

Signification fonctionnelle des séquences signal : leur rôle dans le transfert des lactoprotéines à travers les membranes du réticulum endoplasmique

par P. GAYE, J. C. MERCIER *

Laboratoire de Physiologie de la Lactation
* Laboratoire de Biochimie et Technologie Laitières
I.N.R.A. 78350 Jouy-en-Josas, France.

Summary. *The functional significance of signal sequences : their role in the transfer of lactoproteins across the endoplasmic reticulum membrane.*

The mammary gland of the ruminant synthesizes and secretes six major proteins : 4 caseins (α_{s1} , α_{s2} , β and κ), α -lactalbumin and β -lactoglobulin. The information which determines the translocation of these proteins across the endoplasmic reticulum membrane resides in their transient amino terminal sequences, the so-called « signal sequences ». These sequences are cleaved by a specific membrane protease during the transfer of nascent polypeptide chains to the lumen of microsomal vesicles.

The signal sequences of the various lactoproteins share common features (high hydrophobicity and clustered hydrophobic residues) with other secretory proteins but differ significantly in both length and primary structure. The signal sequences of the 3 « calcium sensitive » caseins (α_{s1} , α_{s2} and β) show a high degree of homology, suggesting that they derive from a common ancestor.

The study of lactoprotein signal sequences from various mammalian species revealed that the primary structures of the short-lived amino terminal sequences have been well conserved in the course of evolution.

Introduction.

Parmi les nombreuses protéines synthétisées dans le cytoplasme certaines vont y demeurer pour être transférées dans différents compartiments cellulaires spécifiques (mitochondries, peroxisomes, chloroplastes) ou insérées dans des membranes (protéines transmembranaires), alors que d'autres vont être acheminées à l'extérieur de la cellule. Pour atteindre leur localisation définitive ces protéines devront obligatoirement franchir une ou deux membranes. Ces destinations différentes supposent l'existence de systèmes de transfert et de reconnaissance sélectifs, et l'on pouvait se demander quelle était la nature des signaux qui permettent à une protéine nouvellement synthétisée d'être aiguillée convenablement vers le site cellulaire pour lequel elle est programmée.

La découverte dans les cellules sécrétrices eucaryotes de nombreuses jonctions

ribosomes membranes, et la démonstration que les chaînes polypeptidiques sécrétées en voie de synthèse étaient transférées unidirectionnellement dans la lumière du réticulum endoplasmique, ont permis de démontrer le rôle fonctionnel du réticulum endoplasmique dans le transfert des protéines à travers les membranes (Palade, 1975).

Ce type d'interaction fournissait les conditions topologiques idéales, puisque dans ce cas les sites de synthèse et de transfert sont étroitement liés. Toutefois il n'apportait aucun éclaircissement sur les mécanismes impliqués dans l'attachement sur les membranes des polyribosomes responsables de la synthèse des protéines sécrétées. Dans ce domaine le progrès conceptuel majeur peut être attribué au modèle proposé par Blobel et Sabatini (1971) et développé ultérieurement par Milstein *et al.* (1972) et Blobel et Dobberstein (1975). Selon ce modèle connu sous le nom d'hypothèse du signal, toute protéine sécrétée serait synthétisée sous la forme d'une préprotéine qui contiendrait à son extrémité amino-terminale une séquence d'acides aminés supplémentaire, la séquence signal. Cette région pourrait reconnaître des sites membranaires spécifiques et induire une jonction ribosome membrane, qui permettrait le transfert unidirectionnel de la chaîne polypeptidique naissante dans la lumière du réticulum endoplasmique. Cette séquence signal serait excisée très rapidement, avant même la finition de la chaîne polypeptidique, par une protéase spécifique localisée sur la face interne de la membrane microsomiale.

