

# **La caséine $\kappa$ et la famille multigénique des trois caséines « sensibles au calcium » Polymorphisme, biosynthèse et évolution**

**Jean-Claude Mercier, François Grosclaude,  
Patrice Martin**

**Société Française  
de Génétique**

**Président**

E. Moustacchi

**Vice-présidents**

R. Devoret

A. Bernheim

F. Grosclaude

**Secrétaire général**

R. Motta

*Prière d'adresser toute correspondance au  
Secrétariat général de la SFG,  
R. Motta, laboratoire d'immunogénétique,  
centre Marcel-Délépine - Cnrs, 1,  
rue Haute, 45071 Orléans Cedex 2*

**Comité de rédaction**

A. Bernheim

M. Fellous

J. Générumont

F. Minvielle

R. Motta

A. Nicolas

S. Sommer

D. de Vienne

**Secrétaire**

M.-L. Prunier

La glande mammaire synthétise et sécrète en grandes quantités diverses protéines spécifiques, notamment les caséines. Ces phosphoprotéines, associées entre elles sous forme de micelles, représentent pondéralement près de 80 % (27 g/l de lait chez les bovins) des lactoprotéines chez les ruminants et 40 % (4 g/l) chez la femme. La caséine, définie comme la fraction du lait précipitable à pH 4,6 et à 20 °C, est une source essentielle d'acides aminés, de phosphore et de calcium pour le jeune mammifère, et elle constitue l'essentiel de la matière première protéique des yaourts et fromages. Quelques dérivés peptidiques auraient également une faible activité biologique [1]. En gel d'électrophorèse et en milieu dénaturant et réducteur, la caséine bovine apparaît comme un mélange d'une trentaine de constituants, mais les études structurales et génétiques ont clairement établi l'existence de seulement 4 types de caséine contrôlés par 4 *loci* étroitement liés. L'hétérogénéité est due à 3 facteurs : le polymorphisme génétique, la teneur variable en groupements prosthétiques et une protéolyse partielle par la plasmine présente dans le lait. Ces résultats ont facilité l'analyse des caséines d'autres espèces, et conduit tout naturellement à l'étude des ARNm et des gènes.

Le but de cet article est de présenter une synthèse des connaissances acquises sur le polymorphisme, la biosynthèse et l'évolution des caséines, en particulier de celles sensibles au calcium,  $\alpha$ 1,  $\alpha$ 2 et  $\beta$ , dont les gènes sont apparentés.

## **Structure et polymorphisme des caséines bovines**

Les caséines  $\alpha$ 1,  $\alpha$ 2,  $\beta$  et  $\kappa$ , sont des phosphoprotéines constituées respectivement de 199, 207, 209 et 169 résidus d'acides aminés dans l'espèce bovine, qui contiennent 6 à 7, 9 à 13, 5, et 1 à 2 groupes phosphates fixés sur des résidus séryles et, dans un cas, thréonyles. La caséine  $\kappa$  comporte en outre des chaînes glycanes fixées sur des résidus thréonyles de la région C-terminale. Les trois premières caséines précipitent en présence de calcium sauf lorsqu'elles sont associées à la caséine  $\kappa$  sous forme de copolymères. La forma-

J.-C. Mercier : directeur de recherche à l'INRA.  
F. Grosclaude : directeur de recherche à l'INRA.  
P. Martin : directeur de recherche à l'INRA.  
Laboratoire de génétique biochimique, Institut national de la recherche agronomique, INRA-CRJ, 78350 Jouy-en-Josas, France.

tion de micelles stables (figure 1) résulte de l'établissement de liaisons phosphocalciques entre des submicelles, elles-mêmes stabilisées par des interactions hydrophobes et ioniques.

Les variants génétiques de chaque caséine (figure 2a, p. IV), détectés dans diverses races bovines françaises et étrangères, ont servi de marqueurs pour étudier la ségrégation mendélienne des allèles. Dès 1964-1965, l'observation que seuls 2 des 4 types de gamètes attendus étaient transmis à

la descendance de taureaux doubles hétérozygotes pour des allèles muets aux loci  $Cn-\alpha1$  et  $Cn-\beta$  :  $Cn-\alpha1$  et  $Cn-\kappa$  ou  $Cn-\beta$  et  $Cn-\kappa$ , démontrait l'étroite liaison de ceux-ci. En 1977, la détection de variants de la caséine  $\alpha2$  permettait d'établir que le loci  $Cn-\alpha2$  était également lié aux 3 autres loci  $Cn$  [2]. La fréquence des allèles et des haplotypes (association d'allèles appartenant à des loci différents) diffère selon les races considérées (Tableau I), et l'analyse statistique des déséquilibres de

liaison ( $\alpha1-\beta > \beta-\kappa > \alpha1-\kappa$ ) observés entre allèles des loci couramment polymorphes  $Cn-\alpha1$ ,  $Cn-\beta$  et  $Cn-\kappa$ , a conduit à proposer un modèle donnant leur disposition relative, soit  $\alpha1-\beta-\kappa$ .

