

## Les mucines humaines : pourquoi une telle hétérogénéité peptidique ?

Nicole Porchet  
José Dufossé  
Pierre Degand  
Jean-Pierre Aubert

Les mucines sont des glycoprotéines, très riches en sucre, jouant un rôle de protection des muqueuses contre les agressions de nature variée. L'importance de la copule glucidique et la nature chimique de l'axe peptidique ont constitué des obstacles majeurs à l'étude structurale de ces molécules dont la partie protéique peut maintenant être analysée grâce aux techniques de l'ADN recombinant. Il semble exister un nombre encore indéterminé de gènes pouvant être constitués d'un très grand nombre de mini-exons épissés en une multitude d'ARN messagers, engendrant donc une très grande diversité des précurseurs protéiques dont la synthèse est spécifique ou non du tissu. La signification de cet extrême polymorphisme reste incertaine. Peut-être correspond-il à la nécessité pour l'organisme de produire une énorme diversité de mucines permettant de faire face, en différents sites, à la multiplicité des agressions possibles.

### ADRESSES

N. Porchet : maître de conférences, praticien hospitalier. J. Dufossé : stagiaire de recherche. Laboratoire de biochimie, hôpital Cl.-Huriez, CHRU Lille, France.

P. Degand : professeur de biochimie, faculté de médecine Lille, directeur de l'U. 16 de l'Inserm. J.-P. Aubert : directeur de recherche à l'Inserm. Inserm U. 16, place de Verdun, 59045 Lille Cedex, France.

**L**e mucus est une sécrétion douée de propriétés rhéologiques (viscoélasticité, filance), élaborée au niveau des muqueuses des tractus respiratoire, digestif et génital par des cellules épithéliales spécialisées. Il assume de nombreuses fonctions dont la finalité est la protection de l'épithélium de surface, à l'égard de multiples agressions potentielles :

- de nature chimique sous la forme des pH extrêmes des sucs digestifs qui, en outre, renferment des activités enzymatiques puissantes et variées

- et contre lesquels les mucines exercent un pouvoir tampon ;
- de nature mécanique, par exemple lors de la progression du contenu luminal digestif ;
- de nature bactérienne, les muqueuses digestives et respiratoires étant protégées de la colonisation bactérienne grâce au « piège » muqueux ;
- de nature physique, enfin, comme la déshydratation ou la toxicité des particules et gaz inhalés.

Les mucines sont les constituants biochimiques majeurs de ce mucus et se définissent comme des populations de

macromolécules complexes. Ce sont des O-glycoprotéines renfermant jusqu'à 80 % en poids de chaînes glycaniques très diversifiées en longueur et en structure oligosaccharidique [1, 2]. Ces glycoprotéines particulières confèrent au mucus ses propriétés physico-chimiques et biologiques spécifiques.

Les mucines coexistent au sein du mucus sous la forme de différentes populations que la chromatographie d'échange ionique permet grossièrement de subdiviser en trois catégories : mucines neutres, mucines acides, mucines très acides. Chaque muqueuse élabore une sécrétion où se définit un équilibre qualitatif et quantitatif de ces trois catégories de mucines. Dans de nombreuses pathologies (bronchite chronique, maladies inflammatoires du tube digestif, ulcère gastrique...), cet équilibre sera rompu et s'accompagnera d'altérations de la muqueuse.

Les mucines sont des molécules très interactives et cette propriété d'interagir avec l'environnement chimique et cellulaire est jusqu'ici essentiellement attribuée à la fraction glycanique, qui peut être considérée comme une résine biologique, un piège moléculaire à l'égard de divers constituants protéiques ou lipidiques, un support sérologique, une barrière bactériostatique [3].

Or le rôle joué par le support peptidique est encore jusqu'ici inconnu. La connaissance structurale même de cet axe peptidique est restée complètement dans l'ombre jusqu'aux premières publications en 1989 de séquences partielles d'ADNc. L'analyse des résultats, même partiels, de séquence peptidique devrait susciter l'élaboration de modèles peptidiques de synthèse et l'obtention d'anticorps de nature à faire progresser l'étude de la biosynthèse de ces molécules. Parallèlement, d'autres molécules rattachées au groupe des mucines sur des critères physico-chimiques et dénommées glycoprotéines *mucin-like* ont suscité également des développements de recherche, et cela dans le cadre de la définition de nouveaux marqueurs biologiques permettant de détecter certains cancers mucosécrétants et de porter un jugement pronostique.

