

Régulation du gène CFTR : beaucoup de questions, peu de réponses

La mucoviscidose est une maladie génétique de transmission autosomique récessive qui touche l'enfant et l'adulte jeune [1]. La fréquence de la maladie varie selon les groupes ethniques et est plus élevée dans les populations d'Europe du Nord (un enfant atteint pour 2 500 naissances). Le portage de l'hétérozygotie atteint le chiffre remarquable de 1 sur 25 individus. La symptomatologie clinique de la mucoviscidose est dominée par l'atteinte respiratoire avec une obstruction des voies aériennes par un mucus épais et collant, suivie d'une infection par *Pseudomonas aeruginosa*. Il y a également une atteinte gastro-intestinale chez la majorité des patients avec, dans 85 % des cas, une insuffisance pancréatique externe. Cinq à 10 % des nouveau-nés naissent avec une forme d'obstruction intestinale appelée iléus méconial et 2 % à 5 % développent une maladie hépatique durant l'évolution de la maladie. Chez l'adulte, l'infertilité est quasi constante chez les hommes (*m/s* n° 3, vol. 9, p. 344) et fréquente chez les femmes.

Le gène responsable de cette maladie a été cloné chez l'homme en 1989 par clonage positionnel et dénommé CFTR (*cystic fibrosis transmembrane conductance regulator*). Situé sur le chromosome 7, il s'étend sur 250 kb. L'ARN messager a une taille de l'ordre de 6,5 kb et code pour une protéine de 1 480 acides aminés appartenant à la superfamille des protéines transporteuses ABC (*ATP-binding cassette*). Ces molécules ont en commun un ou deux domaines transmembranaires hydrophobes et un ou deux domaines NBF (*nucleotide-binding fold*) qui fixent et cliquent l'ATP, fournissant l'énergie nécessaire au transport. De plus, le

CFTR a un domaine central (*regulatory domain*) riche en sérine, qui est la cible de la phosphorylation relayée par la protéine kinase A (PKA). Il est maintenant certain que le CFTR a, au minimum, une fonction de canal chlore directement activable par la PKA. La délétion de trois bases dans l'exon 10 entraînant la perte d'une phénylalanine en position 508 ($\Delta F 508$) au niveau du premier NBF est responsable de 70 % des allèles mutés [1, 2]. Il semble que cette mutation entraîne un mauvais repliement de la protéine qui ne peut subir une maturation normale et ne parvient donc pas à la membrane. Il faut toutefois noter que lorsque celle-ci y parvient, elle possède une certaine activité [3].

La régulation du gène CFTR est à l'évidence extrêmement complexe. La transcription du gène du CFTR est limitée aux tissus épithéliaux. Chez l'embryon, il existe une expression contrôlée durant le développement au niveau des tissus [4] avec, par exemple, un fort niveau d'expression d'ARNm du CFTR au niveau de l'épithélium bronchique chez le fœtus, qui n'est pas retrouvé après la naissance. Chez l'adulte, les expériences d'hybridation *in situ* chez l'homme ou le rongeur ont clairement mis en évidence une spécificité tissulaire tout à fait particulière de l'ARNm et de la protéine. Ainsi, au niveau des organes exprimant le CFTR (essentiellement le poumon, l'intestin, le pancréas, les glandes salivaires, la vésicule biliaire, les glandes sudoripares et le tractus génital), seule une population de cellules épithéliales est concernée. A titre d'exemple, c'est au niveau des glandes submucoales que l'expression du CFTR est la plus importante dans le poumon [5]. De même, au

