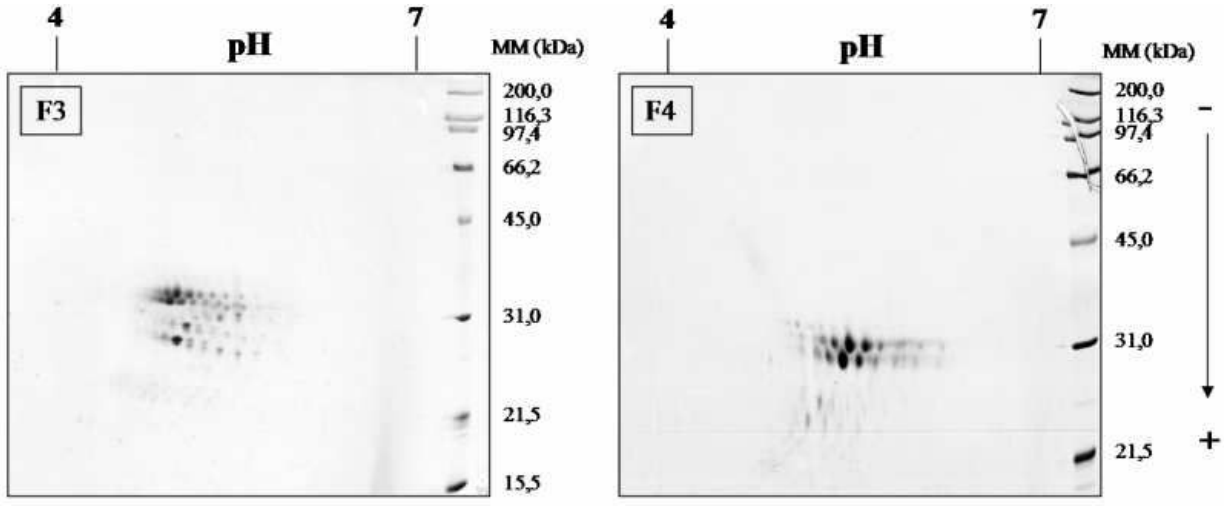


Figure 33 : Electrophorèse bidimensionnelle (première dimension : isoélectrofocalisation, deuxième dimension : PAGE-SDS) des fractions F3 et F4 issues de la séparation du caséinate de sodium par FPLC d'interactions hydrophobes. Des quantités protéiques de 50 µg ont été déposées sur des bandelettes constituées d'un gel à 4% de polyacrylamide contenant des ampholines créant un gradient de pH de 4 à 7. Les protéines ont été révélées dans le gel par le bleu de Coomassie. MM, masse moléculaire.

Afin de caractériser les variants d'épissage présents, il s'avère nécessaire d'obtenir les apoformes (formes déphosphorylées) de la CN-α_{s1}. En effet, le ou les niveau(x) de phosphorylation de cette caséine étant inconnus, il semble difficile de déterminer précisément quels sont les résidus d'acides aminés absents chez une protéine traduite à partir d'un ARNm ayant perdu des exons par épissage alternatif. Les masses théoriques des différentes

apoformes possibles peuvent être calculées avec le logiciel ProtParamTool disponible en ligne (http://expasy.org/tools/protparam.html). Par comparaison des masses observées par spectrométrie de masse avec les masses théoriques, les exons manquants peuvent ainsi être identifiés.

3.1.Déphosphorylation de la CN-α_{s1}

3.1.1. Utilisation de la phosphatase alcaline ou de la phosphatase acide

Les fractions F3 et F4 sont hydrolysées par la phosphatase alcaline. Les hydrolysats obtenus sont analysés par 2-DE et ne montrent qu'une déphosphorylation partielle des composants des fractions, notamment de F3 (Figure 34). En effet, il reste de nombreux spots visibles sur le gel alors que l'on cherche à n'obtenir, par groupe de masse moléculaire donnée, qu'un spot majoritaire avec un pHi le moins acide possible.

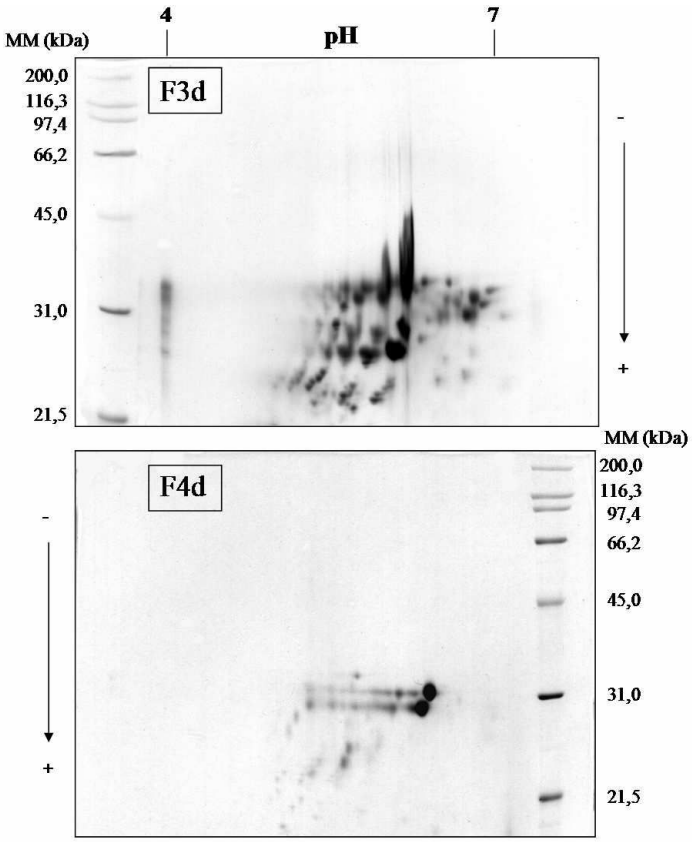


Figure 34 : Electrophorèse bidimensionnelle (première dimension : isoélectrofocalisation, deuxième dimension : PAGE-SDS) des fractions F3 et F4 hydrolysées par la phosphatase alcaline. Des quantités protéiques de 50 µg ont été déposées sur des bandelettes constituées d'un gel à 4% de polyacrylamide contenant des ampholines créant un gradient de pH de 4 à 7. Les protéines ont été révélées dans le gel par le bleu de Coomassie. F3d et F4d, fractions F3 et F4 déphosphorylées ; MM, masse moléculaire.

Un nouvel essai de déphosphorylation est alors mené en utilisant la phosphatase acide. Le profil obtenu par 2-DE montre également une déphosphorylation partielle de la CN-α_{s1} contenue dans la fraction F4 (Figure 35).

Ces résultats semblent suggérer que la CN-α_{s1} équine, contrairement à la CN-β équine, est moins sensible à l'action des phosphatases alcaline et acide. La phosphatase alcaline semble néanmoins plus efficace que la phosphatase acide dans le cas de la fraction F4 si l'on compare les profils obtenus avec ces deux enzymes (Figures 34 et 35).

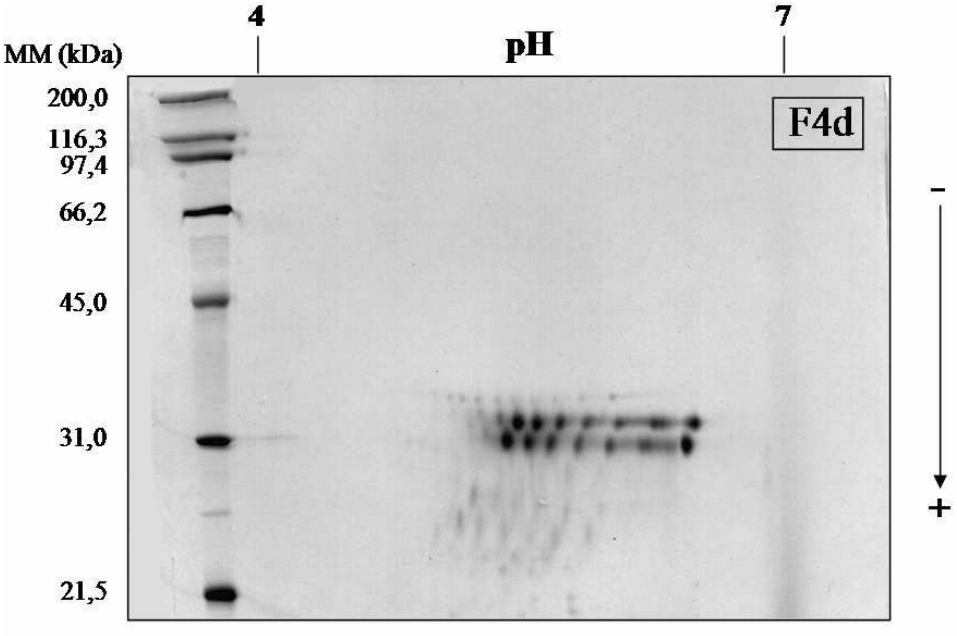


Figure 35 : Electrophorèse bidimensionnelle (première dimension : isoélectrofocalisation, deuxième dimension : PAGE-SDS) de la fraction F4 hydrolysée par la phosphatase acide. Une quantité protéique de 50 µg a été déposée sur une bandelette constituée d'un gel à 4% de polyacrylamide contenant des ampholines créant un gradient de pH de 4 à 7. Les protéines ont été révélées dans le gel par le bleu de Coomassie. F4d, fraction F4 déphosphorylée ; MM, masse moléculaire.

3.1.2. Utilisation de la phosphatase alcaline suivie de la phosphatase acide

Du fait de l'hydrolyse incomplète obtenue avec les phosphatases utilisées séparément, nous avons choisi d'associer les deux phosphatases pour tenter d'obtenir une déphosphorylation de la CN-α_{s1} la plus complète possible.

Dans un premier temps, la fraction F3 est hydrolysée par la phosphatase acide puis par la phosphatase alcaline. Les résultats obtenus par 2-DE ne s'avèrent pas concluants

puisqu'une déphosphorylation incomplète des constituants de la fraction F3 est obtenue. En effet, de nombreux spots restent visibles pour chacun des groupes de masse moléculaire (Figure 36).

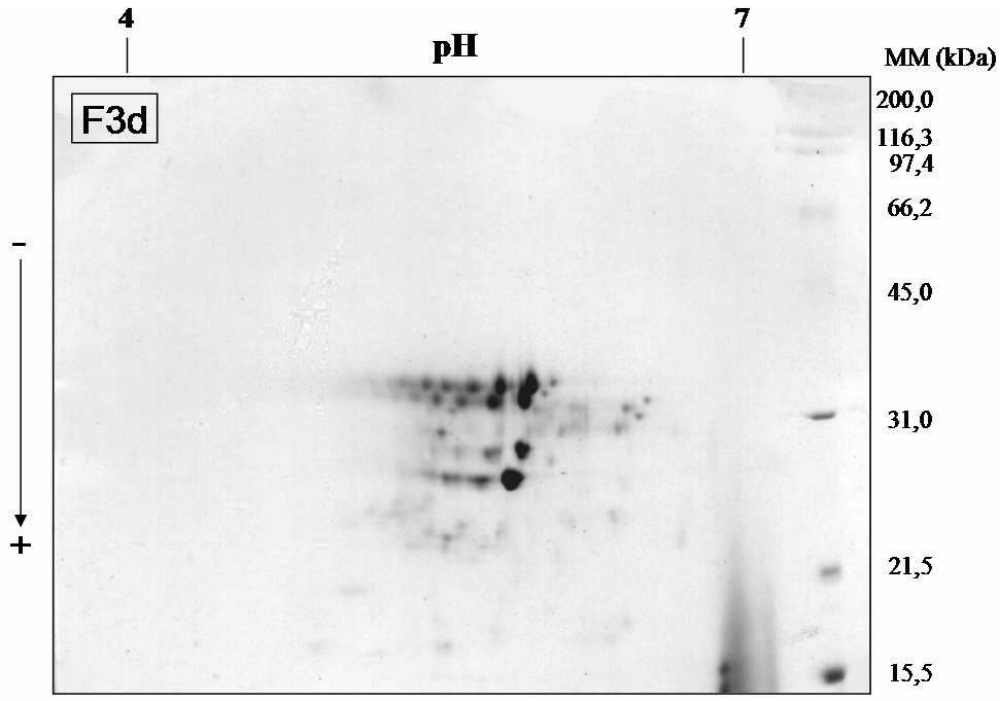


Figure 36 : Electrophorèse bidimensionnelle (première dimension : isoélectrofocalisation, deuxième dimension : PAGE-SDS) de la fraction F3 hydrolysée par la phosphatase acide puis par la phosphatase alcaline. Une quantité protéique de 50 µg a été déposée sur une bandelette constituée d'un gel de polyacrylamide contenant des ampholines créant un gradient de pH de 4 à 7. Les protéines ont été révélées dans le gel par le bleu de Coomassie. F3d, fraction F3 déphosphorylée ; MM, masse moléculaire.

Dans un second temps, l'ordre d'utilisation des phosphatases est inversé. Ainsi, la CN-α_{s1} est d'abord hydrolysée par la phosphatase alcaline puis par la phosphatase acide. Dans ce cas, une déphosphorylation quasi-complète est obtenue, et cela pour les deux fractions F3 et F4. En effet, chaque groupe de masse moléculaire contient alors un spot majoritaire dont le pHi correspond au pHi le moins acide dans sa catégorie et semble donc correspondre à une forme totalement déphosphorylée (Figure 37). La méthode de double déphosphorylation semble cependant plus efficace pour la fraction F4 que pour la fraction F3. Le profil bidimensionnel de la fraction F3 montre effectivement une présence résiduelle de spots mineurs plus marquée que sur le profil bidimensionnel de la fraction F4.

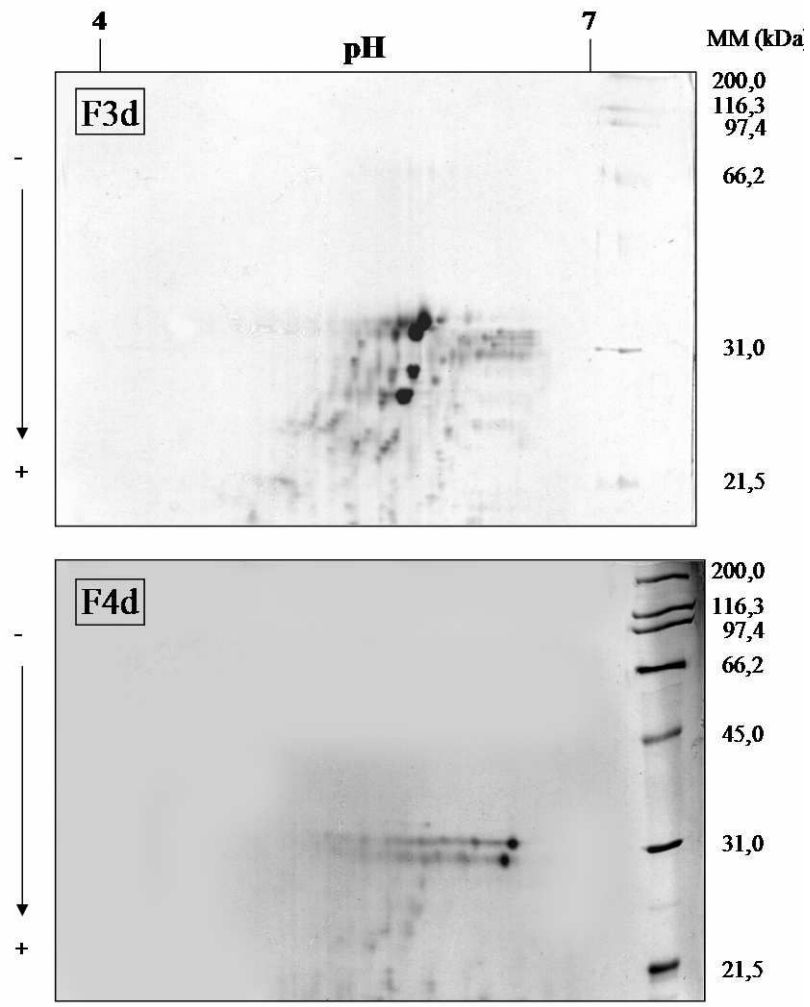


Figure 37 : Electrophorèse bidimensionnelle (première dimension : isoélectrofocalisation, deuxième dimension : PAGE-SDS) des fractions F3 et F4 hydrolysées par la phosphatase alcaline puis par la phosphatase acide. Des quantités de 50 µg ont été déposées sur des bandelettes constituées d'un gel à 4% de polyacrylamide contenant des ampholines créant un gradient de pH de 4 à 7. Les protéines ont été révélées dans le gel par le bleu de Coomassie. F3d et F4d, fractions F3 et F4 déphosphorylées ; MM, masse moléculaire.

L'association des deux phosphatases doit donc être faite dans un ordre précis : il faut d'abord faire agir la phosphatase alcaline puis la phosphatase acide et non l'inverse, méthode moins efficace. Lors de l'utilisation des phosphatases seules, la phosphatase alcaline semblait être plus efficace que la phosphatase acide pour l'obtention d'une déphosphorylation partielle (moins de spots natifs présents sur les profils bidimensionnels). Lors de l'association des deux phosphatases, on peut émettre l'hypothèse que la phosphatase alcaline ne cliverait que certains groupements phosphates de la CN-α_{s1}, ce qui changerait la conformation de la protéine et éventuellement l'accessibilité de l'autre enzyme à certains sites. Dans un second

temps, le changement de conformation rendrait alors accessible les autres groupements phosphates à la phosphatase acide qui complèterait la déphosphorylation amorcée par la phosphatase alcaline.

3.2. Caractérisation des variants d'épissage de la CN-α_{s1}

Les différentes formes déphosphorylées obtenues ont été identifiées par analyse de spectrométrie de masse LC/ESI-MS des fractions F3d et F4d à l'aide des spectromètres de masse API III$^+$ et Q/Star XL. Une meilleure séparation chromatographique est obtenue avec une colonne narrow-bore connectée à l'API III$^+$ plutôt qu'avec une nano-colonne connectée au Q/Star XL. De plus, une quantité de matériel plus importante peut être déposée sur la colonne narrow-bore. Ainsi, les fractions F3d et F4d sont séparées sur une colonne narrow-bore puis ionisées dans la source électrospray, ce qui permet de mesurer des masses avec une précision de ± 0,01% seulement, ce qui est lié au quadripôle. Les fractions majoritaires détectées avec l'API III$^+$ sont collectées et sont infusées à l'aide d'une nano-seringue électrospray sur le Q/Star XL pour obtenir avec plus de précision (± 0,001%) les valeurs des masses.

Les différents résultats présentés ont donc été obtenus avec le Q/Star XL qui possède une plus grande précision que l'API III$^+$ dans la mesure de masse (Tableau 4, Figure 38).

Tableau 4 : Masses moléculaires déterminées par LC/ESI-MS sur l'appareil Q/Star-XL des fractions F3d et F4d obtenues après double déphosphorylation par la phosphatase alcaline puis la phosphatase acide des fractions F3 et F4 issues de la séparation par FPLC d'interactions hydrophobes du caséinate de sodium équin.

Masse observée[a]	Identification[b]	Masse théorique[c]	Fraction[d]
24613	CN-α_{s1} 0P	24614	F3d
24486	ΔQ 0P	24486	F3d
23257	Δ14 0P	23257	F3d
23127	Δ14,Q 0P	23128	F3d
23751	Δ7 0P	23751	F4d
23624	Δ7,Q 0P	23623	F4d
22394	Δ7,14 0P	22393	F4d
22266	Δ7,14,Q 0P	22265	F4d

[a] Masse ± 1 Da.
[b]ΔX: délétion de la région code par l'exon X; ΔQ: délétion du résidu N-terminal (Gln) de la région codée par l'exon 11; CN, caséine ; P, groupement phosphate.
[c]Masse moyenne calculée avec ProtParam tool (expasy.org/tools/protparam.html).
[d]F3d et F4d: fractions F3 et F4 déphosphorylées.

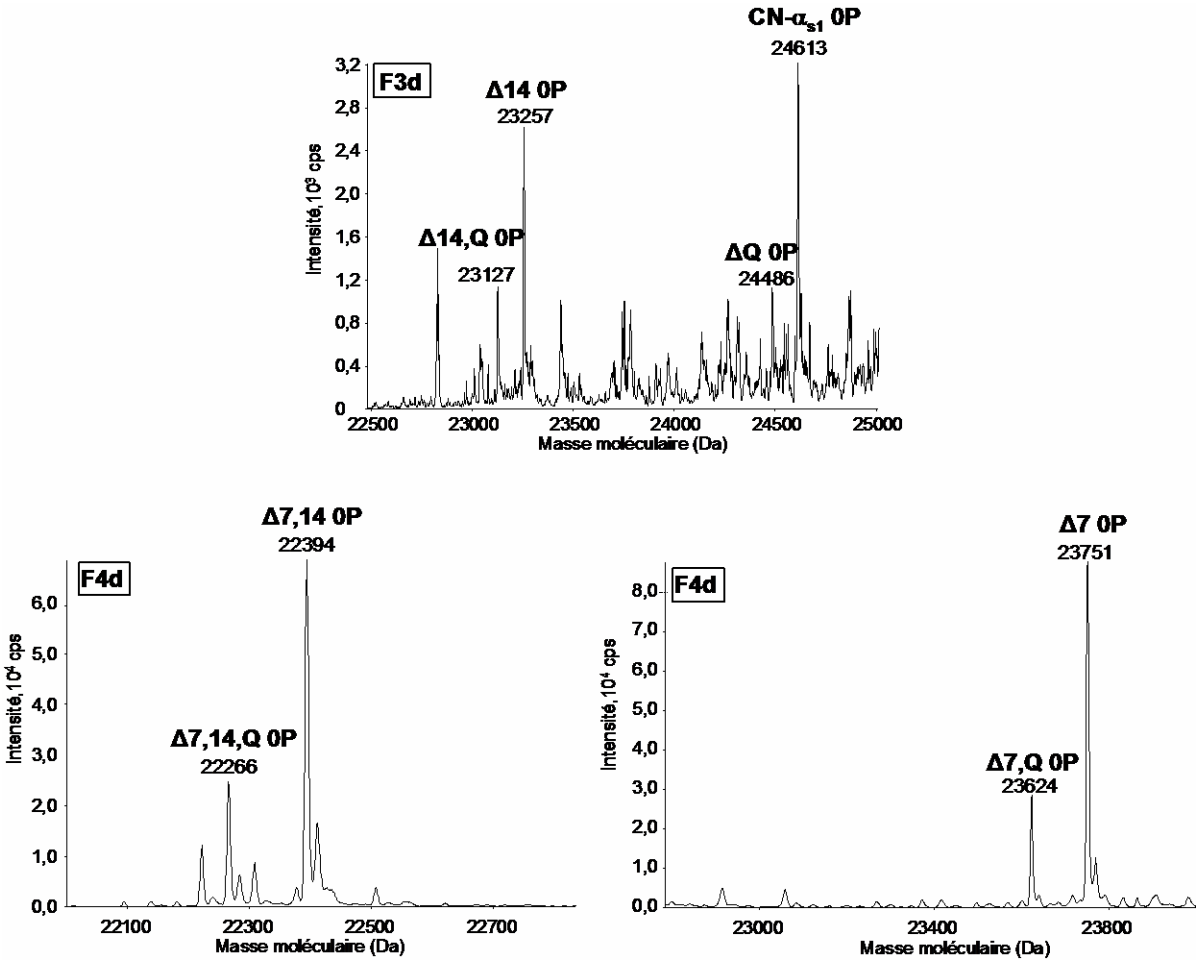


Figure 38 : Masses reconstruites obtenues par chromatographie liquide couplée à la spectrométrie de masse à ionisation électrospray (LC/ESI-MS) sur appareillage Q/Star XL des fractions F3 et F4 déphosphorylées par la phosphatase alcaline puis la phosphatase acide. CN-αs$_1$, caséine αs$_1$; F3d et F4d, fractions F3 et F4 déphosphorylées ; ΔX, délétion de la séquence de résidus d'acides aminés codés par l'exon X ; ΔQ, délétion du résidu N-terminal (Gln) de la région codée par l'exon 11 ; 0P, absence de groupement phosphate ; cps, coup par seconde.

Dans la fraction F3d, malgré un spectre des ions multi-chargés complexe, quatre masses moléculaires (24613 Da, 24486 Da, 23257 Da et 23127 Da) ont été identifiées comme correspondant à une CN-α_{s1} complète déphosphorylée (n'ayant subi aucun épissage alternatif d'exon ; masse théorique de 24614 Da) et à trois isoformes étant probablement des variants d'épissage d'exons. La fraction F4d, quant à elle, contient quatre autres masses moléculaires (23752 Da, 23624 Da, 22394 Da, 22266 Da) qui correspondraient à d'autres variants d'épissage d'exons. Les masses théoriques des variants d'épissage alternatif potentiels (Δ7, Δ14 et Δ7,14) sont comparées aux masses observées. L'existence des trois variants d'épissage alternatif est ainsi mise en évidence au niveau protéique. Ainsi, la masse de 23257 Da obtenue

dans la fraction F3d peut être expliquée par l'épissage alternatif de l'exon 14 qui conduit à la traduction du variant α_{s1}-CN$^{\Delta 14}$ (masse théorique de l'apoforme de 23257 Da). De la même façon, les apoformes des variants CN-$\alpha_{s1}^{\Delta 7}$ (masse théorique de l'apoforme de 23751 Da) et CN-$\alpha_{s1}^{\Delta 7,14}$ (masse théorique de l'apoforme de 22394 Da) ont été identifiées dans la fraction F4d. Les autres masses déterminées (24486, 23127, 23624 et 22266 Da) sont plus faibles de 128 ± 1 Da par rapport aux masses déterminées pour la CN-α_{s1} complète et pour chacun de ses variants d'épissage alternatif d'exons. Cela pourrait s'expliquer par la délétion d'un résidu Lys ou Gln. Lenasi *et al.* (2003) ont fortement suspecté une délétion des trois premiers nucléotides de l'exon 11 codant pour le résidu Gln^{91} dans un transcrit qu'ils ont isolé. La présence dans notre étude de variants protéiques montrant une diminution de masse de 128 ± 1 Da semble confirmer leur hypothèse et montrer que ces transcrits sont traduits. La délétion du résidu Gln^{91} s'explique par un mécanisme d'épissage alternatif impliquant un site d'épissage cryptique présent au niveau des trois premiers nucléotides en position 5' de l'exon 11. Selon Martin *et al.* (2002), la délétion du résidu Gln pourrait s'expliquer par une erreur d'épissage par le splicéosome. Ainsi, les deux nucléotides AG contenus dans le codon CAG, codant pour le résidu Gln et présent en position 5' de l'exon 11, pourraient être confondus avec le site d'épissage accepteur AG de l'intron adjacent et ils seraient utilisés alternativement (Figure 39 ; Smith *et al.,* 1993).

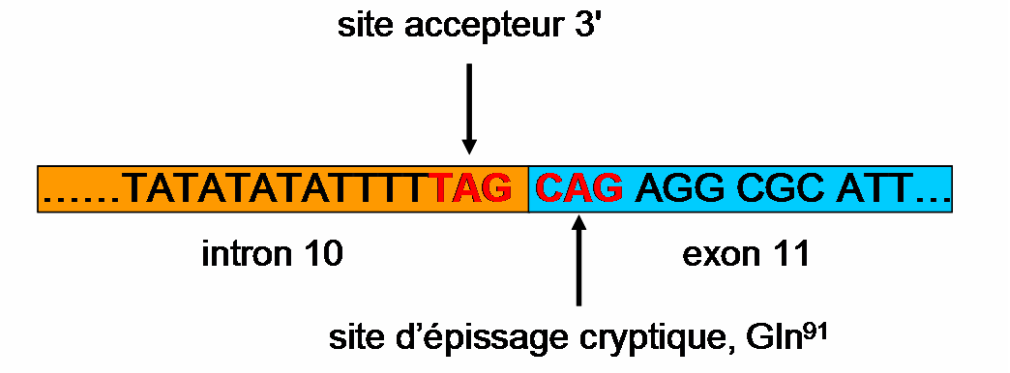


Figure 39 : Schéma représentant la séquence nucléotidique à la jonction intron 10/exon 11 de la CN-α_{s1} équine pouvant induire l'utilisation du site cryptique d'épissage correspondant au premier codon de l'exon 11 qui code pour le résidu Gln^{91}.

Finalement, la CN-α_{s1} complète et sept variants d'épissage alternatif de cette CN-α_{s1}, impliquant l'épissage d'exons ou l'utilisation d'un site d'épissage cryptique, ont été caractérisés dans le lait de jument Haflinger : CN-$\alpha_{s1}^{\Delta Q}$ (où Gln^{91} est délété), CN-$\alpha_{s1}^{\Delta 14}$, CN-$\alpha_{s1}^{\Delta 14,Q}$, CN-$\alpha_{s1}^{\Delta 7}$, CN-$\alpha_{s1}^{\Delta 7,Q}$, CN-$\alpha_{s1}^{\Delta 7,14}$ et CN-$\alpha_{s1}^{\Delta 7,14,Q}$.

Les ARNm des variants CN-$\alpha_{s1}^{\Delta 7}$ et CN-$\alpha_{s1}^{\Delta 14}$ ont été isolés chez la jument hanovrienne (Lenasi *et al.,* 2003) et chez le poney Gallois (Milenkovic *et al.,* 2002), respectivement. De même chez le poney gallois, la présence au niveau protéique des variants CN-$\alpha_{s1}^{\Delta 7}$ et CN-$\alpha_{s1}^{\Delta 7,14}$ a été fortement suggérée par Miranda *et al*. (2004). Cependant, au niveau protéique, le variant CN-$\alpha_{s1}^{\Delta 14}$ et la protéine complète n'avaient jamais été observés. Nos résultats montrent avec précision que, dans le lait de jument Haflinger, les trois types d'ARNm produits par l'épissage alternatif des exons 7 et 14 sont traduits en protéines. De plus, la CN-α_{s1} existe également dans le lait sous sa forme protéique complète, ce qui jusqu'ici n'avait pas été démontré.

L'existence de nombreuses isoformes de CN-α_{s1} due à l'épissage alternatif d'exons ou à la présence de site d'épissage cryptique est un phénomène fréquemment observé chez les ruminants mais aussi chez d'autres mammifères (Martin *et al.,* 2002). Chez la chèvre, les exons 13 et 16 sont épissés alternativement ainsi que le résidu Gln^{78} qui est le premier résidu codé par l'exon 11 (isoformes CN-α_{s1}, CN-$\alpha_{s1}^{\Delta 13}$, CN-$\alpha_{s1}^{\Delta 16}$, CN-$\alpha_{s1}^{\Delta Q}$; Ferranti *et al.,* 1997). Chez la brebis, l'exon 16 peut subir un épissage alternatif et le résidu Gln^{78} peut aussi être délété (Ferranti *et al.,* 1995 ; Passey *et al.,* 1996). Chez la femme, l'exon 7 est alternatif, et il existe également un site d'épissage cryptique correspondant au premier codon de l'exon 6 codant pour le résidu Gln^{37} (Johnsen *et al.,* 1995).

On remarquera que, d'après ce travail, le lait de jument semble posséder beaucoup plus d'isoformes de CN-α_{s1} que le lait d'autres espèces. Ceci semble être dû à l'épissage alternatif simultané d'exons et à l'utilisation fréquente d'un site cryptique d'épissage. En effet, les exons 7 et 14 peuvent être simultanément exclus, ce qui n'a pas été observé chez la chèvre, par exemple, pour laquelle deux exons sont également impliqués dans un processus d'épissage alternatif. De même, pour chaque variant d'épissage alternatif d'exons ainsi que pour la protéine complète, il existe un variant impliquant l'utilisation d'un site cryptique d'épissage. Ce n'est pas le cas pour les autres espèces étudiées où l'utilisation d'un site cryptique d'épissage n'est jamais observée en même temps que l'épissage alternatif d'exons. Il se peut cependant que toutes les isoformes de la CN-α_{s1} n'aient pas été caractérisées chez ces espèces. Bien que nous ayons travaillé avec un lait de petit mélange provenant de plusieurs juments, la variabilité protéique observée n'est probablement pas le fait d'une variabilité génétique due à la présence de laits provenant de plusieurs juments. En effet, la race Haflinger est très homogène. Tous les ascendants paternels de tous les étalons remontent sans discontinuer jusqu'à Folie, premier étalon de la race né en 1874 à Hafling, en Autriche,

d'un pur-sang arabe, El Bedawy, et d'une jument tyrolienne autochtone (Les Haras Nationaux). Afin de vérifier cette hypothèse, il aurait été intéressant d'étudier les caséines d'un lait de jument individuel, mais son obtention ne nous a pas été possible.

4. Caractérisation des variants de phosphorylation de la CN-α_{s1} équine

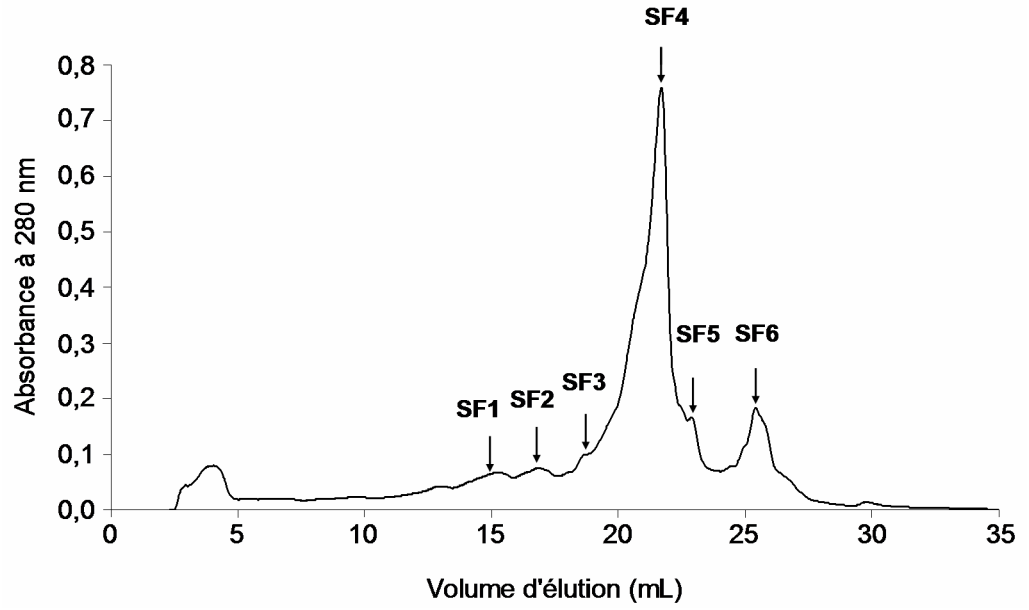
Connaissant maintenant la structure primaire des différents variants d'épissage de la CN-α_{s1} présents dans le lait de jument, il est possible de déterminer leur degré de phosphorylation. Pour cela, les fractions F3 et F4 sont analysées directement par LC/ESI-MS sur le spectromètre Q/Star XL sans subir de traitement de déphosphorylation. La différence de masse entre l'apoforme et les formes phosphorylées doit donc être un multiple de 80 Da (masse de HPO_3). Cependant, les spectres obtenus avec les fractions F3 et F4 sont très complexes et difficiles à interpréter ce qui est probablement dû à la présence des nombreux variants de phosphorylation potentiels et variants d'épissage contenus dans chaque fraction. Il faut donc réaliser une autre approche permettant de simplifier la détermination des masses des variants post-traductionnels.

4.1. Fractionnement de la CN-α_{s1} par chromatographie FPLC d'échange d'ions

Les fractions F3 et F4 obtenues par FPLC d'interactions hydrophobes sont rassemblées afin d'obtenir une seule fraction contenant tous les variants de CN-α_{s1}. La fraction ainsi obtenue est ensuite sous-fractionnée par chromatographie FPLC d'échange d'anions (IEC) et permet d'obtenir six sous-fractions SF1 à SF6 avec SF4 comme sous-fraction majoritaire (Figure 40a). Chaque sous-fraction est collectée puis analysée par PAGE-urée, les bandes principales étant numérotées selon la nomenclature proposée par Egito *et al.* (2002 ; Figure 40b). En chromatographie d'échange d'anions, les constituants les moins chargés négativement sont élués les premiers tandis que les constituants les plus chargés négativement sont élués en fin de gradient de force ionique. De manière similaire, en PAGE-urée, plus une protéine est chargée négativement, plus sa distance de migration sera importante. C'est pourquoi nous observons que l'ordre d'élution chromatographique des protéines est corrélé avec leur distance de migration électrophorétique (comparer Figures 40a et 40b). Ceci semble montrer que l'ordre d'élution serait corrélé au nombre de groupements

phosphates portés par la protéine : plus la protéine est phosphorylée, et plus elle sera éluée tardivement et plus sa distance de migration sera grande.

a)



b)

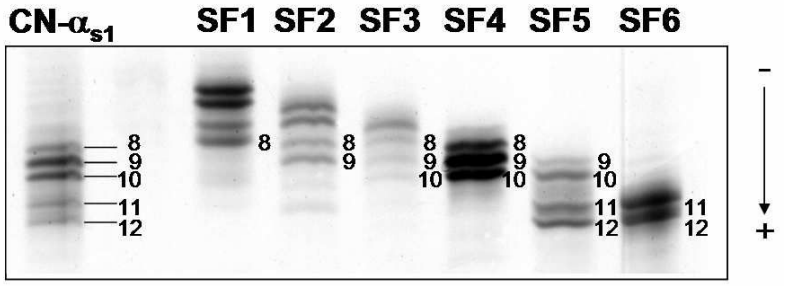


Figure 40 : a) Fractionnement par FPLC d'échanges d'ion du mélange des fractions F3 et F4 précédemment obtenus par FPLC d'interactions hydrophobes du caséinate de sodium. Des quantités protéiques de 60 mg sont déposées sur une colonne TSK DEAE 5-PW semi-préparative. Un gradient linéaire croissant de 0,08–0,28 M de NaCl à pH 8,0 est réalisé pendant 80 min, suivi d'une étape isocratique à 0,28 M de NaCl de 15 min. Le débit est de 4 mL.min^{-1} et l'absorbance est mesurée à 280 nm.

b) Caractérisation électrophorétique (PAGE-urée) des sous-fractions SF1 à SF6 purifiées par FPLC d'échanges d'ion. Une quantité de 40 µg de protéines est déposée dans le gel. Les protéines sont révélées dans le gel par le bleu de Coomassie. Les bandes sont numérotées selon la nomenclature proposée par Egito *et al.* (2002). CN-α_{s1}, caséine α_{s1} correspondant au mélange de F3 et F4 ; SF, sous-fraction.

Chaque sous-fraction a été caractérisée par 2-DE (Figure 41). Les profils bidimensionnels sont sensiblement moins complexes que celui du mélange des fractions F3 et F4 (Figure 42).

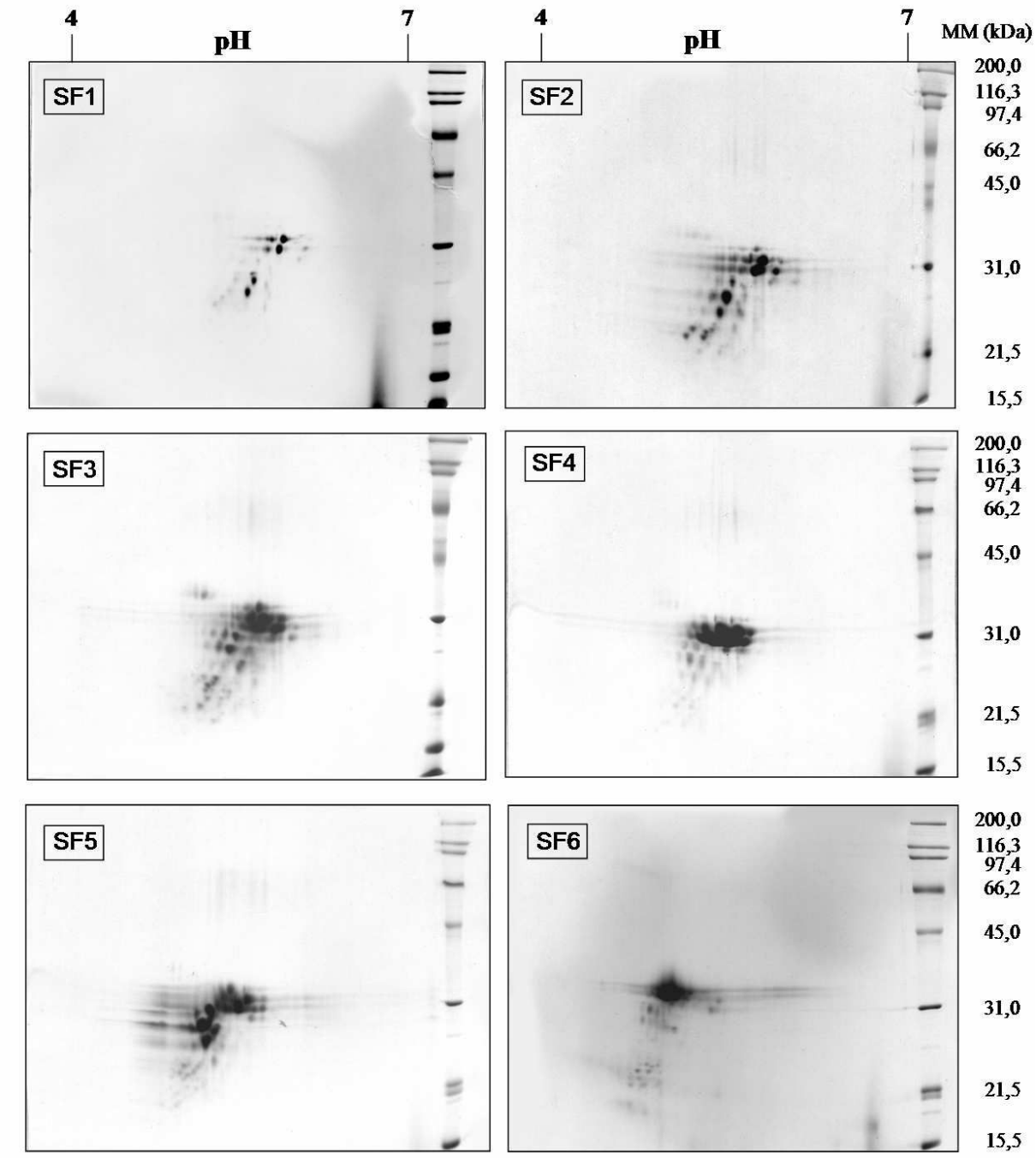


Figure 41 : Electrophorèse bidimensionnelle (première dimension : isoélectrofocalisation, deuxième dimension : PAGE-SDS) des sous-fractions SF1 à SF6 obtenues par FPLC d'échanges d'ions du mélange des fractions F3 et F4. Des quantités protéiques de 50 µg ont été déposées sur des bandelettes constituées d'un gel à 4% de polyacrylamide contenant des ampholines créant un gradient de pH de 4 à 7. Les protéines ont été révélées dans le gel par le bleu de Coomassie. SF, sous-fraction ; MM, masse moléculaire.