### **Connaissances biochimiques actuelles concernant les mucines**

La fraction peptidique des mucines humaines est minoritaire puisqu'elle ne rend compte que de 10 à 20 % de leur masse moléculaire.

Les connaissances structurales acquises ces dix dernières années dans le domaine des mucines concernent essentiellement les chaînes oligosaccharidiques. Plus d'une centaine de structures glycaniques distinctes ont été relatées. L'extrême hétérogénéité des mucines a donc jusqu'ici été attribuée à la copule glycanique. L'étude structurale de l'axe peptidique des mucines est très récente et a bénéficié depuis 1989 de l'apport des techniques de génétique moléculaire. Les premiers résultats tendent à démontrer un deuxième niveau d'hétérogénéité, cette fois-ci peptidique.

#### **• Les mucines « typiques »**

Les méthodes d'étude des protéines — qui consistent classiquement en l'obtention de peptides grâce à des coupures enzymatiques choisies, puis en la détermination de leur séquence par la dégradation d'Edman — se sont révélées inopérantes dans l'étude des mucines.

Jusqu'en 1989, n'ont pu être obtenues que de très courtes séquences, toujours inférieures à une trentaine d'acides aminés [4].

Cela est dû à la présence de très nombreuses chaînes glycaniques mais aussi à la caractéristique de cet axe peptidique que de renfermer 30 à 40 % d'acides aminés hydroxylés, sérine et thréonine engendrant des dérivés labiles lors du séquençage. L'axe peptidique offre certains sites particuliers de sensibilité aux enzymes protéolytiques. Cette observation est à la base de l'hypothèse généralement admise dans la littérature selon laquelle la partie polypeptidique des mucines serait constituée d'une alternance de régions très fortement glycosylées hydrophiles couvertes de centaines de chaînes glycaniques très hétérogènes, et de régions dites « nues », moins glycosylées, plus hydrophobes et plus sensibles à l'attaque de certaines protéases.

Différents modèles d'organisation structurale des mucines ont été pro-

posés et cela pour rendre compte également d'une masse moléculaire élevée (plusieurs millions de daltons) réductible après action d'agents tels que le 2-mercaptoéthanol ou le dithiothréitol.

Deux modèles principaux, associant à cette structure alternée hydrophile-hydrophobe des pontages — soit intermoléculaires, soit intramoléculaires par des ponts disulfures —, ont été proposés respectivement par Allen [5] et par Carlstedt [6].

Le modèle d'Allen dit en « moulin à vent » propose, pour les mucines gastriques de porc, quatre sous-unités réunies à une protéine de liaison grâce à des ponts disulfures.

Le modèle de Carlstedt propose une structure linéaire où alternent des domaines « nus » et glycosylés enchaînés les uns aux autres par des ponts disulfures. Ce modèle est en accord avec les données de la microscopie électronique, qui montre de longs filaments de mucines cervicales, gastriques ou bronchiques, de longueur variant entre 200 et 1 000 nm [7-10].

Cependant, qu'il s'agisse de la détermination de la masse moléculaire, ou de la présence éventuelle de ponts disulfures participant à l'organisation moléculaire des mucines, des résultats contradictoires émaillent la littérature. Ils peuvent être en relation avec une origine tissulaire variée des molécules étudiées (modèles structuraux différents pour les muqueuses gastrique et respiratoire ?), mais aussi avec la grande diversité des protocoles de préparation des mucines.

La grande polydispersion de longueur des filaments de mucines observée en microscopie électronique ne résulte pas d'une dégradation protéolytique. L'utilisation d'anticorps polyclonaux, dirigés spécifiquement contre l'axe peptidique de mucines trachéobronchiques humaines déglycosylées, a permis d'isoler et de précipiter les précurseurs polypeptidiques des mucines synthétisés *in vitro* dans des cultures d'explants de trachée [11].