Le développement de systèmes acellulaires capables de traduire efficacement différents types d'ARNs messagers a permis de démontrer la validité de ce modèle pour une grande variété de protéines sécrétées et plus récemment pour certaines protéines de membranes (Blobel *et al.*, 1979). Ce mécanisme n'est d'ailleurs pas limité aux cellules eucaryotes puisque plusieurs précurseurs ont également été caractérisés chez les procaryotes (Davis et Tai, 1980). Parmi les nombreux systèmes étudiés jusqu'à présent, seule l'ovalbumine, principale protéine sécrétée par l'oviducte de poule, n'est pas synthétisée sous forme de précurseur (Palmiter, Gagnon et Walsh, 1978). Dans ce cas ce serait une séquence signal interne qui déterminerait sa translocation à travers les membranes du réticulum endoplasmique (Lingappa, Lingappa et Blobel, 1979).

Les cellules sécrétrices de la glande mammaire qui synthétisent en grande quantité six protéines majeures : les caséines (α_{s1} , α_{s2} , β et κ), la β -lactoglobuline et l' α -lactalbumine, constituaient un système particulièrement approprié pour étudier les processus de sécrétion. Les ARNs messagers codant pour ces différentes protéines avaient été trouvés essentiellement dans les polyribosomes liés aux membranes du réticulum endoplasmique (Houdebine et Gaye, 1975 ; Gaye et Houdebine, 1975), ce qui suggérait l'existence de précurseurs potentiels pour les lactoprotéines. Toutefois le comportement électrophorétique des caséines natives et des chaînes polypeptidiques homologues synthétisées *in vitro* ne révélait aucune différence de taille apparente en gel de SDS (Gaye et Houdebine, 1975 ; Rosen, 1976). La preuve définitive de l'existence de précurseurs ne pouvait donc être apportée que par la détermination de la séquence amino terminale des lactoprotéines synthétisées *in vitro*. Ce type d'analyse nous a permis de démontrer que les 6 lactoprotéines majeures synthétisées par la glande mammaire de Brebis étaient synthétisées sous forme de précurseurs (Gaye *et al.*, 1977 ; Mercier *et al.*, 1978a, b). Des expériences de traduction *in vitro* réalisées dans des systèmes acellulaires supplémentés avec des membranes microsomiales de glande mam-

maire ont permis de démontrer que les lactoprotéines étaient « ségréguées » dans la lumière des vésicules microsomiales et converties en lactoprotéines authentiques seulement si les membranes sont présentes dès le début de la traduction (Gaye *et al.*, 1979). Synthèse et translocation sont donc des phénomènes étroitement couplés.

Afin de déterminer si le peptide signal d'une protéine donnée a conservé une certaine structure au cours de l'évolution, nous avons récemment abordé l'étude des séquences amino-terminales de pré-lactoprotéines de différentes espèces, et c'est l'ensemble de ces résultats qui seront présentés dans cette communication. Mais auparavant nous rappellerons brièvement les caractéristiques essentielles des séquences signal des six pré-lactoprotéines ovines.

Caractéristiques des séquences signal des six principales lactoprotéines synthétisées par la glande mammaire de brebis.

Les séquences signal des trois caséines sensibles au Ca^{++} (α_{s1} , α_{s2} et β), de la caséine κ , de la β -lactoglobuline et de l' α -lactalbumine, contiennent respectivement 15, 21, 18 et 19 résidus d'acides aminés (fig. 1). A l'exception des pré-peptides des caséines α_{s1} , α_{s2} et β qui présentent une grande homologie, les séquences signal des lactoprotéines diffèrent à la fois dans leur longueur et leur structure primaire. Ces résultats indiquent que les protéines synthétisées par une même cellule ne requièrent pas nécessairement le même signal pour être transférées à travers des membranes du réticulum endoplasmique. Ces différentes séquences signal possèdent cependant les mêmes résidus d'acides aminés en position N et C terminales. La méthionine amino-terminale correspond à la méthionine initiatrice puisqu'elle est insérée dans cette position exclusivement par le tARN met_f (Mercier *et al.*, 1978b ; Gaye *et al.*, 1979). En conséquence ces pré-curseurs représentent bien les produits primaires de traduction de leurs ARNs messagers.