Ainsi, l'analyse structurale et génétique des caséines bovines et ultérieurement d'autres espèces a apporté de nombreuses informations : (1) simplicité du système génique (4 loci) ; (2) similitudes de séquences intra- (40 % entre 2 régions de 60 résidus de la caséine  $\alpha2$ ) et inter-chaînes (sites majeurs de

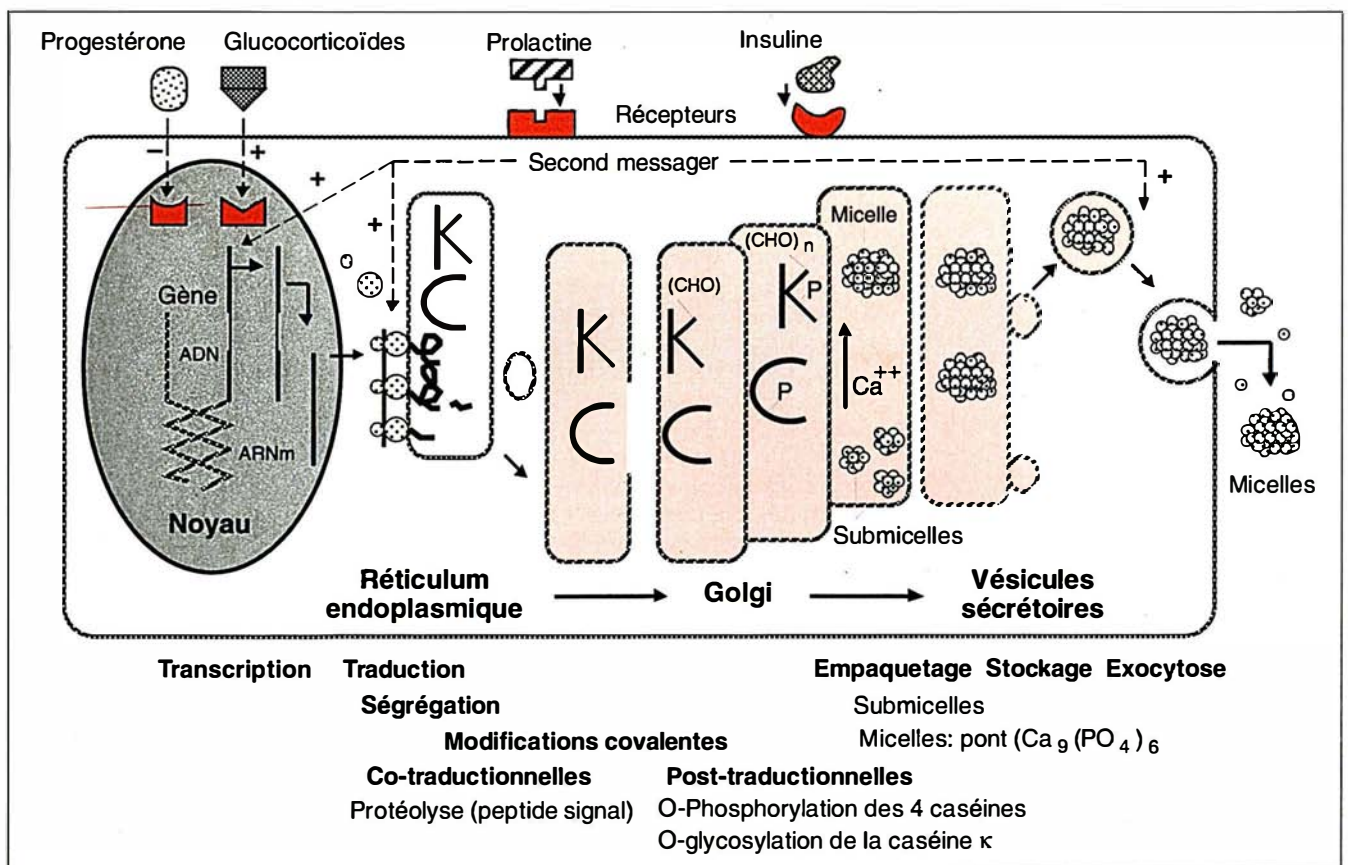


Figure 1. **Biosynthèse et transport intracellulaire des caséines dans les cellules épithéliales mammaires.** L'expression des gènes codant pour les caséines est sous contrôle multihormonal complexe : induction par la prolactine (en présence d'insuline in vitro) et modulation positive par les glucocorticoïdes, négative par la progestérone. L'accumulation relativement coordonnée des ARNm de caséines, en fin de gestation et jusqu'à l'optimum de la lactation, résulte de l'augmentation du taux de transcription des gènes ( $\times 2-4$ ) et de la durée de demi-vie des ARNm ( $\times 25$ ) [18]. Les chaînes polypeptidiques naissantes sont amputées du peptide signal au cours de leur translocation dans le réticulum endoplasmique, puis d'autres modifications covalentes ont lieu dans les divers compartiments golgiens. L'assemblage des caséines en micelles aurait lieu au plus tard au niveau du réseau transgolgien. Les micelles, véhiculées par les vésicules sécrétoires jusqu'au pôle apical de la cellule, sont libérées dans la lumière des acini par exocytose. La prolactine intervient également au niveau de la sécrétion des caséines et cet effet lactogène rapide met en jeu l'activation des phospholipases A2 et la libération d'acide arachidonique [19]. C et K symbolisent respectivement les caséines « sensibles au calcium » et la caséine  $\kappa$ . CHO : chaîne glycane ; P : groupe phosphate.

phosphorylation ; peptide signal) suggérant que les gènes des caséines  $\alpha 1$ ,  $\alpha 2$  et  $\beta$  ont dérivé d'un gène ancestral par duplications successives ; (3) phosphorylation des caséines impliquant la reconnaissance enzymatique d'une séquence code Ser/Thr-X-A, où A représente un résidu glutamyle ou séryle déjà phosphorylé, et X un résidu d'acide aminé quelconque [3] ; (4) localisation des mutations différenciant les variants génétiques et établissement d'un modèle de filiation des allèles ; (5) extrême rapidité d'évolution des caséines.