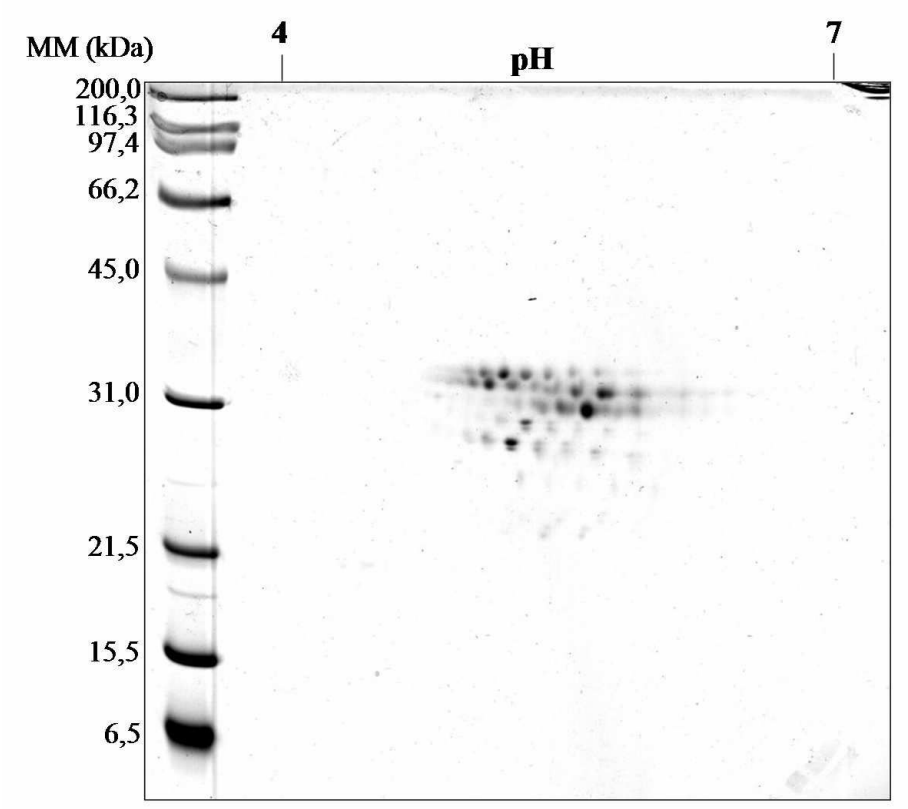


Figure 42 : Electrophorèse bidimensionnelle (première dimension : isoélectrofocalisation, deuxième dimension : PAGE-SDS) du mélange des fractions F3 et F4 obtenues par FPLC d'interactions hydrophobe du caséinate de sodium équin. Des quantités protéiques de 50 µg ont été déposées sur des bandelettes constituées d'un gel à 4% de polyacrylamide contenant des ampholines créant un gradient de pH de 4 à 7. Les protéines ont été révélées dans le gel par le bleu de Coomassie. MM, masse moléculaire.

4.2. Caractérisation du degré de phosphorylation

Les six sous-fractions (SF1 à SF6) obtenues par IEC sont analysées par LC/ESI-MS et chaque spectre obtenu peut être correctement interprété (Figure 43).

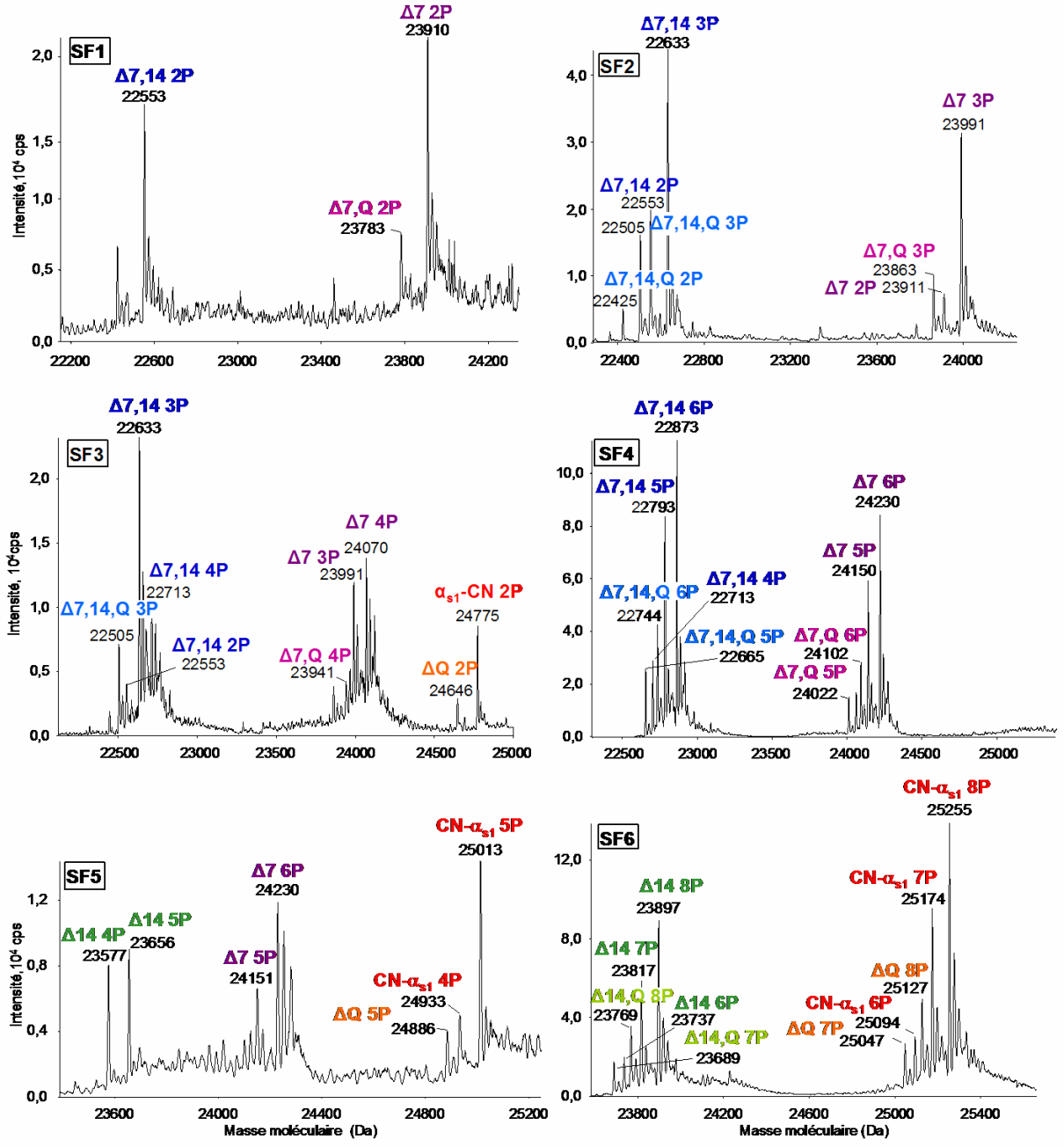


Figure 43 : Masses reconstruites obtenues par chromatographie liquide couplée à la spectrométrie de masse à ionisation électrospray (LC/ESI-MS) sur appareillage Q/Star XL des sous-fractions SF1 à SF6 obtenues par FPLC d'échanges d'ion du mélange F3 et F4. CN-α_{s1}, caséine α_{s1} ; SF, sous-fraction; ΔX, délétion de la séquence de résidus d'acides aminés codés par l'exon X ; ΔQ, délétion du résidu N-terminal (Gln) de la région codé par l'exon 11 ; XP, présence de X groupements phosphate ; cps, coup par seconde.

La synthèse des résultats est rapportée dans le Tableau 5.

Tableau 5 : Masses moléculaires déterminées par LC/ESI-MS sur l'appareil Q/Star XL des différentes sous-fractions obtenues par FPLC d'échanges d'ions de la CN-α_{s1} équine.

Masse observée[a]	Identification[b]	Masse théorique[c]	Fraction[d]
24775	CN-α_{s1} 2P	24774	SF3
24933	CN-α_{s1} 4P	24934	SF5
25013	CN-α_{s1} 5P	25014	SF5
25094	CN-α_{s1} 6P	25094	SF6
25174	CN-α_{s1} 7P	25174	SF6
25255	CN-α_{s1} 8P	25254	SF6
24646	ΔQ 2P	24646	SF3
24886	ΔQ 5P	24886	SF5
25047	ΔQ 7P	25046	SF6
25127	ΔQ 8P	25126	SF6
23577	Δ14 4P	23577	SF5
23656	Δ14 5P	23657	SF5
23737	Δ14 6P	23737	SF6
23817	Δ14 7P	23817	SF6
23897	Δ14 8P	23897	SF6
23689	Δ14,Q 7P	23688	SF6
23769	Δ14,Q 8P	23768	SF6
23911	Δ7 2P	23911	SF1, SF2
23991	Δ7 3P	23991	SF2, SF3
24070	Δ7 4P	24071	SF3
24150	Δ7 5P	24150	SF4, SF5
24230	Δ7 6P	24231	SF4, SF5
23783	Δ7,Q 2P	23783	SF1
23863	Δ7,Q 3P	23863	SF2
23941	Δ7,Q 4P	23943	SF3
24022	Δ7,Q 5P	24023	SF4
24102	Δ7,Q 6P	24103	SF4
22553	Δ7,14 2P	22553	SF1, SF2, SF3
22633	Δ7,14 3P	22633	SF2, SF3
22713	Δ7,14 4P	22713	SF3, SF4
22793	Δ7,14 5P	22793	SF4
22873	Δ7,14 6P	22873	SF4
22425	Δ7,14,Q 2P	22425	SF2
22505	Δ7,14,Q 3P	22505	SF2, SF3
22665	Δ7,14,Q 5P	22665	SF4
22744	Δ7,14,Q 6P	22745	SF4

[a] Masse ± 2 Da.
[b] ΔX: délétion de la région code par l'exon X; ΔQ: délétion du résidu N-terminal (Gln) de la région codée par l'exon 11; CN, caséine ; P, groupement phosphate.
[c] Masse moyenne calculée avec ProtParam tool (expasy.org/tools/protparam.html).
[d] SF: sous-fractions de CN-α_{s1}

Ainsi, les variants CN-$\alpha_{s1}^{\Delta 7}$ et CN-$\alpha_{s1}^{\Delta 7,14}$, qui ont été identifiés dans la fraction F4, possèdent de 2 à 6 groupements phosphates (2P à 6P) et sont retrouvés dans les sous-fractions SF1 à SF5. La CN-α_{s1} complète et le variant CN-$\alpha_{s1}^{\Delta 14}$, identifiés précédemment dans la fraction F3, sont retrouvés principalement dans les sous-fractions SF5 et SF6 et ont un degré de phosphorylation allant de 4P à 8P. La CN-α_{s1} complète possède en plus un variant 2P retrouvé dans la sous-fraction SF3. Il est à noter que les variants qui possèdent l'exon 7 ne présentent pas de variant 3P au contraire des variants pour lesquels cet exon a été épissé. En effet, la CN-α_{s1} complète ne semble pas posséder de variant 3P puisque aucun signal à 24854 Da n'a été détecté. De même, pour la CN-$\alpha_{s1}^{\Delta 14}$ aucun signal n'a été détecté à 23497 Da indiquant qu'elle ne porterait jamais 3P. Il est à noter également qu'aucun variant 0P ou 1P n'a été identifié.

On remarque que les variants ne possédant pas l'exon 7 ont un degré de phosphorylation maximal de 6P et non de 8P. En effet, d'après la structure primaire de la CN-α_{s1} (Figure 25), la région codée par l'exon 7 contient trois sites potentiels de phosphorylation (deux résidus Ser et un résidu Thr). Quand cet exon est épissé, la CN-α_{s1} perd deux degrés de phosphorylation, ce qui signifierait que la région codée par l'exon 7 porte deux résidus phosphorylés et non trois. Les protéines ayant une séquence Thr-X-Glu/SerP sont reconnues pour être de faibles substrats de la caséine kinase de la glande mammaire (Mercier, 1981). Ainsi, nous pouvons émettre l'hypothèse que les deux résidus phosphorylés dans la région codée par l'exon 7 seraient plutôt les résidus Ser^{58} et Ser^{61}. Les résultats obtenus supportent l'idée que la présence de cette région dans la protéine augmente de 2P le nombre minimal et maximal de phosphorylations dans la protéine. En effet, le nombre minimal passe de 2P à 4P et le nombre maximal de 6P à 8P quand l'exon 7 n'a pas été épissé. Cela suggère que la région 56-63 codée par l'exon serait toujours phosphorylée quand celui-ci est traduit. Néanmoins, la CN-α_{s1} complète présente un nombre minimal de 2P au lieu de 4P attendu dans le cadre de notre hypothèse. La localisation précise des sites de phosphorylation pour chaque degré de phosphorylation de chaque variant protéique sera à déterminer afin de vérifier cette hypothèse.

Concernant les variants d'épissage de site cryptique pour lesquels le résidu Gln^{91} est délété, tous les degrés de phosphorylation ne sont pas mis en évidence. La faible proportion de variants ΔQ dans le lait équin (intensité spectrale des variants ΔQ systématiquement plus faible comparée à celle des isoformes correspondantes contenant Gln ; Figure 43) pourrait expliquer le fait que certains variants de phosphorylation ne puissent pas être détectés par

spectrométrie de masse. Néanmoins, plusieurs variants de phosphorylation ont pu être mis en évidence pour les variants ΔQ identifiés : α_{s1}-CN$^{\Delta Q}$ 2P, 5P, 7P et 8P, α_{s1}-CN$^{\Delta 7,Q}$ 2P à 6P, α_{s1}-CN$^{\Delta 14,Q}$ 7P et 8P, α_{s1}-CN$^{\Delta 7,14,Q}$ 2P, 3P, 5P et 6P.

Pour l'ensemble des variants identifiés, le nombre maximal de phosphorylation est de 8 alors qu'il existe dix sites potentiels de phosphorylation dans la structure primaire de la CN-α_{s1}. Les huit résidus pouvant être phosphorylés sont probablement les résidus Ser^{18}, Ser^{58}, Ser^{61}, Ser^{75}, Ser^{77}, Ser^{79}, Ser^{80} et Ser^{81} contenus dans les séquences de type -Ser-X-Glu/SerP (Figure 25).

Si l'on s'attache au contenu des différentes sous-fractions analysées par spectrométrie de masse, on voit que les variants ne contenant pas l'exon 7 sont élués par IEC avant la CN-α_{s1} complète et la CN-$\alpha_{s1}^{\Delta 14}$ (Figure 43). Ceci peut s'expliquer par le fait que les variants Δ7 et Δ7,14 ont des pHi (pHi théoriques des apoformes de 5,86 et 5,85 respectivement) moins acides que ceux de la protéine complète et du variant Δ14 (pHi théoriques des apoformes de 5,47 et 5,45 respectivement). De même, pour un variant protéique donné, son temps de rétention observé par IEC dépend de son degré de phosphorylation, les variants les plus phosphorylés s'éluant le plus tard.

Il semble difficile de mettre en relation les résultats observés pour chaque sous-fraction par PAGE-urée avec les résultats de 2-DE et de spectrométrie de masse. En effet, nous espérions pouvoir attribuer grâce aux résultats de LC-ESI/MS un variant donné à une bande de PAGE-urée ou à un spot de 2-DE. Les différents spectres de masse obtenus nous montrent que chaque sous-fraction s'avère riche en variants. Ainsi, un spot de 2-DE ou une bande de PAGE-urée semble contenir plusieurs variants qui co-migreraient ensembles. Cependant, nous pouvons extraire quelques informations à partir des différents résultats obtenus. La fraction SF4, qui est la fraction majoritaire obtenue par IEC, contient les bandes électrophorétiques 8, 9 et 10 (PAGE-urée, Figure 40). Cette fraction est constituée des variants CN-$\alpha_{s1}^{\Delta 7}$, CN-$\alpha_{s1}^{\Delta 7,Q}$, CN-$\alpha_{s1}^{\Delta 7,14}$ et CN-$\alpha_{s1}^{\Delta 7,14,Q}$ avec la plus forte intensité en spectrométrie de masse pour les variants CN-$\alpha_{s1}^{\Delta 7}$ et CN-$\alpha_{s1}^{\Delta 7,14}$ notamment les plus phosphorylés, c'est-à-dire avec 5P et 6P. La seconde sous-fraction majoritaire est SF6 qui est principalement constituée des bandes 11 et 12. Cette sous-fraction est composée des plus hauts degrés de phosphorylation (6P, 7P et 8P) de la CN-α_{s1} complète, de la CN-$\alpha_{s1}^{\Delta 14}$ et de leurs variants respectifs ΔQ. Dans le lait de jument Haflinger, il semble donc que les variants ne possédant pas l'exon 7 soient les isoformes présentes majoritairement. L'épissage alternatif d'exons serait donc un phénomène crucial dans la production de la CN-α_{s1}. L'absence de cet

exon 7 pourrait avoir un rôle sur la structure de la protéine qui pourrait acquérir une conformation différente et modifier son comportement notamment vis-à-vis de la chélation de minéraux comme le calcium ce qui pourrait influencer entre autres la micelle de caséines. A la lumière des résultats obtenus, nous pouvons également constater que les formes les plus phosphorylées semblent être les plus abondantes pour la CN-α_{s1} complète et pour tous ses variants d'épissage. Ceci n'est pas le cas pour la CN-β équine où le variant majoritaire est CN-β 5P alors que le variant le plus phosphorylé possède 7P (Girardet *et al.,* 2006). La phosphorylation élevée de la CN-α_{s1} et le fait que celle-ci soit variable pourrait jouer un rôle important notamment pour le transport de minéraux chez le jeune et dans la formation de la micelle.

Nous pouvons également nous apercevoir que le profil PAGE-urée du caséinate de sodium est très complexe, car les différentes caséines sont difficiles à individualiser par cette méthode. En effet, Egito *et al.* (2002) ont montré que les bandes de CN-κ s'intercalent entre les bandes de CN-β. Dans le présent travail, nous montrons que des bandes bien focalisées dans la zone de migration de la CN-κ semblent correspondre à la CN-α_{s1} (Figure 31a). Cela signifie qu'il existe une zone sur le profil PAGE-urée où des variants de caséines α_{s1}, β et κ migrent ensemble, ce qui rend l'interprétation du profil électrophorétique particulièrement difficile.

Par comparaison, la CN-α_{s1} bovine est totalement phosphorylée sur ses huit sites potentiels de phosphorylation. Elle peut posséder également un neuvième résidu de sérine phosphorylé contenu dans le motif Ser^{41}-Xxx-Asp^{43} (Farrell Jr. *et al.,* 2004) donnant alors une protéine anciennement connue sous le nom de CN-α_{s0}. Malgré ce fait, la phosphorylation de la CN-α_{s1} bovine ne peut pas être considérée comme variable. Chez les autres ruminants, le degré de phosphorylation est variable comme dans le cas de la CN-α_{s1} équine étudiée. La CN-α_{s1} caprine possède de 7P à 10P (Ferranti *et al.,* 1997) et la CN-α_{s1} ovine, de 5P à 11P (Chianese *et al.,* 1996) selon les variants génétiques considérés. La CN-α_{s1} humaine quant à elle, possède neuf sites potentiels de phosphorylation (Johnsen *et al.,* 1995). Cependant, actuellement il n'a été identifié que quatre sites phosphorylés (Ser^{18}, Ser^{26}, Ser^{73} et Ser^{75} ; Kjeldsen *et al.,* 2007). La CN-α_{s1} équine semble donc, de par son haut degré de phosphorylation, plus se rapprocher de celle du lait des ruminants que de celle du lait humain bien que la jument soit tout comme l'Homme une espèce monogastrique.

5. Conclusions et perspectives

Ce travail sur la CN-α_{s1} équine nous a orienté vers une méthode de fractionnement du caséinate de sodium par chomatographie FPLC d'interactions hydrophobes. Cela a permis l'acquisition de plus grandes quantités de caséines par rapport à la technique HPLC en phase inversée. Comme avec cette dernière technique, la CN-κ co-élue avec la CN-α_{s1} lors de la séparation HIC, ces deux protéines doivent donc posséder une hydrophobicité du même ordre de grandeur.

L'analyse du comportement de la CN-κ équine par chromatographie d'affinité sur WGA et sucWGA immobilisées a permis une approche quant à sa structure glycannique. Elle comporterait des résidus NeuAc mais ne contiendrait pas de résidu GlcNAc. Elle aurait ainsi une structure glycannique proche de celle de la CN-κ bovine et très différente de la CN-κ humaine. Des études complémentaires de spectrométrie de masse ou de résonance magnétique nucléaire seraient à envisager pour déterminer de manière plus précise la nature et la localisation de ces résidus glycanniques.

Pour permettre la caractérisation des différents variants post-transcriptionnels, il s'est avéré nécessaire de recourir à une déphosphorylation quasi-complète de la CN-$\alpha_{s1.}$ Pour cela, il nous a fallu mettre au point une méthode de déphosphorylation en deux étapes. En effet, les méthodes classiquement utilisées pour déphosphoryler les caséines, soit par la phosphatase alcaline, soit par la phosphatase acide, n'ont permis qu'une déphosphorylation partielle de la CN-α_{s1} équine. La méthode mise au point dans ce travail pour déphosphoryler la CN-α_{s1} consiste à utiliser dans un premier temps la phosphatase alcaline puis, dans un second temps, la phosphatase acide. L'inversion de l'ordre d'utilisation des phosphatases est moins efficace, car une déphosphorylation partielle est alors obtenue comme dans le cas de l'utilisation d'une seule phosphatase.

Cette étude nous a permis de mettre en évidence 36 isoformes différentes de CN-α_{s1} dans le lait de juments Haflinger. Pour la première fois, la CN-α_{s1} complète ainsi que les variants d'épissage de site cryptique impliquant le résidu Gln^{91} et le variant d'épissage impliquant l'exon 14 ont été mis en évidence dans le lait équin de petit mélange ; les autres variants d'épissage (Δ7 et Δ7,14) ayant déjà été trouvés dans le lait de poney gallois (Miranda *et al.,* 2004). Le degré de phosphorylation de la CN-α_{s1} équine varie de 2P à 6P ou 8P selon l'absence ou la présence de la région codée par l'exon 7. Les 36 différentes isoformes caractérisées sont récapitulées ci-dessous :

- CN-α_{s1} 2P, 4P, 5P, 6P, 7P, 8P
- CN-$\alpha_{s1}^{\Delta Q}$ 2P, 5P, 7P, 8P
- CN-$\alpha_{s1}^{\Delta 14}$ 4P, 5P, 6P, 7P, 8P
- CN-$\alpha_{s1}^{\Delta 14,Q}$ 7P, 8P
- CN-$\alpha_{s1}^{\Delta 7}$ 2P, 3P, 4P, 5P, 6P
- CN-$\alpha_{s1}^{\Delta 7,Q}$ 2P, 3P, 4P, 5P, 6P
- CN-$\alpha_{s1}^{\Delta 7,14}$ 2P, 3P, 4P, 5P, 6P
- CN-$\alpha_{s1}^{\Delta 7,14,Q}$ 2P, 3P, 5P, 6P

Afin de caractériser plus précisément la structure des isoformes de CN-α_{s1}, il sera intéressant d'identifier les sites de phosphorylation de chacun des variants. Cependant, ce travail semble pour le moment assez difficile à entreprendre étant donné la complexité de cette protéine. En effet, le fractionnement par FPLC d'échange d'ions nous a permis de préparer des quantités appréciables de variants afin de pouvoir les identifier par spectrométrie de masse. Néanmoins, chaque sous-fraction demeure complexe et ne peut permettre d'envisager une analyse MS/MS pour localiser les sites de phosphorylation directement. De même, le prélèvement quantitatif des spots obtenus par séparation 2-DE de chaque sous-fraction semble difficile, car plusieurs isoformes peuvent co-migrer dans un même spot. De plus, beaucoup de variants sont en trop faibles quantités pour être identifiés par spectrométrie de masse à partir d'un spot prélevé. Les efforts devront se focaliser, dans un premier temps, sur les variants majoritaires que nous chercherons à purifier par chromatographie de façon plus poussée.

PARTIE 2

Cartographie bidimensionnelle de la caséine β équine établie après séparation de ses variants de phosphorylation - Étude de sa désamidation spontanée.

La CN-β équine (Figure 44) possède un degré de phosphorylation variant de 3P à 7P, le variant de phosphorylation majoritaire étant celui contenant 5P (Girardet *et al.,* 2006). La CN-β possède deux variants d'épissage identifiés au niveau protéique : un variant impliquant l'épissage de l'exon 5 (CN-$\beta^{\Delta 5}$) et un variant impliquant l'utilisation d'un site cryptique d'épissage à l'intérieur de l'exon 7 (low-M_r CN-β ; Lenasi *et al.,* 2003 ; Miranda *et al.,* 2004 ; Lenasi *et al.,* 2006 ; Miclo *et al.,* 2007). Dans le lait de jument Haflinger, la présence du variant CN-$\beta^{\Delta 5}$ n'a pas encore été démontrée mais elle est suspectée (Girardet *et al.,* 2006). L'ARNm d'un autre variant d'épissage ne contenant pas les exons 5 et 8 (CN-$\beta^{\Delta 5,8}$) a été mis en évidence par Lenasi *et al.* (2006) chez la jument lippizane mais ce variant n'a toutefois jamais été retrouvé au niveau protéique dans le lait de jument.

Le résidu Asn^{135} de la CN-β équine peut subir un phénomène de désamidation non-enzymatique spontané lors d'une incubation à 37°C et à un pH basique ou proche de la neutralité (Girardet *et al.,* 2006). Ce résidu est situé dans une séquence Asn-Gly qui est susceptible de se désamider facilement, comme cela a été montré pour de nombreuses protéines contenant cette séquence et soumises à des conditions physiologiques de température et de pH (Robinson et Robinson, 2004).

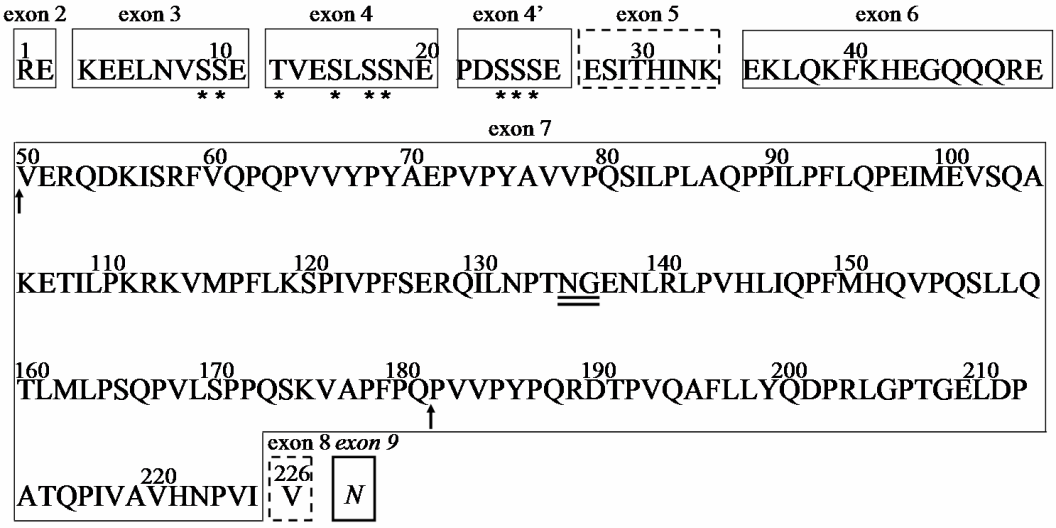


Figure 44 : Structure exonique de la CN-β équine déterminée par Miranda *et al.* (2004). Les sites potentiels de phosphorylation sont indiqués à l'aide d'un astérisque selon Miclo *et al.* (2007). Le site de désamidation non-enzymatique identifié par Girardet *et al.* (2006) est souligné d'un double trait. Les exons pouvant être alternativement épissés sont entourés en pointillés. Les sites d'épissage impliqués dans la production du variant low M_r CN-β sont indiqués par une flèche (perte de la région 50-181). L'exon 9, qui n'est traduit que lors de l'épissage de l'exon 8 est indiqué en italique.

Les profils électrophorétiques de PAGE-urée et de 2-DE de la CN-β équine restent difficiles à interpréter à cause de la microhétérogénéité moléculaire de cette protéine. En effet, il existe des variants d'épissage alternatif, des variants de phosphorylation et tous ces variants pourraient en outre subir une modification chimique de désamidation non-enzymatique, transformant un résidu Asn par un résidu Asp ou iso-Asp, ce qui générerait alors une charge négative supplémentaire. Ainsi, un spot détecté par 2-DE ou une bande visible par PAGE-urée pourraient correspondre à un mélange de différentes formes de la CN-β. Il est donc nécessaire d'élucider la complexité structurale de la CN-β équine.

Afin d'étudier la complexité des profils électrophorétiques de la CN-β de lait de jument Haflinger, la séparation des différentes isoformes de phosphorylation de la CN-β équine s'avère nécessaire et une méthode doit donc être mise au point pour séparer chaque variant phosphorylé. De plus, la séparation de ces isoformes nous permettra plus tard de déterminer la localisation des différents sites de phosphorylation de cette caséine, condition indispensable avant d'entamer toute étude sur l'influence du degré et de la séquence de phosphorylation sur la chélation de minéraux et/ou oligo-élements d'intérêt nutritionnel. L'isolement de chaque variant phosphorylé nous permettra également de déterminer si la désamidation non-enzymatique est susceptible d'intervenir dans le lait ou si elle intervient lors des étapes de préparation de la CN-β.

1. Purification de la CN-β équine par chromatographie HIC

La CN-β équine est obtenue à partir du fractionnement du caséinate de sodium par chromatographie HIC comme expliqué dans la partie précédente. Cette méthode permet rapidement l'obtention d'une grande quantité de CN-β , puisque lors de chaque fractionnement, 100 mg de caséinate de sodium sont déposés sur la colonne semi-préparative. Ainsi, l'injection d'un total de 8,3 g de caséinate de sodium nous a permis d'obtenir environ 3 g de CN-β. Le fractionnement HIC permet l'obtention de 4 fractions F1 à F4, la fraction F1 contenant le variant low-M_r CN-β et la fraction F2, la CN-β complète (Figure 28).

Dans notre étude, nous nous intéresserons uniquement à la fraction F2 contenant la CN-β complète. Sur le profil 2-DE de cette fraction, deux groupes de spots de masses moléculaires apparentes de 32 kDa et 30 kDa sont visibles (Figure 45). Le groupe à 32 kDa est constitué de 5 spots qui ont été identifiés par Girardet *et al.* (2006) comme étant les

variants de phosphorylation 3P à 7P de la CN-β complète (masse théorique de l'apoforme : 25511 Da). Cependant, l'attribution précise des spots à chacun des variants n'est, pour l'heure, pas réalisée. Girardet *et al.* (2006) ont émis l'hypothèse que le groupe de spots à 30 kDa pourrait correspondre à des variants de phosphorylation de la CN-$\beta^{\Delta 5}$ (masse théorique de l'apoforme : 24588 Da).

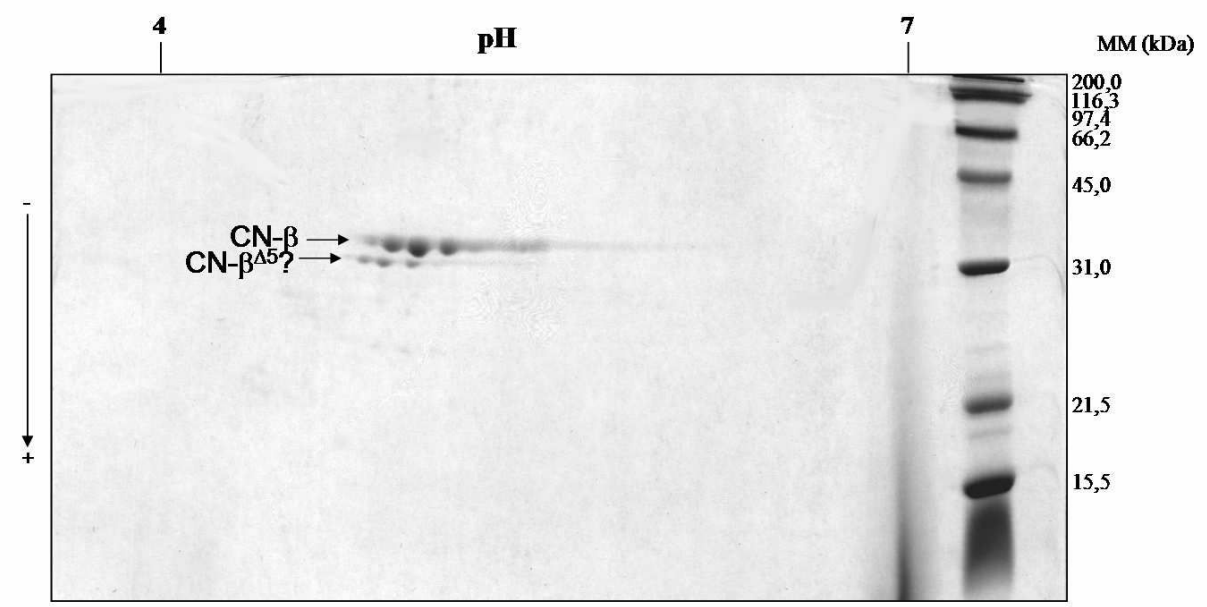


Figure 45 : Electrophorèse bidimensionnelle (première dimension : isoélectrofocalisation, deuxième dimension : PAGE-SDS) de la fraction F2 issue de la séparation du caséinate de sodium par FPLC d'interactions hydrophobes. Une quantité protéique de 50 µg a été déposée sur une bandelette constituée d'un gel à 4% de polyacrylamide contenant des ampholines créant un gradient de pH allant de 4 à 7. Les protéines ont été révélées dans le gel par le bleu de Coomassie. Les deux groupes de spots sont identifiés selon Girardet *et al.* (2006) mais l'attribution du variant CN-$\beta^{\Delta 5}$ est hypothétique. MM, masse moléculaire ; Δ5, délétion de l'exon 5 due à l'épissage de cet exon.

2. Purification chromatographique des variants de phosphorylation de la CN-β équine

2.1. Essai de purification par chromatographie d'affinité aux métaux immobilisés (IMAC)

La chromatographie d'affinité aux métaux immobilisés ou IMAC a été introduite pour la première fois par Porath *et al.* (1975) pour le fractionnement des protéines de sérum humain grâce à des ions Cu^{2+} et Zn^{2+} immobilisés. Le principe est basé sur l'affinité spécifique des protéines pour certains ions métalliques qui sont immobilisés sur une matrice hydrophile par chélation. Le mécanisme d'interaction entre la protéine et l'ion métallique est la formation d'une liaison de coordination entre un doublet électronique d'un atome appartenant à un résidu d'acide aminé de la protéine et une orbitale vide d'un ion métallique de transition accepteur (Fe^{3+}, Ga^{3+}, Cu^{2+}, Co^{2+}, Zn^{2+}, Ni^{2+}). L'atome donneur d'électrons de la protéine peut être un atome d'azote de la chaîne latérale imidazole d'un résidu His, un atome d'oxygène du groupement phosphate d'un résidu Ser phosphorylé ou encore du groupement carboxylique d'un résidu acide. La séparation IMAC utilise l'affinité des charges négatives des groupements phosphates des phosphoprotéines pour des ions métalliques chargés positivement (Fe^{3+}, Ga^{3+}, Cu^{2+}, Co^{2+}, Zn^{2+}, Ni^{2+}). L'ion le plus souvent utilisé est Fe^{3+}, car c'est celui qui a l'une des plus fortes affinités pour les phosphoprotéines et phosphopeptides (Anderson et Porath, 1986). L'affinité pour les ions métalliques est d'autant plus importante que le nombre de groupements phosphates est grand (Muszyńska *et al.,* 1986 ; Lin *et al.,* 2000). De plus, l'ion ferrique immobilisé agit également comme un échangeur d'ions permettant l'élution sélective pH-dépendante des protéines selon leur degré de phosphorylation (Erickson et Johnson, 1993).

D'après les données bibliographiques, cette technique nous a semblé appropriée pour séparer les différents variants de phosphorylation de la CN-β équine avec l'ion Fe^{3+} immobilisé. Pour cela, la CN-β a été solubilisée dans une solution d'acide acétique 0,1 M à pH 3,0 contenant 30% d'acétonitrile et déposée sur une colonne préalablement chargée avec des ions ferriques à pH 3,0. Plusieurs méthodes d'élution ont été testées : (i) application d'un gradient de pH de 7,0 à 8,5 avec des tampons phosphate de sodium 0,02 M à pH 7,0, pH 7,5, pH 8 et pH 8,5, (ii) application d'un gradient de pH 6,0 à 7,0 avec un tampon phosphate de sodium 0,02 M pH 6,0 et un tampon phosphate de sodium 0,02 M pH 7,0, (iii) application

d'un gradient linéaire de concentration avec un tampon phosphate de sodium 0,1 M pH 7,5 et un tampon phosphate de sodium 1 M pH 7,5. Cependant, ces essais n'ont pas conduit à séparer les différentes isoformes de phosphorylation de façon satisfaisante. En effet, la CN-β s'adsorbe fortement à la colonne mais les variants de phosphorylation sont systématiquement élués dans une seule fraction.

2.2. Fractionnement de la CN-β par chromatograhie IEC

Aoki *et al.* (1992) ont utilisé la chromatographie d'échange d'anions pour séparer les différents variants de phosphorylation de la CN-β humaine. Nous avons donc utilisé cette méthode pour essayer de séparer les différents variants de phosphorylation de la CN-β équine selon leurs charges négatives.

La fraction F2 obtenue par HIC est sous-fractionnée par IEC et cinq sous-fractions (F2.1 à F2.5) sont obtenues (Figure 46).

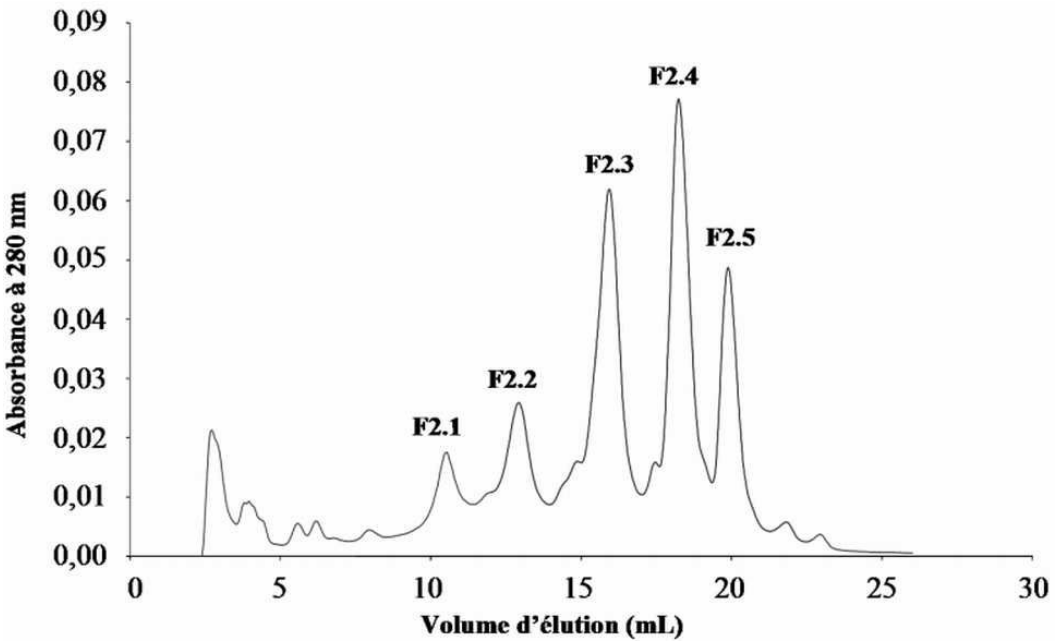


Figure 46 : Fractionnement par FPLC d'échange d'ions du mélange de la fraction F2 précédemment obtenue par FPLC d'interactions hydrophobes du caséinate de sodium. Le profil présenté ici est obtenu avec une une colonne TSK DEAE 5-PW analytique. Des quantités protéiques de 3 mg sont déposées sur la colonne. Un gradient linéaire croissant de 0,08–0,28 M de NaCl à pH 8,0 est réalisé pendant 70 min. Le débit est de 0,5 mL.min^{-1} et l'absorbance est mesurée à 280 nm. Le profil obtenu avec une colonne TSK DEAE 5-PW semi-préparative dans les mêmes conditions est similaire.

A partir d'un bilan massique effectué sur des « pools » issus de plusieurs collectes, l'ordre d'importance des sous-fractions est le suivant : F2.3 (38,1%) > F2.4 (30,1%) > F2.5 (18,9%) > F2.2 (7,4%) > F2.1 (5,5%). Les pourcentages donnés correspondent aux masses des lyophilisats des sous-fractions collectées après dialyse. L'analyse par PAGE-urée montre une seule bande visible, après révélation au bleu de Coomassie, dans chacune des sous-fractions (Figure 47), ce qui signifierait que chaque bande correspondrait à un variant de phosphorylation de la CN-β. La distance de migration dans le gel de la bande correspondant à une sous-fraction est corrélée avec le temps de rétention observé par IEC pour cette sous-fraction.

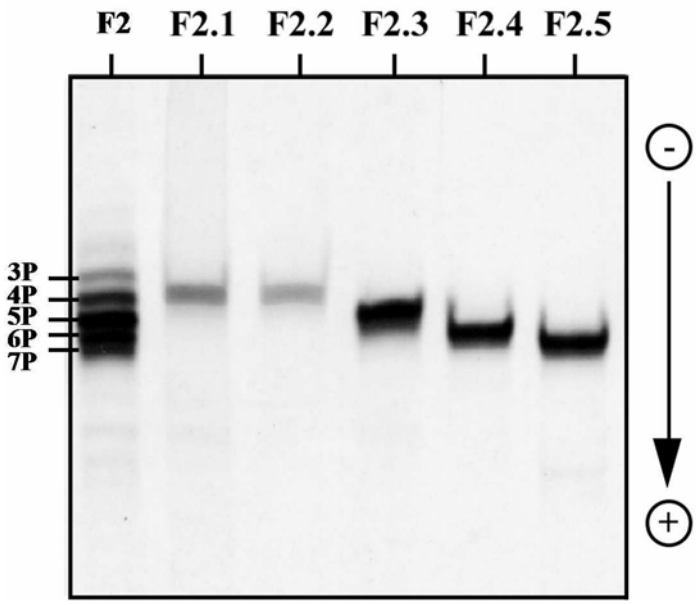


Figure 47 : Caractérisation électrophorétique (PAGE-urée) des sous-fractions F2.1 à F2.5 purifiées par FPLC d'échange d'ions. Une quantité de 40 µg de protéines est déposée dans le gel. Les protéines sont révélées par le bleu de Coomassie. Les bandes sont attribuées selon le degré de phosphorylation croissant de 3P à 7P déterminé par spectrométrie de masse de chaque sous-fraction. F2, fraction F2 contenant la CN-β et obtenue par FPLC d'interactions hydrophobes ; XP, présence de X groupements phosphate.

Par chromatographie d'échange d'anions, les constituants les plus chargés négativement éluent en fin de gradient, c'est-à-dire pour une force ionique croissante de la phase. Ainsi, dans notre cas, plus une protéine est phosphorylée plus elle est chargée négativement et plus elle devrait se désorber tardivement et avoir une distance de migration électrophorétique, déterminée par PAGE-urée, importante. Nous pouvons raisonnablement émettre l'hypothèse que les sous-fractions F2.1, F2.2, F2.3, F2.4 et F2.5 correspondent aux

isoformes 3P, 4P, 5P, 6P et 7P, respectivement. De plus, cette hypothèse serait en accord avec les résultats de Girardet *et al.* (2006) qui ont mis en évidence la prédominance de l'isoforme 5P suivi de l'isoforme 6P, ce qui semblerait être le cas ici, si nous supposons que F2.3 (38,1% du total en masse) contient l'isoforme 5P et F2.4 (30,1% du total en masse), l'isoforme 6P.

2.3. Caractérisation des isoformes de CN-β par LC/ESI-MS

Afin d'identifier les constituants des sous-fractions F2.1 à F2.5, celles-ci ont été analysées par LC/ESI-MS (Figure 48, Tableau 6). Les sous-fraction F2.1 et F2.2 contiennent principalement l'isoforme CN-β 4P tandis que l'isoforme CN-β 3P est retrouvé à l'état de traces dans F2.1. Les sous-fractions F2.3, F2.4 et F2.5 contiennent majoritairement les isoformes CN-β 5P, CN-β 6P et CN-β 7P, respectivement. Ces résultats confirment que l'ordre d'élution observé par IEC est bien corrélé à un degré croissant de phosphorylation de la CN-β. Nous pouvons remarquer cependant une légère contamination des sous-fractions F2.3, F2.4 et F2.5 par le contenu de la sous-fraction précédente (par exemple : F2.3 contaminé par des traces de CN-β 4P, composé principal de F2.2). Ce phénomène est probablement dû à une résolution chromatographique non optimale.