La conversion de ces différents pré-curseurs en lactoprotéines authentiques implique le clivage d'une liaison peptidique contenant un résidu alanyle. Cet acide aminé représente sans doute l'un des signaux reconnus par l'(s) enzyme(s) de clivage, mais étant donné l'existence de liaisons peptidiques similaires dans d'autres régions de ces molécules, il est vraisemblable que c'est une conformation particulière de cette région plus qu'un élément de structure primaire qui est reconnu par les enzymes de clivage.

Bien que ces séquences signal soient présumées contenir les mêmes sites de reconnaissance, il est clair que leur structure diffère notablement. Elles possèdent cependant certaines caractéristiques structurales communes : 1) Une double structure amphipatique comportant un segment très hydrophobe (de 9 à 17 résidus d'acides aminés), précédé par une région polaire chargée positivement ; 2) un fort potentiel à former des structures périodiques en hélice α et à un moindre degré un repliement en structure β étendue (Austen, 1979 ; Garnier *et al.*, 1980). Selon certains auteurs ce type de conformation pourrait permettre l'insertion spontanée de cette séquence hydrophobe dans la bicouche lipidique de la membrane, et 3 modèles ont été récemment proposés (Chan *et al.*, 1979 ; Garnier *et al.*, 1980 ; von Heijne, 1980).

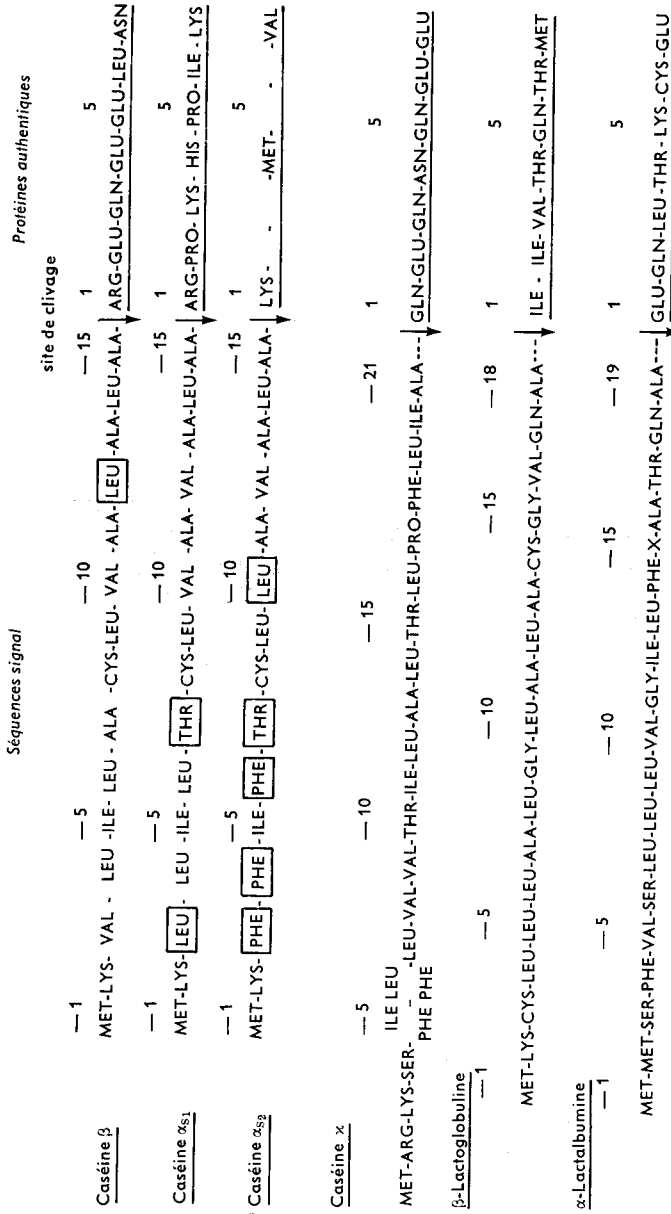


FIG. 1. — Séquences signal des 6 principales lactoprotéines ovines. Dans les séquences signal des 3 caséines sensibles au calcium (α_{S1} , α_{S2} et β) les résidus d'acides aminés qui diffèrent sont encadrés.