Certains variants génétiques diffèrent par leurs propriétés physico-chimiques et technologiques [4]. Par exemple, le variant  $\kappa B$ , associé à une vitesse de coagulation du lait plus rapide et à la formation d'un caillé plus ferme, semble préférable en transformation fromagère. Du fait de cet intérêt techno-

logique et de la possibilité actuelle de déterminer le génotype dès la naissance par simple analyse de l'ADN génomique, le typage systématique des animaux tend à se développer, ce qui permettra d'inclure ce nouveau paramètre dans les schémas de sélection. L'haplotype le plus favorable sur le plan fromager serait la combinaison Cn- $\beta$  B - Cn- $\kappa$  B. La technique RFLP (polymorphisme de longueur de fragments de restriction), utilisée initialement, est progressivement remplacée par la combinaison de deux techniques, l'amplification d'ADN *in vitro* (PCR) et l'emploi d'oligonucléotides spécifiques d'allèles, ce qui réduit le délai de typage à 24-48 heures.

Les caséines des autres espèces sont moins bien connues. Les micelles du lait humain — qui sont de petite taille, ce qui leur conférerait une bonne digestibilité — contiennent les

caséines  $\beta$  et  $\kappa$  et de très faibles quantités de caséine  $\alpha 1$ . Le polymorphisme qualitatif et quantitatif de la caséine  $\alpha 1$  de chèvre et celui, non allélique, des caséines  $\alpha 2$  ovine et caprine méritent un commentaire en raison de leur intérêt sur les plans académique et appliqué.

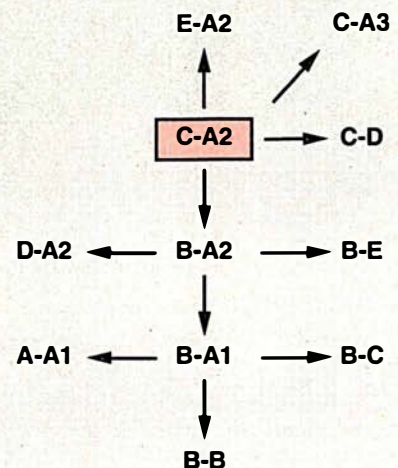
### Polymorphisme des caséines $\alpha 1$ caprine et $\alpha 2$ ovine et caprine

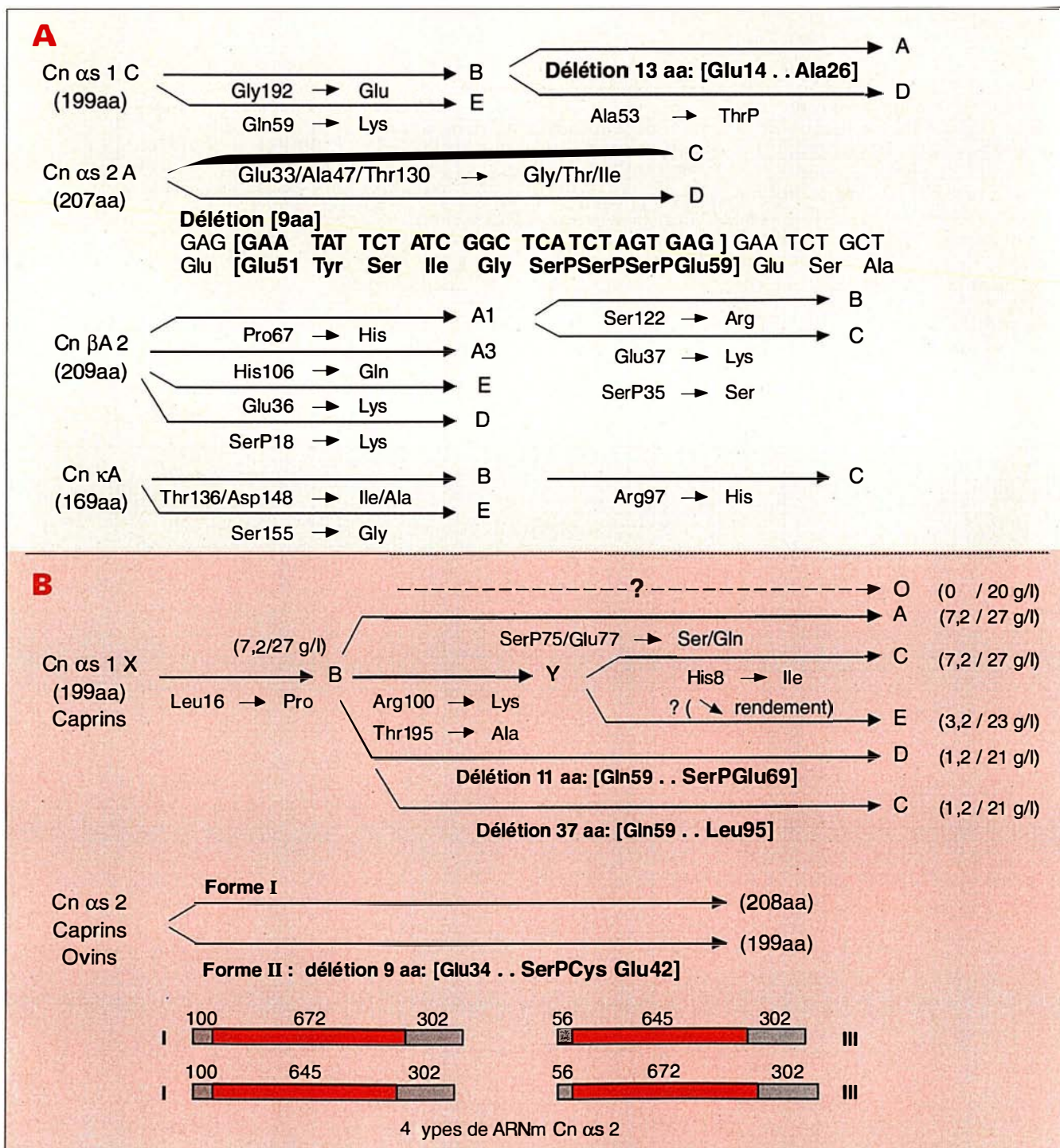
Actuellement, 7 allèles ont été détectés au locus Cn- $\alpha 1$  par analyse électrophorétique de laits de chèvres : l'un d'entre eux est un allèle nul (O) et 3 autres — E, D et F — sont associés à une moindre teneur en caséine  $\alpha 1$  (figure 2b). Les altérations structurales différenciant les 6 variants de caséine  $\alpha 1$  ont été caractérisées, D et F présentant respectivement des délé-