niveau de l'épithélium intestinal, il existe un gradient d'expression décroissant de la crypte vers la villosité et de la partie proximale à la partie distale du tube digestif [6]. Enfin, au niveau du tractus génital, l'expression du CFTR est contrôlée durant le cycle de l'épithélium séminifère et le cycle œstroprogestatif [7]. De plus, le profil d'expression du gène CFTR est complémentaire de celui du gène MDR1 (*multi-drug resistance P-glycoprotein*), un gène codant pour une protéine de la famille des transporteurs ABC [8]. Au niveau de l'intestin, par exemple, le MDR1 remplace progressivement le CFTR lorsque les cellules migrent de la crypte vers la villosité. Le but de cette *mini-synthèse* est de faire le point sur l'état actuel des connaissances sur les mécanismes responsables de la régulation transcriptionnelle du gène CFTR. Ces notions sont essentielles car, outre l'intérêt de comprendre la physiopathologie de la maladie, la possibilité d'isoler les éléments d'ADN agissant *in cis* et conférant un niveau d'expression élevé, spécifique de tissu, serait particulièrement importante afin de les associer à l'ADN complémentaire normal en thérapie génique. De plus, l'identification de facteurs agissant *in trans*, et de leur régulation, pourrait permettre de rechercher le moyen de stimuler l'expression du gène et la synthèse de la protéine, ce qui pourrait constituer une perspective thérapeutique dans les cas, telle la mutation $\Delta F 508$, où une protéine mutée conservant une certaine activité est produite chez les malades.

Séquence, sites d'initiation de la transcription, exons alternatifs

La région d'ADN en amont du gène

CFTR est riche en GC avec de nombreux sites potentiels de fixation des facteurs de transcription Sp1 et AP-1 et dépourvue de boîte TATA, aussi bien chez l'homme [9] que chez la souris [10]. Toutefois, les seules homologies de séquence entre les deux espèces concernent un fragment de 34 pb en amont de l'exon 1, et un fragment de 20 + 12 pb au niveau du premier intron.

Il existe chez l'homme plusieurs sites d'initiation de la transcription. Le site majeur semble être situé, pour les cellules exprimant un niveau élevé de CFTR, à 72 pb en amont de l'ATG, les autres se trouvant entre 60 pb en amont et 60 pb en aval du site majeur. Il faut toutefois noter que les résultats sont un peu discordants selon les équipes [9, 11, 12].

De nombreuses isoformes de l'ARNm du CFTR, dues à un épissage alternatif, ont été identifiées, suggérant un niveau de régulation supplémentaire dans l'expression du gène CFTR. Deux exons, préalablement non caractérisés, ont été détectés entre 900 et 500 pb en amont de l'exon 1 (exon -1a et exon 1a) et il a été possible de montrer par amplification de l'ADNc par PCR que l'épissage pouvait se faire des exons -1a à 1a et à 2 ou de -1a à 2 [12]. Toutefois, la signification physiologique de l'initiation de la transcription au niveau des exons -1a et +1a du CFTR reste obscure.

Méthylation des cytosines

Il existe chez les mammifères deux types de promoteurs différant par la richesse en îlots CpG, aussi appelés îlots HTF [13], ceux qui sont riches en CpG (*CpG - island promoters*) et dont les cytosines ne sont pas méthylées, et ceux qui sont relativement pauvres en CpG et dont les cytosines sont méthylées dans la plupart des tissus. Ces derniers sont invariablement trouvés dans les gènes spécifiques de tissus, suggérant une fonction de la méthylation dans la suppression de l'activité basale du gène là où celui-ci n'est pas exprimé. Les premiers sont en règle générale trouvés au niveau des

gènes ubiquitaires (*housekeeping genes*) [13, 14].

L'étude de la méthylation des îlots CpG par *Southern blot* en utilisant des enzymes de restriction sensibles à la méthylation [12, 15] ou par *ligation mediated PCR** [15] a permis de montrer qu'il existe, chez l'homme comme chez la souris, une corrélation inverse entre le degré de méthylation et l'expression du gène CFTR au niveau des lignées cellulaires. En revanche, au niveau des tissus, aucune méthylation n'est retrouvée [15]. La méthylation de ces promoteurs est donc spécifique des cellules cultivées; dans l'organisme vivant, leur répression, là où elle survient, doit donc être due à d'autres mécanismes que la méthylation.