Le spectre de masse de chaque sous-fraction laisse apparaître des masses moléculaires caractéristiques possédant une différence de 923 Da ± 1 Da (± 7 Da dans le cas de la fraction F2.5) avec la masse de l'isoforme majoritaire de chacune des sous-fractions. Cette différence correspond exactement à la masse moléculaire de la région codée par l'exon 5. Ainsi, nous pouvons attribuer ces différentes masses aux isoformes de phosphorylation 4P à 7P du variant d'épissage CN-$\beta^{\Delta 5}$. Seule l'isoforme 3P ne possédant pas la région codée par l'exon 5 n'a pas pu être détectée par spectrométrie de masse. Nous constatons que les variants de phosphorylation de CN-$\beta^{\Delta 5}$ ne sont présents qu'en faible quantité par rapport à ceux de la protéine complète. Etant donné que l'isoforme CN-β 3P est peu abondante, il semble raisonnable de penser que l'isoforme CN-$\beta^{\Delta 5}$ 3P existe, mais qu'elle est en trop faible quantité pour être détectée.

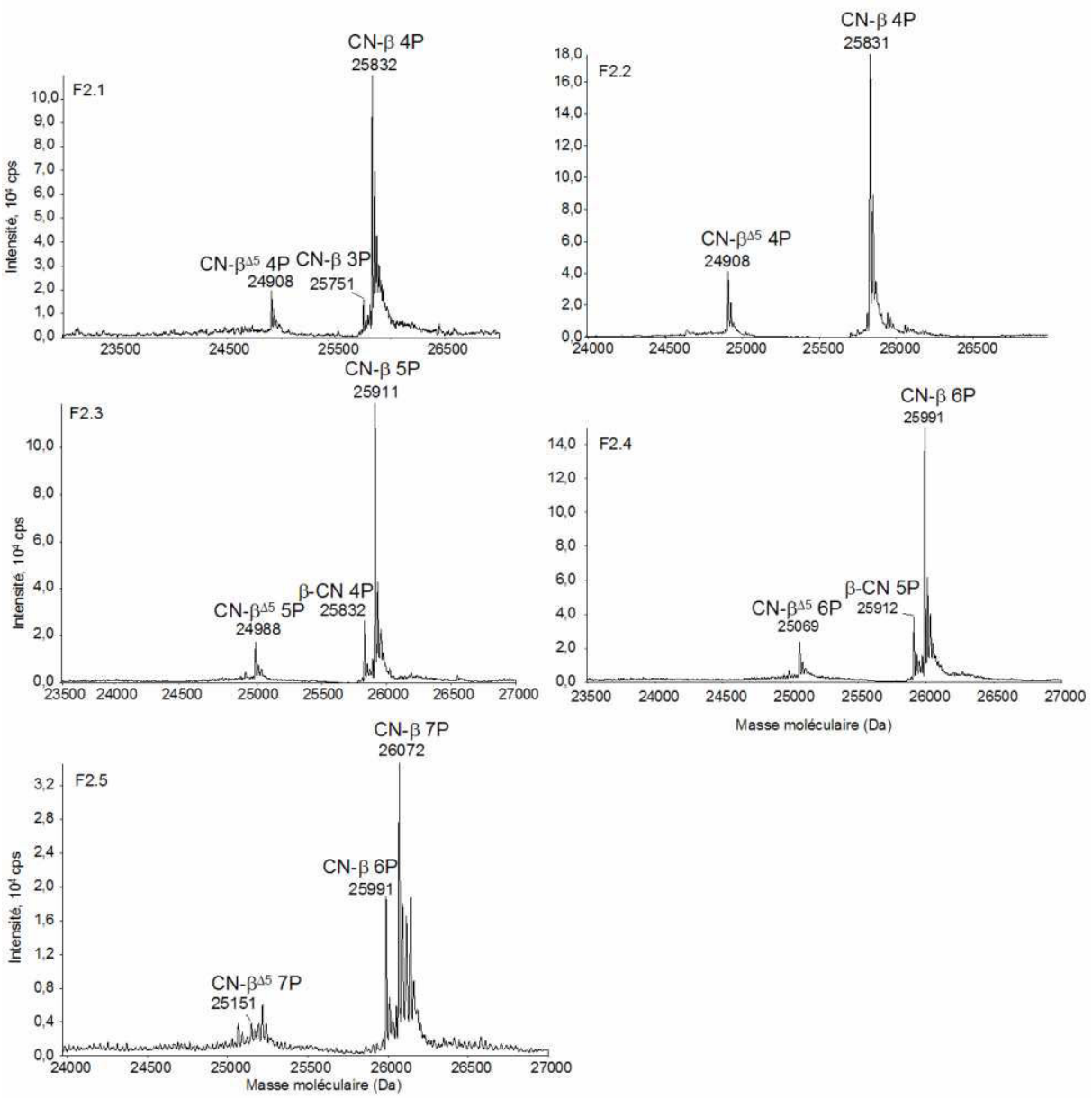


Figure 48 : Masses reconstruites déterminées par chromatographie liquide couplée à la spectrométrie de masse à ionisation électrospray (LC/ESI-MS) des sous-fractions F2.1 à F2.5 obtenues par FPLC d'échange d'ions. ΔX, délétion de la séquence de résidus d'acides aminés codés par l'exon X ; XP, présence de X groupements phosphates ; cps, coup par seconde.

Tableau 6 : Masses moléculaires des différentes sous-fractions de CN-β equine obtenues par chromatographie d'échange d'ions, déterminées par LC/ESI-MS sur l'appareil Q/Star XL.

Sous-fraction	Masse observée[a]	Identification[b]	Masse théorique[c]
F2.1	25832	CN-β 4P	25831
	25751	CN-β 3P	25751
	24908	CN-β $^{\Delta5}$ 4P	24908
F2.2	25831	CN-β 4P	25831
	24908	CN-β $^{\Delta5}$ 4P	24908
F2.3	25911	CN-β 5P	25911
	25832	CN-β 4P	25831
	24988	CN-β $^{\Delta5}$ 5P	24988
F2.4	25991	CN-β 6P	25991
	25912	CN-β 5P	25911
	25069	CN-β $^{\Delta5}$ 6P	25068
F2.5	26072	CN-β 7P	26071
	25991	CN-β 6P	25991
	25151	CN-β $^{\Delta5}$ 7P	25148

[a]Masse ± 1 Da (± 3 Da pour CN-β$^{\Delta5}$ 7P).
[b]ΔX: délétion de la région code par l'exon X; CN, caséine; P, groupements phosphates.
[c]Masse moyenne calculée avec ProtParam tool (expasy.org/tools/protparam.html).

Pour un degré de phosphorylation donné, la CN-β complète et la CN-β$^{\Delta5}$ co-éluent dans la même sous-fraction chromatographique. Sur le profil de PAGE-urée (Figure 47), les isoformes de CN-β$^{\Delta5}$ ne sont pas mises en évidence et pourraient co-migrer dans une bande unique avec l'isoforme de CN-β complète présentant le même degré de phosphorylation. L'épissage de l'exon 5 entraîne chez la protéine traduite la perte d'une charge négative (Glu^{27}) et d'une charge positive (Lys^{34}), ce qui ne change pas la charge globale de la protéine à pH 8,9, c'est-à-dire au pH de la migration électrophorétique. Ceci expliquerait la co-migration des deux variants dans une même bande par PAGE-urée et leur co-élution dans une même sous-fraction obtenue par IEC. Une autre hypothèse serait que les isoformes CN-β$^{\Delta5}$ sont en très faibles quantités et ne seraient donc pas révélées par la coloration au bleu de Coomassie.

Ainsi, la séparation des différents degrés de phosphorylation de la CN-β par IEC a permis la mise en évidence au niveau protéique de l'existence du variant CN-$\beta^{\Delta5}$ dans le lait de jument Haflinger comme l'avait suspecté Girardet *et al.* (2006). Cependant, le variant CN-$\beta^{\Delta5,8}$ dont l'ARNm a été mis en évidence par Lenasi *et al.* (2006) dans le lait de jument Lipizzane, qui ne diffère du variant CN-$\beta^{\Delta5}$ que par le dernier résidu d'acide aminé (V^{226} devient N^{226}, ce qui représente une augmentation de la masse moléculaire de 15 Da), n'a pas été identifié dans ce travail.

3. Etude de la désamidation non-enzymatique de la CN-β équine

Le phénomène de désamidation des protéines et des peptides concerne les résidus Asn et Gln. Il s'agit d'un phénomène courant qui a été rapporté chez plus de 200 types de protéines et peptides naturels (Robinson et Robinson, 2004). Le résidu Gln peut être désamidé par voie enzymatique par l'action de glutaminases ou de transglutaminases. Cependant et à notre connaissance, il n'existe pas d'exemple connu de désamidation enzymatique du résidu Asn au sein d'un peptide ou d'une protéine et la désamidation non-enzymatique est l'unique voie de désamidation pour ce résidu (Robinson et Robinson, 2004). La réaction de désamidation non-enzymatique se déroule en deux grandes étapes (Figure 49).

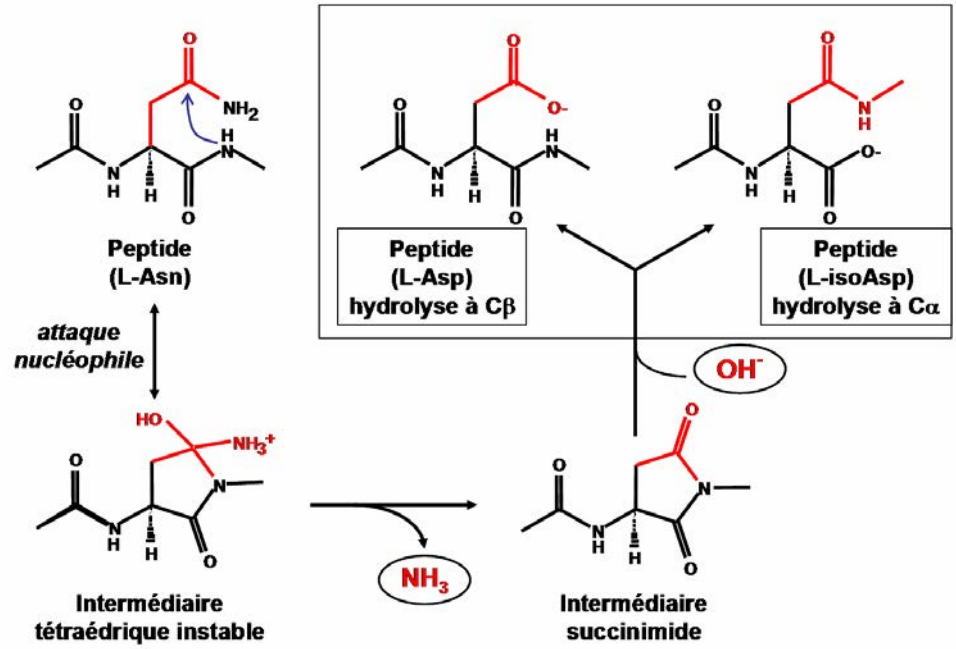


Figure 49 : Schéma du mécanisme de la réaction de désamidation non-enzymatique d'un résidu Asn à pH neutre ou basique (Xie et Schowen, 1999).

La première consiste en la formation d'un cycle imide (succinimide) résultant de l'attaque nucléophile du doublet électronique d'un atome d'azote de la chaîne carbonée principale sur l'atome de carbone du groupement amide de la chaîne latérale du résidu Asn. Le succinimide formé est un intermédiaire de réaction stable qui peut être observé dans une zone de pH variant de 4 à 5, mais qui se décompose aussi vite qu'il s'est formé pour des pH plus élevés. La deuxième étape consiste en l'hydrolyse du cycle imide, ce qui génère au sein de la chaîne peptidique la formation d'un résidu L-isoAsp (75% de probabilité de formation) ou L-Asp (25% de probabilité de formation ; Xie et Schowen, 1999). L'apparition d'un résidu isoAsp introduit une liaison β-amide branchée sur le groupement β-carboxylique du résidu isoAsp, ce qui provoque un changement dans la structure des peptides et des protéines qui subissent une désamidation. Ainsi, la désamidation interviendrait comme une horloge moléculaire dans la régulation de processus physiologiques, notamment le « turn-over » des protéines (comme par exemple le cytochrome c ou l'aldolase ; Robinson et Robinson, 2004). Pour beaucoup de protéines, la formation de résidu isoAsp peut aussi conduire à une perte d'activité (comme c'est le cas pour la calmoduline, le facteur de croissance épidermique et la ribonucléase) tandis que l'activité d'autres protéines ne semble pas être affectée (comme pour l'hormone de croissance humaine par exemple ; Teshima *et al.,* 1995).

Il existe une enzyme de réparation dans la cellule, la PIMT (L-isoaspartyl méthyltransférase), qui permet de limiter l'accumulation des résidus isoAsp générés lors du phénomène de désamidation non-enzymatique (Chavous *et al.,* 2001). En effet, l'accumulation incontrôlée de ces résidus isoAsp pourrait induire une forte stimulation du système immunitaire. Ceci peut mener à l'apparition de maladies auto-immunes (c'est le cas du lupus érythémateux) lorsqu'il s'agit d'une molécule du soi qui est désamidée (Mamula *et al.,* 1999) ou à l'apparition d'allergies lorsqu'il s'agit d'une molécule étrangère (la maladie cœliaque est due à la désamidation enzymatique excessive de la gliadine du blé entraînant une intolérance au gluten ; Van der Wal *et al.,* 1998 ; Vader *et al.,* 2002).

Quelques protéines laitières possèdent des sites de désamidation potentiels. Ainsi, la lactoferrine humaine peut subir une désamidation non-enzymatique (Belizy *et al.,* 2001). La modification de structure faisant suite à la désamidation de cette protéine entraîne une augmentation de ses propriétés anti-oxydantes *via* la stimulation d'une interaction non spécifique de la protéine désamidée avec les ions ferriques, ce qui confère une protection supplémentaire pour le nouveau-né. L'α-lactalbumine équine possède un site de désamidation potentiel en position Asn^{45}-Gly^{46} qui se désamide à pH 7,4 et à 37°C (Girardet *et al.,* 2004). Il se pourrait que cette protéine modifiée soit à l'origine des quelques cas d'allergie aux

protéines lactosériques équines décrits (Gall *et al.*, 1996). De même, la CN-β équine, possédant un résidu Asn, contenu dans une séquence Asn-Gly connue pour pouvoir être désamidée facilement (Robinson et Robinson, 2004), pourrait être potentiellement allergisante. Il semble donc important d'étudier le phénomène de désamidation non-enzymatique de la CN-β équine afin de déterminer notamment si cette désamidation non-enzymatique intervient dans le lait et pourrait conférer un potentiel allergisant au lait de jument.

La cinétique de désamidation des fonctions amides dépend en partie des conditions expérimentales telles que la température, le pH, la force ionique et les ions composant le tampon (Robinson, 2002). Par exemple, un tampon phosphate de sodium 150 mM augmente de 2 à 3 fois le taux de désamidation comparé à un tampon Tris-HCl 150 mM à pH 7,4 et à 37°C (Robinson et Robinson, 2001). Dans ce travail, une première expérience de désamidation a été réalisée sur les isoformes principales de la CN-β, les variants de phosphorylation CN-β 5P, CN-β 6P et CN-β 7P. Pour cela, les sous-fractions F2.3, F2.4 et F2.5 ont été solubilisées dans un tampon phosphate de sodium à 150 mM, pH 7,4 et incubées à 37°C. Au bout de 48 h d'incubation, les échantillons ont été analysés par PAGE-urée (Figure 50).Une nouvelle bande est apparue pour chaque échantillon avec une distance de migration légèrement plus importante que celle de la bande originelle. Chaque nouvelle bande correspond vraisemblablement à la fraction de l'isoforme considérée ayant subi une désamidation. Le profil électrophorétique met également en évidence la présence de nombreuses autres bandes, résultant probablement d'une dégradation des protéines. En effet, il est rapporté que la désamidation des chaînes latérales tend à ouvrir la structure des protéines et par conséquent celles-ci deviennent plus sensibles à la protéolyse (Robinson et Robinson, 2004).

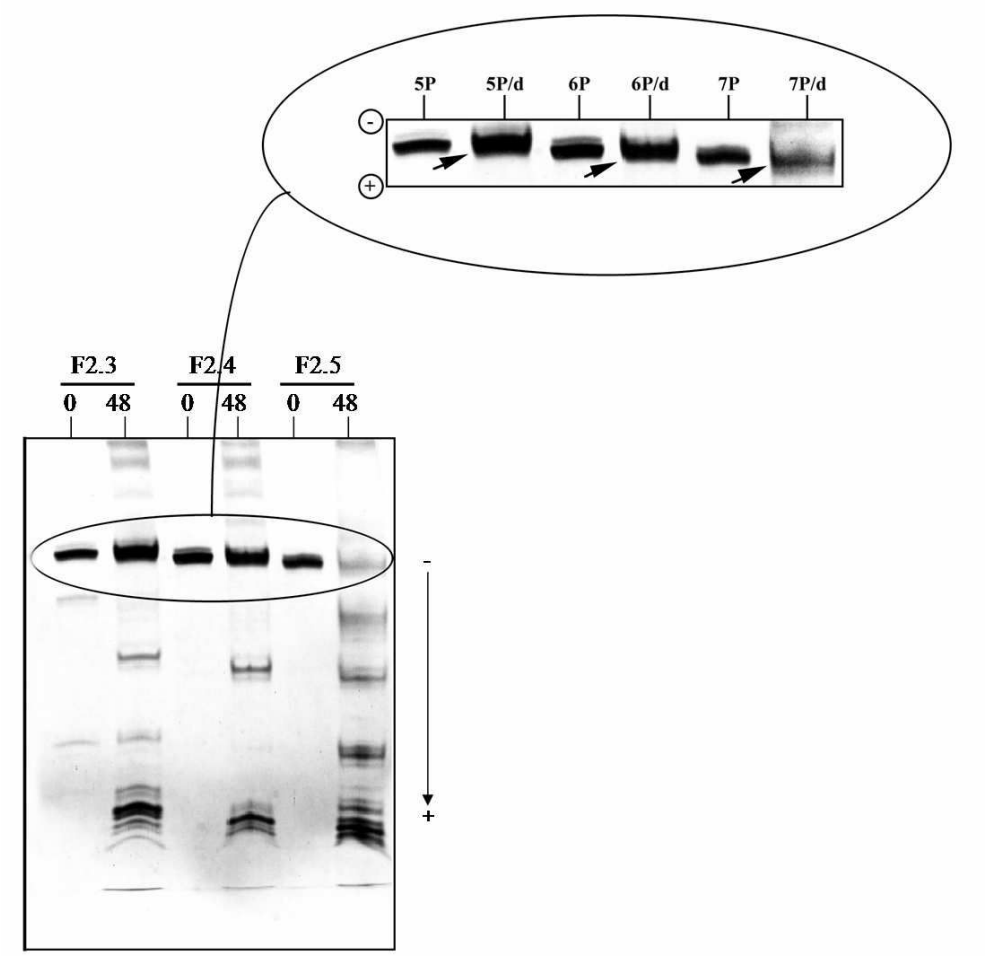


Figure 50 : Suivi électrophorétique (PAGE-urée) de la réaction de désamidation non-enzymatique des isoformes CN-β 5P, CN-β 6P et CN-β 7P contenues dans les sous-fractions F2.3, F2.4 et F2.5, respectivement. Les échantillons ont été incubés à 37°C dans un tampon phosphate de sodium 150 mM à pH 7,4 contenant 0,02% d'azoture de sodium. La réaction se déroule pendant 48 h et des prélèvements sont effectués initialement (0) et au bout de 48 h (48). Les protéines sont révélées dans le gel par le bleu de Coomassie. La flèche indique les bandes des variants désamidés. CN, caséinate de sodium ; XP, présence de X groupements phosphate ; d, désamidé.

Une seconde expérience est alors réalisée en présence de fortes concentrations d'agents bactériostatiques et d'inhibiteurs de protéases. Le pH a été augmenté à 8,4 pour favoriser la réaction, car un pH plus basique accélère la désamidation des résidus Asn (Robinson et Robinson, 2004). Les sous-fractions F2.3 et F2.4 sont incubées dans ces nouvelles conditions. Trois prélèvements sont effectués : juste avant la mise en incubation, au bout de 24 h d'incubation et au bout de 48 h d'incubation. Lors du dépôt des différents prélèvements dans le gel PAGE-urée, l'isoforme CN-β 7P natif contenue dans la sous-fraction F2.5 a également été déposé pour nous servir de repère lors de l'analyse des résultats (Figure 51).

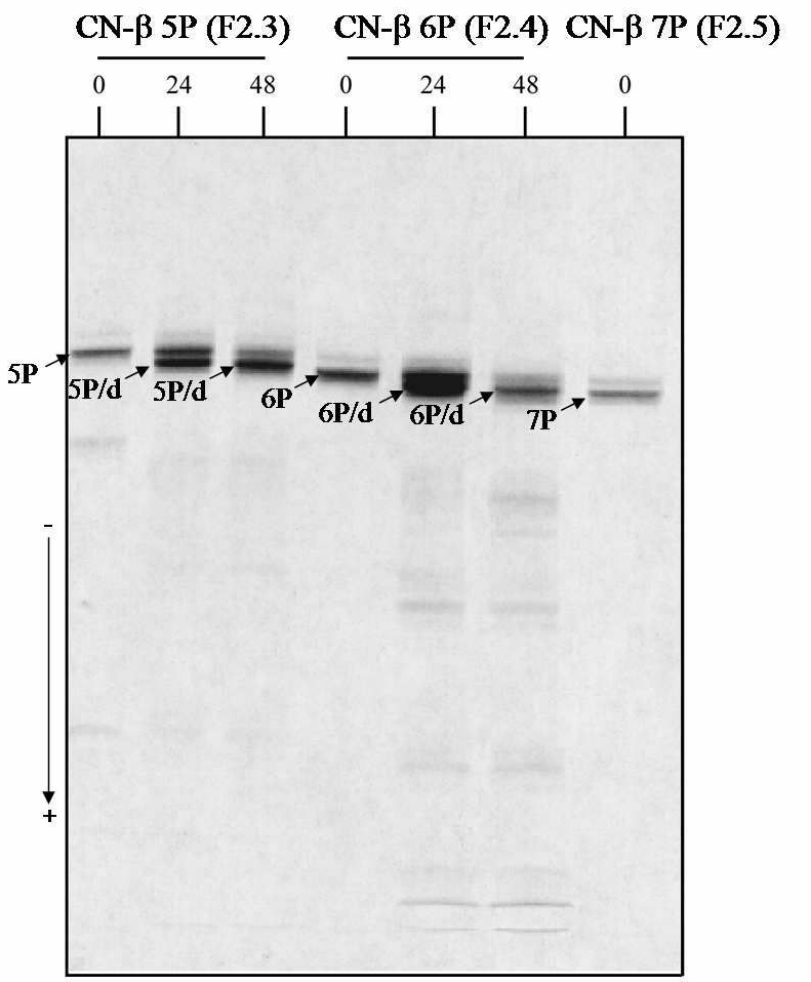


Figure 51 : Suivi électrophorétique (PAGE-urée) de la réaction de désamidation non-enzymatique des isoformes CN-β 5P et CN-β 6P contenues dans les sous-fractions F2.3 et F2.4, respectivement. Les échantillons ont été incubés à 37°C dans un tampon phosphate 150 mM à pH 8,4 contenant 0,02% d'azide de sodium et 10% d'un cocktail d'inhibiteurs de protéases. La réaction se déroule pendant 48 h et des prélèvements sont effectués initialement (0), au bout de 24h (24) et au bout de 48 h (48). Les protéines sont révélées dans le gel par le bleu de Coomassie. XP, présence de X groupements phosphates ; d, désamidé.

Les premiers prélèvements ne montrent qu'une seule bande par PAGE-urée. Au bout de 24 h d'incubation une seconde bande apparaît pour chacun des variants, comme lors de la première expérience. Au bout de 48 h, la plus grande partie des variants de CN-β est désamidée, car c'est la seconde bande qui prédomine. Ces nouvelles bandes générées semblent donc correspondre aux isoformes désamidées des CN-β 5P et CN-β 6P que l'on nommera CN-β 5P/d et CN-β 6P/d. Ces dernières migrent plus rapidement dans le gel de PAGE-urée, car la désamidation génère une charge négative supplémentaire à la protéine. De plus, ces nouvelles bandes ne peuvent pas être confondues avec les bandes des autres isoformes natives. L'isoforme CN-β 5P/d migre effectivement en une bande distincte, qui se situe entre les

bandes de CN-β 5P et de CN-β 6P. De même, l'isoforme CN-β 6P/d migre entre les bandes de CN-β 6P et de CN-β 7P.

Malgré l'utilisation d'un agent bactériostatique et d'inhibiteurs de protéases, une légère protéolyse est observée après 48 h d'incubation.

Les bandes correspondant aux isoformes CN-β 5P/d et CN-β 6P/d sont absentes avant la mise en incubation, ce qui signifie que le processus de désamidation ne se produit pas spontanément dans le lait de jument lors de son stockage. De même, il ne se produit pas pendant les différentes étapes successives de chromatographie et de dialyse bien que des tampons phosphate à un pH relativement basique aient été utilisés. La désamidation semble être initiée par une élévation de la température à 37°C pour altérer de façon sensible la CN-β.

4. Etablissement de la cartographie bidimensionnelle de la CN-β équine

Les variants de phosphorylation de chaque sous-fraction obtenue par chromatographie IEC ne sont pas désamidés lors de leur préparation à partir de caséinate de sodium. De ce fait, l'attribution des spots observée par 2-DE devrait être facilitée. Pour cela, les sous-fractions F2.3 à F2.5 sont analysées par 2-DE (Figure 52).

Au sein de chaque sous-fraction, un groupe de masse moléculaire apparente de 32 kDa et un autre de masse moléculaire apparente de 30 kDa sont observés. Ces deux groupes correspondent respectivement à la CN-β complète et à la CN-$\beta^{\Delta 5}$, la présence du variant d'épissage Δ5 ayant été mise en évidence par spectrométrie de masse. Pour chaque sous-fraction, de nombreux spots avec des masses moléculaires apparentes inférieures à celle de la CN-β complète et de la CN-$\beta^{\Delta 5}$ sont observés reflétant une très probable protéolyse des échantillons. Dans chacun des deux groupes correspondant aux masses moléculaires apparentes 32 et 30 kDa, deux spots bien individualisés sont visibles alors qu'un seul était attendu. Nous émettons alors l'hypothèse qu'un phénomène de désamidation non-enzymatique se produit, ce qui provoque l'apparition d'un spot correspondant à l'isoforme désamidée et l'apparition d'une protéolyse non-enzymatique comme cela a déjà été constaté précédemment par PAGE-urée. Cette désamidation peut s'expliquer par la technique utilisée. En effet, pour réaliser la première dimension de l'électrophorèse 2-DE (isoélectrofocalisation) la bandelette « ReadyStrip » est réhydratée pendant 15 h à 22°C et à pH 7,7 puis la migration dans le gradient de pH se déroule également à 22°C pendant 8 h. Ainsi, l'échantillon reste au

contact d'une solution basique à température ambiante, conditions favorisant le phénomène de désamidation non-enzymatique.

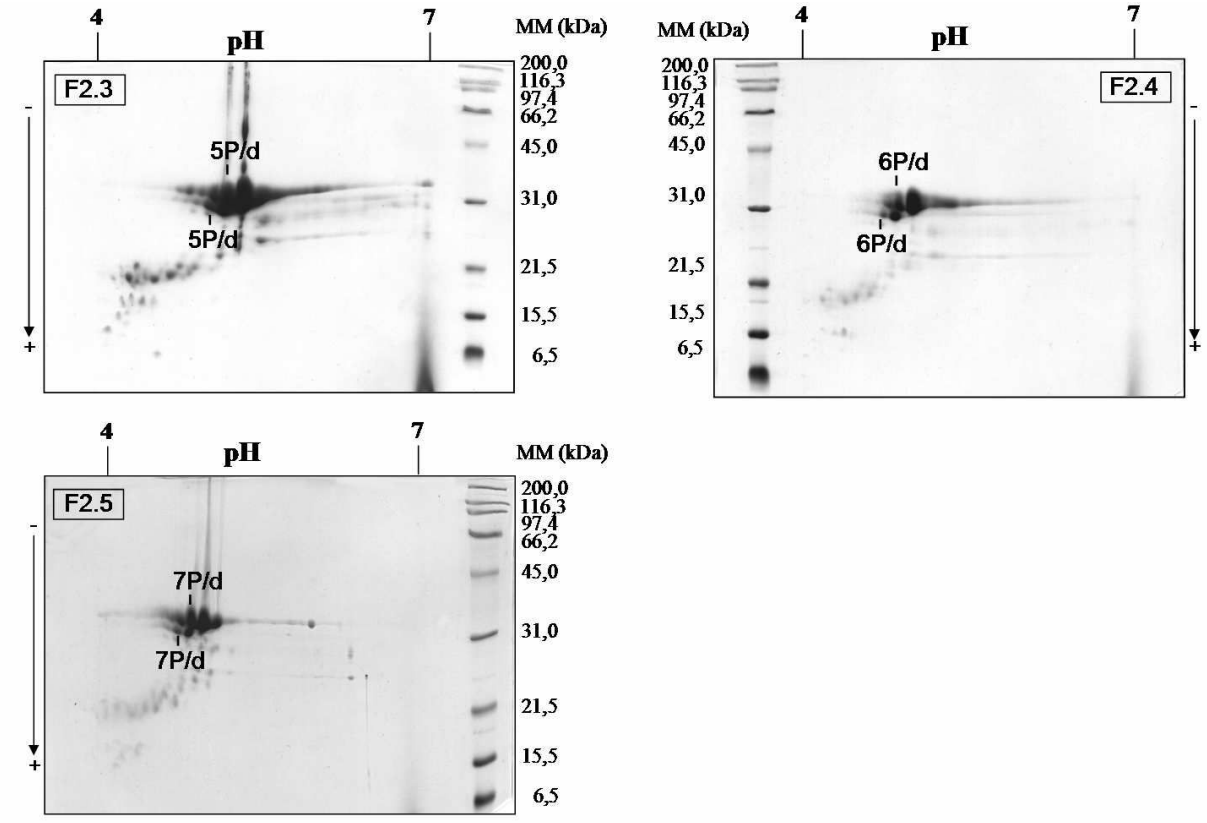


Figure 52 : Electrophorèse bidimensionnelle (première dimension : isoélectrofocalisation, deuxième dimension : PAGE-SDS) des sous-fractions F2.3 à F2.5 obtenues par FPLC d'échange d'ions de la fraction F2. Des quantités protéiques de 50 µg ont été déposées sur des bandelettes constituées d'un gel à 4% de polyacrylamide contenant des ampholines créant un gradient de pH de 4 à 7. La réhydratation des bandelettes ainsi que la première dimension de l'électrophorèse se déroulent à 20°C. Les protéines ont été révélées dans le gel par le bleu de Coomassie. MM, masse moléculaire ; XP, présence de X groupements phosphates ; d, désamidé.

Pour tenter de minimiser ce phénomène supposé, nous ne pouvons faire varier qu'un facteur : la température. En effet, nous allons essayer d'influer sur ce facteur car nous ne pouvons pas modifier les conditions de pH ni réduire les temps de réhydratation et de migration au risque d'altérer en qualité l'analyse 2-DE. Ainsi, la température a été réduite à 10°C pendant la réhydratation et la durée de la première dimension. Les résultats obtenus à 10°C ne montrent plus qu'un spot principal par groupe de masses moléculaires de 30 et 32 kDa pour chacune des sous-fractions F2.2 à F2.5 ce qui indique qu'un phénomène de désamidation non-enzymatique s'est bien produit lors de l'essai à 22°C (Figure 53). Ce phénomène de désamidation est par conséquent fortement ralenti à 10°C.

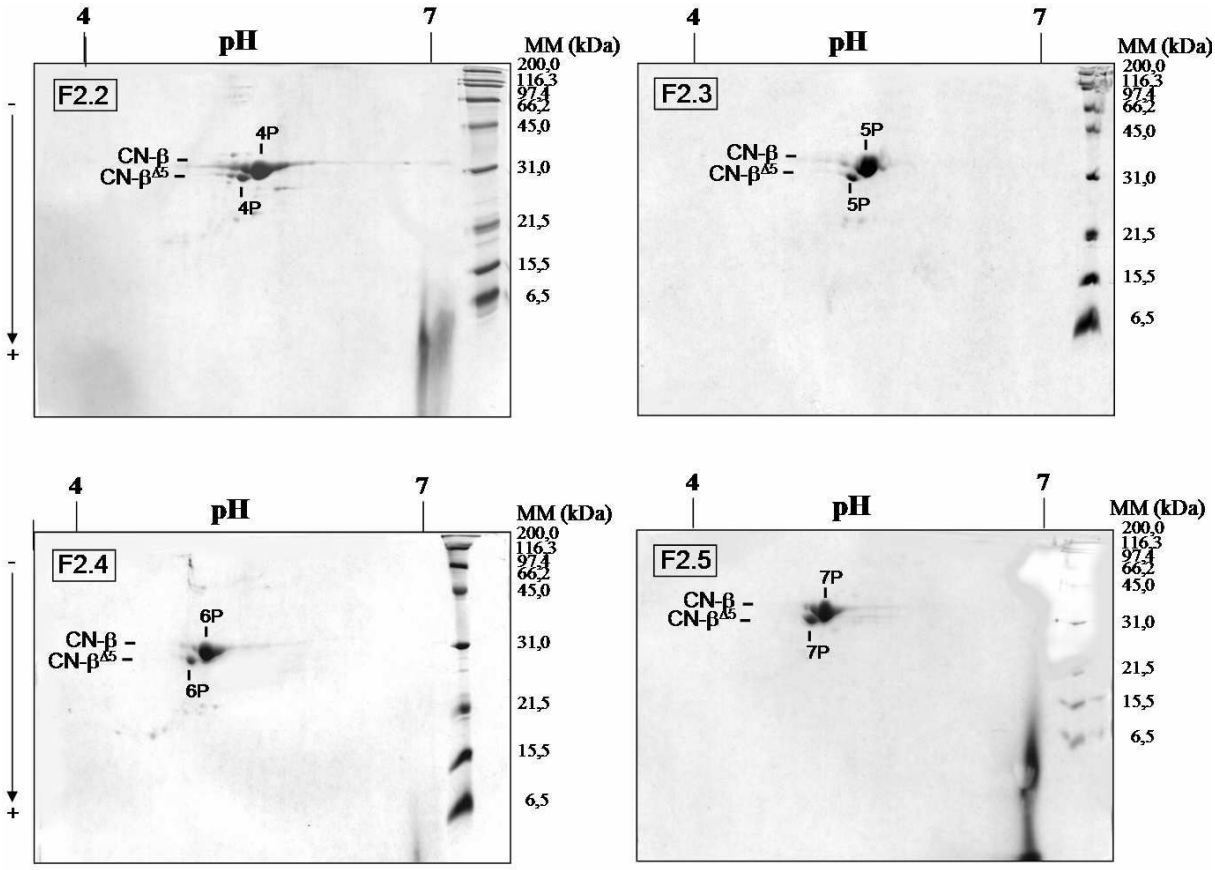


Figure 53 : Electrophorèse bidimensionnelle (première dimension : isoélectrofocalisation, deuxième dimension : PAGE-SDS) des sous-fractions F2.2 à F2.5 obtenues par FPLC d'échange d'ions de la fraction F2. Des quantités protéiques de 50 µg ont été déposées sur des bandelettes constituées d'un gel à 4% de polyacrylamide contenant des ampholines créant un gradient de pH de 4 à 7. La réhydratation des bandelettes ainsi que la première dimension de l'électrophorèse se déroulent à 10°C. Les protéines ont été révélées dans le gel par le bleu de Coomassie. Δ5, délétion de l'exon 5 par épissage alternatif ; XP, présence de X groupements phosphates ; MM, masse moléculaire.

Les profils bidimensionnels obtenus à 10°C ont été comparés à ceux des sous-fractions F2.3/d et F2.4/d contenant les variants 5P et 6P de la CN-β désamidés *in vitro* par incubation à pH 8,4 et 37°C pendant 48 h (la réhydratation des bandelettes et la migration IEF étant réalisées à 10°C ; Figure 54). Les CN-β complète et CN-β^{Δ5} se résolvent en deux spots dans les fractions désamidées alors qu'un seul spot leur correspond lorsque les fractions n'ont pas été incubées (Figure 54).

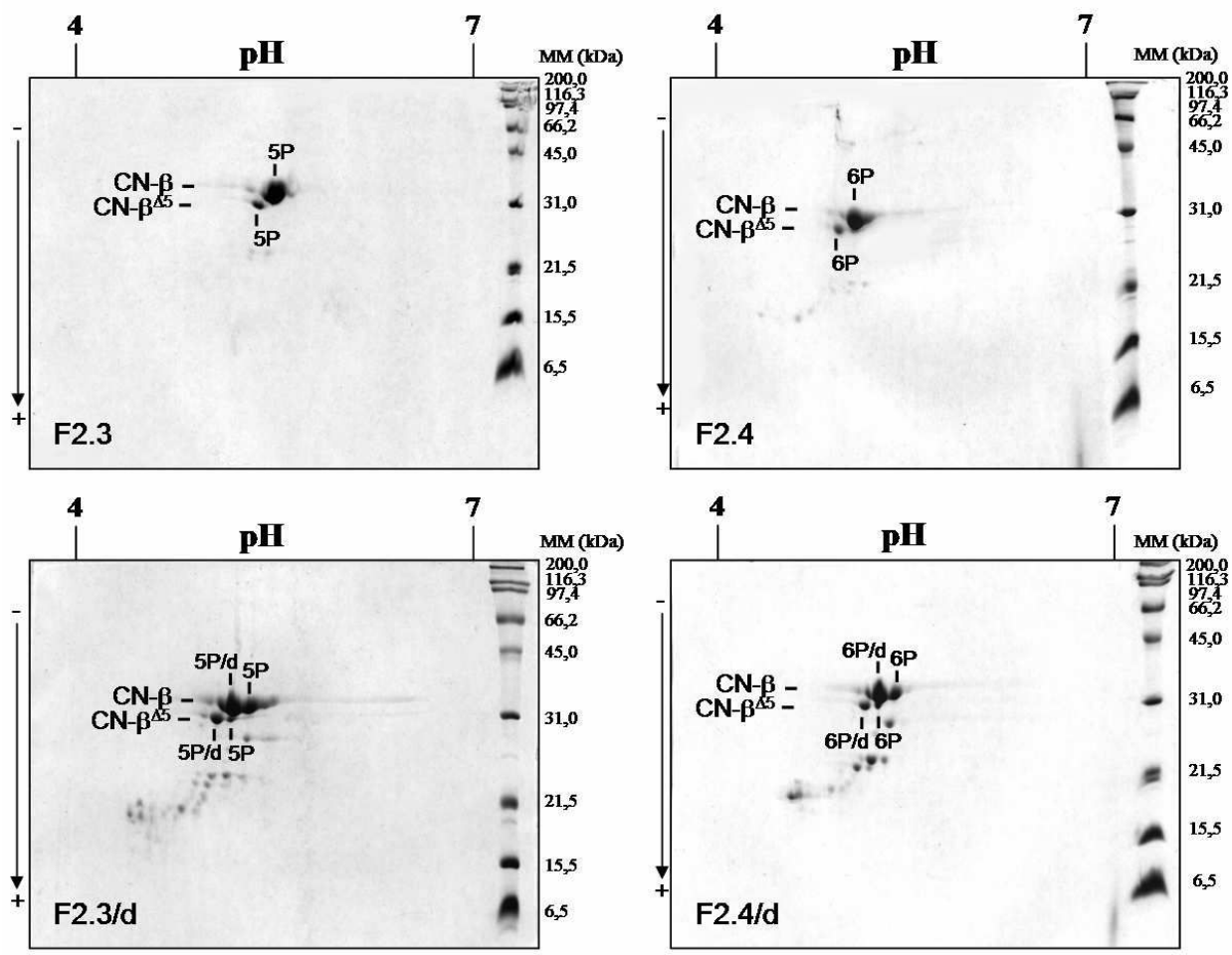


Figure 54 : Electrophorèse bidimensionnelle (première dimension : isoélectrofocalisation, deuxième dimension : PAGE-SDS) des sous-fractions F2.3 à F2.4 (obtenues par FPLC d'échange d'ions de la fraction F2) avant et après désamidation non enzymatique. Les échantillons ont été incubés à 37°C dans un tampon phosphate de sodium 150 mM à pH 8,4 contenant 0,02% d'azoture de sodium et 10% d'un cocktail d'inhibiteurs de protéases pendant 48 h. Des quantités protéiques de 50 µg ont été déposées sur des bandelettes constituées d'un gel à 4% de polyacrylamide contenant des ampholines créant un gradient de pH de 4 à 7. La réhydratation des bandelettes ainsi que la première dimension de l'électrophorèse se déroulent à 10°C. Les protéines ont été révélées dans le gel par le bleu de Coomassie. Δ5, délétion de l'exon 5 par épissage alternatif ; d, désamidé ; XP, présence de X groupements phosphates ; MM, masse moléculaire.

Les spots générés *in vitro* sont plus acides que les spots correspondant aux protéines natives et peuvent être attribués au phénomène de désamidation qui apporte une charge négative. On observe donc le même phénomène que lorsque la réhydratation des bandelettes par les fractions F2.3 et F2.4 est réalisée à 22°C. Ceci confirme le phénomène de désamidation non-enzymatique du résidu Asn^{135} dû aux conditions utilisées lors de l'analyse 2-DE.

En outre, si l'on superpose les profils bidimensionnels des fractions F2.3 et F2.4 avec ceux des fractions respectives F2.3/d et F2.4/d, on remarque que les spots correspondant aux isoformes CN-β 5P/d et CN-β 6P se superposent parfaitement ainsi que les spots correspondant aux isoformes CN-β 6P/d et CN-β 7P (Figure 55).

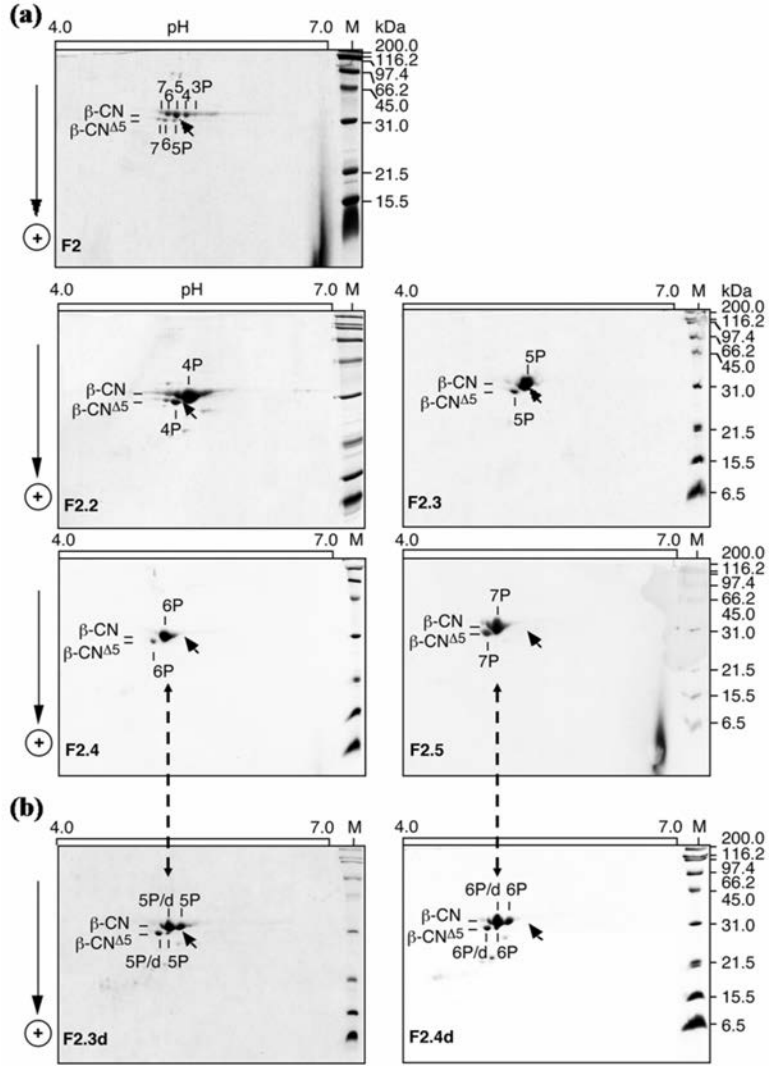


Figure 55 : Electrophorèse bidimensionnelle (première dimension : isoélectrofocalisation, deuxième dimension : PAGE-SDS) de la fraction F2, des sous fractions F2.2 à F2.5 (obtenues par FPLC d'échange d'ions de la fraction F2) et des sous-fractions F2.3 et F.4 obtenues après désamidation non enzymatique. Les échantillons ont été incubés à 37°C dans un tampon phosphate de sodium 150 mM à pH 8,4 contenant 0,02% d'azoture de sodium et 10% d'un cocktail d'inhibiteurs de protéases pendant 48 h. Des quantités protéiques de 50 µg ont été déposées sur des bandelettes constituées d'un gel à 4% de polyacrylamide contenant des ampholines créant un gradient de pH de 4 à 7. La réhydratation des bandelettes ainsi que la première dimension de l'électrophorèse se déroulent à 10°C. Les protéines ont été révélées dans le gel par le bleu de Coomassie. La flèche pleine indique la position du spot correspondant à l'isoforme 5P de la CN-β complète. La double flèche en pointillés indique la superposition des spots natifs et désamidés. Δ5, délétion de l'exon 5 par épissage alternatif ; d, désamidé ; XP, présence de X groupements phosphates ; M, masse moléculaire.