Et là encore, une grande hétérogénéité des précurseurs peptidiques a été démontrée en électrophorèse dénaturante (gel de polyacrylamide en présence de sodium dodecyl sulfate) sous la forme d'un continuum de taille variant de 200 à 400 kDa.

## RÉFÉRENCES

1. Podolsky DK. Oligosaccharide structures of human colonic mucin. *J Biol Chem* 1985 ; 260 : 8262-71.
  2. Roussel P, Lamblin G, Lhermitte M, et al. The complexity of mucins. *Biochimie* 1988 ; 70 : 1471-82.
  3. Ramphal R, Houdret N, Koo L, Lamblin G, Roussel P. Differences in adhesion of *Pseudomonas aeruginosa* to mucin glycopeptides from sputa of patients with cystic fibrosis and chronic bronchitis. *Infect Immun* 1989 ; 57 : 3066-71.
  4. Rose MC, Kaufman B, Martin BM. Proteolytic fragmentation and peptide mapping of human carboxyamidomethylated tracheo-bronchial mucin. *J Biol Chem* 1989 ; 264 : 8193-9.
  5. Allen A, Bell A, Mantle M, Pearson JP. The structure and physiology of gastrointestinal mucus. In : Chantler EH, Elder JB, Elstein M, eds. *Mucus in Health and Disease, Part II, Advances in Experimental Medicine and Biology*. New York : Plenum Press, 1982 ; 144 : 115-33.
  6. Carlstedt I, Sheehan JK, Corfield AP, Gallagher JT. Mucus glycoproteins : a gel of a problem. *Essays in Biochemistry* 1985 ; 20 : 40-76.
  7. Slayter HS, Lamblin G, Le Treut A, et al. Complex structure of human bronchial mucus glycoprotein. *Eur J Biochem* 1984 ; 142 : 209-18.
  8. Rose MC, Voter WA, Brown CF, Hauffman B. Structural features of human tracheo-bronchial mucin glycoproteins. *Biochem J* 1984 ; 222 : 371-7.
  9. Mikkelsen A, Stokke BT, Christensen BE, Ekgsaeter A. Flexibility and length of human bronchial mucin studied using low-shear viscometry, birefringence, relaxation analysis and electron microscopy. *Biopolymers* 1985 ; 24 : 1683-704.
  10. Sheehan JK, Oates K, Carlstedt I. Electron microscopy of cervical, gastric and bronchial mucus glycoproteins. *Biochem J* 1986 ; 239 : 147-53.
  11. Marianne T, Périni JM, Lafitte JJ, et al. Peptides of human bronchial mucus glycoproteins size determination by electron microscopy and by biosynthesis experiments. *Biochem J* 1987 ; 248 : 189-95.
  12. Shimizu M, Yamauchi K. Isolation and characterization of mucin-like glycoprotein in human milk fat globule membrane. *J Biochem* 1982 ; 91 : 515-24.
  13. Swallow DM, Griffiths B, Brahwel HE, Burchell J. Detection of the urinary « P<sub>um</sub> » polymorphism by the tumor binding monoclonal antibodies Ca1, Ca23, HMFG-1 and HMFG-2. *Dis Markers* 1986 ; 4 : 247-54.
- A côté donc de l'hétérogénéité glycanique sur laquelle est fondée la classification en mucines « neutres », « sialylées », « sulfatées », il existe donc également une très large hétérogénéité peptidique.
- **Les mucin-like**  
Les glycoprotéines *mucin-like* ont été décrites initialement dans le lait humain et l'urine [12, 13]. Elles sont élaborées par de nombreux épithéliums sécrétoires (mammaire, ovarien, prostatique, pancréatique, colorectal) et ont fait l'objet d'études immunologiques grâce à l'obtention de sérums immuns polyclonaux ou monoclonaux. Contrairement aux mucines « typiques », elles ne sont pas sécrétées par des cellules épithéliales spécialisées. Il s'agirait plutôt de « mucines membranaires ». Des molécules apparentées immunologiquement, sinon identiques, sont exprimées à la fois par les cellules normales et les cellules tumorales de ces épithéliums. Néanmoins, des modifications qualitatives ou quantitatives de leur biosynthèse dans les cellules tumorales peuvent expliquer des concentrations très élevées de ces glycoprotéines dans le sérum et les liquides d'épanchement. A ce titre, ces glycoprotéines sont actuellement très étudiées comme marqueurs tumoraux. Cependant, l'utilisation en diagnostic médical des anticorps monoclonaux anti-glycoprotéines *mucin-like* a largement anticipé la connaissance fondamentale des antigènes eux-mêmes. Leur masse moléculaire relative est généralement comprise entre 300 et 400 000 daltons. La composition chimique de l'une de ces molécules, la PAS-O glycoprotéine [12] montre environ 50 % en poids de chaînes glycaniques et, sur l'axe peptidique, une teneur élevée en cinq acides aminés : sérine, thréonine, proline, glycine, alanine, qui représentent, comme dans les mucines, 70 % du total des acides aminés. Des tentatives de caractérisation des épitopes (glucidiques et/ou peptidiques) ont été réalisées par des techniques d'inhibition de la fixation des anticorps monoclonaux, après action de la neuraminidase ou de lectines, ou après déglycosylation chimique de ces molécules [14, 15]. Des variations
- quantitatives et qualitatives des épitopes entre *mucin-like* de différentes origines tumorales ont été ainsi décrites, concernant surtout le processus de glycosylation. A titre d'exemple, citons l'épisialine, la plus étudiée des glycoprotéines *mucin-like*. Cette glycoprotéine de tumeur mammaire comporte 50 % en poids de sucre. La séquence de son axe peptidique a été complètement déterminée à partir de celle de l'ADNc. La même structure peptidique (plus de 99 % d'homologie de séquence nucléotidique) est retrouvée dans une tumeur pancréatique, mais cette fois-ci associée à une copule glycanique beaucoup plus volumineuse, 80 % en poids de la molécule [16]. Des études portant sur la biosynthèse ont pu montrer, comme pour les mucines, que la O-glycosylation était précédée d'une N-glycosylation précoce [17].