Structure chromatinienne de la région du promoteur CFTR

Chez les eucaryotes, les modifications de structure chromatinienne constituent un premier et fondamental niveau de régulation. Cinq sites hypersensibles à la DNase I ont été identifiés dans la région du premier exon dans la lignée de cellules T84 dérivée d'un adénocarcinome colique exprimant CFTR à fort niveau, localisés à -900, -800, -400, -200 et +200 par rapport au site majeur de début de la transcription [12]. De façon intéressante, le site hypersensible à la DNase I au niveau de l'intron se situe au niveau d'une portion conservée entre l'homme et la souris [10]. Un autre travail a mis en évidence quatre sites hypersensibles dans des cellules HT-29, également dérivées d'un adénocarcinome colique exprimant CFTR, alors qu'il n'en existe pas dans la lignée HFLI qui n'exprime pas CFTR [16]. Toutefois, il semble que seulement deux de ces sites soient communs entre les deux études.

* Le principe de cette technique est de cliver l'ADN génomique à l'aide d'enzymes agissant, ou n'agissant pas, sur les CpG méthylés, de lier aux extrémités clivées un oligonucléotide, et d'amplifier par PCR un fragment d'ADN entre cet oligonucléotide et une autre amorce située à proximité.

Activité promotrice des régions flanquant le premier exon dans des lignées cellulaires transfectées transitoirement

Plusieurs équipes ont étudié l'activité promotrice des régions situées en 5' du premier exon du gène CFTR humain, à l'aide de constructions délétionnelles insérées en amont d'un gène rapporteur. Il n'est pas toujours aisé de comparer les résultats car les cellules utilisées pour les transfections (HT-29, Caco 2, Panc-1, HeLa), les gènes rapporteurs (luciférase, chloramphénicol acétyl transférase (CAT)) ainsi que les constructions étaient différents selon les équipes [9, 11, 12]. Des études semblables ont également été réalisées en utilisant les régions promotrices murines dans des cellules murines [10] ou humaines [17].

Deux points principaux peuvent être tirés de ces travaux : le promoteur CFTR est de faible efficacité (de l'ordre de 1 % à 2 % de celle du promoteur-*enhancer* du virus SV40), et le promoteur minimal est de l'ordre de 250 pb. Pour le reste, les résultats sont quelque peu contradictoires. Chou *et al.* [11] et Chehab *et al.* [10] ont mis en évidence deux éléments agissant en *cis* (un positif et un négatif) au niveau du promoteur humain [11] et murin [10]. Chez la souris, l'élément agissant négativement en *cis* est une séquence de 320 pb polypurine-polypyrimidine qui pourrait agir en imprimant à l'ADN une conformation spéciale (structure H de l'ADN), et l'élément agissant positivement est une boîte Y (séquence de reconnaissance du facteur de transcription boîte Y de 12 pb comprenant un motif CCAAT inversé). Un semblable motif (boîte Y) est noté dans le promoteur du gène MDR1 humain dont le profil d'expression est symétrique de celui du gène CFTR, et intervient dans son activité basale [18]. Cela signifierait que, bien que les séquences aient divergé, les promoteurs humain et murin du gène CFTR fonctionneraient sur le même principe. Il faut toutefois rester prudent et ne pas tirer de conclusions définitives car ces résultats n'ont pas été

retrouvés par d'autres équipes. Koh *et al.* [12] ont trouvé que la force du promoteur du gène CFTR diminuait progressivement avec les délétions effectuées, indiquant qu'il est constitué de multiples éléments actifs en *cis* dont chacun est individuellement faible. Les résultats d'empreintes *in vitro* à la DNase I (*in vitro footprint*), montrant que la région promotrice fixe une population complexe de protéines nucléaires séquences spécifiques, sont conformes à cette interprétation [12].

Une discordance entre l'expression de l'ARN messager du CFTR par les cellules transfectées et le niveau d'expression du gène rapporteur a été retrouvée chez l'homme [12] et chez la souris [10], suggérant que d'autres séquences que les régions amont testées pourraient jouer un rôle, ou que des niveaux post-transcriptionnels de régulation existent, par exemple la stabilité différentielle du message CFTR dans différents tissus.