Il est donc indispensable de réaliser la première dimension à 10°C, afin de limiter la désamidation pendant l'analyse 2-DE qui pourrait mener à une contamination des spots de variants natifs par des spots de variants désamidés.

La légère protéolyse produite pendant l'incubation à 37°C des sous-fractions F2.3 et F2.4 et observée par PAGE-urée est également mise en évidence par 2-DE, que la réhydratation soit réalisée à 22°C ou 10°C. En effet, plusieurs spots sont présents pour des masses moléculaires apparentes comprises entre 17 et 28 kDa. Ces spots pourraient correspondre à des produits d'autolyse des isoformes 5P et 6P des CN-β et CN-$\beta^{\Delta 5}$, dans la mesure où les éventuelles protéases d'origine endogène ou exogène seraient inhibées par le cocktail d'inhibiteurs ajouté dans le milieu. Il a été montré que la réaction de désamidation d'un résidu Asn peut générer le clivage de la liaison peptidique adjacente et plus spécifiquement de la liaison Asn-Pro (Robinson et Robinson, 2004). Le clivage de la liaison Asn-Pro peut se dérouler à pH acide (Landon, 1977) aussi bien qu'au pH physiologique (Geiger et Clarke, 1987 ; Tarelli et Corran, 2003). Dans une moindre mesure, le clivage d'autres liaisons peptidiques peut également avoir lieu dans un tampon phosphate de sodium à pH neutre et cela concerne essentiellement la séquence Asn-Leu (Geiger et Clarke, 1987 ; Tyler-Cross et Schirch, 1991). La CN-β équine possède deux séquences Asn-Pro, Asn^{132}-Pro^{133} et Asn^{222}-Pro^{223}, ainsi qu'une séquence Asn-Leu, Asn^{138}-Leu^{139}. Ces séquences pourraient être partiellement clivées pendant le processus de désamidation. Donc, les produits de dégradation observés par 2-DE et par PAGE-urée pourraient être générés par clivages de liaisons peptidiques au niveau des ces sites Asn-Pro et/ou Asn-Leu. Un séquençage des spots de 2-DE ou des bandes de PAGE-urée permettrait de vérifier cette hypothèse.

La CN-β bovine ne contient pas de séquence Asn-Gly et Asn-Pro mais juste une séquence Asn-Leu et semble être très stable. En effet, après 96 h d'incubation à pH 7,4 et à 37°C, le dosage de l'ammonium libéré lors de la désamidation montre que moins de 2% de la protéine est désamidée alors que dans le même temps 80% de la CN-β équine l'est (Girardet *et al.,* 2006).

Pour attribuer un spot de 2-DE à une isoforme de phosphorylation de manière précise, nous avons utilisé une méthode qui consiste à mélanger dans un rapport massique de 1:1 chacune des sous-fractions F2.2 à F2.5 avec la fraction F2 contenant toutes les isoformes de la CN-β. Le but est de renforcer l'intensité d'un spot donné sur le profil global de la fraction F2 par l'isoforme contenue dans la sous-fraction ajoutée. Il sera alors aisé d'attribuer les différents spots aux variants de phosphorylation considérés. Grâce à cette méthode, nous

avons ainsi pu localiser de manière précise les isoformes CN-β 4P à 7P et les isoformes CN-$\beta^{\Delta5}$ 4P à 7P (Figure 56).

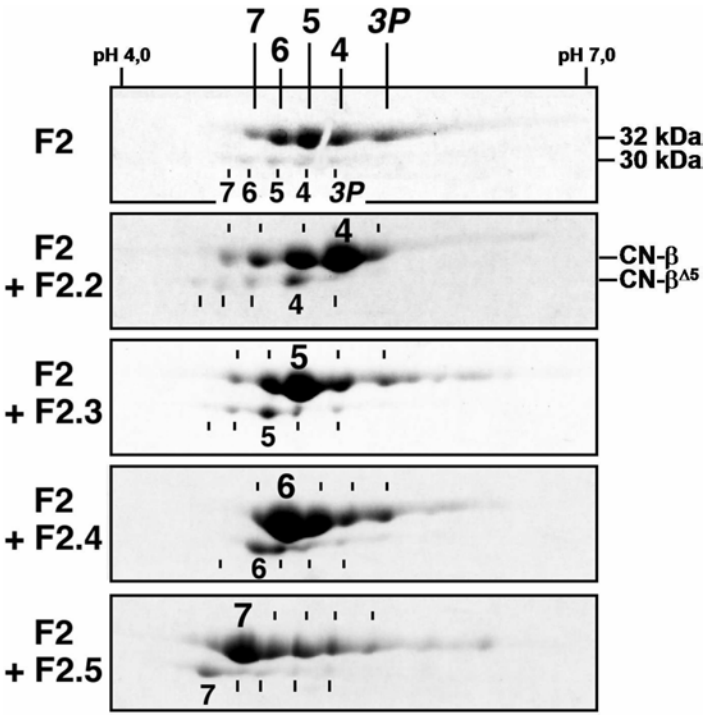


Figure 56 : Electrophorèse bidimensionnelle (première dimension : isoélectrofocalisation, deuxième dimension : PAGE-SDS) des sous-fractions F2.2 à F2.5 obtenues par FPLC d'échange d'ions en mélange (ratio massique 1 :1) avec l'ensemble des isoformes de la CN-β contenues dans la fraction F2 (obtenue par FPLC d'interactions hydrophobes du caséinate de sodium). Des quantités protéiques de 50 µg ont été déposées sur des bandelettes constituées d'un gel à 4% de polyacrylamide contenant des ampholines créant un gradient de pH de 4 à 7. La réhydratation des bandelettes ainsi que la première dimension de l'électrophorèse se déroulent à 10°C. Les protéines ont été révélées dans le gel par le bleu de Coomassie. La localisation potentielle des isoformes 3P de la CN-β et de la CN-$\beta^{\Delta5}$ est indiquée en italique. Δ5, délétion de l'exon 5 par épissage alternatif ; XP, présence de X groupements phosphates ; MM, masse moléculaire.

Comme il est attendu, nous observons que plus l'isoforme est phosphorylée plus elle migre vers les pH acides. Les isoformes 3P de CN-β et CN-$\beta^{\Delta5}$ n'étant pas présentes en grande quantité dans la fraction F2.1, leur localisation sur le profil bidimensionnel n'a pas pu être déterminée de façon précise par cette méthode. Cependant, deux spots de faible intensité (un pour la CN-β et un pour la CN-$\beta^{\Delta5}$) migrant vers les pH les moins acides sont visibles dans la fraction F2 et ne correspondent pas aux isoformes 4P. Ils correspondraient donc probablement aux isoformes 3P de la CN-β et de la CN-$\beta^{\Delta5}$.

5. Conclusions et perspectives

L'individualisation de chaque variant de phosphorylation de la CN-β équine par chromatographie IEC a été nécessaire pour permettre d'établir avec précision la cartographie bidimensionnelle de cette protéine.

L'analyse LC/ESI-MS de chacune des sous-fractions obtenues par IEC a permis la mise en évidence de la présence des isoformes 4P à 7P du variant d'épissage CN-$\beta^{\Delta 5}$. La présence du variant d'épissage CN-$\beta^{\Delta 5}$ suggérée dans le lait de jument Haflinger par Girardet *et al.* (2006) a ainsi été confirmée dans nos travaux. Cependant, la présence de l'isoforme CN-$\beta^{\Delta 5}$ 3P, non détectée par spectrométrie de masse, semble être montrée par 2-DE.

Nous avons démontré que le phénomène de désamidation non-enzymatique du résidu Asn^{135} ne se produisait ni dans le lait lors de son stockage, ni lors des différentes étapes impliquées dans la préparation de CN-β à partir de ce lait. Cette désamidation doit être provoquée sous certaines conditions pour avoir lieu, notamment par une incubation à 37°C et à pH basique mais peut survenir également au cours de certaines étapes expérimentales. Ceci nous a amené à montrer qu'un phénomène de désamidation a une forte probabilité de se produire lors de la mise en oeuvre d'une analyse de 2-DE, si les premières phases de l'expérience se déroulent à température ambiante. Dans ce cas, des spots « natifs » et des spots « désamidés » de la CN-β sont confondus. Les conditions de pH, de température, de tampon et de temps d'incubation sont effectivement réunies pour qu'une telle réaction ait lieu de façon sensible. Ce phénomène peut être fortement ralenti en abaissant la température de l'étape d'isoélectrofocalisation à 10°C. Il est donc nécessaire de bien caractériser le phénomène de désamidation pour élucider le profil bidimensionnel de la CN-β équine.

Une légère protéolyse se produit lors de la réaction de désamidation non-enzymatique même en présence de fortes concentrations d'agent bactériostatique et d'inhibiteurs de protéases. Celle-ci pourrait être due au clivage spontané des séquences Asn-Pro et Asn-Leu de la CN-β équine qui serait induit par le processus chimique de désamidation au niveau de ces sites (Asn-Gly se désamide mais n'entraîne aucun clivage). Il serait nécessaire de réaliser un séquençage des spots de 2-DE et/ou des bandes de PAGE-urée pour vérifier cette hypothèse.

Concernant la désamidation non-enzymatique, il serait intéressant de savoir si celle-ci peut avoir lieu lors de la digestion des protéines du lait de jument dans le tractus digestif notamment au niveau de l'intestin grêle dont le pH pourrait être favorable au processus. Pour cela, il faudrait réaliser des expériences mimant les conditions que rencontrerait la CN-β lors

de son passage dans l'estomac puis dans l'intestin. Il serait également intéressant de connaître les conséquences de cette désamidation potentielle sur les propriétés physico-chimiques et biologiques de la protéine, notamment du point de vue du potentiel allergisant du lait de jument.

La localisation de chaque isoforme de phosphorylation sur un profil bidimensionnel de CN-β a pu être réalisée de manière précise. La méthode utilisée a consisté à augmenter l'intensité de chaque spot grâce à un enrichissement de chaque sous-fraction avec de la CN-β native. Un prélèvement des spots directement à partir du gel pourra donc être envisagé puisque la maîtrise des conditions expérimentales permet d'obtenir des spots purs et non plus susceptibles d'être contaminés par des formes désamidées. En effet, l'attribution des spots de masse moléculaire apparente de 30 kDa à l'isoforme CN-$\beta^{\Delta 5}$ doit encore être démontrée de façon directe.

PARTIE 3
Préparation des caséinophosphopeptides de la CN-β équine et détermination des sites de phosphorylation en fonction du degré de phosphorylation

L'étude de l'influence du degré et de la séquence de phosphorylation des CPP équins sur leur capacité de chélation de minéraux ou d'oligo-éléments d'intérêt nutritionnel ne pourra être envisagée que si les CPP issus de la CN-β et de la CN-α_{s1} sont purifiés individuellement afin d'identifier précisément les résidus phosphorylés.

Dans ce travail, nous allons nous focaliser sur la CN-β équine plutôt que sur la CN-α_{s1} équine, car nous avons développé une méthode efficace de séparation des variants de phosphorylation de la CN-β équine par FPLC d'échange d'anions (voir partie précédente). Ainsi, nous allons pouvoir générer des CPP de degré de phosphorylation différent par hydrolyse trypsique des variants 4P à 7P de la CN-β. En ce qui concerne la CN-α_{s1}, nos travaux nous ont permis de caractériser ses variants post-transcriptionnels (dus à l'épissage alternatif d'exons et à la présence d'un site d'épissage cryptique) et ses variants post-traductionnels (dus à la phosphorylation variable). Cependant, l'isolement de variants possédant un degré de phosphorylation différent ne semble pas aisé au premier abord de par la microhétérogénéité de cette protéine. En effet, des essais ont été réalisés au moyen de la chromatographie d'échange d'anions mais ceux-ci ne se sont pas avéré concluants.

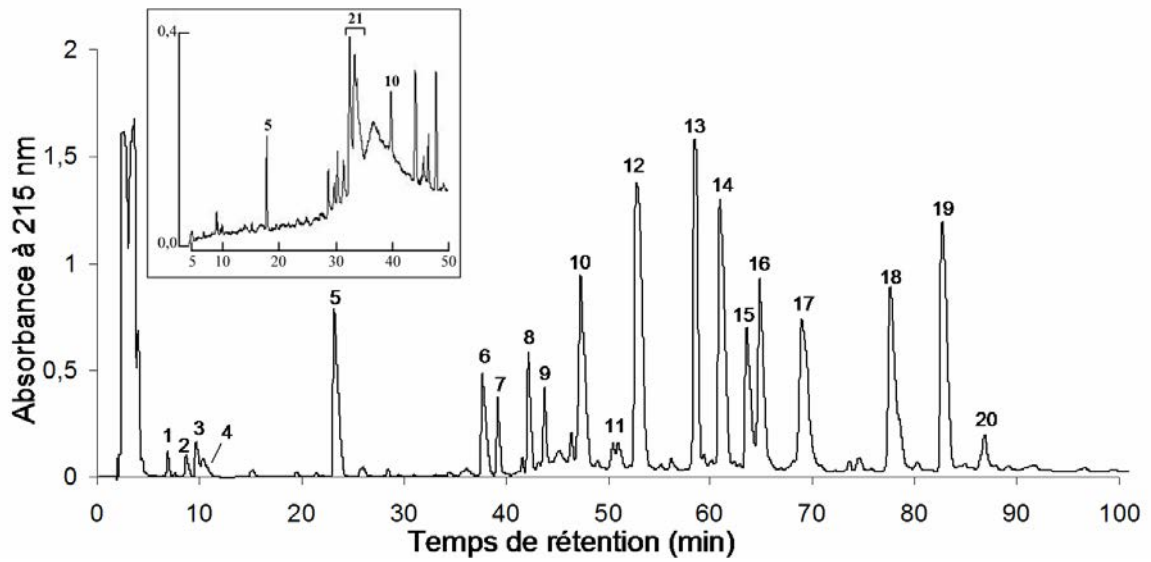
1. Hydrolyse des différents variants de phosphorylation de la CN-β équine isolés par FPLC d'échange d'ions

Chacune des fractions obtenues par fractionnement FPLC d'échange d'anions de la CN-β contient majoritairement un variant de phosphorylation. Les fractions contenant les variants 4P, 5P, 6P et 7P seront utilisées pour préparer les CPP correspondants. Nous n'utiliserons pas la fraction F2.1 contenant le variant 3P, car celui-ci est en très faible quantité et se trouve en mélange avec le variant 4P.

Ainsi, chaque fraction est hydrolysée avec de la trypsine immobilisée sur billes d'agarose. La trypsine immobilisée a pour avantages par rapport à la trypsine libre d'éliminer quasiment l'autolyse de l'enzyme et la possibilité de contamination de l'échantillon par la protéase. De plus, le contrôle de la réaction est possible par le simple retrait de l'enzyme. En outre, la trypsine est plus thermostable lorsqu'elle est immobilisée que lorsqu'elle est à l'état libre. La trypsine est une endoprotéase qui clive les liaisons peptidiques impliquant la fonction carboxylique des résidus Lys et Arg. Ainsi, l'hydrolyse de la CN-β équine devrait générer, si elle est totale, le CPP [4-34], tandis que l'hydrolyse du variant CN-$\beta^{\Delta 5}$ devrait

produire le CPP [4-28]. Girardet *et al.* (2006) ont identifié les peptides trypsiques de la CN-β par spectrométrie de masse. Les auteurs ont montré qu'il pouvait y avoir des coupures aspécifiques au niveau de liaisons impliquant la fonction carboxylique de résidus Tyr et Phe, probablement dues à une activité α-chymotrypsique résiduelle contaminant la trypsine commerciale (Figure 57). En effet, la chymotrypsine est une endoprotéase clivant les liaisons peptidiques impliquant la fonction carboxylique des résidus Tyr, Trp, Phe et Met (Folk et Schirmer, 1965). Les CPP qui ont été mis en évidence sont les fragments [4-34] 3P, 4P, 5P et 6P et les fragments [1-34] 3P, 4P, 5P, 6P et 7P. Les CPP ont été préparés par précipitation sélective au chlorure de calcium selon la méthode de Reynolds *et al.* (1994) et par HPLC en phase inversée en présence de 0,9% de TFA au lieu de 0,1% de TFA, car cela permet de diminuer le pH à 1,0 (pH inférieur au pKa le plus acide) afin de protoner les CPP pour qu'ils soient mieux retenus sur la colonne C_{18} utilisée. L'hydrolyse de la liaison 3-4 n'était donc pas totale puisque des fragments [1-34] ont été trouvés. Par ailleurs, les auteurs n'ont pas identifié de CPP provenant du variant CN-$\beta^{\Delta 5}$.

Dans notre cas, la méthode de précipitation au chlorure de calcium pour purifier les CPP ne semble pas être en adéquation avec nos objectifs. En effet, cette méthode n'est pas très sélective, puisque le chromatogramme rapporté montre la présence de nombreux peptides non phosphorylés (Girardet *et al.,* 2006 ; Figure 57). De plus, les CPP purifiés par cette méthode sont susceptibles de chélater le calcium ce qui n'est pas compatible avec les études de chélation envisagées par la suite, car cela induirait une étape supplémentaire pour dissocier le minéral des CPP et il faudrait s'assurer que celui-ci a été éliminé totalement. Il faut donc trouver un autre moyen pour purifier les CPP de la CN-β équine.



Numéro du peptide	Fragment peptidique correspondant	Séquence du peptide correspondant
1	[1-3]	REK
2	[37-39]	LQK
3	[49-52]	EVER
4	[56-58]	ISR
5	[40-48]	FKHEGQQQR
6	[106-111]	ETILPK
7	[197-203]	LLYQDPR
8	[129-140]	QILNPTNGENLR
9	[190-196]	DTPVQAF
10	[120-128]	SPIVPFSER
11	[113-119]	KVMPFLK
12	[176-189]	VAPFPQPVVPYPQR
13	[59-75]	FVQPQPVVYPYAEPVPY
14	[204-226]	LGPTGELDPATQPIVAVHNPVIV
15	[141-149]	LPVHLIQPF
16	[190-203]	DTPVQAFLLYQDPR
17	[150-175]	MHQVPQSLLQTLMLPSQPVLSPPQSK
18	[141-175]	LPVHLIQPFMHQVPQSLLQTLMLPSQVLSPPQSK
19	[76-105]	AVVPQSILPLAQPPILPFLQPEIMEVSQAK
20	[59-105]	FVQPQPVVYPYAEPVPYAVVPQSILPLAQPPILPFLQPEIMEVSQAK
21	[1-34] 3P, 4P, 5P, 6P, 7P	REKEELNVSSETVESLSSNEPDSSSEESITHINK
	+ [4-34] 3P, 4P, 5P, 6P	EELNVSSETVESLSSNEPDSSSESITHINK

Figure 57 : HPLC en phase inversée d'un hydrolysat trypsique de CN-β équine obtenu par Girardet *et al.* (2006). Les auteurs ont utilisé une colonne analytique C_{18} (250 x 4 mm) sur laquelle a été déposé une quantité protéique de 400 µg. Après une étape isocratique à 5% d'acétonitrile pendant 5 min, un gradient linéaire de 5 à 50% d'acétonitrile a été appliqué pendant 100 min en présence de 0,1% de TFA à un débit de 1 mL.min^{-1}. En encadré, le fractionnement des CPP sélectivement précipités au $CaCl_2$ est réalisé en présence de TFA à 0,9%. Les peptides correspondant aux pics numérotés ont été identifiés par LC/ESI-MS. XP, présence de X groupements phosphates.

2. Préparation d'une fraction enrichie en CPP par chromatographie d'affinité au dioxyde de titane et HPLC en phase inversée

2.1. Chromatographie d'affinité au dioxyde de titane ou MOAC (Metal Oxide Affinity Chromatography)

2.1.1. Choix du métal immobilisé

Il existe plusieurs techniques pour enrichir une fraction en phosphopeptides à partir d'un milieu complexe. La plus ancienne et néanmoins la plus couramment utilisée en phosphoprotéomique est l'IMAC utilisant l'affinité des charges négatives des groupements phosphates des phosphopeptides pour des ions métalliques chargés positivement (Fe^{3+}, Ga^{3+}, Cu^{2+}, Co^{2+}, Zn^{2+}, Ni^{2+}) qui sont immobilisés. Cependant, les peptides non phosphorylés riches en résidus acides Asp et Glu peuvent interagir avec l'ion métallique *via* leurs groupements carboxyliques et sont alors souvent co-purifiés avec les phosphopeptides (Ficarro *et al.,* 2002). D'autres résidus tels que Cys et His peuvent également interagir, ce qui peut rendre la sélectivité pour le groupement phosphate difficile (Porath *et al.,* 1975). Ce désavantage peut être partiellement surmonté par la conversion des groupements carboxyliques des résidus acides en esters de méthyle (Ficarro *et al.,* 2002). Cependant, bien que grâce à cette estérification la majorité des groupements carboxyliques soient neutralisés, cette réaction peut entraîner d'autres modifications telles que la désamidation et la méthylation de résidus Asn et Gln. Une perte des phosphopeptides très hydrophiles peut également se produire pendant l'élution. De plus, pour certains métaux (Cu^{2+} et Zn^{2+}), la liaison avec les phosphopeptides est irréversible entraînant de faibles rendements (Nühse *et al.,* 2003).

Plus récemment, une nouvelle technique de chromatographie a été mise au point utilisant le dioxyde de titane TiO_2 (Pinkse *et al.,* 2004 ; Larsen *et al.,* 2005). A pH acide, le TiO_2 présente des sites acides de Lewis grâce aux ions titane insaturés Ti^{4+}. Un acide de Lewis est une entité moléculaire dont un des atomes possède une orbitale vide pouvant accepter un doublet électronique et donc former une liaison covalente avec une base de Lewis qui possède un doublet d'électrons non liant. Ainsi sur les sites acides de Lewis de Ti^{4+}, des bases de Lewis telles que le phosphate, le fluorure et le carboxylate peuvent se lier avec une forte affinité (Nawrocki *et al.,* 2004). Une fraction peut donc être enrichie en phosphopeptides grâce à cette méthode. La molécule de TiO_2 lie les phosphopeptides à pH acide et ils peuvent

être élués à pH basique. Il a été montré que le pH d'élution doit être supérieur à 10,5 pour éluer des peptides multiphosphorylés (Larsen *et al.,* 2005). Cependant, les résidus acides montrent également, comme cela a déjà été observé par IMAC, une affinité pour TiO_2 notamment les résidus Asp et Glu (Pinkse *et al.,* 2004 ; Klemm *et al.,* 2006 ; Pocsfalvi *et al.,* 2007). Il a été démontré que l'utilisation d'acide 2,5-dihydroxybenzoïque (DHB) à forte concentration (> 20 $mg.mL^{-1}$) dans la phase mobile améliore la rétention sélective des CPP par TiO_2 et diminue les liaisons non spécifiques dans le cas de mélanges peptidiques simples (Larsen *et al.,* 2005). Il a été montré que les acides carboxyliques aromatiques (DHB, acide phtalique et acide salicylique) sont de meilleurs compétiteurs que les acides carboxyliques non aromatiques (acide acétique et acide benzoïque) pour prévenir la liaison de peptides non phosphorylés sur le TiO_2 (Larsen *et al.,* 2005). Des études en spectroscopie infrarouge ont mis en évidence que les acides carboxyliques aromatiques se lient fortement par coordination à TiO_2, tandis que les acides carboxyliques non aromatiques interagissent très faiblement avec TiO_2 (Dobson et Mc Quillan, 2000). Le phosphate se lie avec TiO_2 avec une constante d'affinité (K_A= 4.10^4 M^{-1}) similaire à celle des acides carboxyliques aromatiques (K_A de 10^4 à 10^5 M^{-1} ; Connor et Mc Quillan, 1999) mais l'utilisation d'acide phosphorique à la place de DHB est moins efficace dans la prévention de la liaison des peptides non phosphorylés (Larsen *et al.,* 2005). Ainsi, le mode de liaison des acides carboxyliques aromatiques et du phosphate à TiO_2 diffère (Figure 58).

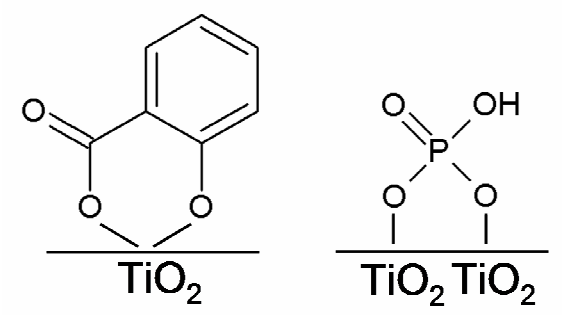


Figure 58 : Schéma représentant les modes de liaison différents des acides carboxyliques aromatiques et du phosphate adsorbés sur le dioxyde de titane. Reproduit d'après Connor et McQuillan (1999) et Dobson et McQuillan (2000).

Il existe une hétérogénéité de surface au niveau des particules de TiO_2 amenant à deux types de sites de liaison. Les acides carboxyliques aromatiques se lient à une molécule de TiO_2 par chélation en formant des coordinats bidentes (Tunesi et Anderson, 1991 ; Dobson et Mc Quillan, 2000). L'anion phosphate, quant à lui, forme un pont entre deux molécules de TiO_2

(Connor et Mc Quillan, 1999). Dans ce cas, le phosphate inorganique serait alors en compétition directe avec les CPP pour le site de liaison à TiO_2, alors que le DHB cible d'autres sites de liaison qui seraient similaires à ceux utilisés par les peptides non phosphorylés (Larsen *et al.,* 2005). La technologie MOAC utilisant TiO_2 présente de nombreux avantages par rapport à l'IMAC. En effet, la prévention de la liaison des peptides non phosphorylés ne nécessite pas de modification chimique des peptides (ce qui est le cas pour la technologie IMAC), ce qui permet d'obtenir des CPP non modifiés pour des études de chélation. Pinkse et Heck (2006) rapportent dans leurs travaux que la liaison des phosphopeptides à TiO_2 apparaît être relativement plus rapide qu'en IMAC qui est généralement connue pour être lente et requérant de longs temps d'incubation. Ces mêmes auteurs relatent que la capacité de chargement du TiO_2 est meilleure que celle de l'IMAC et que le TiO_2 possède une meilleure compatibilité avec les systèmes chromatographiques de phase inversée.

Nous allons donc utiliser la technique MOAC sur TiO_2 immobilisé dans notre étude plutôt que l'IMAC avec Fe^{3+} (Li et Dass, 1999 ; Stensballe *et al.,* 2000) ou Cu^{2+} (Lin *et al.*, 2000) pour séparer les CPP des peptides non phosphorylés issus des hydrolysats trypsiques des différents variants de phosphorylation de la CN-β équine. Nous disposons de microcolonnes consistant en un cône de 200 µL adaptable sur une pipette automatique et contenant une phase de silice greffée avec des molécules de TiO_2. Nous avons adapté notre protocole à partir de ceux utilisés par Larsen *et al.* (2005) et Hsieh *et al.* (2007). Ainsi, l'échantillon a été solubilisé dans une solution contenant 80% d'acétonitrile et 0,1% de TFA, cet acide permettant une amélioration de l'enrichissement en phosphopeptides en diminuant les liaisons des peptides non phosphorylés contenant des résidus acides (Asp ou Glu) avec TiO_2, contrairement à l'acide acétique (Larsen *et al.,* 2005 ; Thingholm *et al.,* 2006). Afin d'essayer de diminuer, voire de supprimer, la co-élution de peptides non phosphorylés, nous ajoutons dans la phase mobile du DHB à 20 $mg.mL^{-1}$, comme proposé par Larsen *et al.* (2005). L'échantillon est déposé sur la colonne, l'éluat est collecté et constitue une première fraction appelée « fraction non liée à TiO_2 ». Puis deux lavages sont effectués afin d'éliminer les peptides non phosphorylés. Le premier lavage est réalisé avec la solution de chargement de l'échantillon contenant le DHB et le TFA tandis que le deuxième lavage consiste en une solution contenant 50% d'acétonitrile. Les deux lavages sont collectés, mélangés et représentent la fraction dite « lavage TiO_2». Les CPP sont élués grâce à un tampon d'hydroxyde d'ammonium à 0,25%, pH 10,5 et la fraction collectée est nommée « fraction liée à TiO_2 ».

2.1.2. Analyse des fractions liées à TiO_2 par spectrométrie de masse

L'hydrolyse trypsique de la CN-β équine et du variant CN-$\beta^{\Delta5}$ peut générer théoriquement, pour un degré de phosphorylation donné, quatre CPP différents susceptibles d'avoir interagit avec TiO_2 : CPP [1-34] et CPP [4-34] pour la CN-β et CPP [1-28] et CPP [4-28] pour la CN-$\beta^{\Delta5}$. En effet, Girardet *et al.* (2006) ont montré que la liaison 3-4 potentiellement hydrolysable par la trypsine présente une certaine résistance vis-à-vis de cette enzyme.

Dans un premier temps, nous voulons déterminer si le DHB a bien une influence sur la rétention de peptides non phosphorylés. Pour cela, une analyse de spectrométrie de masse a été menée sur les fractions « liées à TiO_2 » issues de l'hydrolyse trypsique de l'isoforme 5P (sous-fraction F2.3) et préparées de deux façons soit avec, soit sans utilisation de DHB. Dans le cas de l'utilisation de DHB, les résultats montrent la présence des CPP [1-34] 5P et [4-34] 5P qui ont été fixées sur les molécules de TiO_2 (Figure 59).

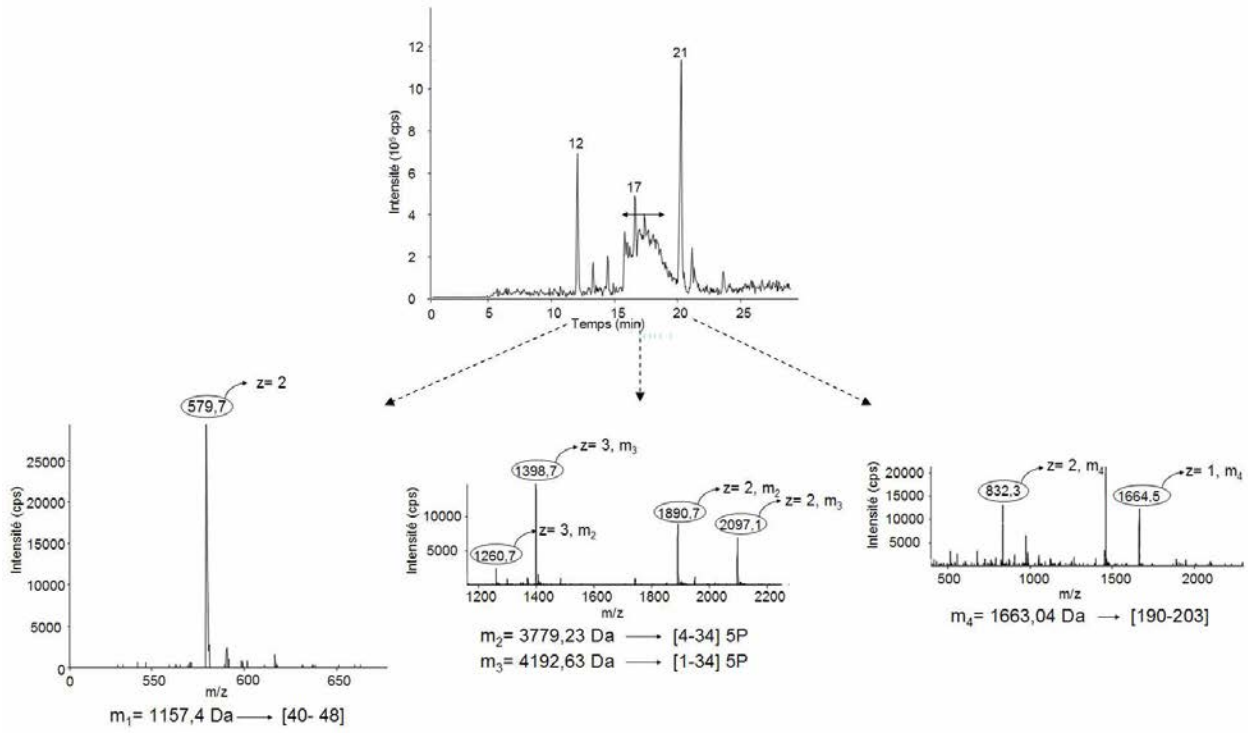


Figure 59 : Chromatographie liquide couplée à la spectrométrie de masse à ionisation électrospray (LC/ESI-MS) de la fraction liée à TiO_2 et issue d'un hydrolysat trypsique de CN-β 5P fractionné par MOAC (Metal Oxide Affinity Chromatography) en présence de DHB (acide 2,5-dihydroxybenzoïque). Seules les fractions majeures 12, 17 et 21 (fractions éluées à 12 min, 15 à 18 min et 21 min, respectivement) ont été analysées par ESI-MS. Les masses des ions extraits sont calculées grâce au logiciel Biomultiview 1.3.1. cps, coups par seconde.

Les CPP [1-28] et [4-28] correspondant aux CPP potentiellement générés à partir de la CN-$\beta^{\Delta 5}$ ne sont pas mis en évidence. Néanmoins, des peptides non phosphorylés ont été identifiés dans la fraction : le peptide de masse 1157,40 Da a été identifié comme étant le fragment [40-48] (masse théorique : 1156,57 Da) et le peptide de masse 1663,05 Da a été identifié comme étant le fragment [190-203] (masse théorique : 1662,86 Da), ces deux peptides étant des fragments trypsiques de la CN-β. Ainsi, malgré la présence de DHB, des peptides non phosphorylés (dont les deux principaux ont été identifiés) ont été co-élués avec les CPP de CN-β. Ces peptides non phosphorylés ont en commun de posséder au moins un résidu acide dans leur séquence. Cependant, d'autres peptides possèdent des résidus acides et n'ont pas été retenus pour autant sur la microcolonne. D'autres facteurs doivent donc intervenir pour entraîner la rétention préférentielle de ces peptides non phosphorylés.

Dans la fraction « liée à TiO_2 » obtenue en n'utilisant pas de DHB, les CPP [1-34] 5P et [4-34] 5P sont identifiés ainsi que les fragments [40-48] et [190-203] (Figure 60).

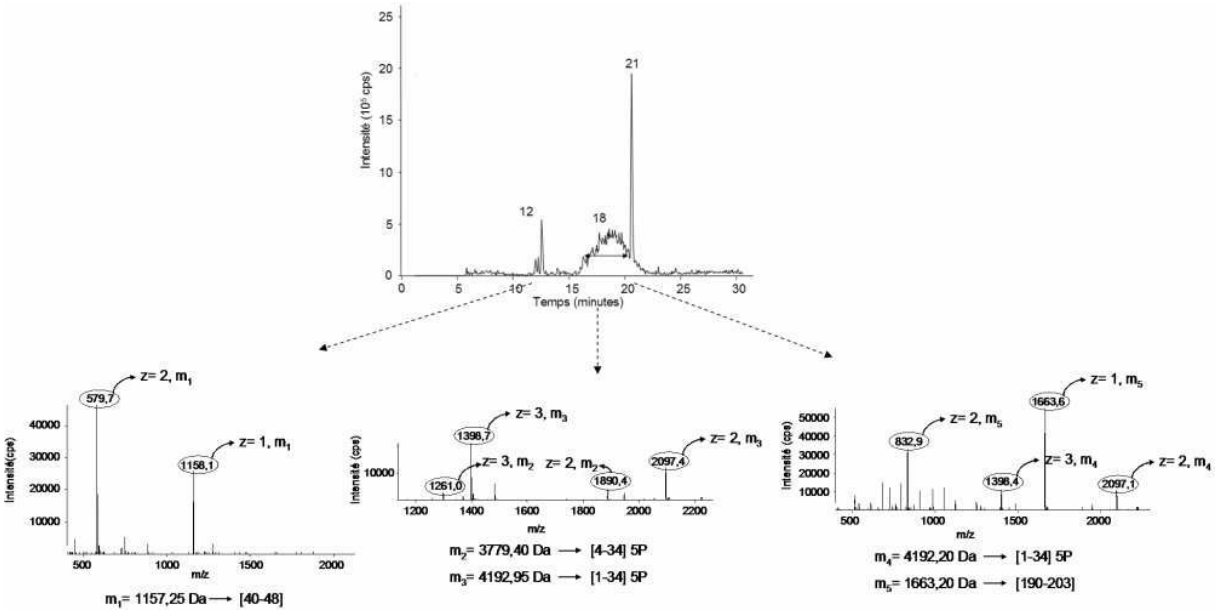


Figure 60 : Chromatographie liquide couplée à la spectrométrie de masse à ionisation électrospray (LC/ESI-MS) de la fraction liée à TiO_2 et issue d'un hydrolysat trypsique de CN-β 5P fractionné par MOAC (Metal Oxyde Affinity Chromatography) en absence de DHB (acide 2,5-dihydroxybenzoïque). Seules les fractions majeures 12, 17 et 21 (fractions éluées à 12 min, 15 à 18 min et 21 min, respectivement) ont été analysées par ESI-MS. Les masses des ions extraits sont calculées grâce au logiciel Biomultiview 1.3.1. cps, coups par seconde.

Ainsi, nous nous rendons compte que l'utilisation de DHB, dans notre cas, n'évite pas la rétention aspécifique de peptides non phosphorylés. De plus, il a été rapporté que le DHB

peut contaminer les systèmes de chromatographie liquide ainsi que les spectromètres de masse, ce qui requiert un nettoyage plus fréquent des instruments et une durée de vie des colonnes plus courte (Jensen et Larsen, 2007). Devant l'absence d'avantage lié à l'utilisation du DHB et aux possibles inconvénients que celui-ci peut entraîner nous ne l'avons plus utilisé par la suite lors des séparations MOAC des isoformes 4P, 6P et 7P.

La technologie MOAC s'est révélée être beaucoup plus spécifique des phosphopeptides que la méthode de précipitation par le chlorure de calcium utilisée par Girardet *et al.* (2006). En effet, beaucoup moins de peptides non phosphorylés sont retrouvés dans la fraction « liée à TiO_2 » que dans la fraction précipitée par le chlorure de calcium (comparer Figures 57 et 60).

Dans le cas des isoformes 6P et 7P (fractions IEC F2.4 et F2.5, respectivement), les CPP [1-34] et [4-34] sont retrouvés dans les fractions « liées à TiO_2 » respectives (Tableau 7), mais le CPP [1-36] 7P, résultant d'une hydrolyse trypsique partielle, est également identifié dans le cas de l'hydrolysat de la CN-β 7P. Dans la fraction IEC F2.2 contenant l'isoforme 4P, les CPP [1-34] 4P et [4-34] 4P ont été identifiés mais également les CPP [1-34] et [4-34] de l'isoforme 3P (Tableau 7).

Tableau 7 : Masses moléculaires des différents caséinophosphopeptides (CPP) déterminées par LC/ESI-MS à partir des fractions « liées à TiO_2 » obtenues par chromatographie d'affinité au TiO_2 (MOAC) des isoformes de phosphorylation 3P à 7P contenues dans les fractions chromatographiques d'échange d'ions (IEC) F2.2 à F2.5, respectivement.

Fraction IEC	Caséinophosphopeptide	Masse théorique[a)]	Masse observée
F2.2	CPP[1-34] 3P	4031,95	4031,97
	CPP[4-34] 3P	3618,47	3617,45
F2.2	CPP[1-34] 4P	4111,95	4111,93
	CPP[4-34] 4P	3698,47	3698,68
F2.3	CPP[1-34] 5P	4191,95	4192,20
	CPP[4-34] 5P	3778,47	3779,25
F2.4	CPP[1-34] 6P	4271,95	4271,60
	CPP[4-34] 6P	3858,47	3859,93
F2.5	CPP[1-34] 7P	4351,95	4350,55
	CPP[4-34] 7P	3938,47	3939,88
	CPP[1-36] 7P	4609,20	4610,52

[a)] Masse moyenne calculée avec ProtParam tool (expasy.org/tools/protparam.html).

Il s'avère donc que la fraction F2.2 contient une partie de l'isoforme 3P. Ceci semble dû à la faible résolution de la séparation des fractions F2.1 et F2.2 (voir partie 2). En ce qui concerne les peptides non phosphorylés, seul le fragment [190-203] a été retrouvé dans les fractions issues des isoformes 6P et 7P. Cependant, d'autres masses mineures ont été mises en évidence mais nous n'avons pas pu les attribuer à des peptides qui résulteraient soit d'une coupure trypsique, soit d'une coupure α-chymotrypsique.

Les CPP pouvant être générés par l'hydrolyse trypsique de la CN-$\beta^{\Delta 5}$ n'ont été mis en évidence dans aucune des fractions étudiées, ce qui est vraisemblablement dû à la faible présence de ce variant d'épissage.

L'étude de la chélation des minéraux par les CPP en fonction de leur degré de phosphorylation nécessite de préparer chaque isoforme de CN-β avec une grande pureté. Il s'agit donc d'éliminer les peptides non phosphorylés contaminant les fractions « liées à TiO_2 » par HPLC en phase inversée. Cependant, les CPP étant très hydrophiles (ce sont notamment les seuls peptides possédant des charges négatives au pH utilisé lors du fractionnement par HPLC, c'est-à-dire pH 2,0), leur rétention dans une colonne de phase inversée est problématique et leur purification par HPLC nécessite une mise au point préalable.

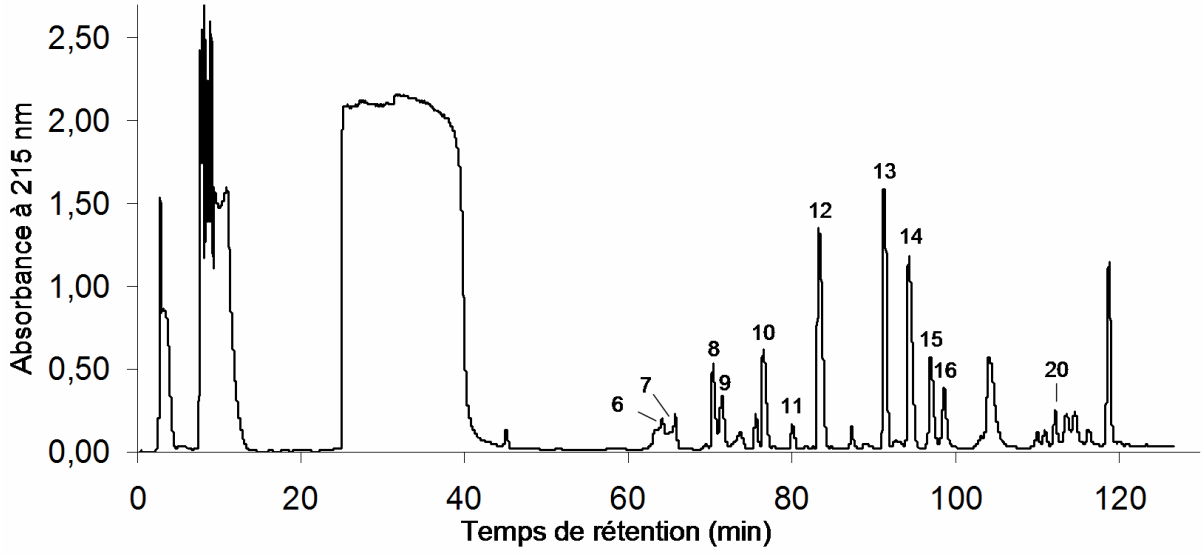
2.2. Mise au point d'une technique de purification des CPP par HPLC en phase inversée

2.2.1. Essai avec une colonne C_{18}

Dans un premier temps, nous avons utilisé les CPP issus de l'isoforme 5P de la CN-β équine, qui est la plus abondante, pour mettre au point une technique permettant leur purification par HPLC en phase inversée.