### Stratégies d'étude de l'axe peptidique des mucines ou des mucin-like par la biologie moléculaire

Des difficultés technologiques majeures ont constitué des obstacles insurmontables lors de la détermination de la séquence peptidique des mucines ou des *mucin-like* :

- accessibilité limitée des enzymes à leur cible peptidique en raison de la densité des chaînes glycaniques ;
- déglycosylation enzymatique trop partielle ;
- déglycosylation chimique efficace mais agressive pour le peptide (coupure des résidus méthionine, hydrolyse acide) ;
- fractionnement des peptides impossible en raison des très fortes homologies existant entre eux ;
- et, surtout, grande labilité des dérivés phénylthiohydantoïnes engendrés par la sérine et la thréonine lors de la dégradation récurrente d'Edman, qui rend impossible l'interprétation des résultats au-delà d'une vingtaine de cycles. Quelques laboratoires se sont donc tournés vers les techniques de l'ADN recombinant.

Schématiquement, toutes les stratégies employées comportaient les étapes suivantes :

- préparation d'un anticorps poly-

<b>Episialine</b>	
Gendler et al. 1989	Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Val-Thr-Ser-Ala
<b>Mucines intestinales</b>	
Gum et al. 1989	Pro-Thr-Thr-Thr-Pro-Ile-Thr-Thr-Thr-Thr-Val-Thr-Pro-Thr-Pro-Thr-Gly-Thr-Gln-Thr
Gum et al. 1990	His-Ser-Thr-Pro-Ser-Phe-Thr-Ser-Ser-Ile-Thr-Thr-Thr-Glu-Thr-Thr-Ser
<b>Mucines trachéobronchiques</b>	
Aubert et al. 1991	Thr-Thr-Ser-Thr-Thr-Ser-Ala-Pro
Porchet et al. 1991	Thr-Ser-Ser-Ala-Ser-Thr-Gly-His-Ala-Thr-Pro-Leu-Pro-Val-Thr-Asp

Figure 1. **Domaines répétitifs en acides aminés de différentes mucines humaines.**

clonal ou monoclonal dirigé contre l'axe peptidique, déglycosylé chimiquement par l'acide fluorhydrique ou l'acide trifluorométhane sulfonique, d'une mucine ou d'une *mucin-like* préparée à partir d'un tissu sain ou tumoral ;

- confirmation de la qualité des anticorps par immuno-transfert ou par immunohistochimie ;
- préparation d'une banque d'ADNc construite dans un vecteur d'expression  $\lambda$ gt 11 à partir de la muqueuse ou de la lignée cellulaire étudiée ;
- criblage immunologique de ces banques et extraction-purification des clones positifs ;
- séquençage et localisation chromosomique.