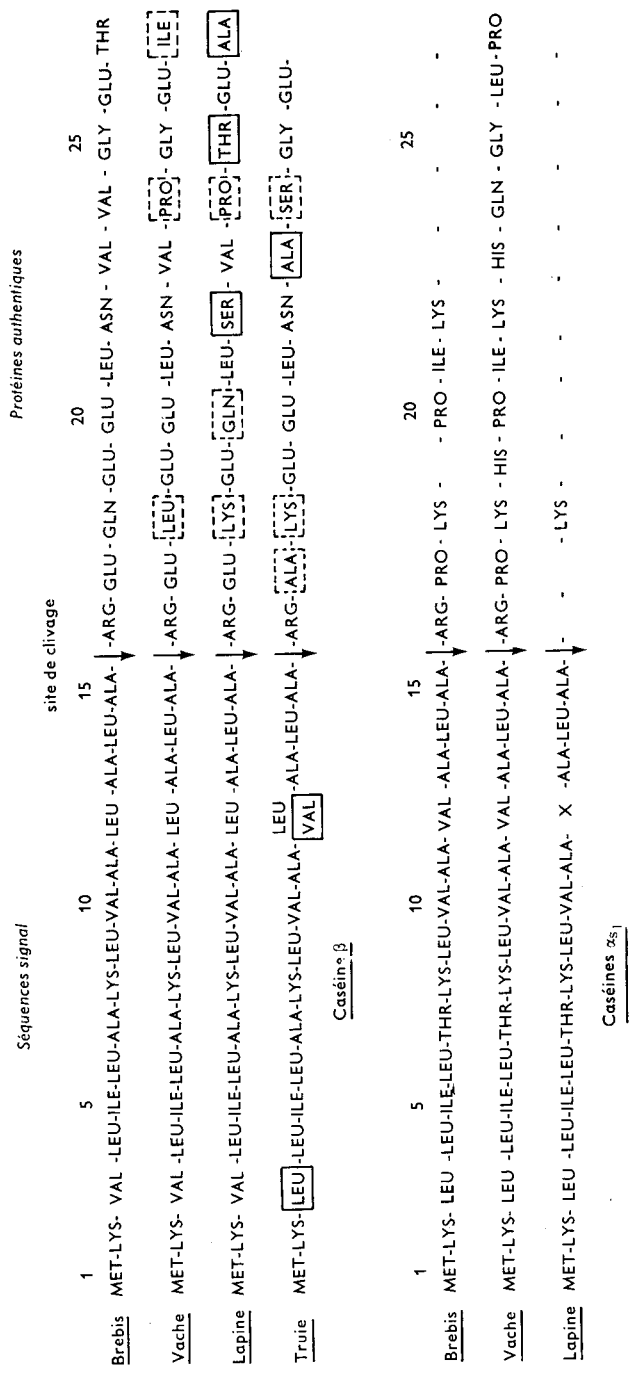


FIG. 2. — Séquences amino-terminales des précurseurs des caséines β et α_{s1} de différentes espèces. Les structures primaires des pré-caséines β et α_{s1} ovine ont été choisies arbitrairement comme référence. Les résidus d'acides aminés encadrés correspondent aux résidus qui ont muté. mutations conservatives ; mutations non conservatives.

Comparaison des structures primaires des séquences signal des lactoprotéines de différentes espèces.