L'hydrolysat trypsique de CN-β 5P (issu de la fraction IEC F2.3), solubilisé dans la solution de chargement contenant du DHB, a été analysé par HPLC en phase inversée sur une colonne analytique C_{18} et comparé avec le profil obtenu par Girardet *et al.* (2006). Les peptides ont été identifiés grâce à l'analyse des spectres UV et des temps de rétention selon la méthode de Perrin *et al.* (1995). Le chromatogramme obtenu montre la présence d'un composé ayant une très forte absorbance à 215 nm éluant entre 25 et 40 minutes (Figure 61a) et qui ne semble pas être de nature peptidique de par son spectre UV (Figure 61b).

a)



b)

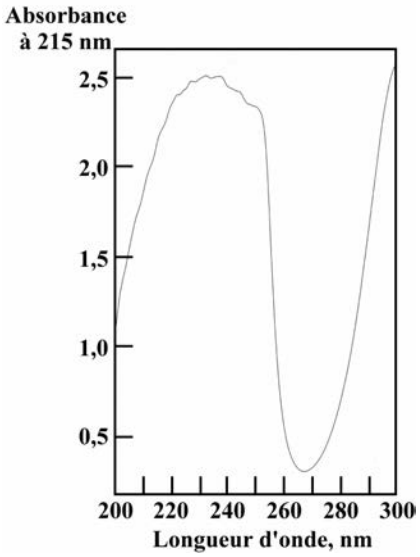


Figure 61 : HPLC en phase inversée d'un hydrolysat trypsique de l'isoforme 5P de la CN-β équine en présence de DHB. Une quantité protéique de 1 mg est déposée sur une colonne C_{18} microbore (150 x 2 mm). Après une étape isocratique à 5% d'acétonitrile pendant 5 min, un gradient linéaire de 5 à 50% d'acétonitrile a été appliqué pendant 100 min en présence de TFA à 0,1%. Le débit est de 0,2 mL.min^{-1}. a) Les numéros de pic sont attribués selon Girardet *et al.* (2006) grâce à l'analyse des spectres UV et des temps de rétention. L'absorbance importante observée entre 20 et 40 min serait celle du DHB élué. b) spectre UV du pic qui correspondrait au DHB.

Nous pouvons remarquer que l'hydrolysat possède un profil chromatographique proche de celui obtenu par Girardet *et al.* (2006). Les peptides aromatiques 7, 9, 10, 11, 12, 13, 15, 16 et 20 ont pu être identifiés par l'analyse de leur spectre UV. Les autres peptides, non aromatiques, sont assignés arbitrairement selon leur ordre d'élution par rapport aux peptides

identifiés. La présence des peptides 7, 9, 13 et 15 atteste de coupures de type α-chymotrypsique comme l'ont déjà observé Girardet *et al.* (2006). Ainsi, même si la trypsine utilisée est traitée au TPCK, une activité α-chymotrypsique résiduelle semble être présente. La fraction « non liée à TiO_2 » est également analysée par HPLC en phase inversée sur colonne analytique C_{18}. Le chromatogramme obtenu est très proche de celui de l'hydrolysat avant passage sur la microcolonne de TiO_2 (Figure 62).

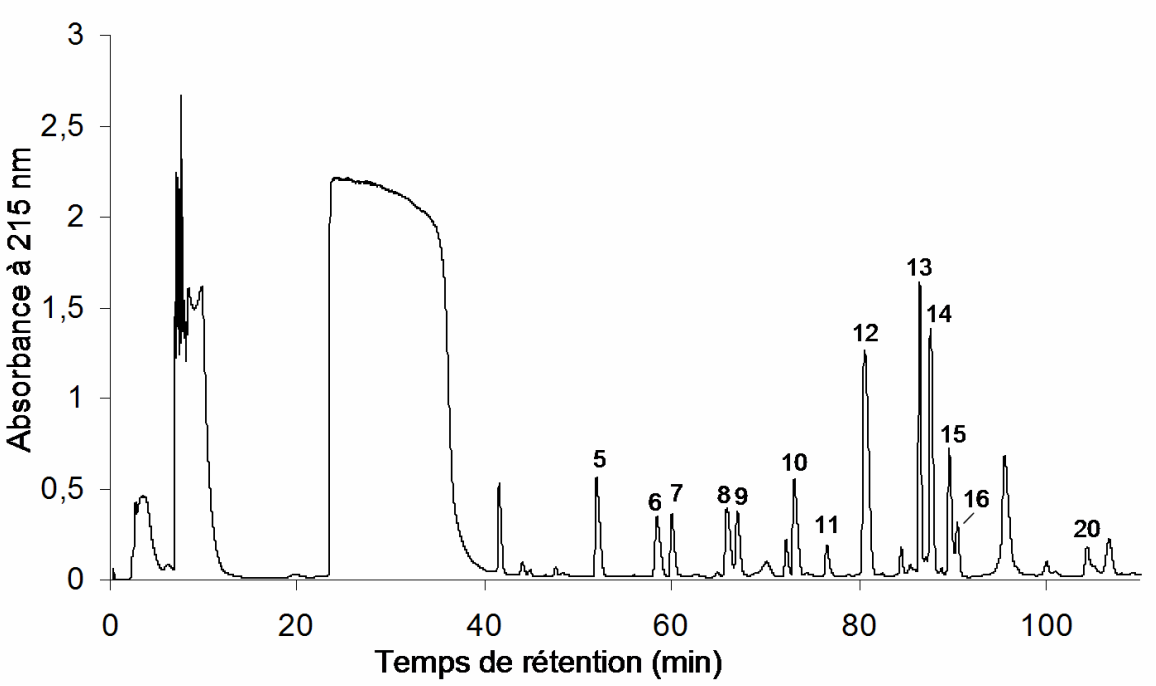
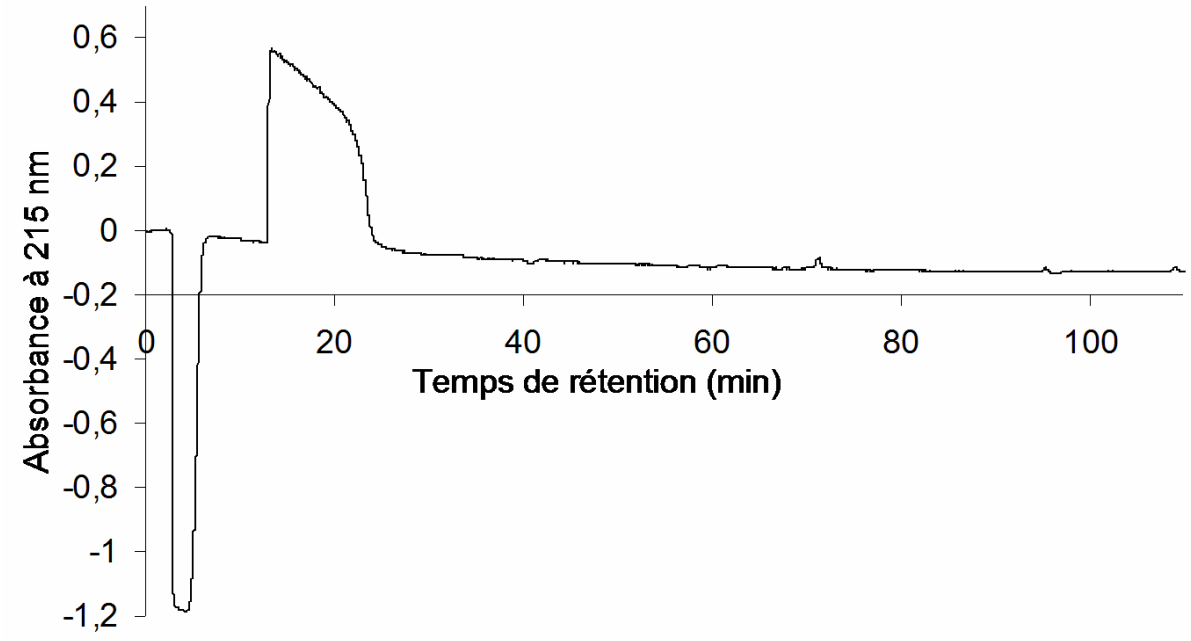


Figure 62 : HPLC en phase inversée de la fraction « non liée à TiO_2 » de l'isoforme 5P de la CN-β équine en présence de DHB. Une quantité protéique de 1 mg est déposée sur une colonne C_{18} microbore (150 x 2 mm). Après une étape isocratique à 5% d'acétonitrile pendant 5 min, un gradient linéaire de 5 à 50% d'acétonitrile a été appliqué pendant 100 min en présence de TFA à 0,1%. Le débit est de 0,2 mL.min^{-1}. Les numéros de pic sont attribués selon Girardet *et al.* (2006) grâce à l'analyse des spectres UV et des temps de rétention. L'absorbance importante observée entre 20 et 40 min serait celle du DHB élué.

Le peptide 5 a pu être identifié grâce à l'analyse spectrale UV. Le DHB éluant entre 25 et 40 minutes avec une très forte absorbance est toujours présent. Son absorbance élevée est liée au fait que le DHB possède un noyau aromatique. De plus, le DHB étant une molécule légèrement hydrophobe, cela lui permet d'interagir avec la phase stationnaire C_{18}. Pour vérifier que ce pic important correspond bien au DHB, nous avons comparé son spectre UV (Figure 61b) à celui du DHB injecté seul dans la colonne (Figure 63) et montré que le pic

important était bien celui du DHB. Néanmoins, nous constatons que le DHB seul élue plus tôt que lorsqu'il est en mélange avec des peptides (comparer Figures 62 et 63).

a)



b)

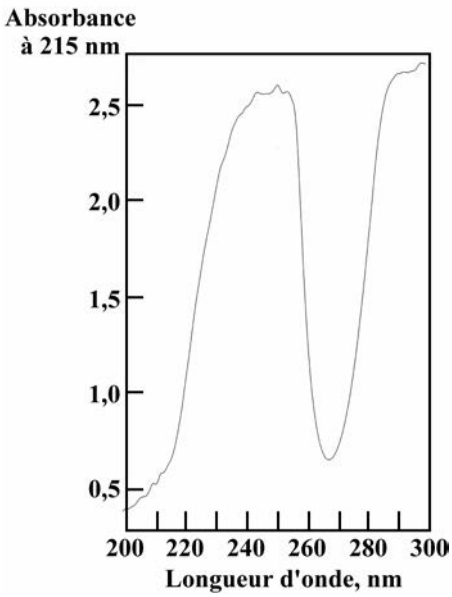


Figure 63 : HPLC en phase inversée d'une solution de DHB. Un volume de 500 µL ([DHB] = 20 mg.mL^{-1}) a été injecté sur une colonne analytique C_{18}. Après une étape isocratique à 5% d'acétonitrile pendant 5 min, un gradient linéaire de 5 à 50% d'acétonitrile a été appliqué pendant 100 min en présence de TFA à 0,1%. Le débit est de 0,2 mL.min^{-1}. a) chromatogramme obtenu. b) spectre UV du pic correspondant au DHB.

En parallèle, l'hydrolysat trypsique de l'isoforme 5P ne contenant pas de DHB a également été analysé (Figure 64) et nous observons bien une absence du pic attribué au DHB. De plus, nous constatons que le temps de rétention de chaque peptide est plus court en l'absence de DHB qu'en sa présence et la séparation des peptides est de meilleure résolution.

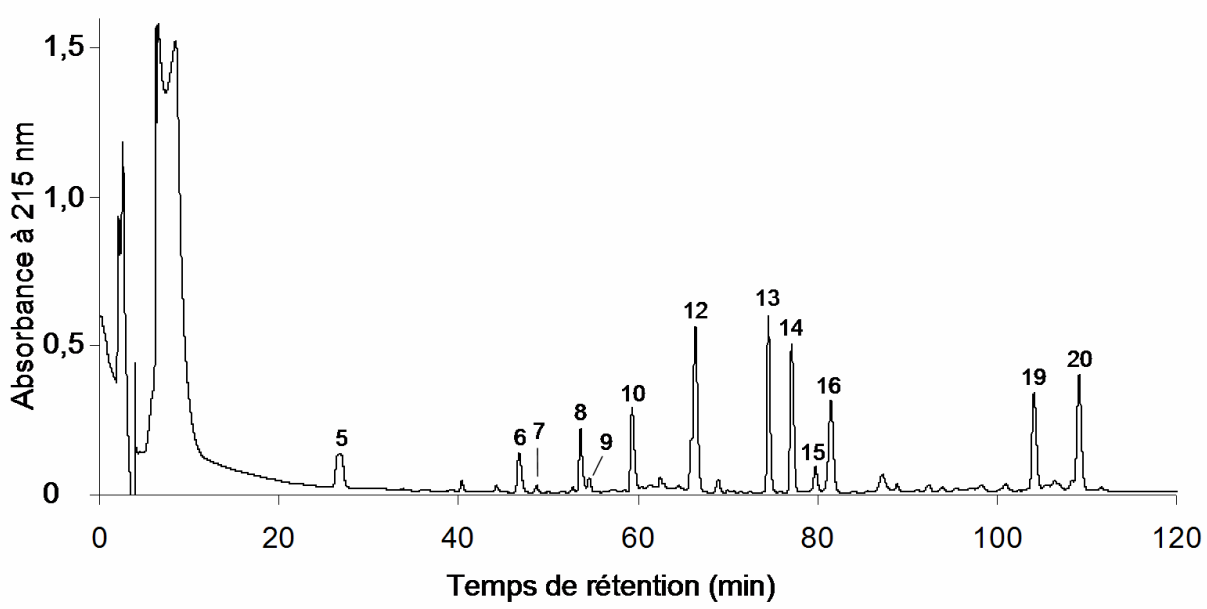


Figure 64 : HPLC en phase inversée d'un hydrolysat trypsique de l'isoforme 5P de la CN-β équine en absence de DHB. Une quantité protéique de 1 mg est déposée sur une colonne C_{18} microbore (150 x 2 mm). Après une étape isocratique à 5% d'acétonitrile pendant 5 min, un gradient linéaire de 5 à 50% d'acétonitrile a été appliqué pendant 100 min en présence de TFA à 0,1%. Le débit est de 0,2 mL.min^{-1}. Les numéros de pic sont attribués selon Girardet *et al.* (2006 ; voir Figure 57) grâce à l'analyse des spectres UV et des temps de rétention.

La fraction « liée à TiO_2 » issue de l'hydrolysat trypsique de l'isoforme 5P a été analysée par HPLC sur colonne analytique C_{18} en présence de TFA à 0,9% selon le protocole proposé par Girardet *et al.* (2006) pour retenir les CPP sur la colonne. Cette concentration de TFA amène le pH aux environs de 1,0, ce qui a pour effet de protoner totalement les groupements phosphate des CPP qui deviennent donc moins hydrophiles et interagissent plus efficacement avec la phase stationnaire C_{18}. Cependant, une très faible quantité de CPP semble être retenue sur la colonne et la préparation de CPP à une plus grande échelle semble peu envisageable par cette méthode. De plus, l'utilisation de TFA à 0,9% est dommageable pour la stabilité de la silice qui compose la phase stationnaire, ce qui est problématique en vue de purifier en grande quantité des CPP.

2.2.2. Purification des CPP par HPLC en phase inversée sur colonne C_4

Afin d'éviter l'utilisation de TFA à 0,9% avec la colonne C_{18}, nous avons testé l'utilisation d'une colonne analytique C_4 en présence de TFA à 0,1%.

Comme lors de l'essai sur colonne C_{18}, l'isoforme 5P a été utilisée en première intention pour réaliser les analyses sur colonne C_4. L'hydrolysat total, la fraction « non liée à TiO_2 », la fraction « lavage TiO_2 » et la fraction « liée à TiO_2 » correspondants à cette isoforme hydrolysée ont été analysés (Figure 65).

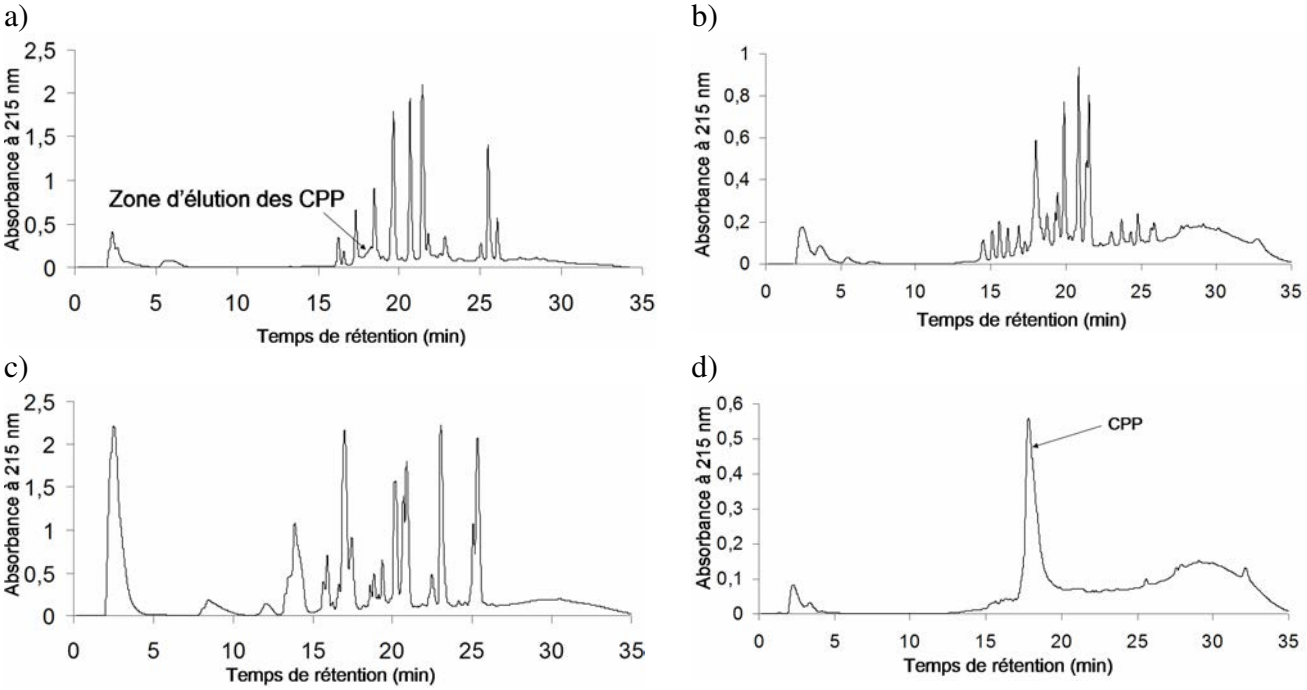


Figure 65 : HPLC en phase inversée des peptides trypsiques de l'isoforme 5P issus de la CN-β équine (fraction F2.3) avant et après chromatographie d'affinité au TiO_2 (MOAC). Une quantité protéique de 0,5 mg est déposée sur une colonne analytique Nucleosil C_4 (150 x 2 mm). Après une étape isocratique à 4% d'acétonitrile pendant 5 min, un gradient linéaire de 4 à 48% d'acétonitrile a été appliqué pendant 15 min puis une étape isocratique à 64% d'acétonitrile est réalisée pendant 5 min en présence de TFA à 0,1%. Le débit est de 250 µL.min^{-1}. a) hydrolysat trypsique ; b) fraction « non liée à TiO_2 » ; c) fraction « lavage TiO_2 » ; d) fraction « liée à TiO_2 » ; CPP : caséinophosphopeptides.

Le chromatogramme obtenu pour l'hydrolysat total montrant plusieurs pics correspondant aux différents peptides trypsiques. Les fractions « non liées à TiO_2 » et « lavage TiO_2 » présentent chacune un chromatogramme montrant également plusieurs pics correspondant à des peptides qui n'ont pas été retenus sur la microcolonne de TiO_2. Le chromatogramme obtenu pour la

fraction « liée à TiO_2 » ne présente qu'un pic majoritaire contenant les CPP [1-34] et [4-34] de l'isoforme 5P (identifiés par spectrométrie de masse).

Pour les isoformes 6P et 7P, les mêmes types de profils que ceux concernant l'isoforme 5P sont obtenus (Figures 66 et 67). Dans les fractions « liées à TiO_2 », la présence de quelques pics correspondant à des peptides non phosphorylés est observée, cependant les pics correspondant respectivement aux CPP 6P et 7P (identifiés par spectrométrie de masse) sont largement majoritaires.

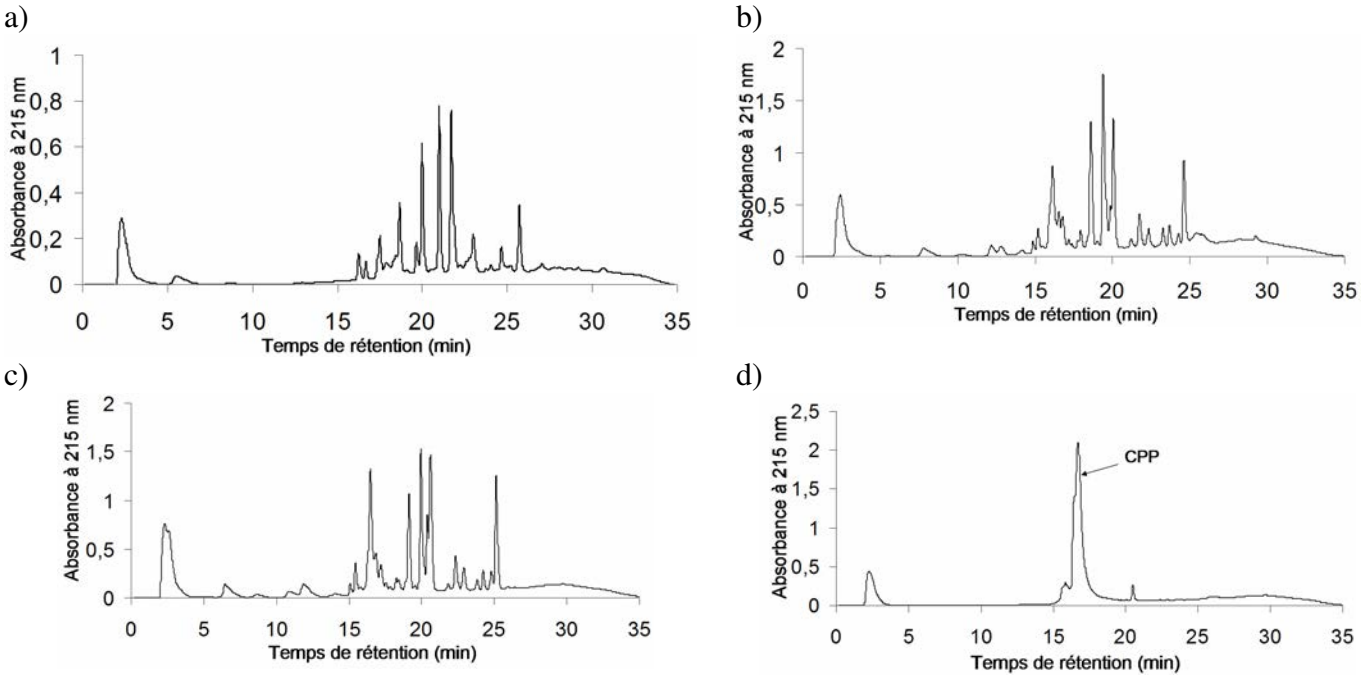


Figure 66 : HPLC en phase inversée des peptides trypsiques de l'isoforme 6P issus de la CN-β équine (fraction F2.4) avant et après chromatographie d'affinité au TiO_2 (MOAC). Une quantité protéique de 0,5 mg est déposée sur une colonne Nucleosil C_4 (150 x 2 mm). Après une étape isocratique à 4% d'acétonitrile pendant 5 min, un gradient linéaire de 4 à 48% d'acétonitrile a été appliqué pendant 15 min puis une étape isocratique à 64% d'acétonitrile est réalisée pendant 5 min en présence de TFA à 0,1%. Le débit est de 250 µL.min^{-1}. a) hydrolysat trypsique ; b) fraction « non liée à TiO_2 » ; c) fraction « lavage TiO_2 » ; d) fraction « liée à TiO_2 » ; CPP : caséinophosphopeptides.

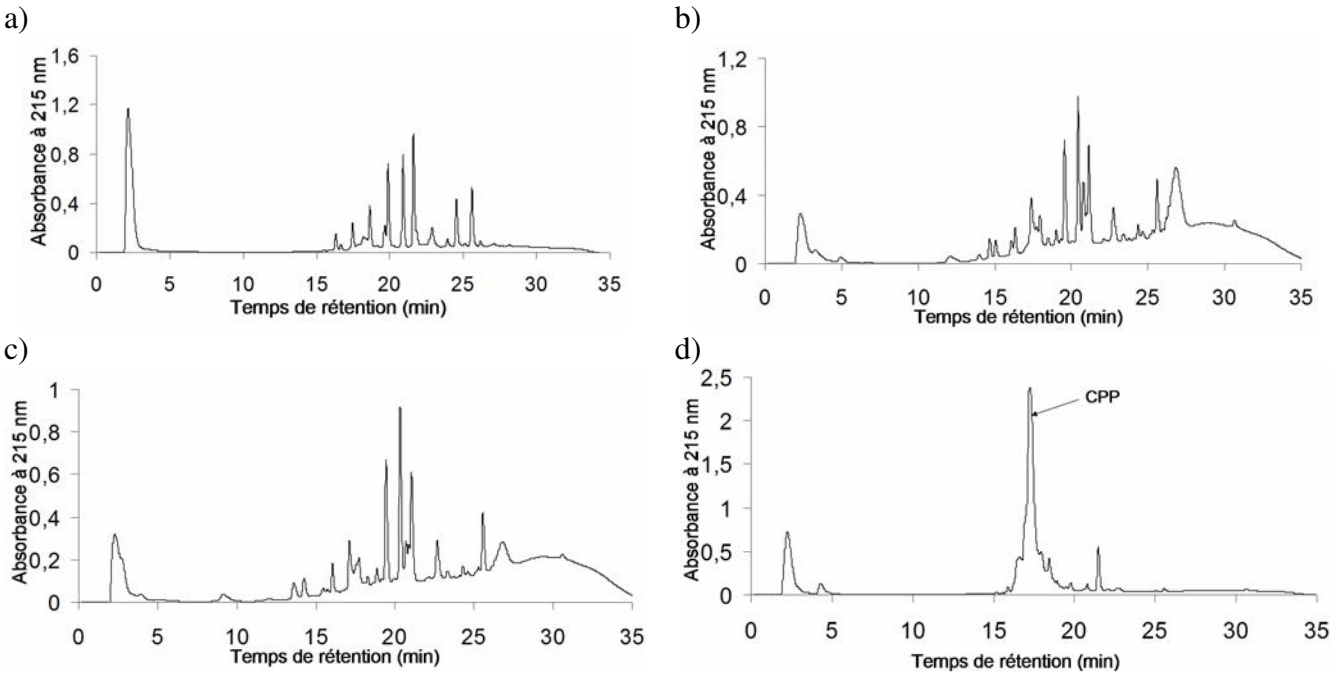


Figure 67 : HPLC en phase inversée des fragments peptides de l'isoforme 7P issus de la CN-β équine (fraction F2.5) avant et après chromatographie d'affinité au TiO_2 (MOAC). Une quantité protéique de 0,5 mg est déposée sur une colonne Nucleosil C_4 (150 x 2 mm). Après une étape isocratique à 4% d'acétonitrile pendant 5 min, un gradient linéaire de 4 à 48% d'acétonitrile a été appliqué pendant 15 min puis une étape isocratique à 64% d'acétonitrile est réalisée pendant 5 min en présence de TFA à 0,1%. Le débit est de 250 µL.min^{-1}. a) hydrolysat trypsique ; b) fraction « non liée à TiO_2 » ; c) fraction « lavage TiO_2 » ; d) fraction « liée à TiO_2 » ; CPP : caséinophosphopeptides.

En ce qui concerne l'isoforme 4P, le chromatogramme de la fraction « liée à TiO_2 » est un peu différent avec la présence de deux pics majoritaires, le premier avec un temps de rétention de 15 minutes et le second de 18 minutes (Figure 68). Dans le cas des isoformes 5P, 6P et 7P, les CPP ont un temps de rétention environ de 17 minutes. Etant donné que l'isoforme 4P est la plus hydrophobe, car la moins phosphorylée, elle devrait être mieux retenue sur la colonne que les autres isoformes. Ceci pourrait expliquer que le temps de rétention des CPP [1-34] et [4-34] de l'isoforme 4P est supérieur à celui des CPP 5P, 6P et 7P. Afin de vérifier cela, la fraction correspondant au pic de temps de rétention 18 min a été analysée par spectrométrie de masse et ce sont effectivement les CPP [1-34] et [4-34] de l'isoforme 4P qui ont été identifiés (résultat non montré). Le pic de temps de rétention 15 minutes semble correspondre à un peptide non phosphorylé.

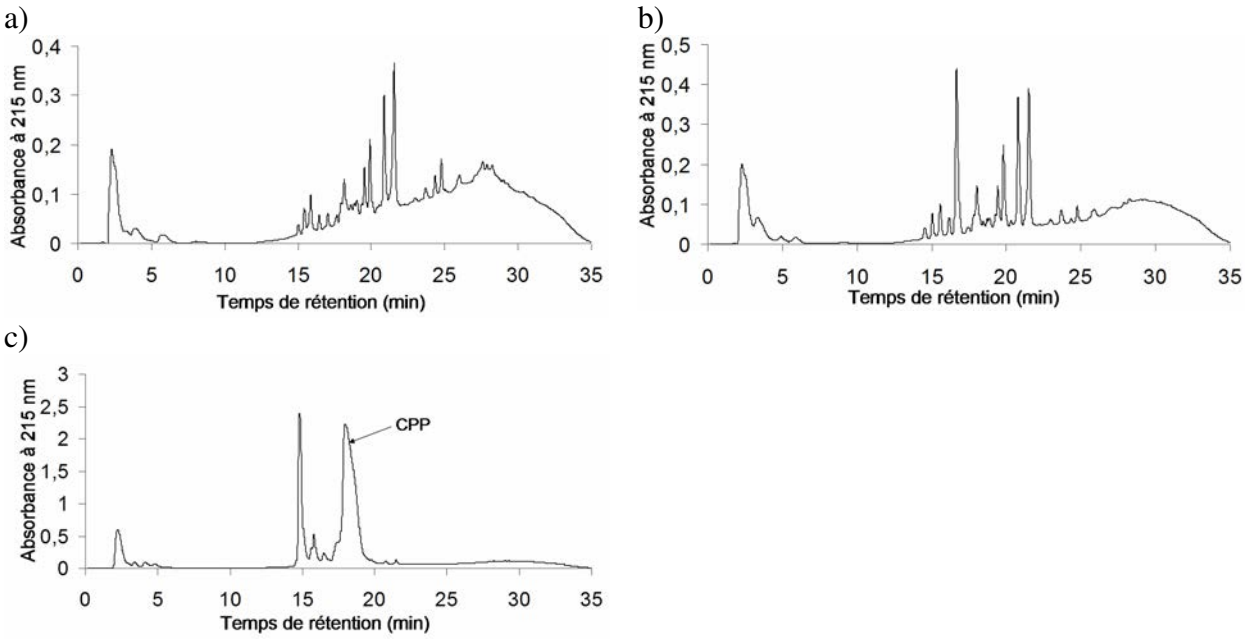


Figure 68 : HPLC en phase inversée des peptides trypsiques de l'isoforme 4P issus de la CN-β équine (fraction F2.2) avant et après chromatographie d'affinité au TiO_2 (MOAC). Une quantité protéique de 0,5 mg est déposée sur une colonne Nucleosil C_4 (150 x 2 mm). Après une étape isocratique à 4% d'acétonitrile pendant 5 min, un gradient linéaire de 4 à 48% d'acétonitrile a été appliqué pendant 15 min puis une étape isocratique à 64% d'acétonitrile est réalisée pendant 5 min en présence de TFA à 0,1%. Le débit est de 250 µL.min^{-1}. a) fraction « non liée à TiO_2 » ; b) fraction « lavage TiO_2 » ; c) fraction « liée à TiO_2 » ; CPP : caséinophosphopeptides.

Ainsi, les CPP de chaque isoforme de phosphorylation de la CN-β équine peuvent être isolés grâce à un enrichissement sur microcolonne d'affinité au TiO_2 suivi d'une purification par HPLC en phase inversée sur colonne C_4 permettant d'éliminer les éventuels peptides non phosphorylés qui se seraient liés de manière non spécifique à TiO_2. Néanmoins, les CPP résultant de l'hydrolyse du variant CN-$\beta^{\Delta 5}$ n'ont pas été mis en évidence, ce qui est probablement dû à la faible abondance de ce variant. Grâce à notre méthode, les CPP de la CN-β équine pourraient être préparés avec une grande pureté pour réaliser des études de chélation avec les minéraux ou oligo-élements d'intérêt nutritionnel.

De même, la purification des CPP de chaque isoforme de phosphorylation nous permet désormais d'envisager de localiser les sites de phosphorylation de chacune de ces isoformes, ce qui est indispensable avant de débuter les études sur la chélation.

3. Localisation des sites de phosphorylation de la CN-β équine par spectrométrie de masse

La CN-β équine possédant un degré de phosphorylation variable, la détermination de la localisation des sites de phosphorylation ne peut pas être menée directement sur un hydrolysat total contenant tous les CPP de 4P à 7P. La localisation des sites phosphorylés d'une isoforme donnée nous permettra de savoir si la phosphorylation de ces sites est aléatoire ou séquentielle. Cette connaissance est indispensable avant d'entreprendre des travaux visant à étudier l'influence du degré et de la séquence de phosphorylation sur la chélation de minéraux ou d'oligo-éléments d'intérêt nutritionnel. Ainsi, il s'avère nécessaire d'obtenir, pour chaque isoforme de phosphorylation, les CPP purifiés. Ceci a pu être réalisé grâce à l'hydrolyse trypsique de chaque isoforme isolée par IEC et grâce à la purification des CPP par chromatographie d'affinité au TiO_2 (MOAC) et HPLC en phase inversée sur colonne C_4.

Pour chaque isoforme de phosphorylation 4P à 7P, la fraction HPLC contenant les CPP a été collectée et hydrolysée par l'endoprotéinase GluC. Cette enzyme clive les liaisons peptidiques impliquant la fonction carboxylique des résidus Glu, et dans une moindre mesure, les résidus Asp. Les CPP, riches en résidus Glu, sont donc hydrolysés afin de permettre l'obtention de peptides de masses moléculaires plus petites comprises entre 500 et 2000 Da. La taille de ces peptides est en effet plus compatible avec une fragmentation MS/MS. Les hydrolysats obtenus avec l'endoprotéinase GluC sont analysés en LC/ESI-MS, les différents peptides permettant de déterminer la position des sites phosphorylés dans la séquence sont collectés puis infusés avec une nanopointe pour être analysés en mode nanospray par ESI-MS/MS.

Comme dans toutes nos expériences, nous avons débuté les analyses avec l'isoforme 5P. L'hydrolyse par l'endoprotéinase GluC des CPP [1-34] et [4-34] de l'isoforme 5P génère notamment les peptides [6-11], [6-14], [15-27] et [28-34]. Les peptides [6-14] et [15-27] sont alors fragmentés par MS/MS et les sites phosphorylés sont identifiés : Ser^{9}, Ser^{18}, Ser^{23}, Ser^{24} et Ser^{25} (Figure 69).

a)

b)

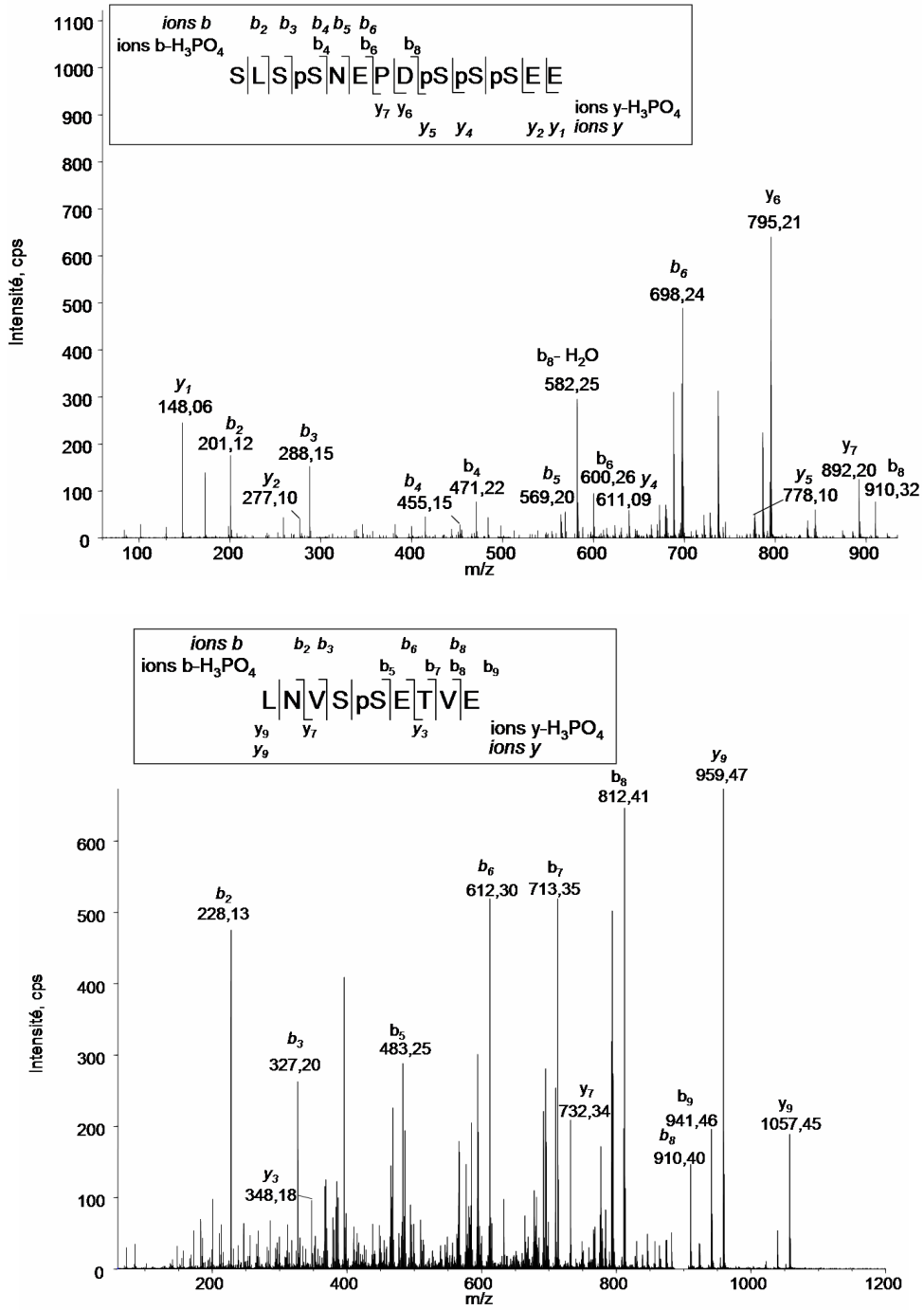


Figure 69 : Localisation des sites phosphorylés de l'isoforme 5P de la CN-β équine par LC/ESI-MS/MS. a) spectre de fragmentation de l'ion m/z 1687,3 correspondant au peptide [15-27] 4P obtenu par hydrolyse des CPP 5P par l'endoprotéinase GluC ; b) spectre de fragmentation de l'ion m/z 1057,0 correspondant au peptide [6-14] 1P obtenu par hydrolyse des CPP 5P par l'endoprotéinase GluC. Les séquences en résidus d'acides aminés correspondantes sont indiquées. pS, résidu de Ser phosphorylé ; cps, coups par seconde.

L'hydrolyse des CPP [1-34] et [4-34] de l'isoforme 4P génère notamment les peptides [5-14], [6-14], [15-27] et [28-34]. La fragmentation MS/MS du peptide [6-14] a permis d'identifier que Ser^{9} est phosphorylé. En revanche, le peptide [15-27] n'est phosphorylé qu'au niveau des résidus Ser^{23}, Ser^{24} et Ser^{25}, le résidu Ser^{18} n'étant pas quant à lui phosphorylé (Figure 70).

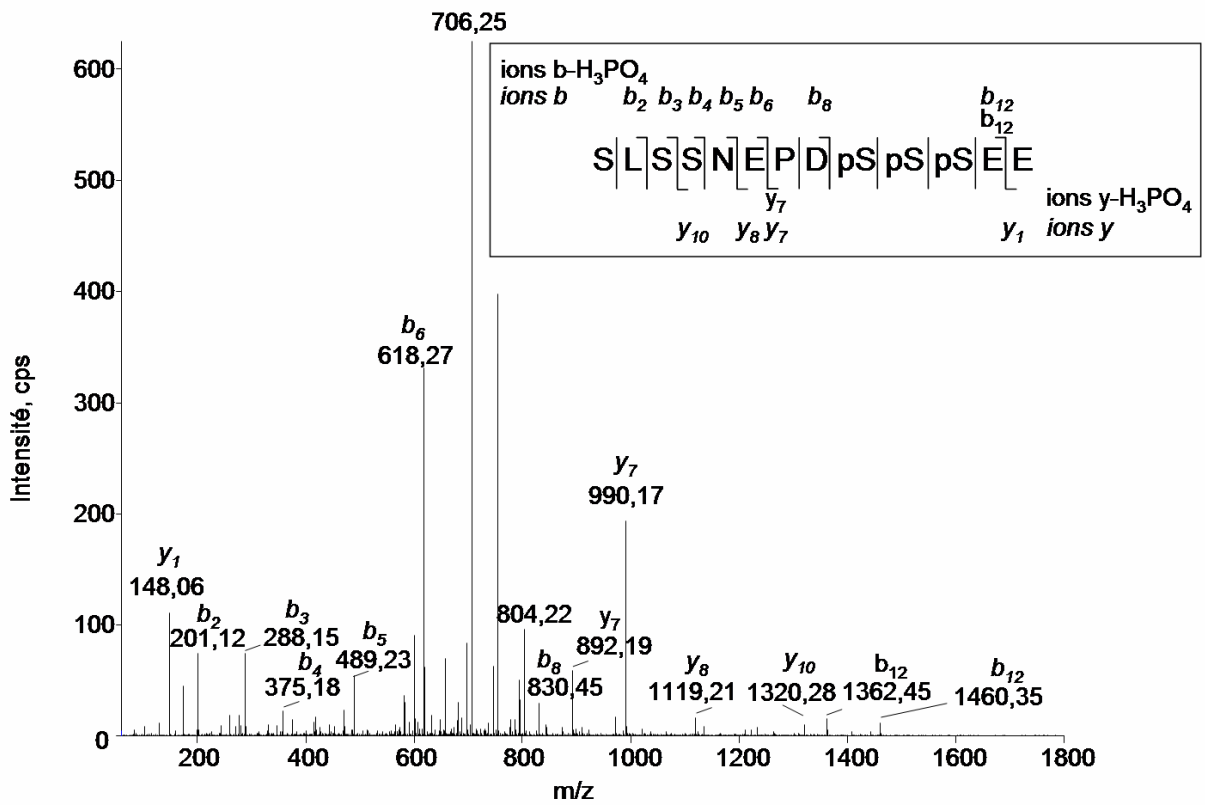


Figure 70 : Localisation des sites phosphorylés de l'isoforme 4P de la CN-β équine par LC/ESI-MS/MS. Le spectre de fragmentation présenté est celui de l'ion m/z 1606,8 correspondant au peptide [15-27] 3P obtenu par hydrolyse des CPP 4P par l'endoprotéinase GluC. La séquence en résidus d'acides aminés correspondante est indiquée. pS, résidu de Ser phosphorylé ; cps, coups par seconde.