Cette stratégie a déjà permis l'obtention de séquences partielles de mucines intestinales et de mucines trachéobronchiques humaines. Dans le cas des *mucin-like*, une seule séquence complète, celle de l'épisialine, a été publiée et confirmée par différents groupes.

### **Mise en évidence de l'hétérogénéité peptidique des mucines humaines**

#### **• Les mucines intestinales**

En 1989 et 1990, les travaux de Gum

*et al.* [19-21] ont montré l'implication de deux gènes dans la biosynthèse des mucines intestinales, l'un situé sur le chromosome 11 en p 15 (*MUC2*), l'autre sur le chromosome 7 (*MUC3*). Le premier gène, en 11p15, a pu être identifié grâce à trois clones d'ADNc (SMUC 40-41-42) isolés d'une banque de jéjunum humain grâce à un sérum immun dirigé contre l'axe peptidique des mucines d'une lignée tumorale colique LS 174 T. Ces trois clones possèdent le même agencement répétitif d'un domaine élémentaire de 69 nucléotides, soit 23 acides aminés (*figure 1*).

Ces « tandems » répétés 14 fois dans le plus long clone ne sont pas strictement identiques mais possèdent une analogie de 90 %. Néanmoins, certains triplets sont strictement conservés dans tous les domaines répétitifs. La thréonine est fortement représentée (14 acides aminés sur 23) de même que la proline. Pour rendre compte de la teneur en sérine de la mucine intestinale, il faut supposer que la molécule comporte d'autres régions, répétitives ou non, riches en sérine. Deux sites potentiels de N-glycosylation ont été relevés.

Le deuxième gène localisé sur le chromosome 7 a été identifié grâce à deux clones (SIB 124-139) isolés de la même banque d'ADNc à l'aide d'un

autre sérum immun dirigé contre l'axe peptidique de mucines intestinales humaines. Un agencement en domaines répétitifs est à nouveau observé, le domaine élémentaire étant toutefois différent du précédent et comportant 51 nucléotides soit 17 acides aminés (*figure 1*).

L'analyse par *Northern blot* de l'ensemble de ces sondes intestinales, qu'elles proviennent de l'un ou l'autre gène, montre une très large hétérogénéité des ARN messagers codant pour ces mucines.

#### **• Les mucines trachéobronchiques.**

Les études menées par notre groupe de 1989 à 1991 ont permis de localiser trois gènes de mucines trachéobronchiques respectivement sur les chromosomes 11 (en p15), 13 et 3 [21-24].

Vingt clones ont été extraits d'une banque d'ADNc de muqueuse trachéobronchique humaine grâce à un sérum immun polyclonal dirigé contre un mélange « équimoléculaire » de glycopeptides de mucines neutres, sialylées, sulfatées, tous déglycosylés par l'acide trifluorométhane sulfonique.

La séquence de ces vingt clones a permis de déduire trois types d'organisation peptidique :