Tableau I  
FRÉQUENCES ET PHYLOGÉNIE DES HAPLOTYPES Cn- $\alpha 1$ -Cn- $\beta$  CHEZ LE ZÉBU MALGACHE ET DIVERSES RACES BOVINES

Les races bovines ont été groupées selon leur origine géographique, dans l'ordre : Europe du nord, îles Anglo-Normandes, Alpes et Jura, centre de la France, et péninsule italienne. Les races podoliques, originaires de la steppe ukrainienne, sont les plus archaïques. La première et la seconde lettre désignent respectivement le type d'allèle aux loci Cn- $\alpha 1$  et Cn- $\beta$ . Les fréquences  $\geq 10\%$  sont en caractères gras et CA<sup>2</sup> est considéré comme l'haplotype sauvage. La phylogénie des haplotypes est fondée sur les données de la figure 2. Le déséquilibre de liaison observé entre allèles des loci Cn- $\alpha 1$  et Cn- $\beta$  (absence ou rareté de certaines combinaisons d'allèles individuellement très répandus) suggérerait que ceux-ci sont très proches l'un de l'autre, ce qui est effectivement le cas puisque le groupe des 3 loci Cn- $\alpha 2$ , Cn- $\alpha 1$  et Cn- $\beta$  occuperait moins de 120-140 kb [9] sur le chromosome 6 bovin.

Races	AA1	BA1	BA2	BB	BC	CA1	CA2	CA3
Zébu	-	2,7	13,9	-	-	7,2	76,2	-
Ayrshire	-	55,4	44,0	0,3	-	-	-	0,3
Shorthorn	-	67,1	32,1	0,4	-	-	-	0,4
Holstein	0,3	56,1	40,5	0,7	-	-	1,6	1,1
Pie-noire	-	55	39	3	-	-	<1	3
Rouge Danoise	<1	71	23	6	-	-	<1	-
Normande	-	19	11	47	1	-	18	4
Jersey	-	9	28	37	3	-	24	-
Guernese	-	6	67	1	5	-	21	-
Brune Suisse	-	28	42	30	2	<1	7	-
Simmental	-	26	55	3	6	2	8	-
Charolaise	-	10	68	13	1	-	8	-
Podoliques Italiennes	-	28	32	4	-	3	33	-





tions internes de 11 et 37 résidus d'acides aminés. L'étude des ARNm de l'allèle F a révélé l'existence de transcrits multiples dont la forme majoritaire est déletée de 111 nucléotides dans sa partie codante. Cette séquence, présente au niveau génomique, correspond à 3 exons éliminés au cours du processus d'épissage. Dans le cas des transcrits issus de l'allèle D, seul le premier de ces exons serait exclu. Le taux réduit de synthèse protéique est probablement lié à une instabilité et/ou à l'altération du cadre de lecture de certains transcrits. Une insertion de 0,5 kb, située dans la région 3' non codante, pourrait être responsable de la plus faible expression de l'allèle E.

Dans les espèces ovine et caprine coexistent 2 formes de caséine  $\alpha_2$  et 4 types d'ARNm caractérisés par la présence ou l'absence d'un ou deux exons de 44 et 27 nucléotides dans la région 5' non codante et dans le cadre de lecture (figure 2b). Ces 4 types d'ARNm résultent d'un processus de maturation aléatoire se traduisant par l'élimination partielle des 2 exons précités. La comparaison des séquences des gènes Cn- $\alpha_2$  ovin, caprin et

bovin, au niveau de l'exon codant fréquemment éliminé, devrait permettre d'identifier la ou les mutations à l'origine de la modification du processus d'épissage dans les espèces ovine et caprine.

### Organisation des ARNm et des gènes codant pour les caséines, et localisation chromosomique

La taille des ARNm, séquence poly(A) exclue, varie de 690 à 1 349 nucléotides, dont 40 % en règle générale ne sont pas traduits (figure 3, p. VII). La comparaison inter-espèces des séquences d'ADNc homologues a confirmé l'évolution rapide des gènes, déjà déduite de l'étude comparative de caséines [5], et montré que le cadre de lecture, à l'exception des régions codant pour le peptide signal et les sites multiples de phosphorylation [6], évolue aussi rapidement (Cn- $\beta$ ) sinon plus (Cn- $\alpha_1$ ,  $\alpha_2$  et  $\kappa$ ) que les parties 5' et 3' non codantes. Cela suggère que la nécessaire conservation des propriétés de copolymérisation des caséines ne requiert pas le maintien

d'une séquence particulière d'acides aminés. La similitude de séquences entre les peptides signaux des 3 caséines « sensibles au calcium » est de l'ordre de 60 à 80 % (de 73 à 82 % au niveau nucléotidique) selon l'espèce considérée ; la plupart des mutations observées sont neutres (sans effet sur la traduction) ou conservatrices (substitution par un acide aminé similaire sur le plan fonctionnel), ce qui traduit une forte contrainte évolutive. Cette remarquable conservation de la structure primaire du peptide signal des caséines « sensibles au calcium » suppose que celle-ci détermine une conformation optimale permettant un transit extrêmement rapide des chaînes naissantes dans le réticulum endoplasmique, ce qui est primordial pour une cellule sécrétrice devant exporter de très grandes quantités de ces protéines.