L'hydrolyse des CPP [1-34] et [4-34] de l'isoforme 6P produit notamment les peptides [3-14], [6-14], [15-27] et [28-34]. La fragmentation MS/MS des peptides [3-14] et [6-14] montre que Ser^{9} est phosphorylée et met en évidence la présence d'un nouveau site phosphorylé par rapport au CPP issu de l'isoforme 5P : la Thr^{12}. La séquence du peptide [15-27] est identique à celle du peptide correspondant issu du CPP 5P (Figure 71).

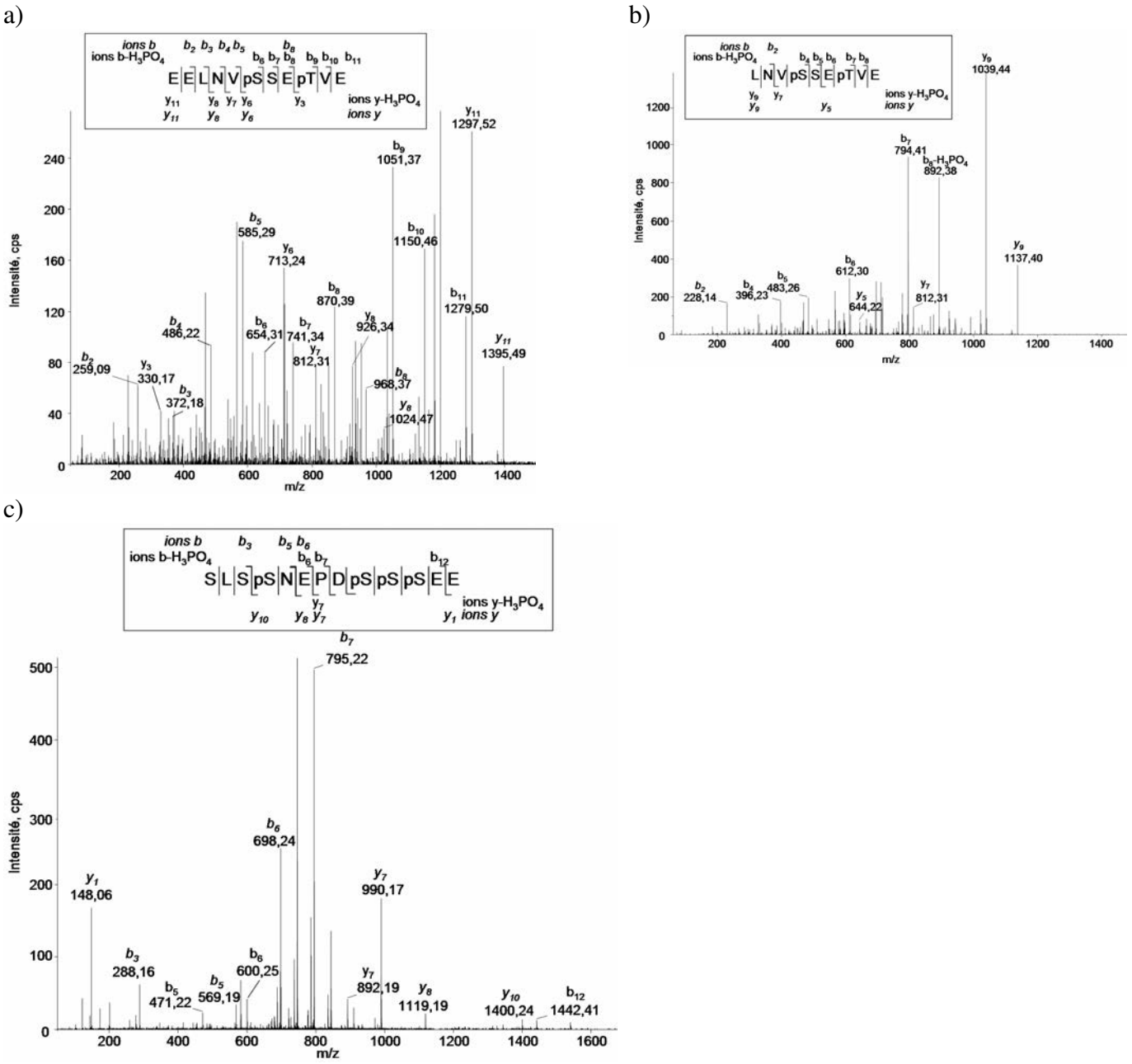


Figure 71 : Localisation des sites phosphorylés de l'isoforme 6P de la CN-β équine par LC/ESI-MS/MS. a) spectre de fragmentation de l'ion m/z 1395 correspondant au peptide [3-14] 2P obtenu par hydrolyse des CPP 6P par l'endoprotéinase GluC ; b) spectre de fragmentation de l'ion m/z 1137 correspondant au peptide [6-14] 2P obtenu par hydrolyse des CPP 6P par l'endoprotéinase GluC ; c) spectre de fragmentation de l'ion m/z 1688 correspondant au peptide [15-27] 4P obtenu par hydrolyse des CPP 6P par l'endoprotéinase GluC. Les séquences en résdus d'acides aminés correspondantes sont indiquées. pS, résidu de Ser phosphorylé ; pT, résidus de Thr phosphorylé ; cps, coups par seconde.

L'hydrolyse des CPP [1-34] et [4-34] de l'isoforme 7P permet l'obtention des peptides [6-14], [15-27] et [28-34]. L'analyse MS/MS du peptide [15-27] établit la même séquence de

phosphorylation que pour le peptide correspondant issu de l'isoforme 6P. La fragmentation du peptide [6-14] permet d'identifier le septième site phosphorylé qui est Ser [10] (Figure 72).

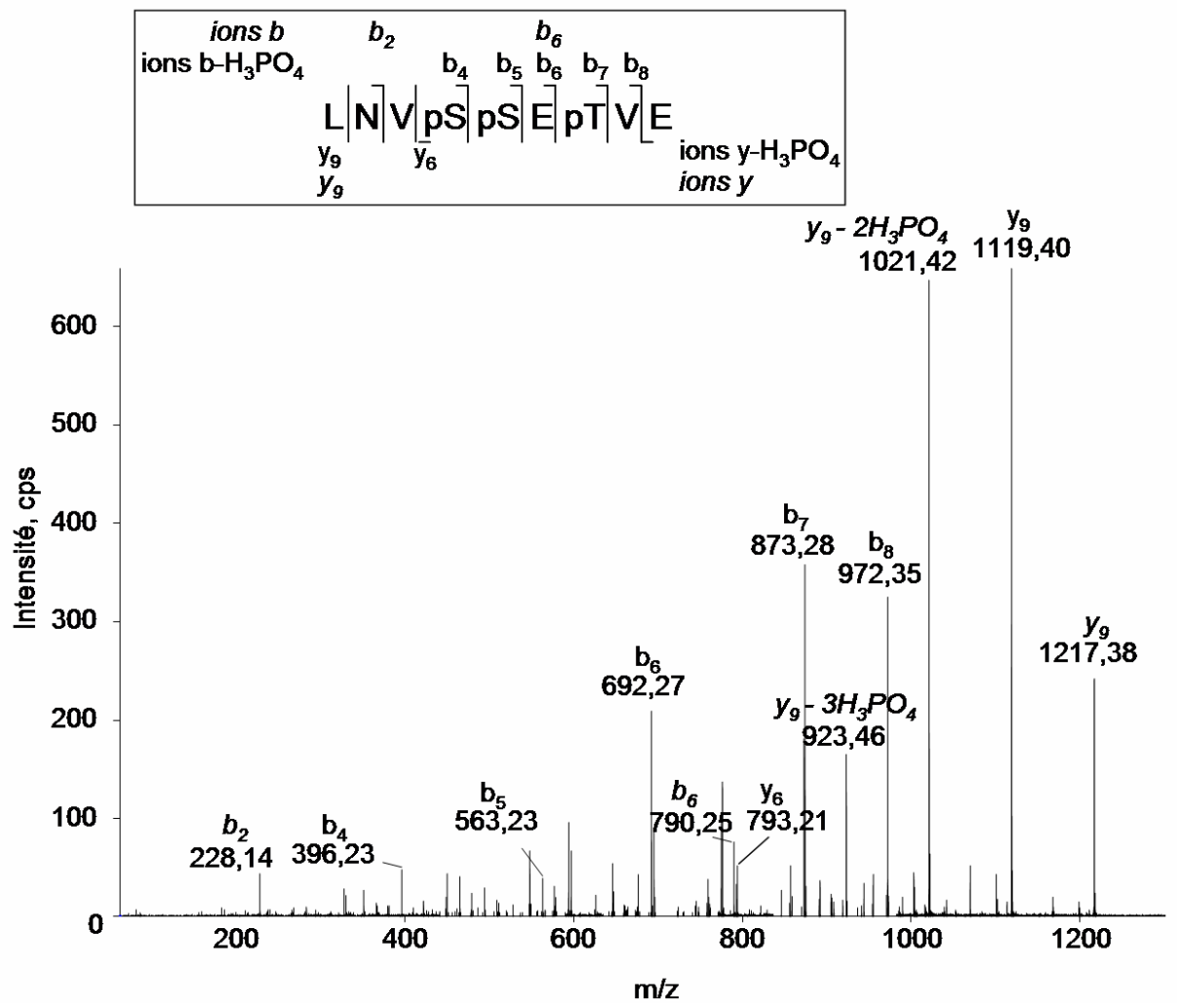


Figure 72 : Localisation des sites phosphorylés de l'isoforme 7P de la CN-β équine par LC/ESI-MS/MS. Le spectre de fragmentation présenté est celui de l'ion m/z 1217 correspondant au peptide [6-14] 3P obtenu par hydrolyse des CPP 7P par l'endoprotéinase GluC. La séquence en résidus d'acides aminés correspondante est indiquée. pS, résidu de Ser phosphorylé ; pT, résidus de Thr phosphorylé ; cps, coups par seconde.

Ainsi, les résidus Ser[9], Ser[23], Ser [24] et Ser [25] sont les résidus phosphorylés de l'isoforme 4P puis le résidu Ser[18] constitue le cinquième site de phosphorylation, Thr[12] est le sixième site et Ser[18] le septième site de phosphorylation (Figure 73). Ces sites phosphorylés sont tous inclus dans des séquences de type Ser/Thr-Xxx-Aaa qui sont des séquences de reconnaissances de la caséine kinase de la glande mammaire. Si Aaa est un résidu d'acide

aminé acide, tel que Glu, il s'agit d'un site primaire de phosphorylation et si Aaa est un résidu SerP alors il s'agit d'un site secondaire de phosphorylation (Greenberg *et al.,* 1984).

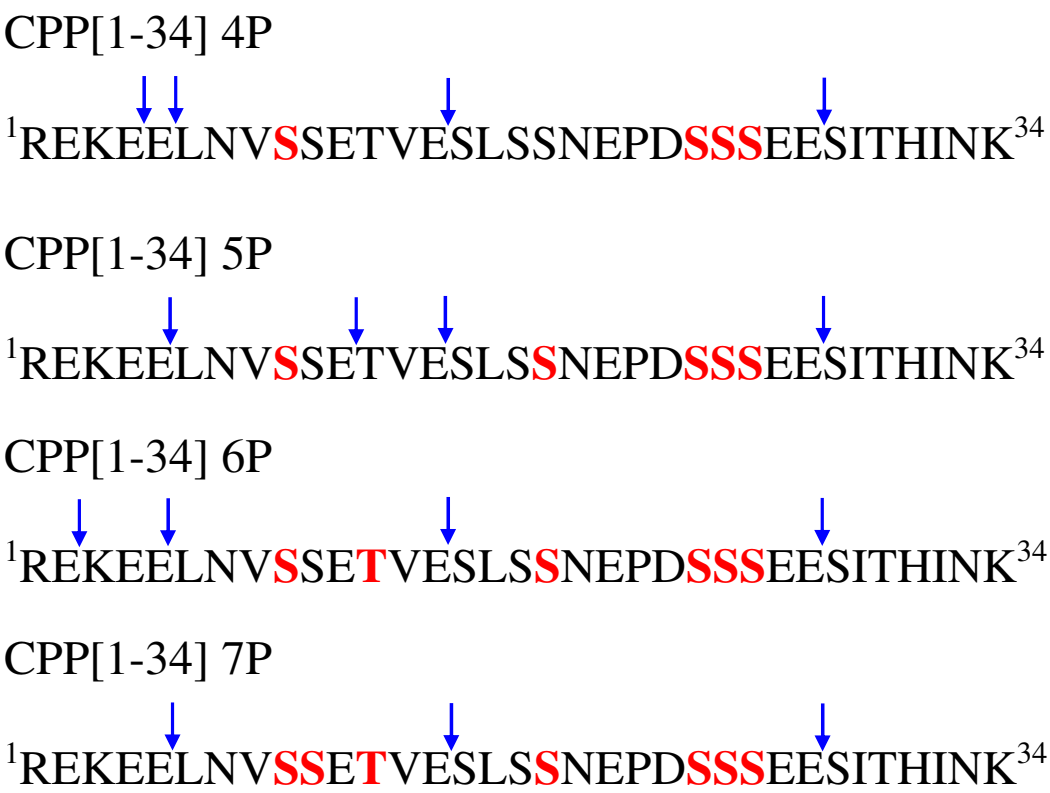


Figure 73 : Structure primaire des CPP [1-34] issus de l'hydrolyse trypsique de la CN-β équine et localisation des sites phosphorylés pour chaque isoforme de phosphorylation. Les résidus phosphorylés sont indiqués en rouge. Les flèches indiquent les sites d'hydrolyse de l'endoprotéinase Glu-C déterminés par LC/ESI-MS/MS.

La CN-β équine comporte donc cinq sites primaires de phosphorylation (Ser^{9}, Ser^{18}, Ser^{24}, Ser^{25} et Thr^{12}) et deux sites secondaires de phosphorylation ($Ser^{10, 23}$). En revanche, les sites potentiels de phosphorylation Ser^{15} et Ser^{17} qui correspondent à des séquences du type Ser-Xxx-Xxx-SerP/Glu reconnues par la caséine kinase II ne sont pas phosphorylés montrant bien que cette enzyme n'est pas impliquée dans la phosphorylation des caséines physiologiquement. La phosphorylation de la CN-β équine est réalisée de manière séquentielle, car aucune isoforme phosphorylée partiellement de manière aléatoire n'a été observée, comme dans le cas de la CN-β humaine. En effet, la CN-β humaine possède également un degré de phosphorylation variable allant de 0P à 5P (Greenberg *et al.,* 1984). L'isoforme 1P est phosphorylée à 50% sur le résidu Ser^{9} et à 50% sur le résidu Ser^{10}, ces deux sites sont effectivement équivalents car ce sont des sites primaires de phosphorylation et la

caséine kinase n'a pas de préférence entre Ser^{9} et Ser^{10}. Les sites phosphorylés de la protéine humaine sont les résidus Ser^{6}, Ser^{8}, Ser^{9}, Ser^{10} et le résidu Thr^{3} qui est le cinquième et dernier site de phosphorylation (Figure 74).

CPP[1-18] 1P

^{1}RETIESLS**SS**EESIPEYK^{18}

CPP[1-18] 2P

^{1}RETIESLS**SS**EESIPEYK^{18}

CPP[1-18] 3P

^{1}RETIESL**SSS**EESIPEYK^{18}

CPP[1-18] 4P

^{1}RETIE**S**L**SSS**EESIPEYK^{18}

CPP[1-18] 5P

^{1}RE**T**IE**S**L**SSS**EESIPEYK^{18}

Figure 74 : Structure primaire des CPP [1-18] issus de l'hydrolyse trypsique de la CN-β humaine et localisation des sites phosphorylés pour chaque isoforme de phosphorylation selon Greenberg *et al.* (1984). Les résidus phosphorylés sont indiqués en rouge. Dans le cas de l'isoforme 1P, les résidus phosphorylés sont indiqués en bleu car dans 50% des cas, le résidu Ser^{9} est phosphorylé et dans 50% des cas c'est le résidu Ser^{10} (Greenberg *et al.,* 1984).

La CN-β bovine ne possède pas de phosphorylation séquentielle. Elle ne possède qu'une isoforme de phosphorylation contenant en moyenne 4,9 groupements phosphates par molécule (Bandura *et al.,* 2002). Les résidus phosphorylés sont les résidus Ser^{15}, Ser^{17}, Ser^{18}, Ser^{19} qui forment le cluster de phosphorylation caractéristique des caséines ainsi que le résidu Ser^{35} (Bandura *et al.,* 2002). Le lait de bufflonne d'Asie (ou buflesse d'eau) contient deux

variants génétiques de CN-β, les variants A et B, qui diffèrent par trois substitutions. Le variant B comporte les mêmes sites de phosphorylation que la CN-β bovine (variants A1 et A2 confondus). Dans le variant A, une des substitutions concerne le résidu Met^{41} qui est remplacé par un résidu Thr^{41}. Cela entraîne la formation de la séquence Thr-Glu-Asp qui est un site de reconnaissance de la caséine kinase de la glande mammaire (Mercier, 1981). Ainsi, le résidu Thr^{41} du variant A est phosphorylé et augmente le degré de phosphorylation à 6P (Ferranti *et al.,* 1998). Les caséines β caprine et ovine possèdent également, comme la bufflonne d'Asie, deux isoformes de phosphorylation 5P et 6P (Richardson et Creamer, 1974 ; Richardson et Mercier, 1979). L'isoforme 5P de la CN-β ovine possède les mêmes sites phosphorylés que la CN-β bovine et la CN-β de bufflonne d'Asie. L'isoforme 6P est due à la phosphorylation de résidu Thr^{12} (Richardson et Mercier, 1979). La CN-β caprine a un schéma de phosphorylation différent. L'isoforme 5P a été retrouvée sous deux formes différentes : selon Trujillo *et al.* (1997), ce sont les résidus Ser^{15}, Ser^{17}, Ser^{18}, Ser^{19}, Ser^{35} (les mêmes sites que ceux de la protéine bovine) qui sont phosphorylés tandis que pour Neveu *et al*. (2002), ce sont les résidus Ser^{15}, Ser^{17}, Ser^{18}, Ser^{19} et le résidu Thr^{12}. Une isoforme 6P ayant en plus le résidu Ser^{35} phosphorylé a de plus été identifiée pour la protéine caprine par Neveu *et al.* (2002). Des isoformes 3P et 4P de cette CN-β caprine ont également été décrites par Chianese *et al.* (1993).

Si l'on compare le motif de phosphorylation de la CN-β équine avec ceux des espèces citées ci-dessus, nous observons que la CN-β équine comporte deux clusters de phosphorylation supplémentaires (^{9}SpSpETpVE^{14} et ^{23}SpSpSpEE^{27} ; Figures 3 et 73) compensant le manque de résidus phosphoséryls de la région codée par l'exon 4 (un seul résidu phosphorylé dans la séquence ^{17}SSpNE^{20}). Cet exon contient habituellement le cluster de phosphorylation des CN-β des autres espèces étudiées. Miranda *et al.* (2004) ont déjà remarqué que l'exon 4', uniquement observé chez la jument, code pour un cluster de résidus phosphoséryls substitutif. Nos travaux ont donc montré qu'un deuxième cluster substitutif s'ajoute à celui codé par l'exon 4' et se trouve dans la région codée par la fin de l'exon 3 et le début de l'exon 4. Le nombre élevé de résidus phosphorylés de la CN-β équine pourrait avoir une incidence sur la capacité de liaison des CPP avec les minéraux et oligo-éléments. Le résidu Thr phosphorylé correspond au sixième site phosphorylé sur les sept de la CN-β équine, ce qui peut être mis en parallèle avec la CN-β caprine (Neveu *et al.,* 2002) où il s'agit du cinquième site phosphorylé sur les six possibles. Cependant, en général, le résidu Thr est le dernier résidu à être phosphorylé de manière séquentielle (brebis, femme, bufflonne d'Asie).

La phosphorylation du résidu Thr^{12} de la CN-β équine est peut-être due au fait que Thr^{12} se trouve entre les résidus $SerP^{9}$ et $SerP^{18}$, cet environnement acide favorisant alors sa phosphorylation. La phosphorylation du résidu Thr^{12} permet par la suite la phosphorylation du résidu Ser^{10}.

4. Conclusions et perspectives

Les CPP des différentes isoformes de phosphorylation de la CN-β équine ont pu être purifiés à partir d'un hydrolysat trypsique de chacune de ces isoformes grâce à l'association d'une chromatographie d'affinité au TiO_2 (appelée MOAC) pour effectuer un enrichissement en CPP et d'une purification par HPLC en phase inversée sur colonne C_4. La stratégie utilisant la méthode MOAC est plus efficace que la précipitation sélective au chlorure de calcium pour purifier les CPP. En effet, les CPP peuvent de plus fixer le calcium, ce qui n'est pas compatible avec les études de chélation envisagées, la méthode MOAC évite donc l'emploi de cations. De plus, la méthode MOAC est plus sélective que la précipitation au chlorure de calcium, car peu de peptides non phosphorylés co-éluent avec les CPP.

Les sites phosphorylés des isoformes 4P, 5P, 6P et 7P ont pu être identifiés par spectrométrie de masse MS/MS à partir des CPP purifiés. Les sites phosphorylés sont les résidus Ser^{9}, Ser^{10}, Ser^{18}, Ser^{23}, Ser^{24}, Ser^{25} ainsi que le résidu Thr^{12} constituant le sixième et avant-dernier site phosphorylé (site phosphorylé dans l'isoforme 6P). La phosphorylation de la CN-β équine est réalisée par la caséine kinase de la glande mammaire de manière séquentielle. Cette protéine possède en outre deux sites qui pourraient être potentiellement phosphorylés par la caséine kinase II, enzyme nucléaire et cytosolique, or ce travail a montré que ces sites ne sont pas phosphorylés indiquant bien que la caséine kinase II n'est pas physiologiquement impliquée dans la phosphorylation de la CN-β équine.

Ainsi, la séquence de phosphorylation des isoformes 4P à 7P de la CN-β équine est connue et désormais des études de chélation avec notamment des minéraux ou oligo-éléments d'intérêt nutritionnel sont envisageables pour étudier l'influence du degré et de la séquence de phosphorylation sur cette chélation. Il est envisagé d'étudier la chélation par microcalorimétrie à titration isotherme. Une mise au point de cette technique est actuellement réalisée au laboratoire par M^{elle} Faiza Zidane avec le CPP [1-25] de la CN-β bovine qui contient quatre résidus SerP.

La méthode de purification des CPP pourrait être optimisée de manière à essayer d'éliminer totalement la rétention de peptides non phosphorylés sur les molécules de TiO_2 immobilisées afin d'éviter la dernière étape par HPLC en phase inversée. Notamment, l'utilisation d'une concentration plus élevée de TFA (6%) permettrait d'empêcher la liaison de peptides non phosphorylés (Iminashi *et al.,* 2007). Le DHB peut également être remplacé par de l'acide glycolique à 1 M qui ne présente pas les inconvénients du DHB tels que les soucis de contamination des systèmes chromatographiques (Jensen et Larsen, 2007). Les peptides phosphorylés peuvent interagir avec les surfaces plastiques des microtubes ou des cônes de micropipettes utilisés (Stewart *et al.,* 2001), ce qui conduit à des pertes de phosphopeptides et à une baisse de rendement. L'utilisation de détergents tel que le SDS à 1% peut empêcher l'adhésion des phosphopeptides sur les parois plastiques et donc augmenter la quantité de CPP récupérée (Jensen et Larsen, 2007).

CONCLUSION GENERALE ET PERSPECTIVES

Le lait de jument est un produit traditionnellement utilisé en Russie et en Asie Centrale pour des vertus thérapeutiques, cosmétiques, voire même « vivifiantes » dont la véracité est encore loin d'être étayée par des données scientifiques. Néanmoins, étant donné l'intérêt que suscite ce produit, notamment dans les pays d'Europe Occidentale, il semble nécessaire d'approfondir les connaissances concernant les composés protéiques de ce lait, notamment les caséines puisqu'elles pourraient être responsables de phénomènes allergiques mais aussi être la source de phosphopeptides d'intérêt nutritionnel potentiel. Si les études sur la caractérisation du lait bovin sont nombreuses, peu se sont intéressées au lait équin.

Il a été montré que le degré de phosphorylation et la séquence des CPP bovins issus des CN-β et CN-α_{s1} peuvent avoir une influence sur l'absorption intestinale du fer (Gaucheron *et al.,* 1997 ; Bouhallab *et al.,* 2002, Ani-Kibangou *et al.,* 2005). Dans le lait équin, Girardet *et al.* (2006) ont mis en évidence une phosphorylation variable de la CN-β allant de 3P à 7P. Les profils électrophorétiques observés pour la CN-α_{s1} équine (Egito *et al.,* 2002) laisseraient également supposer que cette caséine aurait un degré de phosphorylation variable. Il est donc intéressant d'étudier quelle serait l'influence du degré de phosphorylation et de la séquence des CPP issus de ces caséines équines sur la chélation de minéraux ou d'oligo-éléments d'intérêt nutritionnel. De plus, les CPP équins peuvent servir de modèles d'étude de la chélation des minéraux par les CPP humains qui possèdent, eux aussi, un degré de phosphorylation variable. En effet, il est relativement plus simple de se procurer du lait équin plutôt que du lait humain et une valorisation de composés issus du lait équin peut de plus être envisagée. La chélation de certains minéraux et oligo-éléments par les CPP équins pourrait être également comparée avec celle effectuée par des CPP bovins partiellement déphosphorylés, ce qui crée une phosphorylation variable expérimentale. Ceci pourrait permettre l'émergence de séquences bovines avec un plus grand potentiel de chélation et de transport. Afin de mener à bien de telles études, il est d'abord nécessaire de caractériser précisément d'un point de vue structural les caséines équines et notamment les caséines β et α_{s1} qui sont majoritaires. Les travaux ont été effectués sur du lait de jument de race Haflinger provenant d'une jumenterie de l'Est de la France.

La première partie de ce travail a été consacrée à la caractérisation au niveau protéique de l'hétérogénéité de la CN-α_{s1} équine. Le caséinate de sodium équin a été fractionné par chromatographie FPLC d'interactions hydrophobe (HIC) permettant l'obtention de deux fractions contenant de la CN-α_{s1} ; dans l'une de ces deux fractions, la CN-κ a co-éluée avec la CN-α_{s1} (comme cela est aussi le cas lors d'un fractionnement HPLC en phase inversée sur

une colonne C_{18}). Le mélange CN-α_{s1}/CN-κ a été séparé de façon efficace par chromatographie d'affinité à la lectine de germe de blé (WGA). La CN-κ interagit avec la WGA *via* son groupement prosthétique glycannique, tandis que la CN-α_{s1}, non glycosylée, est directement éluée de la colonne. En outre, la CN-κ n'a pas présenté d'affinité lorsque de la WGA succinylée remplaçait la WGA native dans la phase stationnaire, montrant ainsi que la CN-κ est sialylée et ne semble pas contenir de résidu GlcNAc. La CN-κ équine posséderait ainsi des structures glycanniques similaires à celles de la CN-κ bovine qui contient des résidus NeuAc mais pas de résidu GlcNAc. Ces structures glycanniques seraient par conséquent très différentes de celles de la CN-κ humaine qui contient une forte proportion de résidus GlcNAc. Une étude plus approfondie utilisant des outils tels que la spectrométrie de masse ou la résonance magnétique nucléaire pourra être entreprise afin de déterminer précisément les structures glycanniques de la CN-κ équine ainsi que leur localisation sur la séquence polypeptidique.

La caractérisation des variants post-transcriptionnels de la CN-α_{s1} a nécessité une étape préalable de déphosphorylation totale. En effet, le taux de groupements phosphates n'étant pas connu initialement, cela rendait très difficile l'identification par spectrométrie de masse des variants d'épissage. Une méthode de déphosphorylation en deux étapes a été mise au point permettant l'obtention d'une déphosphorylation quasi-totale de la CN-α_{s1}. Cette méthode consiste à traiter la caséine successivement à la phosphatase alcaline puis à la phosphatase acide. Il a été observé que l'ordre d'utilisation des phosphatases est important, car l'inversion de cet ordre ne produit qu'une déphosphorylation partielle de la caséine comme lors de l'utilisation d'une seule phosphatase. L'analyse par spectrométrie de masse des fractions déphosphorylées nous a permis de mettre en évidence plusieurs types d'isoformes de la CN-α_{s1} équine. Ainsi, la protéine complète (correspondant à la forme pour laquelle tous les exons ont été traduits) a été mise en évidence dans le lait équin et la présence de variants d'épissage alternatif d'exons a été montrée puisque les variants Δ7, Δ14 et Δ7,14 (cas d'épissage mutuellement exclusif d'exons) sont retrouvés dans le lait de jument Haflinger. En outre, la présence d'un site d'épissage cryptique a également été mis en évidence au niveau du premier codon de l'exon 11 entraînant la présence de quatre autres isoformes de CN-α_{s1} pour lesquelles le résidu Gln^{91} est délété : ΔQ ; Δ7,Q ; Δ14,Q et Δ7,14,Q. Ainsi, au total il existe huit isoformes de CN-α_{s1} équine dont sept sont dues à des phénomènes d'épissage alternatif soit d'exons, soit de site d'épissage cryptique. Seul les

variants Δ7 et Δ7,14 avaient déjà été retrouvés au niveau protéique dans le lait de poney gallois (Miranda *et al.,* 2004).

Afin de caractériser les variants post-traductionnels (variants de phosphorylation), les deux fractions issues de la chromatographie HIC ont été mélangées puis séparées par chromatographie d'échange d'anions en six sous-fractions qui ont été analysées par spectrométrie de masse. Il a ainsi été montré l'existence d'une phosphorylation variable de la CN-α_{s1} équine allant de 2P à 6P ou 4P à 8P selon l'absence ou la présence de la région codée par l'exon 7 qui contient deux sites potentiels de phosphorylation. Au final, ce sont donc 36 isoformes différentes de CN-α_{s1} qui ont été caractérisées dans un lait de petit mélange issu de juments Haflinger : CN-α_{s1} 2P, 4P, 5P, 6P, 7P, 8P ; CN-$\alpha_{s1}^{\Delta Q}$ 2P, 5P, 7P, 8P ; CN-$\alpha_{s1}^{\Delta 14}$ 4P, 5P, 6P, 7P, 8P ; CN-$\alpha_{s1}^{\Delta 14,Q}$ 7P, 8P ; CN-$\alpha_{s1}^{\Delta 7}$ 2P, 3P, 4P, 5P, 6P ; CN-$\alpha_{s1}^{\Delta 7,Q}$ 2P, 3P, 4P, 5P, 6P ; CN-$\alpha_{s1}^{\Delta 7,14}$ 2P, 3P, 4P, 5P, 6P et CN-$\alpha_{s1}^{\Delta 7,14,Q}$ 2P, 3P, 5P, 6P.

L'étape suivante sera de localiser les différents résidus phosphorylés pour chaque isoforme de CN-α_{s1} afin de pouvoir étudier l'influence de la phosphorylation variable et de la séquence de la CN-α_{s1} sur l'affinité des CPP issus de cette CN-α_{s1} pour les minéraux ou oligo-éléments d'intérêt nutritionnel. Cependant, la localisation des différents sites phosphorylés semble difficile à entreprendre dans l'état actuel car, bien que le fractionnement par IEC ait permis d'identifier les isoformes de phosphorylation, les sous-fractions IEC obtenues demeurent trop complexes pour envisager une analyse directe par MS/MS. La méthode consistant à prélever quantitativement les spots obtenus par électrophorèse bidimensionnelle de chacune des sous-fractions est difficile à mettre en œuvre, car plusieurs isoformes semblent co-migrer dans un seul spot ; de plus, certaines isoformes sont en trop faible quantité pour permettre une identification par spectrométrie de masse à partir d'un spot prélevé. Les travaux devront se focaliser dans un premier temps sur les variants majoritaires de la CN-α_{s1} qu'il faudra essayer de purifier par une méthode chromatographique adaptée.

La deuxième partie du travail s'est focalisée sur la CN-β équine. Girardet *et al.* (2006) ont montré la phosphorylation variable de cette caséine (3P à 7P) ainsi que la présence d'un site de désamidation non enzymatique au niveau du résidu Asn^{135} (montré *in vitro*). Un variant dû à l'épissage alternatif de l'exon 5 (CN-$\beta^{\Delta 5}$) a été mis en évidence chez le poney gallois (Miranda *et al.,* 2004) et un variant dû à l'utilisation d'un site d'épissage cryptique provoquant la délétion de 75% de la région codée par l'exon 7 (low-M_r CN-β) a, quant à lui, été mis en évidence chez la jument Haflinger (Miclo *et al.,* 2007). Cette microhétérogénéité

de la CN-β équine se reflète dans les profils électrophorétiques observés par 2-DE et par PAGE-urée qui s'avèrent être complexes et difficiles à interpréter.

Une partie du travail a donc eu pour objectif d'élucider notamment la cartographie bidimensionnelle de la CN-β. Le phénomène potentiel de désamidation non enzymatique de la CN-β a également été étudié afin d'établir si cette réaction se déroule dans le lait ou si elle a lieu pendant la préparation et/ou l'analyse de la CN-β et si elle peut ainsi contribuer à la complexité des profils électrophorétiques observés. Pour réaliser ce travail, il s'est avéré nécessaire de séparer les différentes isoformes de phosphorylation de la CN-β. La CN-β a donc été préparée par chromatographie HIC à partir de caséinate de sodium puis fractionnée par IEC permettant l'obtention de cinq sous-fractions. Chaque sous-fraction a été analysée par spectrométrie de masse, ce qui a permis de mettre en évidence la présence de chacune des isoformes de phosphorylation de la CN-β dans chaque sous-fraction IEC, mais également cela a révélé la présence des isoformes 4P à 7P de la CN-$\beta^{\Delta 5}$ (la forme 3P n'a pas été trouvée). L'existence de ce variant CN-$\beta^{\Delta 5}$ dans le lait de jument Haflinger a été suggérée par Girardet *et al.* (2006) et nos travaux ont pu confirmer leur hypothèse. L'étude de la désamidation non enzymatique a été menée sur les isoformes majeures CN-β 5P et 6P incubées à 37°C et à pH 7,4 ou 8,4 et la réaction a été suivie au cours du temps par PAGE-urée. Les résultats montrent que les isoformes qui ont subi une désamidation ne migrent pas au même niveau que les isoformes natives dans le gel. Les bandes correspondantes s'intercalent entre des bandes d'isoformes natives (par exemple, CN-β 5P/d migre entre CN-β 5P et CN-β 6P) ; ce qui montre accessoirement que la charge négative apportée par la fonction acide carboxylique est moins forte que celle(s) apportée(s) par le groupement phosphate. Ces résultats montrent également que la réaction de désamidation non enzymatique n'a pas lieu dans le lait ou lors des étapes de préparation de la CN-β, car aucune bande de forme désamidée n'est observée quand la caséine n'est pas volontairement incubée à 37°C et à un pH proche du neutre ou alcalin. La réaction de désamidation non enzymatique est donc provoquée par des conditions d'incubation particulières de température et de pH qui s'apparentent aux conditions physiologiques. Par ailleurs, cette réaction a également lieu lors de la mise en œuvre d'une analyse 2-DE, quand l'étape de réhydratation des bandelettes « IPG-strip » se déroule sous les conditions habituelles, c'est-à-dire à température ambiante (22°C environ) et à pH alcalin. On observe alors une superposition de spots correspondant à des isoformes « natives » avec des spots correspondant à des isoformes « désamidées ». Une diminution de la température à 10°C pendant l'étape d'isoélectrofocalisation permet d'inhiber fortement cette réaction

chimique. Il a de plus été constaté une légère protéolyse lors de la réaction de désamidation non enzymatique malgré la présence de fortes concentrations d'inhibiteurs de protéases et d'agent bactériostatique. Cette protéolyse pourrait être due à un clivage spontané des liaisons peptidiques Asn-Pro et Asn-Leu de la CN-β qui serait induit par la désamidation des résidus Asn. Il serait intéressant de déterminer si les produits de protéolyse sont effectivement dus à un clivage spontané, par exemple en prélevant les spots 2-DE et en les analysant par spectrométrie de masse. Une autre perspective concernant la désamidation serait de voir si celle-ci peut se dérouler lors de la digestion des protéines du lait de jument dans l'intestin, et pour cela mimer les conditions physiologiques rencontrées par la CN-β dans le tractus digestif. Si cette réaction a lieu, il serait également intéressant d'évaluer quelles pourraient être les conséquences de cette désamidation, notamment sur les risques d'allergénicité du lait de jument. La désamidation non enzymatique d'un résidu Asn entraîne la formation de résidu isoAsp et d'une liaison isopeptidique de type β-amide potentiellement non reconnue par les protéases digestives et générant ainsi des peptides pouvant déclencher une réaction immunitaire.

Une fois (i) les différentes formes phosphorylées isolées et (ii) l'aptitude des variants de CN-β à se désamider déterminée, la cartographie bidimensionnelle a pu être établie. Pour attribuer précisément un spot donné à une isoforme phosphorylée de CN-β, nous avons mélangé une sous-fraction IEC contenant majoritairement une isoforme possédant un degré de phosphorylation donné avec la CN-β totale. Nous avons déterminé ensuite quel spot 2-DE voyait son intensité augmenter après coloration au bleu de Coomassie. Ainsi, nos travaux ont permis d'élucider la complexité électrophorétique de la CN-β équine observée par 2-DE. Les conditions expérimentales ont été maîtrisées et nous permettent d'obtenir des spots purs non contaminés par des formes désamidées. Ainsi, un prélèvement direct des spots à partir du gel pourrait être envisagé afin de démontrer de manière directe l'attribution des spots de masse moléculaire apparente de 30 kDa à l'isoforme CN-$\beta^{\Delta 5}$. Cependant, il reste encore à déterminer la localisation précise des isoformes phosphorylées du variant d'épissage CN-$\beta^{\Delta 5}$ sur un profil de PAGE-urée.

Les isoformes de phosphorylation de la CN-β ayant été séparées par IEC, la troisième et dernière partie de ce travail s'est focalisée sur la préparation des CPP issus de l'hydrolyse trypsique de chacune des isoformes phosphorylées et à la localisation des résidus phosphorylés (l'isoforme 3P, présente en trop faible quantité, n'a pas été étudiée). Pour séparer les CPP des peptides non phosphorylés, une méthode récente, la chromatographie

d'affinité au TiO_2 (ou MOAC), a été développée dans cette partie. Cette technique a été préférée à la précipitation sélective au calcium (méthode de Reynolds), car cette dernière peut entraîner la chélation d'ions Ca^{2+} par les CPP, ce qui n'est pas souhaitable pour des études ultérieures de chélation de minéraux ou d'oligo-éléments. De plus, la technique MOAC s'est révélée être plus sélective (même en absence de DHB), car peu de peptides non phosphorylés co-éluent avec les CPP contrairement à ce qui a été observé avec la méthode de Reynolds. Pour éliminer les peptides non phosphorylés contaminant éventuellement la fraction affine pour TiO_2, nous avons purifié les CPP par HPLC en phase inversée sur colonne C_4. Chaque CPP purifié a été ensuite identifié par spectrométrie de masse puis hydrolysé par l'endoprotéinase GluC afin de générer des fragments de plus petites tailles facilement séquençables par spectrométrie MS/MS. De cette manière, les sites phosphorylés de la CN-β équine ont été identifiés pour chaque degré de phosphorylation allant de 4P à 7P. Les résidus phosphorylés de la CN-β 4P sont les résidus Ser^9, Ser^{23}, Ser^{24}, Ser^{25}. Puis la phosphorylation du résidu Ser^{18} permet de générer la CN-β 5P. Le résidu Thr^{12} constitue le sixième site de phosphorylation et le résidu Ser^{10}, le septième. Les sites phosphorylés sont tous inclus dans une séquence de phosphorylation typique de la caséine kinase de la glande mammaire, les résidus Ser^9, Ser^{18}, Ser^{24}, Ser^{25} et Thr^{12} sont des sites primaires de phosphorylation tandis que les résidus Ser^{10} et Ser^{23} sont des sites secondaires. La phosphorylation de la CN-β équine se déroule de manière séquentielle comme c'est le cas pour la CN-β humaine (Greenberg *et al.,* 1984). Aucune isoforme partiellement phosphorylée de manière aléatoire n'a en effet été observée.

Désormais, des études sur l'influence du degré variable et de la séquence de phosphorylation sur la chélation par les CPP de minéraux et d'oligo-éléments d'intérêt nutritionnel peuvent être envisagées avec, comme modèle, les CPP issus de la CN-β équine. L'affinité des CPP équins pour des minéraux ou d'oligo-éléments peut être étudiée par microcalorimétrie à titration isotherme (ITC). L'ITC est la seule méthode directe permettant de mesurer, à température constante, la chaleur libérée ou absorbée lors de l'interaction entre une macromolécule et un ligand. C'est également la seule méthode permettant de déterminer en une seule expérience et de façon simultanée, les paramètres de liaison tels que la constante d'affinité (Ka), l'enthalpie (ΔH), l'entropie (ΔS), l'énergie libre (ΔG) et la stoechiométrie de l'interaction (n) (Perozzo *et al.,* 2004). Toutes ces caractéristiques font de l'ITC une technique de choix pour l'étude de la complexation minéral/CPP. Une mise au point de cette méthode est actuellement conduite au laboratoire avec le CPP [1-25] issu de la CN-β bovine (qui

contient quatre SerP) avant d'être appliquée aux CPP équins. Cette méthode nécessite néanmoins des quantités de CPP équins supérieures à celles qui ont été obtenues jusqu'à présent. La technique de purification des CPP équins pourrait alors être améliorée en empêchant notamment la rétention de peptides non phosphorylés sur le TiO_2, ce qui pourrait permettre de s'affranchir de l'étape de purification des CPP par HPLC en phase inversée sur colonne C_4 qui diminue le rendement. Pour cela, des concentrations plus importantes en TFA, de l'ordre de 6% comme préconisé par Iminashi *et al.* (2007), ou encore de l'acide glycolique à 1 M (Jensen et Larsen, 2007) pourraient être utilisées pour empêcher l'adsorption de peptides non phosphorylés sur la phase stationnaire. De même, il a été montré que les CPP pouvaient adhérer aux surfaces plastiques des divers tubes et cônes de pipettes utilisés lors d'un enrichissement réduisant ainsi la quantité de CPP retenue (Stewart *et al.,* 2001). Pour y remédier, du SDS à 1% peut être utilisé afin d'augmenter le rendement (Jensen et Jensen, 2007) mais son élimination ultérieure peut poser certains problèmes. Obtenus en plus grande quantité, les phosphopeptides équins pourraient permettre de comprendre quels sont les motifs importants pour leur interaction avec un minéral ou oligo-élément et d'anticiper la force de cette liaison. Des phosphopeptides pouvant solubiliser des minéraux d'intérêt tout en permettant un relargage optimal pourraient alors être sélectionnés.

REFERENCES BIBLIOGRAPHIQUES

Achsel T., Brahms H., Kastner B., Bachi A., Wilm M. & Lührmann R. (1999). A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation *in vitro*. *The EMBO Journal*, 18, 5789-5802.

Addeo F., Chianese L., Salzano A., Sacchi R., Cappuccio V. & Ferranti P. (1992). Characterization of the 12% trichloroacetic acid-insoluble oligopeptides of Parmigiano-Reggiano cheese. *Journal of Dairy Research*, 59, 401-411.

Adkins Y., Lepine A.J. & Lönnerdal B. (2001). Changes in protein and nutrient composition of milk throughout lactation in dogs. *American Journal of Veterinary Research,* 62, 1266-1272.

Aït-oukhatar N., Bouhallab S., Bureau F., Arhan P., Maubois J.L., Drosdowsky M. & Bouglé D. (1997). Bioavailability of caseinophosphopeptide bound iron in the young rat. *The Journal of Nutritional Biochemistry*, 8, 190-197.

Aït-oukhatar N., Bouhallab S., Arhan P., Maubois J.L., Drosdowsky M. & Bouglé D. (1999). Iron tissue storage and hemoglobin levels of deficient rats repleted with iron bound to the caseinophosphopeptide 1-25 of β-casein. *Journal of Agricultural and Food Chemistry*, 47, 2786-2790.

Aït-oukhatar N., Pérès J.M., Bouhallab S., Neuville D., Bureau F., Bouvard G., Arhan P. & Bouglé D. (2002). Bioavailability of caseinophosphopeptide-bound iron. *Journal of Laboratory and Clinical Medicine*, 140, 290-294.