Dans cette étude nous avons considéré les précurseurs de lactoprotéines d'espèces proches sur le plan phylogénique (Brebis et Vache) et ceux d'espèces relativement plus éloignées, tels les précurseurs des lactoprotéines de Porc et de Lapin. Cette dernière étude a nécessité au préalable : 1) L'isolement et la caractérisation des différentes lactoprotéines (caséines, β -lactoglobuline et α -lactalbumine) de ces espèces ; 2) l'obtention d'anticorps spécifiques contre ces protéines ; 3) l'isolement et la traduction d'ARNs messagers codant pour ces différentes molécules.

Caséines. — La comparaison de la séquence amino terminale des précurseurs des caséines β de 4 espèces différentes révèle une distribution irrégulière des mutations (fig. 2). Les séquences signal de la caséine β ovine, bovine et de Lapin présentent une parfaite identité structurale en contraste avec la partie amino terminale des caséines natives. Les deux substitutions qui peuvent être observées dans la séquence signal de la caséine β porcine sont de nature conservatrice et n'affectent probablement pas d'une manière significative les caractéristiques conformationnelles de cette région. En outre l'une de ces mutations correspond à une forme allélique. Une homologie structurale tout aussi remarquable a également été observée lors de l'étude des séquences des peptides signaux des précurseurs des caséines α_{s1} bovine, ovine et de Lapin (fig. 2).

β -Lactoglobulines. — La séquence amino-terminale de la β -lactoglobuline porcine native a considérablement muté par rapport à son homologue ovine puisqu'il ne persiste qu'une homologie de 50 p. 100 (fig. 3). Par contre la séquence signal de cette protéine a subi peu de changement. Parmi les trois mutations observées, deux sont conservatrices et la conformation de cette région de la molécule n'est probablement pas affectée.

α -Lactalbumines. — La comparaison des séquences amino-terminales des précurseurs d' α -lactalbumine de 4 espèces révèle l'existence d'une grande variabilité (fig. 4). Les mutations apparaissent toutefois plus nombreuses dans la région amino-terminale des α -lactalbumines natives que dans le peptide signal. Ainsi les homologies calculées pour les précurseurs des α -lactalbumines ovine et porcine sont respectivement de 85 p. 100 pour les séquences signal et de 60 p. 100 pour les 30 premiers résidus de la région amino-terminale des molécules matures. Dans le cas de la pré α -lactalbumine de Lapin, la séquence signal et la partie amino-terminale de la protéine native ont évolué parallèlement.

Conclusion.

L'information qui détermine la translocation de différentes classes de protéines à travers les membranes du réticulum endoplasmique est contenue dans des séquences spécifiques transitoires, localisées dans la partie amino-terminale de ces molécules à l'exception toutefois de l'ovalbumine (Palmiter, Gagnon et Walsh, 1978). La détermination des structures primaires de séquences signal d'une grande variété de protéines



FIG. 3. — Séquences amino-terminales des précurseurs des β -lactoglobulines ovine et porcine. La structure primaire de la pré- β -lactoglobuline ovine a été choisie arbitrairement comme référence. Les résidus d'acides aminés encadrés correspondent aux résidus qui ont muté. mutations conservatives ; mutations non conservatives.