L'analyse encore en cours des gènes de structure des caséines a conforté l'hypothèse d'une origine commune de ceux codant pour les caséines  $\alpha_1$ ,  $\alpha_2$  et  $\beta$  : présence de plusieurs motifs structuraux communs immédiatement en amont de la boîte ATA (figure 3) ; similitude d'organisation du début de l'unité de transcription (figure 4, p. VII) et des séquences codant pour les sites multiples de phosphorylation de type « Ser-Ser-Ser-Glu-Glu » (figure 2), qui sont engendrées par l'épissage de 2 exons. Ces 3 gènes diffèrent toutefois par la taille et le degré de mosaïcisme, ce qui suggère une évolution en plusieurs étapes [7] : duplications intragéniques d'un gène primitif comportant entre autres les éléments précédents, suivies de 3 duplications intergéniques et d'une évolution indépendante des triplicatas ainsi produits. On notera (encadré de la figure 4) la conservation de l'organisation du gène de structure de la caséine  $\beta$  au cours de l'évolution, les seules différences notables portant sur la taille de certains introns où sont souvent localisées des séquences répétitives.

L'organisation très différente du gène de structure de la caséine  $\kappa$  suppose une autre origine et son appartenance à la famille des gènes codant pour les fibrinogènes a été suggérée. Les structures primaires de ces protéines, dont les fonctions sont analogues puisque

**Figure 2. (a). Caractérisation et phylogénie des variants génétiques dans l'espèce bovine.** Les 2 délétions caractérisant les variants  $\alpha_1$  A et  $\alpha_2$  D résultent très probablement de mutations affectant les modalités de l'épissage, ce qui provoque l'élimination d'un exon. La séquence nucléotidique déletée dans l'ARNm  $\alpha_2$  D, mais présente dans le gène correspondant, est donnée à titre d'exemple pour montrer que la séquence codant pour un site multiple de phosphorylation résulte de l'épissage de 2 exons. Les variants  $\beta$ E et  $\beta$ C illustrent la validité du code de phosphorylation : dans la séquence tripeptidique SerP35-Glu-Glu37 (variant A1), la substitution Glu  $\rightarrow$  Lys qui supprime le déterminant Glu37 (C) du site de phosphorylation provoque la perte du groupe phosphate, alors qu'en position 36 (E) elle est sans effet. **(b). Mutations non silencieuses du cadre de lecture des allèles caprins au locus Cn- $\alpha_1$  et formes non alléliques de la caséine  $\alpha_2$  caprine et ovine (4 types d'ARNm).** X et Y désignent des variants hypothétiques au locus  $\alpha_1$ -Cn. Les nombres entre parenthèses indiquent respectivement la teneur en caséine  $\alpha_1$  du lait d'une chèvre homozygote pour l'allèle considéré, et celle de la caséine totale. Les 4 délétions, dont 3 affectent des séquences codant pour un site multiple de phosphorylation, seraient dues à un épissage anormal entraînant la perte d'un (Cn- $\alpha_1$  D, Cn- $\alpha_2$  forme II) ou de plusieurs (Cn- $\alpha_1$  F) exons. Les 4 types d'ARNm codant pour la caséine  $\alpha_2$  résultent de l'excision partielle de 2 exons de 44 et 27 nt affectant respectivement la région 5' non codante et la région codante. Leurs proportions relatives sont respectivement de l'ordre de : I + II : 5 % ; II + III : 30-40 % (ovins) et 10 % (caprins).

leur protéolyse limitée conduit respectivement à la coagulation du lait et du sang, présentent effectivement quelques similitudes, ultérieurement observées au niveau des exons codant pour ces régions [8].

Chez la souris et le lapin, la présence des gènes de caséines sur un même chromosome (synténie) a été démontrée à l'aide d'hybrides somatiques. Dans ces 2 espèces et chez les bovins, le chromosome porteur — respectivement le 5, 12 et 6 — a été identifié par hybridation *in situ*. La disposition relative des 4 *loci* Cn bovins, soit  $\alpha 1$ - $\beta$ - $\alpha 2$ - $\kappa$ , et la taille du complexe ( $\leq 300$  kb), récemment déterminées sur des fragments d'ADN nucléaire isolés par électrophorèse en champ pulsé [9], confirment et précisent le modèle d'organisation de cette unité génique déduit de l'analyse des déséquilibres de liaison.

### Expression des gènes codant pour les caséines

L'expression des gènes de structure des protéines du lait et plus particulièrement des caséines est encore très mal connue au niveau moléculaire. Les résultats obtenus chez les animaux transgéniques exprimant le gène d'une protéine du lactosérum plus ou moins délété en amont de l'unité de transcription et substitué dans le cadre de lecture, indiquent que des séquences en *cis* impliquées dans l'expression temporelle et tissulaire sont proches et en amont de l'unité de transcription ( $\alpha$ -lactalbumine  $\leq 0,2$  kb ;  $\beta$ -lactoglobuline  $\leq 0,8$  kb). Une bonne expression du transgène (de 10 à 40 % du niveau normal) a pu être obtenue avec moins de 1 kb de la région 5' flanquante [10, 11]. La zone

– 300/– 150 contient effectivement divers sites reconnus par des effecteurs nucléaires, mis en évidence par le retard de migration électrophorétique et/ou l'insensibilité vis-à-vis de nucléases [12] de fragments de ces gènes mis au contact d'extraits nucléaires mammaires.