## RÉFÉRENCES

14. Stracker SA, Tjandra JJ, Walker DI, McKenzie FC. Purification and biochemical characterization of a novel breast carcinoma associated mucin-like glycoproteins defined by antibody 3 E1.2. *Br J Cancer* 1989 ; 59 : 544-53.
15. Burchell J, Gendler S, Taylor-Papadimitriou J, Lamport DTA. Development and characterization of breast cancer reactive monoclonal antibodies directed to the core protein of the human milk mucin. *Cancer Res* 1987 ; 47 : 5476-82.
16. Lan MS, Batra SK, Qi WN, Metzgar RS, Hollingworth MA. Cloning and sequencing of a human pancreatic tumor mucin cDNA. *J Biol Chem* 1990 ; 265 : 15294-9.
17. Hilkens J, Buijs F. Biosynthesis of MAM-6 an epithelial sialomucin. *J Biol Chem* 1988 ; 263 : 4215-22.
18. Gum JR, Byrd JC, Hicks JW, Toribara NW, Lamport DTA, Kim YS. Molecular cloning of human intestinal mucin cDNAs. *J Biol Chem* 1990 ; 264 : 6480-7.
19. Gum JR, Hicks JW, Swallow DM, et al. Molecular cloning of cDNAs derived from a novel human intestinal mucin gene. *Biochem Biophys Res Comm* 1990 ; 171 : 407-15.
20. Griffiths B, Gum JR, West LF, Povey S, Swallow DL, Kim YS. Mapping of the gene coding for intestinal mucin to chromosome 11p15. *Cytogenet Cell Genet* 1989 ; 51 : 1008.
21. Aubert JP, Porchet N, Crépin M, et al. Evidence for different human tracheo-bronchial mucin peptides deduced from nucleotide cDNA sequences. *Am J Resp Cell Mol Biol* 1991 ; 5 : 178-85.
22. Crépin M, Porchet N, Aubert JP, Degand P. Diversity of the peptide moiety of human airway mucins. *Biorheology* 1990 ; 27 : 471-84.
23. Nguyen Van Cong J, Aubert JP, Gross MS, Porchet N, Degand P, Frezal J. Assignment of human tracheo-bronchial mucin gene(s) to 11p15 and a tracheo-bronchial mucin-related sequence to chromosome 13. *Hum Genet* 1990 ; 86 : 167-72.
24. Porchet N, Nguyen Van Cong J, Dufossé J, et al. Molecular cloning and chromosomal localization of a novel human tracheo-bronchial mucin cDNA containing tandemly repeated sequences of 48 base pairs. *Biochem Biophys Res Comm* 1991 ; 175 : 414-22.
25. Bhargava AK, Woitach JT, Davidson EA, Bhavanandan VP. Cloning and cDNA sequence of a bovine submaxillary gland mucin-like protein containing two distinct domains. *Proc Natl Acad Sci USA* 1990 ; 87 : 6798-802.
- La famille I, constituée de séquences répétées en tandem :  
- Le domaine de 8 acides aminés TTSTTSAP\* est le plus fréquemment rencontré. Dans certains cas, sont intercalées des zones non répétitives, non typiques d'une séquence de mucine. Ce motif est localisé sur le chromosome 11 en p15 (figure 1) [21-23].  
- Le domaine de 16 acides aminés (TSSASTGHATPLPVTD\*) est localisé sur le chromosome 3 (MUC 4) (figure 1). 39 répétitions de ce domaine constitué de 48 nucléotides ont été relevées. Les domaines ne sont pas rigoureusement analogues, néanmoins 7 positions sont parfaitement conservées d'un domaine à l'autre [24].
- La famille II, localisée à la fois sur les chromosomes 11 et 13, possède un agencement moins structuré, possédant cependant des zones hydrophiles à 100 % de résidus hydroxylés, dont 75 % de thréonine [23].
- La famille III, localisée sur le chromosome 11 en p15, démontre une organisation alternée en domaines hydrophiles glycosylables et en zones nues hydrophobes. L'alignement et la comparaison des séquences de cette famille laisse suspecter :  
- que les exons de ces gènes sont petits ;  
- que les ARNm puissent résulter d'un système d'épissage très complexe.
- De l'ensemble de nos résultats, il ressort qu'il existe une grande hétérogénéité structurale des axes peptidiques de mucines trachéobronchiques humaines. Quelques séquences potentielles de N-glycosylation ont été relevées. Cinq acides aminés C-terminaux distincts ont été mis en évidence, mais aucun acide aminé N-terminal n'a pu être caractérisé comme dans le cas de la glande sous-maxillaire de bœuf [25].
- L'analyse de l'expression des mucines, par Northern blot, démontre une extrême hétérogénéité en nombre et en taille des transcrits (continuum de 20 kb à 0,4 kb) [22-24]. Elle démontre également une spécificité tissulaire variable : les sondes de la famille III reconnaissent des motifs ubiquitaires communs à l'ensemble des mucines étudiées (muqueuses trachéo-bronchique, colique, gastrique, lignées tumorales mammaires MCF7 et VHB1). Les sondes de la famille I répétitive se caractérisent, au contraire, par une spécificité tissulaire d'expression :  
- spécificité strictement limitée à la muqueuse trachéo-bronchique pour le motif de 8 acides aminés ;  
- spécificité limitée à la muqueuse trachéo-bronchique et à la muqueuse colique pour le motif de 16 acides aminés.
- La purification des différents anticorps spécifiques à partir de l'antisérum polyclonal fut réalisée sur les protéines de fusion synthétisées par chacun des 20 clones d'*E. coli*/λgt 11 recombinants. Ces anticorps monospécifiques permirent alors, sur coupe histologique de muqueuse trachéo-bronchique, d'observer une spécificité cellulaire de biosynthèse des mucines. Les cellules caliciformes sont reconnues par tous les anticorps et semblent donc avoir un potentiel de synthèse large alors que les cellules des glandes muqueuses, qui ne sont reconnues que par quelques-uns des anticorps, auraient une capacité sécrétoire plus spécifique.
- Pour conclure sur la connaissance de ces premières séquences d'ADNc de mucines humaines, la caractéristique de ces glycoprotéines est donc d'être structurée en domaines répétitifs et cela quel que soit le tissu considéré. Néanmoins, d'autres organisations structurales coexistent, comme l'alternance de domaines hydrophiles/hydrophobes.
- Les mucin-like de tumeur mammaire. Des difficultés identiques à celles rencontrées avec les « mucines typiques » ont été observées lors de la détermination des séquences des axes peptidiques de mucin-like. Les techniques de génétique moléculaire ont conduit plusieurs équipes [26-28] à publier la première séquence d'ADNc complète d'une mucin-like, l'épisialine ou PEM associée à la tumeur du sein ; son gène *MUC1* est localisé sur le chromosome 1 [29].