Séquences signal

Protéines authentiques

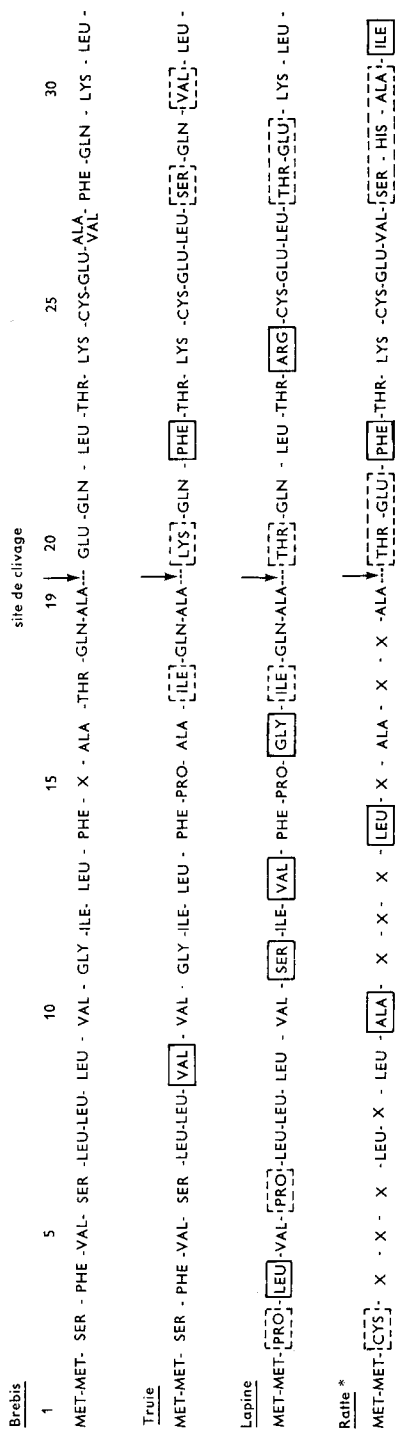


FIG. 4. — Séquences amino-terminales des pré- α -lactalbumines de différentes espèces. La structure primaire de l' α -lactalbumine ovine a été choisie arbitrairement comme référence. Les résidus d'acides aminés encadrés correspondent aux résidus qui ont muté. * : d'après Blobel *et al.* (1979). mutations conservatives ; mutations non conservatives.

de cellules eucaryotes (protéines sécrétées et protéines de membranes) et de quelques protéines d'origine procaryote n'a révélé jusqu'ici qu'une homologie structurale limitée (Blobel *et al.*, 1979). Une grande variabilité dans les structures primaires des séquences signal des protéines appartenant à une même famille (immunoglobulines : Schechter *et al.*, 1979) ou synthétisées par une même cellule (oviducte, glande mammaire) a également été observée. Néanmoins les structures primaires des séquences signal de lactoprotéines homologues d'espèces relativement distantes sur le plan évolutif semblent avoir été remarquablement conservées, puisque parmi les différentes lactoprotéines examinées les homologies sont toujours supérieures à 80 p. 100. De plus les mutations qui affectent cette région sont pour la plupart conservatives et par voie de conséquence elles ne devraient pas altérer de façon notable la conformation de la séquence signal. La comparaison des séquences d'ADN des pré-proinsulines humaine et de Rat indique également l'existence d'une remarquable homologie (> 80 p. 100) dans la structure primaire des séquences signal de ces molécules (Sures *et al.*, 1980). Nos résultats suggèrent que les propriétés de la séquence signal dépendent non seulement de son hydrophobicité mais également d'une conformation particulière, ce qui nécessite une certaine intégrité de la structure primaire de cette région. La séquence signal ne pourrait donc tolérer qu'un nombre restreint d'altérations pour conserver ses propriétés fonctionnelles et des expériences récentes l'ont confirmé. Ainsi, il a été démontré récemment que le remplacement d'un résidu glycyle en position 14 par l'acide aspartique dans la séquence signal d'un mutant de la lipoprotéine de *E. coli* n'affecte pas la translocation de cette protéine, mais abolit complètement le clivage de cette préséquence (Lin *et al.*, 1978). Des altérations qui affectent l'hydrophobicité et probablement la structure secondaire de la séquence signal peuvent interférer irréversiblement avec ses fonctions. La diminution de l'hydrophobicité de la séquence signal de la prolactine et de l'hormone placentaire humaine par le remplacement des résidus leucine par un analogue, la β -hydroxyleucine, réduit considérablement la translocation *in vitro* de ces protéines. Chaque substitution d'un résidu leucine par une β -hydroxyleucine dans la protéine, réduit l'hydrophobicité d'environ 800 cal/mole (Hortin et Boime, 1980).