L'analyse de l'expression de gènes Cn- $\beta$  hybrides, dans des animaux transgéniques et dans des lignées cellulaires mammaires, a montré respectivement que 2 et même 0,5 kb [13, 14] de la région promotrice suffisent pour la spécificité d'expression et que l'induction hormonale subsiste avec 0,3 kb [15, 16]. Toutefois, les faibles niveaux d'expression obtenus ( $< 1\%$ ) supposent la présence de séquences amplificatrices en dehors de ces régions. L'expression d'un gène hybride comportant 21 et 2 kb des régions flanquantes 5' et 3' [17] du gène Cn- $\alpha 1$  bovin où la région codant pour la première moitié de la précaséine a été remplacée par l'unité de transcription du gène de l'urokinase humaine, est effectivement nettement plus élevée :  $\leq 2$  mg d'urokinase humaine/ml de lait de souris transgéniques, soit 1/6 de la teneur en caséine  $\alpha 1$  du lait bovin.

### Conclusion

L'étude structurale et génétique des caséines, puis celle des ARNm et des gènes ont apporté de nombreuses informations sur la structure, les propriétés physico-chimiques, la genèse et l'évolution de ces protéines spécifiques de la glande mammaire. L'étude des allèles associés à des taux de synthèse différents d'une caséine et/ou à une modification du processus d'épissage est une voie d'approche prometteuse pour analyser les régions régulatrices des gènes. L'étude comparative de l'expression de gènes normaux et modifiés est en effet difficile, que ce soit par transgénèse — où le niveau d'expression dépend du site et des modalités d'intégration dans le génome —, ou par transfection dans les lignées disponibles de cellules mammaires, celles-ci ayant perdu une partie de leurs potentialités.

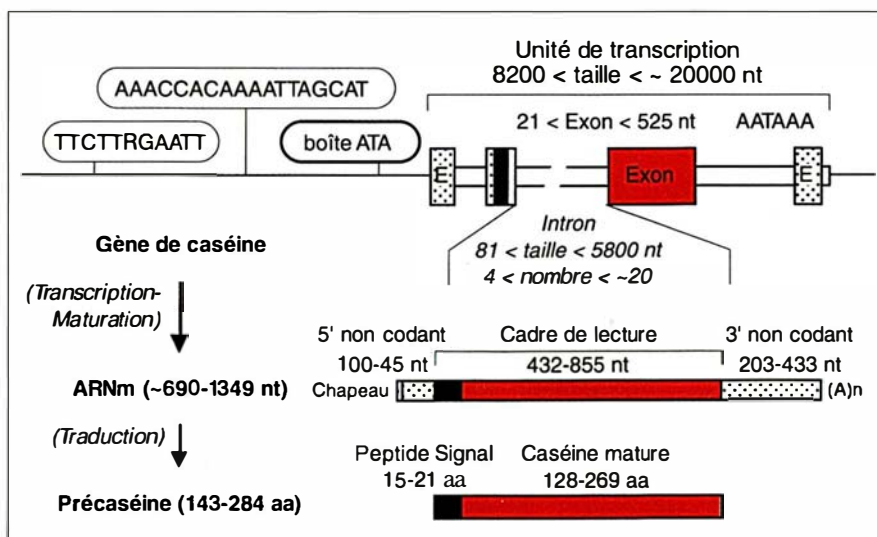


Figure 3. **Représentation schématique d'un gène de caséine, de l'ARNm et de la précaséine.** Les grands rectangles grisés, noirs et rouges symbolisent respectivement les exons ou parties d'exon non traduites et le cadre de lecture du peptide signal et de la protéine mature. Les nombres indiquent les tailles extrêmes de ces molécules et de leurs régions constitutives, dans les 8 espèces de mammifères étudiées. Les valeurs maximales sont celles de la caséine  $\alpha$  (type  $\alpha 1$ ) de rat, et de l'ARNm et du gène correspondants. Les 2 motifs représentés en amont de la boîte ATA, à moins de 90 nucléotides de l'unité de transcription, sont uniquement présents dans les gènes de structure des caséines « sensibles au calcium », mais leur rôle biologique est inconnu. Celui de 18 nucléotides comporte 2 octamères présentant respectivement des analogies avec une séquence amplificatrice du virus simien SV-40 et une séquence reconnue par un facteur nucléaire Oct. nt : nucléotide ; aa : résidu d'acide aminé.