\* Code à une lettre. T : Thr ; S : Ser ; A : Ala ; P : Pro ; G : Gly ; H : His ; L : Leu ; V : Val ; D : Asp.

Elle est constituée, en grande partie, de la répétition en tandem, 40 à 90 fois selon les sujets (polymorphisme interindividuel), d'un motif de 20 acides aminés contenant 25 % de résidus hydroxylés (figure 1).

Ces domaines sont rigoureusement identiques au centre de la structure répétitive, le degré d'analogie diminue au fur et à mesure que l'on s'éloigne vers les extrémités N- et C-terminales. Il existe deux transcrits dont l'un est pourvu d'une région d'ancrage intramembranaire, le deuxième transcrit coderait pour une forme sécrétée de l'épisialine.

### Conclusion

Les acquisitions structurales dans le domaine des mucines, qu'elles concernent les glycanes, les précurseurs peptidiques et maintenant les ARN messagers et les gènes concernés, concourent toutes à démontrer, pour cette famille de glycoprotéines, une hétérogénéité extrême et singulière. La plupart des sondes révèlent les ARN messagers de mucines sous la forme d'un continuum rassemblant plusieurs centaines ? plusieurs milliers ? de transcrits de taille différente.

A ce jour, six gènes codant pour les mucines humaines ont été identifiés : *MUC1* (mucines urinaires, *mucin-like*), *MUC2* et *MUC3* (mucines intestinales), *MUC 6 A, B, C + L* (*JER 57, 58, 47*) et *MUC4* (mucines trachéobronchiques) (figure 2). On peut supposer que la découverte de sondes de mucines gastriques ou de mucines cervicales puisse encore élever ce chiffre. Pourquoi une telle diversité génique ? Existe-t-il une spécificité tissulaire de l'expression de ces gènes ? Les deux gènes localisés en 11p15 ne sont-ils en réalité qu'un seul gène transcrit en ARNm différents selon la muqueuse considérée ? S'agit-il de gènes dupliqués ayant acquis une spécificité d'expression différente ? Il semble évident que la régulation de l'expression des gènes de mucines est un phénomène extrêmement complexe, faisant intervenir différents systèmes dont notamment :

- un système répétitif donnant naissance à un polymorphisme interindividuel important ;
- un système d'épissage très varié et