Les manipulations génétiques ont permis d'obtenir des mutants bactériens dont les fonctions sécrétrices étaient altérées et l'analyse des séquences d'ADN des gènes des protéines exportées a permis de localiser certaines de ces mutations. Ainsi le remplacement d'un seul acide aminé hydrophobe ou apolaire par un acide aminé chargé dans le segment hydrophobe de la séquence signal de la protéine qui lie le maltose chez *E. coli*, inhibe complètement son exportation (Bedouelle *et al.*, 1980). Ces mêmes auteurs ont également pu démontrer qu'une mutation qui altère la structure périodique de la séquence signal de cette protéine était suffisante pour altérer son exportation (Bedouelle et Hofnung, 1980).

Beckwith et ses associés (Bassford et Beckwith, 1979 ; Bassford, Silhavy et Beckwith, 1979) à l'aide d'élégantes techniques de fusion de gènes, ont réussi à obtenir des hybrides de la β -galactosidase (protéine cytoplasmique) qui possèdent dans leur partie amino-terminale soit une séquence signal, soit un fragment plus substantiel d'une protéine sécrétée. L'addition de la séquence signal à la β -galactosidase n'est pas suffisante pour modifier la localisation de cette protéine, par contre les produits

résultant de la fusion de la β -galactosidase avec un fragment plus important de la protéine sécrétée sont répartis à la fois dans le cytoplasme et la membrane.

Ces derniers résultats sont particulièrement intéressants puisqu'ils suggèrent que dans une protéine sécrétée la séquence signal n'est probablement pas le seul domaine spécifique nécessaire à son transfert à travers une membrane.

6^e Réunion du groupe Développement, I.N.R.A.,
Clermont-Ferrand/Theix, 22-23 mai 1980.

Remerciements. — Nous remercions tout particulièrement Mlle Dominique Hue et M. Gérard Hazé pour leur excellente collaboration technique.

Références

- AUSTEN B. M., 1979. Predicted secondary structures of amino-terminal extension sequences of secreted proteins. *FEBS Letters*, **103**, 308-313.
- BASSFORD P. J., BECKWITH J., 1979. *Escherichia Coli* mutants accumulating the precursor of a secreted protein in the cytoplasm. *Nature*, **277**, 538-541.
- BASSFORD P. J., SILHAVY T. J., BECKWITH J. R., 1979. Use of gene fusion to study secretion of maltose binding protein into *Escherichia Coli* periplasm. *J. Bact.*, **139**, 19-31.
- BEDOUELLE H., BASSFORD P. J., FOWLER A. V., ZABIN I., BECKWITH J. R., HOFNUNG M., 1980. Mutations which alter the function of the signal sequence of the maltose binding protein of *Escherichia Coli*. *Nature*, **285**, 78-81.
- BEDOUELLE H., HOFNUNG M., 1980. Communication personnelle.
- BLOBEL G., SABATINI D. D., 1971. Ribosome-membrane interaction in eukaryotic cells. In L. A. MANSON, *Biomembranes*, 193-195, **2**, Plenum Press N. Y., London.
- BLOBEL G., DOBBERSTEIN B., 1975. Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. II. Reconstitution of functional rough microsomes from heterologous components. *J. Cell Biol.*, **67**, 835-851 ; 852-862.
- BLOBEL G., WALTER P., CHANG C. N., GOLDMAN B. M., ERICKSON A. H., LINGAPPA V. R., 1979. Translocation of proteins across membranes : the signal hypothesis and beyond. In C. P. HOPKINS, C. DUNCAN, *Symp. Soc. exp. Biol. (G.-B.)*, **33**, 9-38, Cambridge Univ. Press.
- CHAN S. J., PATZELT C., DUGUID J. R., QUINN P., LABRECQUE A., NOYES B., EIM P., HEINRIKSON R. L., STEINER D. F., 1979. Precursors in the biosynthesis of insulin and other peptide hormones, from gene to information transfer in normal and abnormal cells. In T. R. RUSSEL, K. BREW, H. FARBER, J. SCHULTZ, *Miami Winter Symp.*, **16**, 361-378, Acad. Press.
- DAVIS B. D., TAI P. C., 1980. The mechanism of protein secretion across membranes. *Nature*, **283**, 433-438.
- GARNIER J., GAYE P., MERCIER J. C., ROBSON B., 1980. Structural properties of signal peptides and their membrane insertion. *Biochimie*, **62**, 231-239.
- GAYE P., HOUEBINE L. M., 1975. Isolation and characterization of casein mRNAs from lactating ewe mammary glands. *Nucl. Ac. Res.*, **2**, 707-722.
- GAYE P., GAUTRON J. P., MERCIER J. C., HAZE G., 1977. Amino terminal sequences of the precursors of ovine caseins. *Biochem. biophys. Res. Commun.*, **79**, 903-911.
- GAYE P., HUE D., MERCIER J. C., HAZE G., 1979. Enzymatic processing of precursors of ovine lactoproteins by mammary microsomal membranes and a deoxycholate soluble extract from rough microsomes. *FEBS Letters*, **101**, 137-142.
- HEIJNE Von G., 1980. Trans-membrane translocation of proteins. *Eur. J. Biochem.*, **103**, 431-438.
- HORTIN G., BOIME I., 1980. Inhibition of preprotein processing in ascites tumor lysates by incorporation of a leucin analog. *Proc. nat. Acad. Sci. USA*, **77**, 1356-1360.