Akoh C.C. & Min D.B. (2002). Food Lipids Chemistry, Nutrition and Biotechnology, 2nd edition. New-York, 388 pages.

Alexander L.J., Stewart A.F., Mackinlay A.G., Kapelinskaya T.V., Tkach T.M. & Gorodetsky S.I. (1988). Isolation and characterization of the bovine κ-casein gene. *European Journal of Biochemistry*, 178, 395-401.

Alexander L.J., Das Gupta N.A. & Beattie C.W. (1992). The sequence of porcine α_{s2}-casein cDNA. *Animal Genetics*, 23, 365-367.

Alioto T.S. (2007). U12DB: a database of orthologous U12-type spliceosomal introns. *Nucleic Acids Research*, 35 (Database issue), 110-115.

Allende J.E. & Allende C.C. (1995). Protein kinase CK2: An enzyme with multiple substrates and puzzling regulation. *FASEB Journal*, 9, 313-323.

Aimutis W.R. (2004). Bioactive properties of milk proteins with particular focus on anticariogenesis. *The Journal of Nutrition*, 134, 989-995.

Andersson L. & Porath J. (1986). Isolation of phosphoproteins by immobilized metal (Fe^{3+}) affinity chromatography. *Analytical Biochemistry*, 154, 250-254.

Ani-Kibangou B., Bouhallab S., Mollé D., Henry G., Bureau F., Neuville D., Arhan P & Bouglé D. (2005). Improved absorption of caseinophosphopeptide-bound iron: Role of alkaline phosphatase. *The Journal of Nutritional Biochemistry*, 16, 398-401.

Aoki T., Umeda T. & Kako Y. (1992). The least number of phosphate groups for crosslinking of casein by colloidal calcium phosphate. *Journal of Dairy Science*, 75, 971-975.

Aruoma O.I., Halliwell B., Grootveld M. & Chaudhary S.S. (1989). Binding of iron (II) ions to the pentose sugar 2-deoxyribose. *Journal of Inorganic Biochemistry*, 35, 149-155.

Autelitano D.J., Rajic A., Smith A.I., Berndt M.C., Ilag L.L. & Vadas M. (2006). The cryptome: a subset of the proteome, comprising cryptic peptides with distinct bioactivities. *Drug Discovery Today*, 11, 306-314.

Baines F.M. (1981). Milk substitutes and the hand rearing of orphan puppies and kittens. *Journal of Small Animal Practice*, 22, 555-578.

Bandura D.R., Baranov V.I. & Tanner S.D. (2002). Detection of ultratrace phosphorus and sulfur by quadrupole ICPMS with dynamic reaction cell. *Analytical Chemistry*, 74, 1497-1502.

Barton J.C., Conrad M.E. & Parmley R.T. (1983). Calcium inhibition of inorganic iron absorption in rats. *Gastroenterology*, 84, 90-101.

Beata C., Beaumont-Graff E., Coll V., Cordel J., Marion M., Massal N., Marlois N. & Tauzin J. (2007). Effect of alpha-casozepine (Zylkene) on anxiety in cats. *Journal of Veterinary Behavior*, 2, 40-46.

Beitz D.C., Mohrenweiser H.W., Thomas J.W. & Wod W.A. (1969). Synthesis of milk proteins in a cell-free system isolated from lactating bovine mammary tissue. *Archives of Biochemistry and Biophysics*, 132, 210-222.

Belizy S., Nasarova I.N, Procof'ev V.N., Sorokina I.A. Puschkina N.V. & Lukach A.I. (2001). Changes in antioxidative properties of lactoferrin from women's milk during lactation. *Biochemistry*, 66, 576-580.

Berget S.M. (1995). Exon recognition in vertebrate splicing. *The Journal of Biological Chemistry*, 270, 2411-2414.

Bergström S., Hansson L., Hernell O., Lönnerdal B., Nilsson A.-K. & Strömqvist M. (1992). Cloning and sequencing of human κ-casein cDNA. *Journal of DNA sequencing and mapping*, 3, 245-246.

Beeryl K.E., Hood L.F. & Patton S. (1971). Formation of casein micelles in Golgi vesicles of mammary tissue. *Journal of Dairy Science*, 54, 911-912.

Bevilacqua C., Martin P., Candalh C., Fauquant J., Piot M., Roucayrol A.-M., Pilla F. & Heyman M. (2001). Goats' milk of defective α_{s1}-casein genotype decreases intestinal and systemic sensitization to β-lactoglobulin in guinea pigs. *Journal of Dairy Research*, 68, 217-227.

Bhattacharyya L. & Brewer C.F. (1988). Lectin-carbohydrate interactions. Studies of the nature of hydrogen bonding between D-galactose and certain D-galactose-specific lectins, and between D-mannose and concanavalin A. *European Journal of Biochemistry*, 176, 207-212.

Bingham E.W., Farrell Jr. H.M. & Basch J.J. (1972). Phosphorylation of casein, role of the Golgi apparatus. *The Journal of Biological Chemistry*, 247, 8193-8194.

Bingham E.W. & Farrell Jr. H.M. (1974). Casein kinase from the Golgi apparatus of lactating mammary gland. *The Journal of Biological Chemistry*, 249, 3647-3651.

Bingham E.W. (1976). Modification of casein by phosphatases and protein kinases. *Journal of Agricultural and Food Chemistry*, 24, 1094-1099.

Bingham E.W. & Farrell Jr. H.M. (1977). Phosphorylation of casein by the lactating mammary gland: a review. *Journal of Dairy Science*, 60, 1199-1207.

Black D.L. (2003). Mechanisms of alternative pre-messenger RNA splicing. *Annual Review of Biochemistry*, 72, 291-336.

Boisnard M., Hue D., Bouniol C., Mercier J.C. & Gaye P. (1991). Multiple mRNA species code for two non-allelic forms of ovine α_{s2}-casein. *European Journal of Biochemistry*, 201, 633-641.

Bouhallab S., Léonil J. & Maubois J.L. (1991). Complexation du fer par le phosphopeptide (1-25) de la caséine β : action de l'alcalase et de la phosphatase acide. *Le Lait*, 71, 435-443.

Bouhallab S., Aït-oukhatar N., Mollé D., Henry G., Maubois J.L., Arhan P. & Bouglé D. (1999). Sensitivity of β-casein phosphopeptide-iron complex to digestive enzymes in ligated segment of rat duodenum. *The Journal of Nutritional Biochemistry*, 10, 723-727.

Bouhallab S., Cinga V., Aït-oukhatar N., Bureau F., Neuville D., Arhan P., Maubois J.L. & Bouglé D. (2002). Influence of various phosphopeptides of caseins on iron absorption. *Journal of Agricultural and Food Chemistry*, 50, 7127-7130.

Bouhallab S. & Bouglé D. (2004). Biopeptides of milk: caseinophosphopeptides and mineral bioavailability. *Reproduction Nutrition Development*, 44, 493-498.

Brantl V., Teschemacher H., Henschen A. & Lottspeich F. (1979). Novel opoid peptides derived from casein (β-casomorphins). I. Isolation from bovine casein peptone. *Hoppe-Seyler's Zeitschrift für Physiologie und Chemie*, 360, 1211-1216.

Breathnach R., Benoist C., O'Hare K., Gannon F. & Chambon P. (1978). Ovalbumine gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. *Proceedings of the National Academy of Sciences of the United States of America*, 75, 4853-4857.

Brommage R., Juillerat M.A. & Jost R. (1991). Influence of casein phosphopeptides and lactulose on intestinal calcium absorption in adult female rats. *Le Lait*, 71, 173-180.

Brow D.A. (2002). Allosteric cascade of spliceosome activation. *Annual Review of Genetics*, 36, 333-360.

Brulé G. & Fauquant J. (1982). Interactions des protéines du lait et des oligoéléments. *Le Lait*, 62, 323-331.

Brunati A.M., Marin O., Bisinella A., Salviati A. & Pinna L.A. (2000). Novel consensus sequence for the Golgi apparatus casein kinase, revealed using proline-rich protein-1 (PRP1)-derived peptide substrates. *Biochemical Journal*, 351, 765-768.

Burge C.B., Padgett R.A. & Sharp P.A. (1998). Evolutionary fates and origin of U12-type introns. *Molecular Cell*, 2, 773-785.

Burge C.B., Tuschl T. & Sharp P.A. (1999). Splicing of precursors to mRNAs by the spliceosomes. In: The RNA World, 2nd edition. Gesteland R. F., Cech T. R. & Atkins J. F. Editors. New York: Cold Spring Harbor Laboratory Press, pp. 525-560.

Burnett G. & Kennedy E.P. (1954). The enzymatic phosphorylation of proteins. *The Journal of Biological Chemistry*, 211, 969-980.

Businco L., Giampietro P.G., Lucenti P., Lucaroni F., Pini C., Di Felice G., Iacovacci P., Curadi C. & Orlandi M. (2000). Allergenicity of mare's milk in children with cow's milk allergy. *Journal of Allergy and Clinical Immunology*, 105, 1031-1034.

Byler M.D., Farrell Jr. H.M. & Susi H. (1988). Raman spectroscopic study of casein structure. *Journal of Dairy Science*, 71, 2622-2629.

Carles C. & Ribadeau Dumas B. (1986). Influence de la phosphorylation sur le comportement des peptides trypsiques de caséine β bovine en chromatographie liquide haute performance en phase inverse. *Le Lait*, 66, 247-256.

Carroll R.J., Thompson M.P. & Farrell Jr. H.M. (1970). Formation and structure of casein micelles in the lactating mammary tissue. *28th Annual Proceedings EMSA*, 150.

Cartegni L. & Krainer A.R. (2002). Disruption of an SF2/ASF-dependant exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN. *Nature Genetics*, 30, 377-384.

Cayot F. & Lorient D. (1998). Structures et technofonctions des protéines du lait. *Tec & Doc, Lavoisier, Paris, France,* 363 pages.

Champion-Arnaud P. & Reed R. (1994). The prespliceosome components SAP 49 and SAP 145 interact in a complex implicated in tethering U2 snRNP to the branch site. *Genes Development*, 8, 1974-1983.

Chandra N., Brew K. & Acharya K.R. (1998). Structural evidence for the presence of a secondary calcium binding in human α-lactalbumin. *Biochemistry*, 37, 4767-4772.

Chardot T., Shen H. & Meunier J.C. (1995). Dual specificity of casein kinase II from the yeast Yarrowia lipolytica. *Comptes Rendus de l'Academie des Sciences - Serie III*, 318, 937-942.

Chavous D.A., Jackson F.R. & O'Connor C.M. (2001) Extension of the *Drosophila* lifespan by overexpression of a protein repair methyltransferase. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 14814-14818.

Chianese L., Garro G., Nicolai M.A., Mauriello R., Ferranti P., Pizzano R., Cappuccio U., Laezza P., Addeo F., Ramunno L., Rando A. & Rubino R. (1993). The nature of β-casein heterogeneity in caprine milk. *Le Lait*, 73, 533-547.

Chianese L., Garro G., Mauriello R., Laezza P., Ferranti P. & Addeo F. (1996). Occurrence of five α_{s1}-casein variants in ovine milk. *Journal of Dairy Research*, 63, 49-59.

Chiba H. & Yoshikawa M. (1986). Biologically functional peptides from food proteins. In: Protein Tailoring for Food and Medical Uses. Feeney R.E. & Whitaker J.R. editors. New York: Marcel Dekker, pp. 123-153.

Chiba H., Tani F. & Yoshikawa M. (1989). Opioid antagonist peptides derived from κ-casein. *Journal of Dairy Research*, 56, 363-366.

Coddeville B., Stratil A., Wieruszeski J.M., Strecker G., Montreuil J. & Spik G. (1989). Primary structure of horse serotransferrin glycans. *European Journal of biochemistry*, 136, 583-590.

Connor P.A. & McQuillan A.J. (1999). Phosphate adsorption onto TiO_2 from aqueous solutions: an in situ internal reflection infrared spectroscopic study. *Langmuir*, 15, 2916-2921.

Conrad M.E. (1993). Regulation of iron absorption. In: Essential toxic trace elements in human health and disease: an update. Wiley-Liss editor, Inc, pp. 203-219.

Conti A., Godovac-Zimmermann J., Liberatori J. & Braunitzer G. (1984). The primary structure of monomeric β-lactoglobulin I from horse colostrum (*Equus caballus*, Perissodactyla). *Hoppe-Seyler's Zeitschrift fur Physiologische Chemie*, 365, 1393-1401.

Cramer P., Pesce C.G., Baralle F.E. & Kornblihtt A.R. (1997). Functional association between promoter structure and transcript alternative splicing. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 11456-11460.

Cross K.J., Huq N.L., Bicknell W. & Reynolds E.C. (2001). Cation-dependent structural features of β-casein-(1-25). *Biochemical Journal*, 356, 277-286.

Dalgleish D.G., Horne D.S. & Law A.J.R. (1989). Size-related differences in bovine casein micelles. *Biochimical et Biophysica Acta*, 991, 383-387.

Decker E.A. & McClements D.J. (2001). Transition metal and hydroperoxyde interactions: an important determinant in the oxidation stability of lipid dispersions. *Inform*, 12, 251-255.

DeKruif C.G. & Holt C. (2003). Casein micelle structure, functions and interactions. In: Fox P.F., McSweeney P.L.H., editors. Advanced dairy chemistry, 3rd edition. Proteins, vol. 1. New York: Kluwer Academic/Plenum publishers, pp. 213-276.

De La Mata M., Alonso C.R. Kadener S., Fededa J.P., Balustein M., Pelisch F., Cramer P., Bentley D. & Kornblihtt A.R. (2003). A slow-RNA polymerase II affects alternative splicing in vivo. *Molecular Cell*, 12, 525-532.

De La Mata M. & Kornblihtt A.R. (2006). RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20. *Nature Structural and Molecular Biology*, 13, 973-980.

Dev B.C., Sood S.M., DeWind S. & Slattery C.W. (1993). Characterization of human κ-casein purified by HPLC. *Preparative Biochemistry*, 23, 389-409.

De Waal H.O. (1990). Formulating a milk diet for lion cubs. *Animal Science*, 29, 5-6.

De Waal H.O., Osthoff G., Hugo A., Myburgh J. & Botes P. (2004). The composition of African lion (*Panthera leo*) milk collected a few days postpartum. *Mammalian Biology*, 69, 375-383.

Díaz M., Dunn C.M., McClements D.J. & Decker E.A. (2003). Use of caseinophosphopeptides as natural antioxidants in oil-in-water emulsions. *Journal of Agricultural and Food Chemistry*, 51, 2365-2370.

Díaz M. & Decker E.A. (2004). Antioxidant mechanisms of caseinophosphopeptides and casein hydrolysates and their application in ground beef. *Journal of Agricultural and Food Chemistry*, 52, 8208-8213.

Dietrich R.C., Peris M.J., Seybolt A.S. & Padgett R.A. (2001). Role of the 3' splice site in U12-dependent intron splicing. *Molecular and Cellular Biology*, 21, 1942-1952.

Dobson K.D. & McQuillan A.J. (2000). In situ infrared spectroscopic analysis of the adsorption of aromatic carboxylic acids to TiO_2, ZrO_2, Al_2O_3, and Ta_2O_5 from aqueous solutions. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, 56, 557-565.

Dufour E. & Haertlé T. (1991). Binding of retinoids and β-carotene to β-lactoglobulin. Influence of protein modifications. *Biochimica et Biophysica Acta*, 1079, 316-320.

Dulaney J.T. (1979). Factors affecting the binding of glycoproteins to concanavalin A and *Ricinus communis* agglutinin and the dissociation of their complexes. *Analytical Biochemistry*, 99, 254-267.

Duncan J.S., Wilkinson M.C. & Burgoyne R.D. (2000). Purification of Golgi casein kinase from bovine milk. *Biochemical Journal*, 350, 463-468.

Ebner K.E., Denton W.L. & Brodbeck U. (1966). The substitution of α-lactalbumin for the B protein of lactose synthetase. *Biochemical and Biophysical Research and Communications*, 24, 232-236.

Egito A.S., Girardet J.-M., Miclo L., Mollé D., Humbert G. & Gaillard J.-L. (2001). Susceptibility of equine κ- and β-caseins to hydrolysis by chymosin. *International Dairy Journal*, 11, 885–893.

Egito A.S., Miclo L., Lopez C., Adam A., Girardet J.-M. & Gaillard J.-L (2002). Separation and characterization of mares' milk α_{s1}-, β-, κ-caseins, γ-casein-like, and proteose peptone component 5-like peptides. *Journal of Dairy Science*, 85, 697–706.

Eigel W.N., Butler J.E., Emstrom C.A., Farrell Jr. H.M., Harwalkar V.R, Jenness R. & McWhitney R. (1984). Nomenclature of protein cow's milk: fifth revision. *Journal of Dairy Science*, 67, 1599-1631.

Emery T. (1991). Iron oxidation by casein. *Biochemical and Biophysical Research Communication*, 182, 1047-1052.

Eperon I.C., Ireland D.C., Smith R.A., Mayeda A. & Krainer A.R. (1993). Pathways for selection of 5' splice sites by U1 snRNPs and SF2/ASF. *EMBO Journal*, 12, 3607-3617.

Erba D., Ciappellano S. & Testolin G. (2002). Effect of the ratio of casein phosphopeptides to calcium (w/w) on passive calcium transport in the distal small intestine of rats. *Nutrition*, 18, 743-746.

Erickson A.C. & Johnson G.V.W. (1993). Metal (Fe^{3+}) affinity chromatography: Differential adsorption of tau phosphoproteins. *Journal of Neuroscience Methods*, 46, 245-249.

Farrell Jr. H.M. (1988). Physical equilibria: proteins. In: Fundamentals of Dairy Chemistry. 3rd edition.Wong N. B. editor. New-York: Van Nostrand, pp. 461-510.

Farrell Jr. H.M., Brown E.M. & Kumosinski T.F. (1993). Three dimensional molecular modeling of bovine caseins. *Food Structure*, 12, 235-250.

Farrell Jr. H.M., Jimenez-Flores R., Bleck G.T., Brown E.M., Butler J.E., Creamer L.K., Hicks C.L., Hollar C.M., Ng-Kwai-Hang K.F. & Swaisgood H.E. (2004). Nomenclature of the proteins of cows' milk - Sixth revision. *Journal of Dairy Science*, 87, 1641-1674.

Farrell Jr. H.M., Malin E.L., Brown E.M. & Qi P.X. (2006). Casein micelle structure: What can be learned from milk synthesis and structural biology? *Current Opinion in Colloid and Interface Science*, 11, 135-147.

Ferranti P., Malorni A., Nitti G., Laezza P., Pizzano R., Chianese L. & Addeo F. (1995). Primary structure of ovine α_{s1}-caseins: localization of phosphorylation sites and characterization of genetic variants A, C and D. *Journal of Dairy Research*, 62, 281-296.

Ferranti P., Addeo F., Malorni A., Chianese L., Leroux C. & Martin P. (1997). Differential splicing of pre-messenger RNA produces multiple forms of mature caprine α_{s1}-casein. *European Journal of Biochemistry*, 249, 1-7.

Ferranti P., Scaloni A., Caira S., Chianese L., Malorni A. & Addeo F. (1998). The primary structure of water buffalo α_{s1}- and β-casein: Identification of phosphorylation sites and characterization of a novel β-casein variant. *Journal of Protein Chemistry*, 17, 835-844.

Ferranti P., Lilla S., Chianese L. & Addeo F. (1999). Alternative nonallelic deletion is constitutive of ruminant α_{s1}-casein. *Journal of Protein Chemistry*, 18, 595−602.

Ferrarretto A, Signorile A., Gravaghi C. Fiorilli A. & Tettamanti G. (2001). Casein phosphopeptides influence calcium uptake by cultured human intestinal HT-29 tumor cells. *The Journal of Nutrition*, 131, 1655-1661.

Ferrarretto A, Gravaghi C. Fiorilli A. & Tettamanti G. (2003). Casein-derived bioactive phosphopeptides : role of phosphorylation and primary structure in promoting calcium uptake by HT-29 tumor cells. *FEBS Letters*, 551, 92-98.

Ferretti L., Leone P. & Sgaramella V. (1990). Long range restriction analysis of the bovine casein genes. *Nucleic Acids Research*, 18, 6829-6833.

Fiat A.M., Jollès J., Aubert J.P., Loucheux-Lefebvre M.H. & Jollès P. (1980). Localisation and importance of the sugar part of human casein. *European Journal of Biochemistry*, 111, 333-339.

Fiat A.M. & Jollès P. (1989). Caseins of various origins and biologically active casein peptides and oligosaccharides : structural and physiological aspects. *Molecular and Cellular Biochemistry*, 87, 5-30.

Fiat A.M., Migliore D. & Jollès P. (1993). Biologically active peptides from milk proteins with emphasis on two examples concerning antithrombotic and immunostimulating activities; *Journal of Dairy Science*, 76, 301-310.

Ficarro S.B., McCleland M.L., Stukenberg P.T., Burke D.J., Ross M.M., Shabanowitz J., Hunt D.F. & White F.M. (2002). Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nature Biotechnology*, 20, 301-305.

FitzGerald, R.J. (1998). Potential uses of caseinophosphopeptides. *International Dairy Journal*, 8, 451-457.

Flotow H., Graves P.R., Wang A., Fiol C.J., Roeske R.W. & Roach P.J. (1990). Phosphate groups as substrate determinants for casein kinase I action. *The Journal of Biological Chemistry*, 265, 14264-14269.

Folk J.E. & Schirmer E.W. (1965). Chymotrypsin C. I. Isolation of the zymogen and the active enzyme: Preliminary structure and specificity studies. *Journal of Biological Chemistry*, 240, 181-192.

Fox P.F. (2003). Milk proteins; general and historical behaviour. In: Advanced Dairy Chemistry. 3rd edition. Proteins, vol. 1. Fox P.F., McSweeney P.L.H., editors. New York: Kluwer Academic/Plenum publisher, pp. 1-48.

Fox P.F. & Brodkorb, A. (2008). The casein micelle: Historical aspects, current concepts and significance. *International Dairy Journal*, 18, 677-684.

Fox-Walsh K.L., Dou Y., Lam B.J., Hung S.P., Baldi P.F. & Hertel K.J. (2005). The architecture of pre-mRNAs affects mechanisms of splice-site pairing. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 16176-16181.

Frilander M.J. & Steitz J.A. (1999). Initial recognition of U12-dependent introns requires both U11/5' splice-site and U12/ branchpoint interactions. *Genes Development*, 13, 851-863.

Fu X.D. & Maniatis T. (1992). The 35-kDa mammalian splicing factor SC35 mediates specific interactions between U1 and U2 small nuclear ribonuclearprotein particles at the 3' splice site. *Proceedings of the National Academy of Sciences of the United States*, 89, 1725-1729.

Gall H., Kalveram C.M. & Sick H. (1996). Allergy to the heat labile proteins α-lactalbumin and β-lactoglobulin in mare's milk. *Journal of Allergy and Clinical Immunology*, 97, 1304-1307.

Gallagher D.S., Schelling C.P., Groenen M.M. & Womack J.E. (1994). Confirmation that the casein gene cluster resides on cattle chromosome 6. *Mammalian Genome*, 5, 524-524.

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D. & Bairoch A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In: John M. Walker editor. The ProteomicsProtocolsHandbook, Humana Press, pp. 571-607.

Gaucheron F., Mollé D., Léonil J. & Maubois J.L. (1995). Selective determination of phosphopeptide β-CN (f1-25) in a β-casein digest by adding iron: Characterization by liquid chromatography with on-line electrospray-ionization mass spectrometric detection. *Journal of Chromatography*, 644, 193-200.

Gaucheron F., Le Graet Y., Boyaval E. & Piot M. (1997). Binding of cations to casein molecules: Importance of physicochemical conditions. *Milchwissenschaft*, 52, 322-327.

Gaye P., Houdebine L. & Denamur R. (1973). Isolation of active messenger RNA for α_s casein from bound polyribosomes of mammary gland. *Biochemical and Biophysical Research Communications*, 51, 637-644.

Gaye P., Gautron J.P., Mercier J.C. & Haze G. (1977). Amino terminal sequences of the precursors of the ovine caseins. *Biochemical and Biophysical Research Communications*, 79, 903-911.

Geiger T. & Clarke S. (1987). Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. *The Journal of Biological Chemistry*, 15, 785-794.

Girardet J.-M., N'negue M.-A., Egito A.S., Campagna S., Lagrange A. & Gaillard J.-L. (2004). Multiple forms of equine α-lactalbumin: Evidence for N-glycosylated and deamidated forms. *International Dairy Journal*, 14, 207-217.

Girardet J.-M., Miclo L., Florent S., Mollé D. & Gaillard J.-L. (2006). Determination of the phosphorylation level and deamidation susceptibility of equine β-casein. *Proteomics*, 6, 3707-3717.

Godovac-Zimmermann J., Conti A., Liberatori J. & Braunitzer G. (1985). The amino-acid sequence of β-lactoglobulin II from horse colostrum (*Equus caballus*, perissodactyla): β-lactoglobulins are retinol-binding proteins. *Biological Chemistry Hoppe-Seyler*, 366, 601-608.

Godovac-Zimmermann J., Shaw D., Conti A. & Mckenzie H. (1987). Identification and primary structure of equine α-lactalbumin B and C (*Equus caballus*, *Perissodactyla*). *Biological Chemistry Hoppe-Seyler*, 368, 427-433.

Gordon W.G., Groves M.L., Greenberg R., Jones S.B., Kalen E.B., Peterson R.F. & Townsend R.E. (1972). Probable identification of γ-, TS-, R-, S-caseins as fragments of β−casein. *Journal of Dairy Science*, 55, 261-263.

Gordon P.M., Sontheimer E.J. & Piccirilli J.A. (2000). Metal ion catalysis during the exon-ligation step of nuclear pre-mRNA splicing: Extending the parallels between the spliceosome and group II introns. *RNA*, 6, 199-205.

Görnemann J., Kotovic K.M., Hujer K. & Neugebauer K.M. (2005). Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. *Molecular Cell*, 19, 53-63.

Gozani O., Feld R. & Reed R. (1996). Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. *Genes Development*, 10, 233-243.

Gravaghi C., Del Favero E., Cantu L., Donetti E., Bedoni M., Fiorilli A., Tettamanti G. & Ferraretto A. (2007). Casein phosphopeptide promotion of calcium uptake in HT-29 cells – Relationship between biological activity and supramolecular structure. *FEBS Journal*, 274, 4999-5011.

Graveley B.R (2000). Sorting out the complexity of SR protein functions. *RNA*, 6, 1197-1211.

Graveley B.R. (2001). Alternative splicing: Increasing diversity in the proteomic world. *Trends in Genetics*, 17, 100-107.

Green M.R. (1991). Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. *Annual Review of Cell Biology*, 7, 559-599.

Greenberg R., Groves M.L. & Peterson R.F. (1976). Amino terminal sequence and location of phosphate groups of the major human casein. *Journal of Dairy Science*, 59, 1016-1018.

Greenberg R., Groves M.L. & Dower H.J. (1984). Human β-casein. Amino acid sequence and identification of phosphorylation sites. *The Journal of Biological Chemistry*, 259, 5132-5138.

Groenen M.A., Dijkhof R.J., van der Poel J.J., van Diggelen R. & Verstege E. (1992). Multiple octamer binding sites in the promoter region of the bovine α_{s2}-casein gene. *Nucleic Acids Research*, 20, 4311-4318.

Groenen M.A., Dijkhof R.J., Verstege A.J. & van der Poel J.J. (1993). The complete sequence of the gene encoding bovine α_{s2}-casein. *Gene*, 123, 187-193.

Guesdon B., Messaoudi M., Lefranc-Millot C., Fromentin G., Tomé D. & Even P.C. (2006). A tryptic hydrolysate from bovine milk α_{s1}-casein improves sleep in rats subjected to chronic mild stress. *Peptides*, 27, 1476-1482.

Guthrie C. & Patterson B. (1988). Spliceosomal snRNAs. *Annual Review on Genetics*, 22, 387-419.

Haga M., Yamauchi K. & Aoyagi S. (1983). Conformation and some properties of bovine α_{s2} group casein. *Agricultural and Biological Chemistry*, 47, 1467-1471.

Haj Khelil A., Deguilien M., Morinière M., Ben Chibani J. & Baklouti F. (2008). Cryptic splicing sites are differentially utilized in vivo. *FEBS Journal*, 275, 1150-1162.

Hall S.L. & Padgett R.A. (1994). Conserved sequences in a class rare eukaryotic nuclear introns with non-consensus splice sites. *Journal of Molecular Biology*, 239, 357-365.

Hall S.L. & Padgett R.A. (1996). Requirement of U12 snRNA for in vivo splicing of a minor class of eukaryotic nuclear pre-mRNA introns. *Science*, 271, 1716-1718.

Hansen M., Sandstrom B. & Lonnerdal B. (1996). The effect of casein phosphopeptides on zinc and calcium absorption from high phytate infant diets assessed in rat pups and Caco-2 cells. *Pediatric Research*, 40, 547-552.

Hansson L., Edlund A., Johansson T., Hernell O., Strömqvist M., Lindquist S., Lönnerdal B. & Bergstrom S. (1994). Structure of the human β-casein encoding gene. *Gene*, 139, 193-199.

Hastings M.L., Resta N., Traum D., Stella A., Guanti G. & Krainer A.R. (2005). An LKB1 AT-AC intron mutation causes Peutz-Jeghers syndrome via splicing at noncanonical cryptic splice sites. *Nature Structural and Molecular Biology*, 12, 54-59.

Hata T., Veda J. & Otani H. (1999). Immunostimulatory action of a commercially available casein phosphopeptide preparation, CPP III, in cell cultures. *Milchwissenschaft*, 54, 3-7.

Heald C.W. & Saacke R.G. (1972). Cytological comparison of milk protein synthesis of rat mammary tissue in vivo and in vitro. *Journal of Dairy Science*, 55, 621-628.

Herbert B.R., Molloy M.P., Gooley A.A., Walsh B.J., Bryson W.G. & Williams K.L. (1998). Improved protein solubility in two-dimensional electrophoresis using tributyl phosphine as reducing agent. *Electrophoresis*, 19, 845–851.

Hercberg S.G. & Preziosi P. (2000). La déficience en fer au cours de la grossesse en France. *Cahiers de Nutrition et de Diététique*, 35, 13-23.

Hernández-Ledesma B., Quirós A., Amigo L. & Recio I. (2007). Identification of bioactive peptides after digestion of human milk and infant formula with pepsin and pancreatin. *International Dairy Journal*, 17, 42-49.

Herskovitis T.T. (1966). On the conformation of caseins. Optical rotatory properties. *Biochemistry*, 5, 1018-1026.

Hertel K.J. (2008). Combinatorial control of exon recognition. *The Journal of Biological Chemistry*, 283, 1211-1215.

Hicks M.J., Yang C.-R., Kotlajich M.V. & Hertel K.J. (2006). Linking splicing to Pol II transcription stabilizes pre-mRNAs and influences splicing patterns. *Plos Biology*, 4, 943-951.

Hill J. & Boland M. (2001). Milk protein variants - Relationships with health and disease. *Australian Journal of Dairy Technology*, 56, 117-122.

Hoagland D.P., Unruh J.J., Wickham E.D. & Farrell Jr. H.M. (2001). Secondary structure of bovine α_{s2}-casein: Theoretical and experimental approaches. *Journal of Dairy Science*, 84, 1944-1949.

Holland J.W., Deeth H.C. & Alewood P.F. (2004). Proteomic analysis of κ-casein micro-heterogeneity. *Proteomics*, 4, 743-752.

Holland J.W., Deeth H.C. & Alewood P.F. (2006). Resolution and characterisation of multiple isoforms of bovine κ-casein by 2-DE following a reversible cysteine-tagging enrichment strategy. *Proteomics,* 6, 3087-3095.

Holt C. (1992). Structure and stability of bovine casein micelle. *Advances in Protein Chemistry*, 43, 63-151.

Holt C. & Horne D.S. (1996). The hairy casein micelle: Evolution of the concept and its implication for dairy technology. *Netherlands Milk and Dairy Journal*, 50, 85-111.

Horne D.S. (1998). Casein interactions: Casting light on the *black boxes*, the structure in dairy products. *International Dairy Journal*, 8, 171-177.

Horne D.S. (2006). Casein micelle structure: Models and muddles. *Current Opinion in Colloid and Interface Science*, 11, 148-153.

Hsieh H.-C., Sheu C., Shi F.-K. & Li D.-T. (2007). Development of a titanium dioxide nanoparticle pipette-tip for the selective enrichment of phosphorylated peptides. *Journal of Chromatography A*, 1165, 128-135.

Hurrell R.F., Juillerat M.-A., Reddy M.B., Lynch S.R., Dassenko S.A. & Cook J.D. (1992). Soy protein, phytate and iron absorption in humans. *American Journal of Clinical Nutrition*, 56, 573-578.

Iametti S., Tedeschi G., Oungre E. & Bonomi F (2001). Primary structure of κ-casein isolated from mares' milk. *Journal of Dairy Research*, 68, 53-61.

Imanishi S.Y., Kochin V., Ferraris S.E., de Thonel A., Pallari H.-M., Corthals G.L. & Eriksson J.E. (2007). Reference-facilitated phosphoproteomics: Fast and reliable phosphopeptide validation by μLC-ESI-Q -TOF MS/MS. *Molecular and Cellular Proteomics*, 6, 1380-1391.

Incorvaia R. & Padgett R.A. (1998). Base pairing with U6atac snRNA is required for 5' splice site activation of U12-activation dependent introns *in vivo*. *RNA*, 4, 709-718.

Jackson I.J. (1991). A reappraisal of non-consensus mRNA splice sites. *Nucleic Acids Research*, 19, 3795-3798.

Jasinska B. & Jaworska T. (1988). Comparison of structures of micellar caseins of milk of cows, goats and mares with human milk casein. *Animal Science Papers and Reports*, 7, 45-54.

Jauregui-Adell J. (1975). Heat stability and reactivation of mare milk lysozyme. *Journal of Dairy Science*, 58, 835-838.

Jensen S.S. & Larsen M.R. (2007). Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques. *Rapid Communications in Mass Spectrometry*, 21, 3635-3645.

Johnsen L.B., Rasmussen L.K., Petersen T.E. & Berglung L. (1995). Characterization of three types of human α_{s1}-casein mRNA transcripts. *Biochemical Journal*, 309, 237-242.

Johnson J.M., Castle J., Garett-Engele P., Kan Z., Loerch P.M., Armour C.D., Santos R., Schadt E.E., Stoughton R. & Shoemaker D.D. (2003). Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science*, 302, 2141-2144.

Jollès P. (1975). Structural aspects of the milk clotting process. Comparative features with the blood clotting process. *Molecular and Cellular Biochemistry*, 7, 73-85.

Jollès P. & Henschen A. (1982). Comparison between clotting of blood and milk. *Trends in Biochemical Science*, 7, 325-328.

Jollès P., Levy-Toledano S., Fiat A.M., Soria C., Gillesen D., Thomaidis A., Dunn F.W. & Caen J. (1986). Analogy between fibrinogen and casein: Effect of an undecapeptide isolated from κ-casein on platelet function. *European Journal of Biochemistry*, 158, 379-382.

Jones E.M., Smart A., Bloomberg G., Burgess L. & Millar M.R. (1994). Lactoferricin, a new antimicrobial peptide. *Journal of Applied Bacteriology*, 77, 208-214.

Kajiwara K., Niki R., Urakawa H., Hiragi Y., Donkai N. & Nagura M. (1988). Micellar structure of β-casein observed by small-angle X-ray scattering. *Biochimica et Biophysica Acta*, 955, 128-134.

Kaminogawa S. & Yamauchi K. (1972). Decomposition of β-casein by milk protease similary of the decomposed products to temperature sensitive and R-caseins. *Agricultural and Biological Chemistry*, 36, 255-260.

Kaminogawa S., McKenzie H.A. & Shaw D.C. (1984). The amino acid sequence of equine α-lactalbumin. *Biochemistry International*, 9, 539-546.

Kanekanian A.D., Williams R.J.H., Brownsell V.L. & Andrews A.T. (2008). Caseinophosphopeptides and dental protection: Concentration and pH studies. *Food Chemistry*, 107, 1015-1021.

Kansci G., Genot C., Meynier A., Gaucheron F. & Chobert J.-M. (2004). β-caseinophosphopeptide (f1-25) confers on β-casein tryptic hydrolysate an antioxidant activity during iron/ascorbate-induced oxidation of liposomes. *Le Lait*, 84, 449-462.

Kapitany R.A. & Zebrowski E.J. (1973). A high resolution PAS stain for polyacrylamide gel electrophoresis. *Analytical Biochemistry*, 56, 361-369.

Kashima T. & Manley J.L. (2003). A negative element in SMN2 exon 7 inhibits splicing in spinal muscular atrophy. *Nature Genetics*, 34, 460-463.

Kawahara T., Katayama D. & Otani H. (2004). Effect of β-casein (1-28) on the proliferative responses and secretory functions of human immunocompetent cell lines. *Bioscience Biotechnology and Biochemistry*, 68, 2091-2095.

Kibangou I., Bouhallab S., Henry G., Bureau F., Allouche S., Blais A., Guérin P., Arhan P. & Bouglé D. (2005). Milk proteins and iron absorption: Contrasting effects of different caseinophosphopeptides. *Pediatric Research*, 58, 731-734.

Kim J.H., Desor D., Kim Y.T., Yoon W.J., Kim K.S., Jun J.S., Pyun K.H. & Shim I. (2007). Efficacy of α_{s1}-casein hydrolysate on stress-related symptoms in women. *European Journal of Clinical Nutrition*, 61, 536-541.

Kitts D.D. & Yuan Y.V. (1992). Casein phosphopeptides and calcium bioavailability. *Trends in Food and Science Technology*, 3, 31-35.

Kitts D.D., Yuan Y.V., Nagasawa T. & Moriyama Y. (1992). Effect of casein, caseinophosphopeptides and calcium intake on ileal ^{45}Ca disappearance and temporal systolic blood pressure in spontaneously hypertensive rats. *British Journal of Nutrition*, 68, 765-781.

Kitts D.D. & Weiler K. (2003). Bioactive proteins and peptides from food sources. Applications of bioprocesses used in isolation and recovery. *Current Pharmaceutical Design*, 9, 1309-1323.

Kitts D.D. (2005). Antioxidant properties of casein-phosphopeptides. *Trends in Food Science and Technology*, 16, 549-554.

Kitts D.D. & Nakamura S. (2006). Calcium enriched casein phosphopeptide stimulate release of IL-6 cytokine in human epithelial intestinal cell line. *Journal of Dairy Research*, 73, 44-48.

Kjeldsen F., Savitski M.M., Nielsen M.L., Shi L. & Zubarev R.A. (2007). On studying protein phosphorylation patterns using bottom-up LC-MS/MS: The case of human α-casein. *The Analyst*, 132, 768–776.

Klemm C., Otto S., Wolf C., Haseloff R.F., Beyermann M. & Krause E. (2006). Evaluation of the titanium dioxide approach for MS analysis of phosphopeptides. *Journal of Mass Spectrometry*, 41, 1623-1632.

Koch G., Wiedemann K., Drebes E., Zimmermann W., Link G. & Teschemacher H. (1988). Human β-casomorphin-8 immunoreactive material in the plasma of women during pregnancy and after delivery. *Regulatory Peptides*, 20, 107-117.

Koczan D., Hobom G. & Seyfert H.M. (1991). Genomic organization of the bovine α_{s1}-casein gene. *Nucleic Acids Research*, 18, 5591-5596.

Kohmura M., Nio N., Kubo K., Minoshima E., Munekata E. & Ariyoshi Y. (1989). Inhibition of angiotensin-converting enzyme by synthetic peptide fragments of human β-casein. *Agricultural and Biological Chemistry*, 53, 2107-2114.

Kohmura M., Nio N. & Ariyoshi Y. (1990). Inhibition of angiotensin-converting enzyme by synthetic peptide fragments of human κ-casein. *Agricultural and Biological Chemistry*, 54, 835-836.

Kohtz J.D., Jamison S.F., Will C.L., Zuo P., Luhrmann R., Garcia-Blanco M. A. & Manley J. L. (1994). Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. *Nature*, 368, 119-124.

Kolossova I. & Padgett R.A. (1997). U11 snRNA interacts *in vivo* with the 5' splice site of U12-dependent (AU-AC) pre-mRNA introns. *RNA*, 3, 227-233.

Kopra N., Scholz-Ahrens K.E. & Barth C.A. (1992). Effect of casein phosphopeptides on utilisation of calcium in vitamin D-deficient rats. *Milchwissenschaft*, 47, 488-492.

Korhonen H. & Pihlanto A. (2006). Bioactive peptides: Production and functionality. *International Dairy Journal*, 16, 945-960.

Kornblihtt A.R. (2005). Promoter usage and alternative splicing. *Current Opinion in Cell Biology*, 17, 262-268.

Krawczak M., Thomas N.S., Hundrieser B., Mort M., Wittig M., Hampe J. & Cooper D.N. (2007). Single base-pair substitutions in exon-intron junctions of human genes: Nature, distribution, and consequences for mRNA splicing. *Human Mutation*, 28, 150-158.

Kumosinski T.F., Brown E.M. & Farrell Jr. H.M. (1991). Three dimensional molecular modeling of bovine caseins : κ-Casein. *Journal of Dairy Science*, 74, 2879-2887.

Kumosinski F.C., Brown E.M. & Farrell Jr. H.M. (1993). Three dimensional molecular modeling of bovine caseins: An energy minimized β-casein structure. *Journal of Dairy Science*, 76, 931-945.

Kunz C. & Lönnerdal B. (1989). Casein micelles and casein subunits in milk. In: Proteins and Non-protein Nitrogen in Human Milk. Atkinson S. A. & Lönnerdal B. editors. Boca Raton, USA: CRC Press, pp. 10-27.

Kuwata H., Yip T.-T., Tomita M. & Hutchens T.W. (1998). Direct evidence of the generation in human stomach of an antimicrobial peptide domain (lactoferricin) from ingested lactoferrin. *Biochimica et Biophysica Acta*, 1429, 129-141.

Laemmli U.K. & Favre M. (1973). Maturation of the head of bacteriophage T4. I. DNA packaging events. *Journal of Molecular Biology*, 80, 575-599.

Lahov E. & Regelson W. (1996). Antibacterial and immunostimulating casein-derived substances from milk: Casecidin, isracidin peptides. *Food and Chemical Toxicology*, 34, 131-145.
Landon M. (1977). Cleavage at aspartyl–prolyl bonds. *Methods in Enzymology*, 47, 145-149.

Larsen M.R., Thingholm T.E., Jensen O.N., Roepstorff P. & Jørgensen T.J.D. (2005). Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Molecular and Cellular Proteomics*, 4, 873-886.

Lenasi T., Rogelj I. & Dovc P. (2003). Characterization of equine cDNA sequences for α_{s1}, β- and κ-casein. *Journal of Dairy Research*, 70, 29-36.

Lenasi T., Peterlin B.M. & Dovc P. (2006). Distal regulation of alternative splicing by splicing enhancer in equine β-casein intron 1. *RNA*, 12, 498-507.

Leroux C., Mazure N. & Martin P. (1992). Mutation away from splice site recognition sequences might *cis*-modulate alternative splicing of goat α_{s1}-casein transcript. Structural organization of the relevant gene. *The Journal of Biological Chemistry*, 267, 6147-6157.

Lestradet H. 1988. Laits et immunité. *Cahiers de Nutrition et de Diététique*, 23, 297-300.

Levine A. & Durbin R. (2001). A computational scan for U12-dependent introns in the human genome sequence. *Nucleic Acids Research*, 29, 4006-4013.

Li S. & Dass C. (1999). Iron(III)-immobilized metal ion affinity chromatography and mass spectrometry for the purification and characterization of synthetic phosphopeptides. *Analytical Biochemistry*, 270, 9-14.

Lin F.-Y., Chen W.-Y. & Huang S.-Y. (2000). Selective separation of caseinophosphopeptides through immobilized metal ion affinity chromatography: Effects of pH values, salt concentrations and degree of phosphorylation. *Bioprocess Engineering*, 23, 467-471.

Lorson C.L. & Androphy E.J. (2000). An exonic enhancer is required for inclusion of an essential exon in the SMA-determining gene SMN. *Human Molecular Genetics*, 9, 259-266.