*m/s n° 10, vol. 7, décembre 91*

• Swallow D. M. et al. [29]	1987	Gène de l'épisialine <i>MUC 1: 1q21-q24</i>
• Griffiths B. et al. [20]	1989	Gène de mucine intestinale humaine <i>MUC 2: 11p15</i>
• Gum J. R. et al. [19]	1990	Gène de mucine intestinale humaine <i>MUC 3: 7</i>
• Nguyen Van Cong et al. [23]	1990	Gène de mucine trachéobronchique humaine <i>JER 47:</i> <i>JER 57: 11p15</i> <i>JER 58:</i>  <i>JER 47: 13</i>
• Porchet N. et al. [24]	1991	Gène de mucine trachéobronchique humaine <i>JER 64: 3</i>

Figure 2. **Localisation chromosomique des différents gènes de mucines humaines.**

complexe à partir d'un très grand nombre de petits exons ;  
- une régulation concertée des différents gènes qui, en outre, peuvent posséder de fortes analogies bien que situés sur des chromosomes différents. La question soulevée à la lumière des récents résultats résumés ici est donc : « Pourquoi une telle hétérogénéité peptidique ? » Pourquoi la muqueuse doit-elle disposer de tant de molécules de mucines, à la fois aussi semblables et aussi distinctes ? Pourquoi l'hétérogénéité glycanique se double-t-elle d'une telle hétérogénéité peptidique ? Y a-t-il une relation causale entre l'hétérogénéité glycanique et l'hétérogénéité peptidique ? L'hétérogénéité des mucines est-elle directement nécessitée par la diversité des agressions auxquelles la muqueuse est confrontée ?

Bien évidemment, actuellement, toutes ces questions demeurent soulevées mais sans réponses. Néanmoins, pour

notre part, la notion de régions peptidiques constantes et variables dans les séquences de mucines, la notion d'interactions multiples avec l'environnement, ne sont pas sans évoquer la superfamille des gènes de l'immunité. Et il serait très important de relier « structure et fonction » pour l'étude des maladies impliquant les mucines ■

### Remerciements

Les auteurs tiennent à remercier mesdames Édith Delaleau et Marie-Thérèse Maillard pour la préparation du manuscrit. Ce travail a bénéficié d'un contrat ARC.

### TIRÉS A PART

J.-P. Aubert.

## RÉFÉRENCES

---

26. Wreschner DH, Hareuveni M, Tsarfaty I, *et al.* Human epithelial tumor antigen cDNA sequences. *Eur J Biochem* 1990 ; 189 : 463-73.
27. Ligtenberg MJL, Vos HL, Gennison MC, Hilkens J. Episialin, a carcinoma-associated mucin, is generated by a polymorphic gene encoding splice variant with alternative amino-termini. *J Biol Chem* 1990 ; 265 : 5573-8.
28. Gendler SJ, Lancaster CA, Taylor-Papadimitriou J, *et al.* Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin. *J Biol Chem* 1990 ; 265 : 15286-93.
29. Swallow DM, Gendler S, Griffiths B, *et al.* The hypervariable gene locus PUM which codes for the tumor associated epithelial mucins is located on chromosome 1, within the region 1q21-24. *Ann Hum Genet* 1987 ; 51 : 289-94.

## Summary

### Human mucins : why such a peptide polydispersity ?

The molecules responsible for the viscous and gel forming properties of mucus are carbohydrate-rich glycoproteins designated as « mucins ». Mucins are secreted by *mucosae* and some exocrine glands at the interface with the external environment. « Mucin-like » glycoproteins isolated from the surface of some cancer cells exhibit the same physical and chemical properties as mucins and are regarded as tumor markers.

Due to the difficulty in solubilizing, purifying, deglycosylating and sequencing the mucin polypeptide, the structural organization of the

peptide moiety of these molecules — large mucin monomers ? aggregation of smaller subunits ? — has not been yet established. With molecular genetics techniques, it was easier to deduce the primary sequence of some mucin peptides by first isolating and sequencing some cDNAs from human jejunum or tracheobronchial cDNA libraries. This paper describes molecular biology strategies and recent data related to human mucins and « mucin-like » glycoproteins and raises new questions about mRNAs polydispersity and gene diversity.