- HOUEBINE L. M., GAYE P., 1975. Absence of mRNA for casein in free polysomes of lactating ewe mammary gland. *Nucl. Ac. Res.*, **2**, 165-177.
- LIN J. J. C., KANAZAWA H., OZOLS J., WU H. C., 1978. An *Escherichia Coli* mutant with an amino acid alteration within the signal sequence of outer membrane prolipoprotein. *Proc. nat. Acad. Sci. USA*, **75**, 4891-4895.
- LINGAPPA V. R., LINGAPPA J. R., BLOBEL G., 1979. Chicken ovalbumin contains an internal signal sequence. *Nature*, **281**, 117-121.
- MERCIER J. C., HAZE G., GAYE P., HUE D., 1978a. Amino terminal sequence of the precursor of ovine β -lactoglobulin. *Biochem. biophys. Res. Commun.*, **82**, 1236-1245.
- MERCIER J. C., HAZE G., GAYE P., PETRISSANT G., HUE D., BOISNARD M., 1978b. Amino terminal sequence of the precursor of ovine α -lactalbumin. *Biochem. biophys. Res. Commun.*, **85**, 662-670.
- MILSTEIN C., BROWNLEE G. C., HARRISON T. M., MATHEWS M. B., 1972. A possible precursor of immunoglobulin light chains. *Nature New Biol.*, **239**, 117-120.
- PALADE G., 1975. Intracellular aspects of the process of protein synthesis. *Science*, **189**, 347-358.
- PALMITER R. D., GAGNON J., WALSH K. A., 1978. Ovalbumin : a secreted protein without transient hydrophobic leader sequence. *Proc. nat. Acad. Sci. USA*, **75**, 94-98.
- ROSEN J. M., 1976. Isolation and characterization of purified rat casein messenger ribonucleic acids. *Biochemistry*, **15**, 5263-5271.
- SCHECHTER I., WOLF O., ZEMELL R., BURSTEIN Y., 1979. Structure and function of immunoglobulin genes and precursors. *Fed. Proc.*, **38**, 1839-1845.
- SURES I., GOEDDEL D. V., GRAY A., ULRICH A., 1980. Nucleotide sequence of human preproinsulin complementary DNA. *Science*, **208**, 57-59.
-