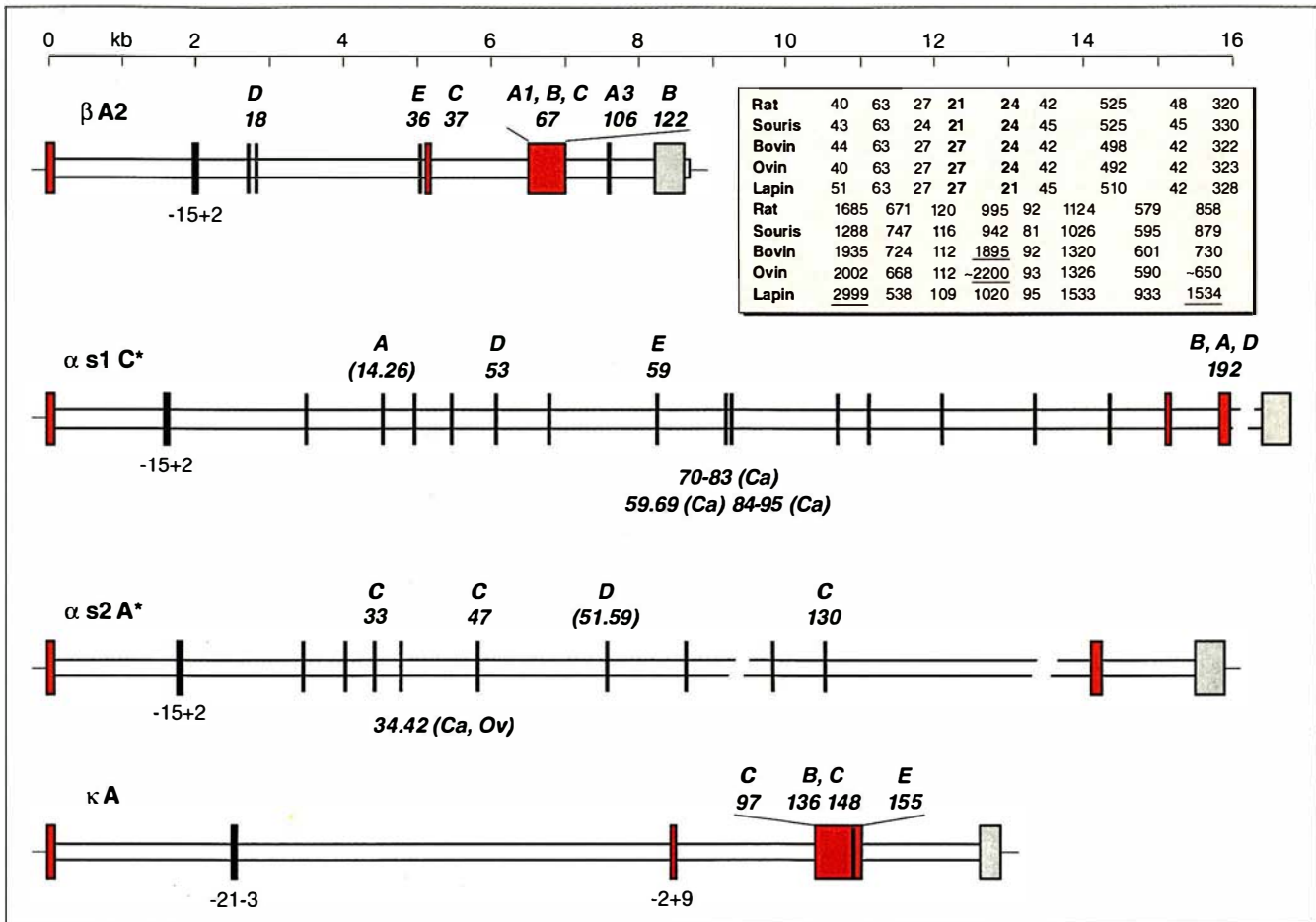


Figure 4. **Structure des gènes codant pour les caséines bovines et localisation des codons mutés (niveau du gène) ou délétés (au cours de l'épissage du pré-ARNm).** Les séquences complète et partielle des gènes codant respectivement pour les caséines  $\beta$  [20] et  $\kappa$  [8] bovines sont connues. La comparaison des gènes de structure de la caséine  $\beta$  de 5 espèces (encadré) montre la conservation de l'organisation générale. Les nombres indiquent les tailles respectives des exons et des introns (en italiques) exprimées en nucléotides. Les introns soulignés sont ceux dont la taille est très supérieure à celle de leurs homologues. Les exons en caractères gras sont ceux dont l'épissage produit un site de phosphorylation multiple. \* L'organisation probable du gène  $Cn-\alpha 1$  bovin est calquée sur celle de son homologue caprin (Leroux et al., résultats non publiés). L'organisation du gène  $Cn-\alpha 2$ , qui comporterait une dizaine d'introns, est hypothétique après l'exon II, car fondée sur des données très fragmentaires obtenues chez le rat et les bovins. Les lettres et les nombres en italiques indiquent respectivement le variant génétique et les codons substitués, ainsi que ceux qui sont délétés au cours de la maturation du pré-ARNm de l'allèle. Les exons éventuellement délétés dans les espèces caprine (Ca) ou ovine (Ov) sont soulignés. Les nombres indiquent la position des codons (- peptide signal ; + protéine mature) délimitant un exon. On notera la similitude d'organisation des unités de transcription des gènes  $Cn-\alpha 1$ ,  $\alpha 2$  et  $\beta$ , au niveau des deux premiers exons.

Ces recherches ont des implications dans les domaines de la technologie laitière et fromagère et de la nutrition, en particulier de l'alimentation infantile. Grâce aux techniques de la biologie moléculaire, il est maintenant possible de caractériser dès la naissance les gènes codant pour les protéines du lait et donc de prendre en compte les allèles ayant un intérêt économique dans les schémas de sélection laitière. Dans l'espèce caprine, par exemple, le choix comme géniteurs d'animaux porteurs des allèles « forts » A, B ou C permettrait de corriger le défaut d'hétérogénéité du lait associé aux allèles défectifs O, E, D et F, et d'augmenter notablement sa valeur nutritionnelle, son rendement fromager et la qualité du caillé. Par ailleurs, les remarquables capacités de synthèse et de sécrétion de la glande mammaire ont été mises à profit pour des essais de production, par des animaux transgéniques, de protéines humaines d'intérêt pharmaceutique telles que le facteur IX de la coagulation sanguine, l'activateur tissulaire du plasminogène, l'urokinase, l' $\alpha$ 1-antitrypsine, l'interleukine 2, etc., et les résultats déjà obtenus sont séduisants. La modification de la composition du lait à des fins nutritionnelles — par exemple, la production d'un lait mieux adapté aux besoins des nourrissons — est également envisageable si l'efficacité et la rentabilité de la transgénèse sont améliorées ■