Enfin, la région conservée de l'intron 1, identique chez l'homme et la souris, et semblant correspondre à un site hypersensible à la DNase I, ne s'est pas révélée active en *cis* [10].

Gènes de fusion chez les souris transgéniques

Un gène rapporteur, contrôlé par des séquences amont, dans les lignées cellulaires, est loin de la réalité d'un *locus* génétique *in vivo*. Pour tenter de se placer dans un contexte plus physiologique, Perraud *et al.* [19] ont créé des souris transgéniques en utilisant 2,2 kb en 5' du gène CFTR placés en amont du gène codant pour l'antigène T de SV40. De manière surprenante, ces animaux ont développé des tumeurs uniquement localisées au niveau des cellules épendymaires des ventricules cérébraux. Bien qu'aucune symptomatologie clinique neurologique n'ait été rapportée chez les patients mucoviscidosiques, il a été montré par hybridation *in situ* que les cellules épendymaires contenaient, de fait, du message CFTR.

A l'encontre de ces résultats, Grier-

senbach *et al.* [20] n'ont pas été capables de montrer la moindre expression du gène *LacZ* lorsque celui-ci était placé en 3' d'un fragment d'ADN représentant jusqu'à 19 kb en amont du gène CFTR.

Conclusions - perspectives

Ces travaux suggèrent que le contrôle de l'expression du gène CFTR résulte de plusieurs mécanismes et que l'on doit être extrêmement méfiant vis-à-vis des modèles qui ne reflètent pas la réalité d'un *locus* entier dans un organisme vivant. Un des moyens de se rapprocher le plus de la situation *in vivo* serait de réaliser des souris transgéniques avec le gène CFTR entouré de ses séquences flanquantes et compris dans un chromosome artificiel de levure (*yeast artificial chromosome*). Une telle approche vient d'être réalisée notamment pour le *locus* humain de globine à l'aide d'un YAC de 248 kb ([21], *m/s* n° 11, vol. 9, p. 1282). Les difficultés pour le gène CFTR sont de deux ordres: premièrement, étant donné la taille du gène CFTR par rapport à celui de globine, le YAC devrait être beaucoup plus grand; deuxièmement, à l'inverse des gènes de globine, on ne sait pas quelles séquences choisir pour faire une mutagenèse dirigée par recombinaison homologue afin de disséquer les mécanismes régulateurs. Enfin, malgré plus de 300 mutations décrites à ce jour chez des patients atteints de mucoviscidose, aucune n'a été trouvée dans des régions régulatrices, alors que, comme cela a été montré dans les thalassémies, de telles mutations représentent un modèle précieux pour l'étude de la régulation d'un gène ■

Erick Denamur

Inserm U. 120, hôpital Robert Debré, 48, boulevard Sérurier, 75935 Paris Cedex 19, France.