Loucheux-Lefebvre M.H., Aubert J.P. & Jollès P. (1978). Prediction of the conformation of the cow and sheep κ-caseins. *Biophysical Journal*, 23, 323-336.

Lovisi P., Jolivet P., Jagic F., Dalgleish D. & Chardot T. (2003). A protein kinase is located in the micellar fraction of fresh pasteurized skimmed farm milk. *Journal of Dairy Science*, 86, 1147-1156.

Loukas S., Varoucha D., Zioudrou C., Streaty R.A. & Klee W.A. (1983). Opioid activities and structures of α-casein-derived exorphins. *Biochemistry*, 22, 4567-4573.

Lozovich S. (1995). Medical uses of whole and fermented mare milk in Russia. *Cultured Dairy Product Journal*, 30, 18-21.

McCann K.B., Shiell B.J., Michalski W.P., Lee A., Wan J., Roginski H. & Coventry M.J. (2006). Isolation and characterisation of a novel antibacterial peptide from bovine α_{s1}-casein. *International Dairy Journal*, 16, 316-323.

McMananon R. (1993). Practicals tips in nursery rearing of exotic cats. *Journal of Small Exotic Animal Medicine*, 2, 137-140.

Madhani H.D. & Guthrie C. (1994). Dynamic RNA-RNA interactions in the spliceosome. *Annual review of genetics*, 28, 1-26.

Maeno M., Yamamoto N. & Takano T. (1996). Identification of an antihypertensive peptide from casein hydrolysate produced by a proteinase from *Lactobacillus helveticus* CP790. *Journal of Dairy Science*, 79, 1316-1321.

Malacarne M., Martuzzy F., Summer A. & Mariani P. (2002). Protein and fat composition of mare's milk: Some nutritional remarks with reference to human and cow's milk. *International Dairy Journal*, 12, 869-877.

Mamula M.J., Gee R.J., Elliott J.I., Sette A., Southwood S., Jones P.-J. & Blier P.R. (1999). Isoaspartyl post-translational modification triggers autoimmune responses to self-proteins. *The Journal of Biological Chemistry*, 274, 22321-22327.

Maniatis T. & Tasic B. (2002). Alternative pre-mRNA splicing and proteome expansion in metazoans. *Nature*, 418, 236-243.

Mao I. L., Buttazozoni L. G. & Aleandri R. (1992). Effects of polymorphic milk protein genes on milk yield and composition traits in Holstein cattle. *Acta Agriculturae Scandinavica Section A (Animal Science)*, 42, 1-7.

Martin P., Brignon G., Furet J.P. & Leroux C. (1996). The gene encoding α_{s1}-casein is expressed in human mammary epithelial cells during lactation. *Le Lait*, 76, 523-535.

Martin P., Szymanowska M., Zwierzchowski L. & Leroux C. (2002). The impact of genetic polymorphisms on the protein composition of ruminant milks. *Reproduction Nutrition Development*, 42, 433-459.

Martinez-Contreras R., Fisette J.-F., Nasim F.-U.H., Madden R., Cordeau M. & Chabot B. (2006). Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing. *Plos Biology*, 4, 172-185.

Maruyama S. & Suzuki H. (1982). A peptide inhibitor of angiotensin I-converting enzyme in the tryptic hydrolysate of casein. *Agricultural and Biological Chemistry*, 46, 1393-1394.

Maruyama S., Mitachi H., Awaya J., Kurono M., Tomizuka N. & Suzuki H. (1987). Angiotensin I-converting enzyme inhibitory activity of the C-terminal hexapeptide of α_{s1}-casein. *Agricultural and Biological Chemistry*, 51, 2557-2561.

Meisel H. & Frister H. (1989). Chemical characterization of bioactive peptides from in vivo digests of casein. *Journal of Dairy Research*, 56, 343-349.

Meisel H. & Schlimme E. (1993). Calcium and iron binding capacity of different fractions from in vitro proteolysis of casein. *Kiel Milchwirtschaftliche Forschungsberichte*, 45, 235-243.

Meisel M. & Schlimme S. (1994). Inhibitors of angiotensin I-converting enzyme derived from bovine casein (casokinins). β-Casomorphins and related peptides: Recent developments. In: Brantl V. & Teschemacher H. editors.VCH-Weinheim, Germany, pp. 27-33.

Meisel H., Goepfert A. & Günther S. (1997). Occurrence of ACE inhibitory peptides in milk products. *Milchwissenschaft*, 52, 307-311.

Meisel H. (1998). Overview on milk protein-derived peptides. *International Dairy Journal*, 8, 363-373.

Meisel H., Bernard H., Fairweather-Tait S., FitzGerald R.J., Hartmann R., Lane C.N., McDonagh D., Teucher B. & Wal J.M. (2003). Detection of caseinophosphopeptides in the distal ileostomy fluid of human subjects. *British Journal of Nutrition*, 89, 351-359.

Mellander O. (1950). The physiological importance of the casein phosphopeptide calcium salt. II. Peroral calcium dosage of infants. *Acta of the Society of Medicine of Uppsala*, 55, 247-255.

Mercier J.-C., Grosclaude F. & Ribadeau Dumas B. (1971). Structure primaire de la caséine α_{s1} B bovine. Séquence complète. *European Journal of Biochemistry*, 23, 41-51.

Mercier J.-C. & Ribadeau Dumas B. (1972). Primary structure of bovine caseins. A review. *Milchwissenschaft*, 27, 402-408.

Mercier J.C., Brignon G. & Ribadeau Dumas B. (1973). Structure primaire de la caséine κ B bovine. Séquence complète. *European Journal of Biochemistry*, 35, 222-235.

Mercier J.-C. & Gaye P. (1980). Study of secretory lactoproteins: Primary structures of the signals and enzymatic processing. *Annals of the New York Academy of Sciences*, 343, 232-251.

Mercier J.-C. (1981). Phosphorylation of caseins, present evidence for an amino acid triplet code posttranslationally recognized by specific kinases. *Biochimie*, 63, 1−17.

Mercier J.-C. & Vilotte J.-L. (1993). Structure and function of milk protein genes. *Journal of Dairy Science*, 76, 3079-3098.

Messaoudi M., Lefranc-Millot C., Desor D., Demagny B. & Bourdon L. (2005). Effects of a tryptic hydrolysate from bovine milk α_{s1}- casein on hemodynamic responses in healthy human volunteers facing successive mental and physical stress situations. *European Journal of Nutrition*, 44, 128-132.

Miclo L., Perrin E., Driou A., Iung C., Boudier J.-F., Linden G. & Gaillard J.-L. (2001). Characterization of α-casozepine, a tryptic peptide from bovine α_{s1}-casein with benzodiazepine-like activity. *Faseb Journal*, 15, 1780-1782.

Miclo L., Girardet J.-M., Egito A.S., Mollé D., Martin P. & Gaillard J.-L. (2007). The primary structure of a low-M_r multiphosphorylated variant of β-casein in equine milk. *Proteomics*, 7, 1327-1335.

Milenkovic D., Martin P., Guerin G. & Leroux C. (2002). A specific pattern of splicing for the horse α_{S1}-casein mRNA and partial genomic characterization of the relevant locus. *Genetics Selection Evolution*, 34, 509-519.

Miranda G., Mahé M.-F., Leroux C. & Martin P. (2004). Proteomic tools to characterize the protein fraction of Equidae milk. *Proteomics*, 4, 2496-2509.

Mohr U., Koczan D., Linder D., Hobom G. & Erhardt G. (1994). A single point mutation results in A allele-specific exon skipping in the bovine α_{s1}-casein mRNA. *Gene*, 143, 187-192.

Mollé D. & Léonil J. (1995). Heterogeneity of the bovine κ-casein caseinomacropeptide, resolved by liquid chromatography on-line with electrospray ionization mass spectrometry. *Journal of Chromatography A*, 708, 223-230.

Moore M.J., Query C.C. & Sharp P.A. (1993). Splicing of precursors to mRNA by the spliceosome. In: The RNA Word, Gesteland R. F. & Atkins J. F. editors. New York: Cold Spring Harbor Laboratory Press, pp. 303-357.

Muszyńska G., Andersson L. & Porath J. (1986). Selective adsorption of phosphoproteins on gel-immobilized ferric chelate. *Biochemistry*, 25, 6850-6853.

Mykkänen H.M. & Wasserman R.H. (1980). Enhanced absorption of casein by casein phosphopeptides in rachitic and normal chicks. *The Journal of Nutrition*, 119, 2141-2148.

Nakamura Y., Yamamoto N., Sakai K. & Takano T. (1995a). Antihypertensive effect of sour milk and peptides isolated from it that are inhibitors to angiotensin I-converting enzyme. *Journal of Dairy Science*, 78, 1253-1257.

Nakamura Y., Yamamoto N., Sakai K., Okubo A., Yamazaki S. & Takano T. (1995b). Purification and characterization of angiotensin I-converting enzyme inhibitors from sour milk. *Journal of Dairy Science*, 78, 777-783.

Nakhasi H.L., Grantham F.H. & Gullino P.M. (1984). Expression of κ-casein in normal and neoplastic rat mammary gland is under the control of prolactin. *The Journal of Biological Chemistry*, 259, 14894-14898.

Nawrocki J., Dunlap C., McCormick A. & Carr P.W. (2004). Part I. Chromatography using ultra-stable metal oxide-based stationary phases for HPLC. *Journal of Chromatography A*, 1028, 1-30.

Narva M., Kärkkäinen M., Poussa T., Lamberg-Allardt C. & Korpela R. (2003). Caseinophosphopeptides in milk and fermented milk do not affect calcium metabolism acutely in postmenopaused women. *Journal of the American College of Nutrition*, 22, 88-93.

Nawar W.W. (1996). Lipids. In: Food Chemistry, Fennema O.R., editor. New-York, pp. 225-319.

Neveu C., Mollé D., Moreno J., Martin P. & Léonil J. (2002). Heterogeneity of caprine β-casein elucidated by RP-HPLC/MS: Genetic variants and phosphorylations. *Journal of Protein Chemistry*, 21, 557-567.

Newman A.J. (1997). The role of U5 snRNP in pre-mRNA splicing. *EMBO Journal*, 16, 5797-5800.

Ng-Kwai-Hang K.F. & Kroeker E.M., (1984). Rapid separation and quantification of major caseins and whey proteins of bovine milk by polyacrylamide gel electrophoresis. *Journal of Dairy Science*, 67, 3053-3056.

Ng-Kwai-Hang K.F (2003). Milk Proteins. Heterogeneity, Fractionation and Isolation. In: Encyclopedia of Dairy Sciences, Roginski H, Fuquay JW & Fox PF, editors. London: Academic Press, pp. 1881–1894.

Nilsen T.W. (1994). RNA-RNA interactions in the spliceosome: Unravelling the ties that bind. *Cell*, 78, 1-4.

Nilsen T.W. (1998). In: RNA Structure and Function, Simons R. & Grunberg-Manago M., editors. New York: Cold spring Harbor Laboratory Press, pp. 279-307.

Nühse T.S., Stensballe A., Jensen O.N. & Peck S.C. (2003). Large-scale analysis of in vivo phosphorylated membrane proteins by immobilized metal ion affinity chromatography and mass spectrometry. *Molecular & Cellular Proteomics: MCP*, 2, 1234-1243.

Ochirkhuyag B., Chobert J.-M., Dalgalarrondo M. & Haertlé T. (2000). Characterization of mare caseins. Identification of α_{s1}- and α_{s2}-casein. *Le Lait*, 80, 223-235.

Ono T. & Obata T. (1989). A model for the assembly of bovine casein micelles from F2 and F3 subunits. *Journal of Dairy Research*, 56, 453-461.

Osthoff G., De Waal H.O., Hugo A., de Wit M. & Botes P. (2005). Milk composition of a free-ranging African elephant *(Loxodonta africana)* cow during early lactation. *Comparative Biochemistry and Physiology - Part A: Molecular and Integrative Physiology*, 141, 223-229.

Osthoff G., Hugo A. & de Wit M. (2006). The composition of cheetah *(Acinonyx jubatus)* milk. *Comparative Biochemistry and Physiology - Part B: Biochemistry and Molecular Biology*, 145, 265-269.

Osthoff G., Hugo A. & de Wit M. (2007). The composition of serval *(Felis serval)* milk during mid-lactation. *Comparative Biochemistry and Physiology - Part B: Biochemistry and Molecular Biology*, 147, 237-241.

Otani H., Watanabe T. & Tashira Y. (2001). Effects of bovine β-casein (1-28) and its chemically synthesized partial fragments on proliferative responses and immunoglobulin production in mouse spleen cell cultures. *Bioscience Biotechnology and Biochemistry*, 65, 2489-2495.

Otani H., Nakana K. & Kawahara T. (2003). Stimulatory effect of a dietary casein phosphopeptide preparation on the mucosal IgA response of mice to orally ingested lipopolysaccharide from *Salmonella typhymurium*. *Bioscience Biotechnology and Biochemistry*, 67, 729-735.

Passey R., Glenn W. & Mackinlay A. (1996). Exon skipping in the ovine α_{s1}-casein gene. *Comparative Biochemistry and Physiology- Part B: Biochemistry and Molecular Biology*, 114, 389-394.

Patel A.A. & Steitz J.A. (2003). Splicing double: Insights from the second spliceosome. *Nature Reviews Molecular Cell Biology*, 4, 960-970.

Patton S. (2004). Milk: Its remarkable contribution to human health and well-being. New Brunswick, NJ: Transaction Publishers, 182 pages.

Pelletier E., Sostmann K. & Guichard E. (1998). Measurement of interactions between β-lactoglobulin and flavor compounds (esters, acids, and pyrazines) by affinity and exclusion size chromatography. *Journal of Agricultural and Food Chemistry*, 46, 1506-1509.

Pérès J.M., Bouhallab S., Bureau F., Maubois J.L., Arhan P. & Bouglé D. (1997). Absorption digestive du fer lié au caséinophosphopeptide (1-25) de la β-caséine. *Le Lait*, 77, 433-440.

Pérès J.M., Bouhallab S., Bureau F., Neuville D., Maubois J.L., Arhan P. & Bouglé D. (1999). Mechanisms of absorption of caseinophosphopeptide bound iron. *The Journal of Nutritional Biochemistry*, 10, 215-222.

Permyakov E.A. & Berliner L.J. (2000). α-Lactalbumin: Structure and function. *FEBS Letters*, 269-274.

Perozzo R., Folkers G. & Scapozza L. (2004). Thermodynamics of protein-ligand interactions: History, presence, and future aspects. *Journal of Receptors and Signal Transduction*, 24, 1-52.

Perpetuo E.A., Juliano L. & Lebrun I. (2003). Biochemical and pharmalogical aspects of two bradykinin-potentiating peptides from hydrolysis of casein. *Journal of Protein Chemistry*, 22, 601-606.

Perrin E., Miclo L., Driou A. & Linden G. (1995). Rapid determination of the ratios of three aromatic residues in peptides by reversed-phase high-performance liquid chromatography with a high-resolution photodiode-array detector. *Journal of Chromatography -B: Biomedical Applications*, 664, 267-276.

Pihlanto-Leppälä A., Antila P., Mäntsälä P. & Hellman J. (1994). Opioid peptides produced by in vitro proteolysis in bovine caseins. *International Dairy Journal*, 4, 291-301.

Pinkse M.W.H., Uitto P.M., Hilhorst M.J., Ooms B. & Heck A.J.R. (2004). Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns. *Analytical Chemistry*, 76, 3935-3943.

Pinkse M.W.H. & Heck A.J.R. (2006). Essential enrichment strategies in phosphoproteomics. *Drug Discovery Today: Technologies*, 3, 331-337.

Pisano A., Redmond J.W., Williams K.L. & Gooley A.A. (1993). Glycosylation sites identified by solid-phase Edman degradation: O-linked glycosylation motifs on human glycophorin A. *Glycobiology*, 3, 429-435.

Pisano A., Packer N.H., Redmond J.W., Williams K.L. & Gooley A.A. (1994). Characterization of O-linked glycosylation motifs in the glycopeptide domain of bovine κ-casein. *Glycobiology*, 4, 837-844.

Pitois E., Chardot T. & Meunier J.-C. (1999). Overphosphorylation of milk caseins by a recombinant protein kinase CK2 catalytic subunit. *Journal of Agricultural Food Chemistry*, 47, 3996-4002.

Plowman J.E., Creamer L.K., Liddell M.J. & Cross J.J. (1999). Structural features of a peptide corresponding to human κ-casein residues 84-101 by ^{1}H-nuclear magnetic resonance spectroscopy. *Journal of Dairy Research*, 66, 53-63.

Pocsfalvi G., Cuccurullo M., Schlosser G., Scacco S., Papa S. & Malorni A. (2007). Phosphorylation of B14.5a subunit from bovine heart complex I identified by titanium dioxide selective enrichment and shotgun proteomics. *Molecular and Cellular Proteomics*, 6, 231-237.

Pointillard A. & Guéguen L. (1989). Absence of effect of the incorporation of a milk phosphopeptide on the utilisation of calcium and phosphorus in the young pig. *Reproduction Nutrition Development*, 29, 477-486.

Porath J., Carlsson J., Olsson I. & Belfrage G. (1975). Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature*, 258, 598-599.

Rasmussen L.K., Hojrup P. & Petersen T.E. (1992). Localization of two interchain disulfide bridges in dimers of bovine α_{s2}-casein. Parallel and antiparallel alignments of the polypeptides chains. *European Journal of Biochemistry*, 203, 381-386.

Rasmussen L.K., Hojrup P. & Petersen T.E. (1994). Disulphide arrangement in bovine caseins: Localization of intrachain disulphide bridges in monomers of κ- and α_{s2}-casein from bovine milk. *Journal of Dairy Research*, 61, 485-493.

Rasmussen L.K., Due H.A. & Petersen T.E. (1995). Human α_{s1}-casein: Purification and characterization. *Comparative Biochemistry and Physiology - Part B: Biochemistry and Molecular Biology*, 111, 75-81.

Recio I. & Visser S. (1999). Identification of two distinct antibacterial domains within the sequence of bovine α_{s2}-casein. *Biochimica et Biophysica Acta*, 1428, 314-326.

Reed R. (1996). Initial splice-recognition and pairing during pre-mRNA splicing. *Current Opinion in Genetics and development*, 6, 215-220.

Remeuf F., Cossin V., Dervin C., Lenoir J. & Tomassone R. (1991). Relation entre les caractères physicochimiques des laits et leur aptitude fromagère. *Lait*, 71, 397-421.

Remillard R.L., Pickett J.P., Thatcher C.D. & Davenport D.J. (1993). Comparison of kittens fed queen's milk with those fed milk replacers. *American Journal of Veterinary Research*, 54, 901-907

Reynolds E.C. (1993). Phosphopeptides for treatment of dental calculus. World Patent WO 93/03707.

Reynolds E.C., Riley P.F. & Adamson N.J. (1994). A selective precipitation purification procedure for multiple phosphoseryl-containing peptides and methods for their identification. *Analytical Biochemistry*, 217, 277-284.

Ribadeau Dumas B., Brignon G., Grosclaude F. & Mercier J.C. (1972). Structure primaire de la caséine βA2 bovine. Séquence complète. *European Journal of Biochemistry*, 25,505-514.

Ribadeau Dumas B. & Grappin R. (1989). Milk protein analysis. *Le Lait*, 69, 357-416.

Richardson B.C. & Creamer L.K. (1974). Comparative micelle structure. III. The isolation and chemical characterization of caprine β1 casein and β2 casein. *Biochimica et Biophysica Acta*, 365, 133-137.

Richardson B.C. & Mercier J.C. (1979). The primary structure of the ovine β-caseins. *European Journal of Biochemistry*, 99, 285-297.

Rijnkels M., Kooiman P.M., de Boer H.A. & Pieper F.R. (1997). Organization of the bovine casein gene locus. *Mammalian Genome*, 8, 148-152.

Robinson N.E. (2002). Protein deamidation. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 5283–5288.

Robinson N.E. & Robinson A.B. (2001). Deamidation of human proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 12409–12413.

Robinson N.E. & Robinson A.B. (2004). In: Molecular Clocks. Althouse Press, Cave Junction. 419 pages.

Roscigno R.F. & Garcia-Blanco M.A. (1995). SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. *RNA*, 1, 692-706.

Rosen J.M., Wyszomierski S.L. & Hadsell D. (1999). Regulation of milk protein gene expression. *Annual Review of Nutrition*, 19, 407-436.

Saito T. & Itoh.T. (1992). Variation and distributions of O-glycosidically linked sugar chains in bovine κ-casein. *Journal of Dairy Science*, 75, 1768-1774.

Saito Y., Lee Y.S. & Kimura S. (1998). Minimum effective dose of casein phosphopeptides (CPP) for enhancement of calcium absorption in growing rats. *International Journal of Vitamine and Nutrition Research*, 68, 335-340.

Sakanaka S., Tachibana Y., Ishihara N. & Juneja L.R. (2005). Antioxidant properties of casein calcium peptides and their effects on lipid oxidation in beef homogenates. *Journal of Agricultural and Food Chemistry*, 53, 464-468.

Salimei E., Fantuz F., Coppola R., Chiofalo B., Polidori P. & Varisco G. (2004). Composition and characteristics of ass's milk. *Animal Research*, 53, 67-78.

Shah N.P. (2000). Effects of milk-derived bioactives: An overview. *British Journal of Nutrition*, 84, 3-10.

Schlimme E. & Meisel H. (1993). Bioactive peptides: Structural, physiological, and analytical aspects. *Newsletters of the International Dairy Federation*, 139.

Schlimme E. & Meisel H. (1995). Bioactive peptides derived from milk proteins. Structural, physiological and analytical aspects. *Die Nahrung*, 39, 1-20.

Schmidt DG. (1982). Association of caseins and casein micelle structure. In: Fox P.F. & Sweeney P.L.H. editors. Development in Dairy Chemistry, vol. 1. London: Applied Science, pp. 61-86.

Sheth N., Roca X., Hastings M.L., Roeder T., Krainer A.R. & Sachidanandam R. (2006). Comprehensive splice-site analysis using comparative genomics. *Nucleic Acids research*, 34, 3955-3967.

Shin K., Yamauchi K., Teraguchi S., Hayasawa H., Tomita M., Otsuka Y. & Yamazaki S. (1998). Antibacterial activity of bovine lactoferrin and its peptides against enterohaemorrhagic *Escherichia coli* O157:H7. *Letters in Applied Microbiology*, 26, 407-411.

Shukla G.C. & Padgett R.A. (2001). The intramolecular stem-loop structure of the U6 snRNA can functionally replace the U6atac snRNA stem-loop. *RNA*, 7, 94-105.

Silva S.V. & Malcata F.X. (2005). Caseins as source of bioactive peptides. *International Dairy Journal*, 15, 1-15.

Singh N.N., Androphy E.J. & Singh R.N. (2004). In vivo selection reveals combinatorial controls that define a critical exon in the spinal muscular atrophy genes. *RNA*, 10, 1291-1305.

Smacchi E. & Gobbetti M. (1998). Peptides from several Italian cheeses inhibitory to proteolytic enzymes of lactic acid bacteria, *Pseudomonas fluorescens* ATCC 948 and to the angiotensin I-converting enzyme. *Enzyme and Microbial Technology*, 22, 687-694.

Smith C.W., Chu T.T. & Nadal-Ginard B. (1993). Scanning and competition between AGs are involved in 3' splice site selection in mammalian introns. *Molecular and Cellular Biology*, 13, 4939-4952.

Snoeren, T.H.M., van Markwijk B. & van Montfort, R. (1980). Some physicochemical properties of bovine α_{s2}-casein. *Biochimica et Biophysica Acta*, 622, 268-276.

Sontheimer E.J., Sun S. & Piccirilli J.A. (1997). Metal ion catalysis during splicing of pre-messenger RNA. *Nature*, 388, 801-805.

Soulier S. & Gaye P. (1981). Enzymatic O-glycosylation of κ-caseinomacropeptide by ovine mammary Golgi membranes. *Biochimie*, 63, 619-628.

Stensballe A., Andersen S. & Jensen O.N. (2001). Characterization of phosphoproteins from electrophoretic gels by nanoscale Fe(III) affinity chromatography with off-line mass spectrometry analysis. *Proteomics*, 1, 207-222.

Sterner D.A., Carlo T. & Berget S.M. (1996). Architectural limits on split genes. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 15081-15085.

Stewart I.I., Thomson T. & Figeys D. (2001). O labeling: A tool for proteomics. *Rapid Communications in Mass Spectrometry*, 15, 2456-2465.

Stoianova L.G., Abramova L.A. & Ladado K.S. (1988). Sublimation-dried mare's milk and the possibility of its use in creating infant and dietary food products. *Voprosy Pitaniia*, 2, 64-67.

Svedberg J., De Haas J. & Leimenstoll G. (1985). Demonstration of β-casomorphin immunoreactive materials in *in vitro* digests of bovine milk and in small intestine contents after bovine milk ingestion in adult humans. *Peptides*, 6, 825-830.

Swaisgood H.E. (1982). Chemistry of milk protein. In: Developments in Dairy Chemistry. Fox P.F., editor. London & New York: Applied Science Publishers, pp. 1-59.

Swaisgood H.E. (1988). Structural changes in milk proteins. In: Milk proteins. Nutritional, Clinical, Functional and Technological Aspects. Barth C.A. and Schlimme E., editors. New York: Steinkoppf Verlab Darmstadt, Springer Verlag, pp. 192-210.

Swaisgood H.E. (1992). Chemistry of the caseins. In: Advance Dairy Chemistry-I: Proteins. Fox P. F., editor. New York: ElsevierApplied Science, pp. 63-110.

Takase K., Niki R. & Arima S. (1980). A sedimentation equilibrium study of the temperature-dependent association of bovine β-casein. *Biochimica et Biophysica Acta*, 622, 1-8.

Talerico M. & Berget S.M. (1994). Intron definition in splicing of small Drosophila introns. *Molecular and Cellular Biology*, 14, 3434-3445.

Tange T.Ø., Damgaard C.K., Guth S., Valcárcel J. & Kjems J. (2001). The hnRNP A1 protein regulates HIV-1 tat splicing via a novel intron silencer element. *The EMBO Journal*, 20, 5748-5758.

Tardiff D.F., Lacadie S.A. & Rosbash M. (2006). A genome-wide analysis indicates that yeast pre-mRNA splicing is predominantly posttranscriptional. *Molecular Cell*, 24, 917-929.

Tarelli E. & Corran P.H. (2003). Ammonia cleaves polypeptides at asparagine proline bonds. *Journal of Peptide Research*, 62, 245–251.

Tarn W.Y. & Steitz J.A. (1995). Modulation of 5' splice site choice in pre-messenger RNA by two distincts steps. *Proceedings of the National Academy of Sciences of the United States of America*, 92, 2504-2508.

Tarn W.Y. & Steitz J.A. (1996a). Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. *Science*, 273, 1824-1832.

Tarn W.Y. & Steitz J.A. (1996b). A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) introns *in vitro*. *Cell*, 84, 801-811.

Tauzin J., Miclo L. & Gaillard J.-L. (2002). Angiotensin-I-converting enzyme inhibitory peptides from tryptic hydrolysate of bovine α_{S2}-casein. *FEBS Letters*, 531, 369-374.

Tauzin J., Miclo L., Roth S., Mollé D. & Gaillard J.-L. (2003). Tryptic hydrolysis of bovine α_{s2}-casein: Identification and release kinetics of peptides. *International Dairy Journal*, 13, 15-27.

Tchorbanov B. & Iliev I. (1993). Limited enzymatic hydrolysis of casein in the presence of ethanol. *Enzyme and Microbial Technology*, 15, 974-978.

Teschemacher H. (2003). Opioid receptor ligands derived from food proteins. *Current Pharmaceutical Design*, 9, 1331-1344.

Teshima G., Hancock W.S. & Canova-Davis E. (1995). Effect of deamidation and isoaspartate formation on the activity of proteins. In: Deamidation and Isoaspartate Formation in Peptides and Proteins. Aswald D.W., editor. Boca Raton: CRC, pp. 167-191.

Teucher B., Majsak-Newman G., Dainty J.R., McDonagh D., FitzGerald R.J. & Fairweather-Tait S.J. (2006). Calcium absorption is not increased by caseinophosphopeptides. *American Journal of Clinical Nutrition*, 84, 162-166.

Thingholm T.E., Jørgensen T.J.D., Jensen O.L. & Larsen M.R. (2006). Highly selective enrichment of phosphorylated peptides using titanium dioxide. *Nature Protocols*, 1, 1929-1935.

Threadgill D.W. & Womack J.E. (1990). Genomic analysis of the major bovine milk protein genes. *Nucleic Acids Research*, 18, 6935-6942.

Thompson M.P., Tarassuk N.P., Jenness R., Lillevik H.A., Ashworth V.S.& Rose D. (1965). Nomenclature of the proteins of cow's milk. Second revision. *Journal of Dairy Science*, 48, 159-169.

Tomita M., Bellamy W., Takase M., Yamauchi K. Wakabayashi H. & Kawase K. (1991). Potent antibacterial peptides generated by pepsin digestion of bovine lactoferrin. *Journal of Dairy Science*, 74, 4137-4142.

Topper Y.J. & Freeman C.S. (1980). Multiple hormone interactions in the developmental biology of the mammary gland. *Physiological Reviews*, 60, 1049-1106.

Tranter H.S. & Board R.G. (1982). The antimicrobial defense of avian eggs: Biological perspective and chemical basis. *Journal of Applied Biochemistry*, 4, 295-338.

Trujillo A.J., Guamis B. & Carretero, C. (1997). Hydrolysis of caprine β-casein by plasmin. *Journal of Dairy Science*, 80, 2258-2263.

Tsuchita H., Goto T., Shimizu T., Yonehara Y. & Kuwata T. (1996). Dietary casein phosphopeptides prevent bone loss in aged ovariectomized rats. *Journal of Nutrition*, 126, 86-93.

Tsuchita H., Suzuki T. & Kuwata T. (2001). The effect of casein phosphopeptides on calcium absorption from calcium-fortified milk in growing rats. *British Journal of Nutrition*, 85, 5-10.

Tunesi S. & Anderson M. (1991). Influence of chemisorption on the photodecomposition of salicylic acid and related compounds using suspended TiO_2 ceramic membranes. *The Journal of Physical Chemistry*, 95, 3399-3405.

Tyler-Cross R. & Schirch V. (1991). Effects of amino acid sequence, buffers, and ionic strength on the rate and mechanism of deamidation of asparagine residues in small peptides. *The Journal of Biological Chemistry*, 266, 22549-22556.

Umbach M., Teschemacher H. & Praetorius K. (1985). Demonstration of a β-casomorphin immunoreactive material in the plasma of newborn calves after milk intake. *Regulatory Peptides*, 12, 223-230.

Umeyama M., Kira A., Nishimura K. & Naito A. (2006). Interactions of bovine lactoferricin with acidic phospholipid bilayers and its antimicrobial activity as studied by solid-state NMR. *Biochimica et Biophysica Acta*, 1758, 1523-1528.

Vader L.W., De Ru A., van der Wal Y., Kooy Y.M.C., Benckhuijsen W., Mearin, M.L., Drijfhout J.W., van Veelen P & Koning, F. (2002). Specificity of tissue transglutaminase explains cereal toxicity in celiac disease. *Journal of Experimental Medicine*, 195, 643-649.

van der Wal Y., Kooy Y., van Veelen P., Peña S., Mearin L., Papadopoulos G. & Koning F. (1998). Cutting edge: Selective deamidation by tissue transglutaminase strongly enhances gliadin-specific T cell reactivity. *Journal of Immunology*, 161, 1585-1588.

Vidal V.P., Verdone L., Mayes A.E. & Beggs J.D. (1999). Characterization of U6 snRNA-protein interactions. *RNA*, 5, 1470-1481.

Visser S. (2000). Is bovine β-casein G true genetic variant? *International Dairy Journal*, 10, 133-134.

Vonderhaar B.K. & Ziska S.E. (1989). Hormonal regulation of milk protein gene expression. *Annual Review of Physiology*, 51, 641-652.

Vreeman H.J., Visser S., Slangen C.J. & Van Riel J.A.M. (1986). Characterization of bovine κ-casein fraction and the kinetics of chymosin-induced macropeptide release from carbohydrate-free and carbohydrate-containing fractions determined by high-performance gel-permeation chromatography. *Biochemical Journal*, 240, 87-97.

Wang Q., Allen J.C. & Swaisgood H.E. (1997a). Binding of retinoids to β-lactoglobulin isolated by bioselective adsorption. *Journal of Dairy Science*, 80, 1047-1053.

Wang Q., Allen J.C. & Swaisgood H.E. (1997b). Binding of vitamin D and cholesterol to β-lactoglobulin. *Journal of Dairy Science*, 80, 1054-1059.

Wassarman K.M. & Steitz J.A. (1992). The low-abundance U11 and U12 small nuclear ribinucleoproteins (snRNPs) interact to form a two-snRNP complex. *Molecular and Cellular Biology*, 12, 1276-1285.

Walstra P. (1997). Casein sub-micelles, Do they exist? *Hannah Symposium Caseins and Caseinates, Structures, Interactions, Networks. Ayr, Écosse, Royaume-Uni,* 21-23 mai.

West D.W. (1986). Structure and function of the phosphorylated residues of casein. *Journal of Dairy Research*, 53, 333-352.

Whitney R.M., Brunner J.R., Ebner K.E., Farrell Jr. H.M., Josephson R.V., Morr C.V. & Swaisgood H.E. (1976). Nomenclature of the proteins of cow's milk: Fourth revision. *Journal of Dairy Science*, 59, 795-815.

Will C.L. & Lührmann R. (1997). snRNP structure and function (1997). In: Eukaryotic mRNA Processing. Krainer A. R., editor. Oxford: IRL Press, pp. 130-173.

Will C.L., Schneider C., Reed R. & Lührmann R. (1999). Identification of both shared and distinct proteins in the major and minor spliceosomes. *Science*, 284, 2003-2005.

Will C.L., Schneider C., MacMillan A.M., Katopodis N.F. Neubauer G., Wilm M., Lührmann R. & Query C.C. (2001). A novel U2 and U11/U12 snRNP protein that associates with the pre-mRNA branch site. *The EMBO Journal*, 20, 4536-4546.

Will C.L., Schneider C., Hossbach M., Urlaub H., Rauhut R., Elbashir S., Tuschl T. & Lührmann R. (2004). The human 18S U11/U12 snRNP contains a set of novel proteins not found in the U2-dependent spliceosome. *RNA*, 10, 929-941.

Will C.L. & Lührmann R. (2005). Splicing of a rare lass of introns by the U12-dependent spliceosome. *Biological Chemistry*, 380, 713-724.

Wu A.M. & Herp A. (1985). A table of lectin carbohydrate specificities. In: Lectins, vol. IV. de Gruiter W. and Co., editors. Berlin, New York. Pp. 629-636.

Wu S.Y., Perez M.P., Puyol P. & Sawyer L. (1999). β-Lactoglobulin binds palmitate within its central cavity. *The Journal of Biological Chemistry*, 274, 170-174.

Xie M. & Schowen R.L. (1999). Secondary structure and protein deamidation. *Journal of Pharmaceutical Sciences*, 88, 8-13.

Yamamoto N., Akino A. & Takano T. (1994). Antihypertensive effect of the peptides derived from casein by an extracellular proteinase from *Lactobacillus helveticus* CP790. *Journal of Dairy Science*, 77, 917-922.

Yean S.-L., Wuenschell G., Termini J. & Lin R.-J. (2000). Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature*, 408, 881-884.

Yeung A.C., Glahn R.P. & Miller D.D. (2001). Dephosphorylation of sodium caseinate, enzymatically hydrolyzed casein and casein phosphopeptides by intestinal alkaline phosphatase: Implications for iron availability. *Journal of Nutritional Biochemistry*, 12, 292-299.

Yoshikawa M., Tani F., Shiota H., Usui H., Kurahashi K. & Chiba H. (1994). Casoxin D, an opioid antagonist/ileum-contracting/vasorelaxing peptide derived from human casein. In: β-Casomorphins and Related Peptides: Recent development. Brantl V. & Teschemacher H., editors. Weinheim: VCH, Germany, pp. 43-48.

Yu Y.T. & Steitz J.A. (1997). Site-specific crosslinking of mammalian U11 and U6atac to the 5' splice site of an AT-AC intron. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 6030-6035.

Yu Y.T., Scharl E.C., Smith C.M. & Steitz J.A. (1997). In: RNA world II. Gesteland R. F., Cech T. R. & Atkins J. F., editors. New York: Cold Spring Harbor Laboratory Press, pp. 487-524.

Zhu J., Mayeda A. & Krainer A.R. (2001). Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Molecular Cell*, 8, 1351-1361.

Zucht H.-D., Raida M., Adermann K., Magert H.-J. & Forssmann W.-G (1995). Casocidin-I: A casein-α_{s2} derived peptide exhibits antibacterial activity. *FEBS Letters*, 372, 185-188.

Zuo P. & Manley J.L. (1994). The human splicing factor ASF/SF2 can specifically recognize pre-mRNA 5' splice sites. *Proceedings of the National Academy of Sciences of the United States of America*, 91, 3363-3367.

Zwierzchowski L. (1996). Structure, expression and engineering of milk protein genes. *Biotechnology (Poland)*, 33, 81-94.

Liste des publications

-Equine α_{s1}-casein : characterization of alternative splicing isoforms and determination of phosphorylation levels of multiple isoforms.
A Matéos, L. Miclo, D. Mollé, J-M Girardet and J-L Gaillard.
Publication soumise à Journal of Dairy Science

-Two-dimensional cartography of equine β-casein variants achevied by isolation of phosphorylation isoforms and control of the deamidation phenomenon.
A Matéos, J-M Girardet, D Mollé, A Dary, L Miclo and J-L Gaillard.
Publication soumise à Journal of Dairy Science

-Caractérisation des isoformes de la caséine α_{s1} équine : variants post-transcriptionnels et post-traductionnels
A Matéos, J-M Girardet, D Mollé, L Miclo et J-L Gaillard
Acte du séminaire de l'école doctorale RP2E, 17 janvier 2008

Liste des communications par voie d'affichage

-Etude structurale des caséines équines et de leurs phosphopeptides
A Matéos
Doctoriales de Lorraine 2007

-Analyse protéomique des variants post-transcriptionnels et post-traductionnels de la caséine α_{s1} équine
A Matéos, L Miclo, D Mollé, J-M Girardet et J-L Gaillard
34ème Forum des Jeunes Chercheurs, Iles des Embiez, 15 au 17 octobre 2007

-Caractérisation des isoformes de la caséine α_{s1} équine : variants post-transcriptionnels et post-traductionnels
A Matéos, J-M Girardet, D Mollé, L Miclo et J-L Gaillard
Séminaire de l'école doctorale RP2E, 17 janvier 2008
-Characterization of isoforms of equine α_{s1}-casein: post-transcriptional variants and phosphorylation post-translational variants
A Matéos, L Miclo, D Mollé, J-M Girardet et J-L Gaillard
First international symposium on minerals and dairy products, Saint-Malo, 1er au 3 octobre 2008

-Phosphorylation variants of equine β-casein: purification of caseinophosphopeptides and determination of the phosphorylation sites
A Matéos, L Miclo, D Mollé, J-M Girardet et J-L Gaillard
First international symposium on minerals and dairy products, Saint-Malo, 1er au 3 octobre 2008

-Preparation of partially dephosphorylated bovine β-casein, a model for the human homologous protein
A Matéos, J-M Girardet, F Bouzobra, L Miclo et J-L Gaillard
First international symposium on minerals and dairy products, Saint-Malo, 1er au 3 octobre 200

Résumé

La caséine β (CN-β) et la caséine α_{s1} (CN-α_{s1}) du lait de jument possèdent un taux variable de phosphorylation et sont de bons modèles d'étude de l'influence du degré de phosphorylation et de la séquence peptidique sur la chélation de caséinophosphopeptides (CPP) avec des minéraux d'intérêt nutritionnel. Avant d'envisager une telle étude, la structure des caséines doit être déterminée précisément. Notre travail a été consacré à la caractérisation de variants post-transcriptionnels et post-traductionnels de CN-α_{s1} et de CN-β. Après fractionnement chromatographique et analyse par spectrométrie de masse, la CN-α_{s1} entière, trois variants d'épissage alternatif des exons 7 et 14 de la CN-α_{s1} et quatre variants délétés du résidu Gln^{91}, résultat de l'utilisation d'un site d'épissage cryptique, ont été identifiés dans le lait équin. Nous avons montré que le degré de phosphorylation de ces isoformes varie de 2 à 8 groupes phosphates. Au total, 36 isoformes différentes de CN-α_{s1} ont été caractérisées. La cartographie bidimensionnelle de la CN-β a été établie avec précision après avoir isolé par chromatographie chacun des variants de phosphorylation (possédant de 3 à 7 groupes phosphates) et après avoir caractérisé les formes de CN-β modifiées par désamidation non enzymatique du résidu Asn^{135}. Des CPP trypsiques de chaque variant de phosphorylation ont été préparés avec une technique récente de chromatographie d'affinité au dioxyde de titane, ce qui a permis de localiser par spectrométrie de masse en tandem les sites phosphorylés de la CN-β (Ser^9, Ser^{10}, Thr^{12}, Ser^{18}, Ser^{23}, Ser^{24}, Ser^{25}) et de montrer que la phosphorylation de la CN-β n'est pas aléatoire mais séquentielle.

Mots-clés : lait de jument / caséines équines / phosphorylation variable / épissage alternatif / caséinophosphopeptide

Abstract

Equine β-casein (β-CN) and α_{s1}-casein (α_{s1}-CN) have a variable phosphorylation degree and are good models for the study of the influence of phosphorylation degree and peptide sequence on chelation of caseinophosphopeptides (CPP) with minerals of nutritional interest. Before considering such study, structure of caseins must be precisely determined. Our work has been devoted to the characterization of post-transcriptional and post-translational variants of α_{s1}-CN and β-CN. Concerning α_{s1}-CN, the full-length protein, three alternative splicing variants involving exons 7 and 14 and four variants involving cryptic splice site resulting in deletion of residue Gln^{91} have been identified in mare's milk after chromatographic fractionation and mass spectrometric analysis. The phosphorylation degree of these variants varies between 2 and 8 phosphate groups. Finally, 36 isoforms of α_{s1}-CN have been identified. Isolation of each phosphorylation variant (having 3 to 7 phosphate groups) of β-CN by chromatography, and characterization of modified forms of β-CN by non enzymatic deamidation of residue Asn135 permits the establishment of bidimensional cartography of β-CN with precision. After hydrolysis by trypsin, CPP of each phosphorylation variant have been prepared by affinity chromatography to titanium dioxide, a recent technology, which allowed to locate by mass tandem spectrometry the phosphorylated sites of β-CN (Ser^9, Ser^{10}, Thr^{12}, Ser^{18}, Ser^{23}, Ser^{24}, Ser^{25}). It was shown that the phosphorylation of β-CN is not a random process but follows a sequential way.

Key words: mare's milk / equine casein / variable phosphorylation / alternative splicing / caseinophosphopeptide

INSTITUT NATIONAL
POLYTECHNIQUE
DE LORRAINE

AUTORISATION DE SOUTENANCE DE THESE
DU DOCTORAT DE L'INSTITUT NATIONAL
POLYTECHNIQUE DE LORRAINE

o0o

VU LES RAPPORTS ETABLIS PAR :
Monsieur Jean-Marc CHOBERT, Directeur de Recherche, INRA, Nantes

Monsieur Saïd BOUHALLAB, Directeur de Recherche, INRA, Rennes

Le Président de l'Institut National Polytechnique de Lorraine, autorise :

Madame MATEOS Aurélie

NANCY BRABOIS
2, AVENUE DE LA
FORET-DE-HAYE
BOITE POSTALE 3
F - 54501
VANDŒUVRE CEDEX

à soutenir devant un jury de l'INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE, une thèse intitulée :

"Etude protéomique de la microhétérogénéité des caséines αs_1 et ß équines : identification des variants transcriptionnels et de phosphorylation ; identification des sites phosphorylés de la caséine ß »

en vue de l'obtention du titre de :

DOCTEUR DE L'INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE

Spécialité : « **Sciences Agronomiques** »

Fait à Vandoeuvre, le 05 novembre 2008
Le Président de l'I.N.P.L.,
F. LAURENT

TEL. 33/03.83.59.59.59
FAX. 33/03.83.59.59.55