## Références

- Migliore-Samour D, Jollès P. Casein, a prohormone with an immunomodulating role for the newborn. *Experientia* 1988 ; 44 : 188-93.
- Grosclaude F. Polymorphism of milk proteins : some biochemical and genetical aspects. In : International Society for Animal Blood Group Research, ed. *Proc 16th Int Conf Anim Blood Grps Biochem. Polymorphisms*. Leningrad : International Society for Animal Blood Group Research, 1979 : 54-92.
- Mercier JC. Phosphorylation of caseins. Present evidence for an amino acid tripeptide code posttranslationally recognized by specific kinases. *Biochimie* 1981 ; 63 : 1-17.
- Grosclaude F. Le polymorphisme génétique des principales lactoprotéines bovines. Relations avec la quantité, la composition et les aptitudes fromagères du lait. *INRA Prod Anim* 1988 ; 1 : 5-17.
- Mercier JC, Chobert JM, Addeo F. Comparative study of the amino acid sequences of the caseinomacropptides from seven species. *FEBS Letters* 1976 ; 72 : 208-14.
- Mercier JC, Gaye P. Milk protein synthesis. In : Mephram TB, ed. *Biochemistry of Lactation*. Amsterdam-New York : Elsevier, 1983 : 177-227.
- Rosen JM. Milk protein gene structure and expression. In : Neville MC, Daniel CW, eds. *The Mammary Gland*. New York : Plenum Publishing Corporation, 1987 : 301-22.
- Alexander LJ, Stewart AF, MacKinlay AG, Kapelinskaya TV, Tkach TM, Gorodetsky SI. Isolation and characterization of the bovine  $\kappa$ -casein gene. *Eur J Biochem* 1988 ; 178 : 395-401.
- Ferretti L, Leone P, Rognoni G, Sgarrella S. Linkage of the four bovine casein genes as demonstrated by pulsed field gel electrophoresis. In : Hill WG, Thompson R, Woolliams JA, eds. *Proceedings of the 4th World Congress on Genetics Applied to Livestock Production*. Edinburgh, 1990 ; XIII : 75-8.
- Vilote JL, Soulier S, Stinnakre MG, Massoud M, Mercier JC. Efficient tissue-specific expression of bovine  $\alpha$ -lactalbumin in transgenic mice. *Eur J Biochem* 1989 ; 186 : 43-8.
- Harris S, McClenaghan M, Simons JP, Ali S, Clark AJ. Gene expression in the mammary gland. *J Reprod Fert* 1990 ; 88 : 707-15.
- Lubon H, Hennighausen L. Conserved region of the rat  $\alpha$ -lactalbumin promoter is a target site for protein binding *in vitro*. *Biochem J* 1988 ; 256 : 391-6.
- Bühler ThA, Bruyère Th, Went DF, Stranzinger G, Bürki K. Rabbit  $\beta$ -casein promoter directs secretion of human interleukin-2 into the milk of transgenic rabbits. *Biotechnology* 1990 ; 8 : 140-3.
- Lee KF, Attee SH, Rosen JM. Differential regulation of rat  $\beta$ -casein chloramphenicol acetyl-transferase fusion gene expression in transgenic mice. *Mol Cell Biol* 1989 ; 8 : 3183-90.
- Doppler W, Groner B, Ball RK. Prolactin and glucocorticoid hormones synergistically induce expression of transfected rat  $\beta$ -casein gene promoter constructs in a mammary epithelial cell line. *Proc Natl Acad Sci USA* 1989 ; 86 : 104-8.
- Yoshimura M, Oka T. Transfection of  $\beta$ -casein chimeric gene and hormonal induction of its expression in primary murine mammary epithelial cells. *Proc Natl Acad Sci USA* 1990 ; 87 : 3670-4.
- Meade H, Gates L, Lacy E, Lonberg N. Bovine  $\alpha$ s1-casein gene sequences direct high level expression of active human urokinase in mouse milk. *Biotechnology* 1990 ; 8 : 443-6.
- Rosen JM, Supowit SC, Gupta P, Yu-Lee LY, Hobbs AA. Regulation of casein gene expression in hormone-dependent mammary cancer. In : Banbury Report 8. *Hormones and Breast Cancer*. Cold Spring Harbor Laboratory, 1981 : 397-424.
- Blachier F, Lacroix MC, Ahmed Ali M, Léger C, Ollivier-Bousquet M. Arachidonic acid metabolism and casein secretion in lactating rabbit mammary epithelial cells : effects of inhibitors of prostaglandins and leukotrienes syntheses. *Prostaglandins* 1988 ; 35 : 259-76.
- Bonsing J, Ring JM, Stewart AF, MacKinlay AG. Complete nucleotide sequence of the bovine  $\beta$ -casein gene. *Aust J Biol Sci* 1988 ; 41 : 527-37.

## Summary

### Polymorphism, biosynthesis and evolution of caseins

Caseins, which amount to 80 % of milk proteins, are phosphoproteins associated as micelles. Structural and genetic studies of heterogeneous whole bovine casein showed the occurrence of only 4 types of caseins, controlled by 4 closely linked loci that have evolved rapidly,  $\alpha$ s1-,  $\alpha$ s2- and  $\beta$ -casein genes might have derived from a common ancestor, while  $\kappa$ -casein gene might be related to fibrinogen encoding genes. Deleted caseins, which often occur at a low yield in milk, seem to result from exon skipping. Current applications of new technologies in the field are genotyping of dairy animals for further selection, and the use of transgenic animals for producing milk containing valuable therapeutic proteins.