RÉFÉRENCES

1. Goossens M. Biologie de la mucoviscidose: progrès récents et perspectives. *médecine/sciences* 1991; 7: 1048-51.
2. Collins FS. Cystic fibrosis: molecular biology and therapeutic implications. *Science* 1992; 256: 774-9.
3. Welsh MJ, Smith AE. Molecular mechanisms of CFTR chloride channel dysfunction in cystic fibrosis. *Cell* 1993; 73: 1251-4.
4. Trezise AEO, Chambers JA, Wardle CJ, Gould S, Harris A. Expression of the cystic fibrosis gene in human foetal tissues. *Hum Mol Genet* 1993; 2: 213-8.
5. Engelhardt JF, Yankaskas JR, Ernst SA, Yang Y, Marino CR, Boucher RC, Cohn JA, Wilson JM. Submucosal glands are the predominant site of CFTR expression in the human bronchus. *Nature Genet* 1992; 2: 240-7.
6. Trezise AEO, Buchwald M. *In vivo* cell-specific expression of the cystic fibrosis transmembrane conductance regulator. *Nature* 1991; 353: 434-7.
7. Trezise AEO, Linder CC, Grieger D, Thompson EW, Meunier H, Griswold MD, Buchwald M. CFTR expression is regulated during both the cycle of the seminiferous epithelium and the oestrous cycle of rodents. *Nature Genet* 1993; 3: 157-64.
8. Trezise AEO, Romano PR, Gill DR, Hyde SC, Sepulveda FV, Buchwald M, Higgins CF. The multidrug resistance and cystic fibrosis genes have complementary patterns of epithelial expression. *EMBO J* 1992; 11: 4291-303.
9. Yoshimura K, Nakamura H, Trapnell BC, Dalemans W, Pavirani A, Lecocq JP, Crystal RG. The cystic fibrosis gene has a «housekeeping»-type promoter and is expressed at low levels in cells of epithelial origin. *J Biol Chem* 1991; 266: 9140-4.
10. Chehab F, Denamur E. Characterization of the murine CFTR promoter region reveals functional but not structural identity to the human CFTR gene. *Am J Hum Genet* 1993; 53 (suppl): 669 (abstr).
11. Chou JL, Rozmahel R, Tsui LC. Characterization of the promoter region of the cystic fibrosis transmembrane conductance regulator gene. *J Biol Chem* 1991; 266: 24471-6.
12. Koh J, Sferra TJ, Collins FS. Characterization of the cystic fibrosis transmembrane conductance regulator promoter region. *J Biol Chem* 1993; 268: 15912-21.
13. Jordan BR. Ilots HTF: le gène annoncé. *médecine/sciences* 1991; 7: 153-60.
14. Bird A. The essentials of DNA methylation. *Cell* 1992; 70: 5-8.

15. Denamur E, Ikuta T, Chehab F. CpG island methylation and CFTR expression in cell lines do not mimic a physiological mechanism. *Am J Hum Genet* 1993; 53 (suppl) 674 (abstr).

16. Yoshimura K, Nakamura H, Trapnell BC, Chu CS, Dalemans W, Pavirani A, Lecocq JP, Crystal RG. Expression of the cystic fibrosis transmembrane conductance regulator gene in cells of non-epithelial origin. *Nucleic Acids Res* 1991; 19: 5417-23.

17. Rozmahel R, Dobbs M, Tsui LC. Characterization of the promoter of the mouse cystic fibrosis transmembrane conductance regulator CFTR gene. *Am J Hum Genet* 1992; 51 (suppl): 135 (abstr).

18. Goldsmith ME, Madden MJ, Morrow CS, Cowan KH. A Y-box consensus sequence is required for basal expression of the human multidrug resistance (*mdr1*) gene. *J Biol Chem* 1993; 268: 5856-60.

19. Perraud F, Yoshimura K, Louis B, Dalemans W, Ali-Hadji D, Schultz H, Claudepierre MC, Chartier C, Danel C, Bellocq JP, Crystal RG, Lecocq JP, Pavirani A. The promoter of the human cystic fibrosis transmembrane conductance regulator gene directing SV40 antigen expression induces malignant proliferation of ependymal cells in transgenic mice. *Oncogene* 1992; 7: 993-7.

20. Griesenbach U, Suen TC, Frumkin A, Olek K, Tsui LC. Study of the promoter of the human cystic fibrosis transmembrane conductance regulator in transgenic mice. *Am J Hum Genet* 1993; 53 (suppl): 69 (abstr).

21. Peterson KR, Clegg CH, Huxley C, Josephson BM, Haugen HS, Furukawa T, Stamatoyannopoulos G. Transgenic mice containing a 248-kb yeast artificial chromosome carrying the human-globin locus display proper developmental control of human globin genes. *Proc Natl Acad Sci USA* 1993; 90: 7593-7.

TIRÉS A PART

E. Denamur.

m/s n° 1 vol. 10, janvier 94