

## Les mécanismes moléculaires du déficit en 21-hydroxylase

Le déficit en 21-hydroxylase est une maladie autosomique récessive relativement fréquente (une naissance sur 12 000). Elle se manifeste par une virilisation excessive et, le plus souvent, un syndrome de perte de sel qui sont la conséquence des désordres de synthèse des stéroïdes hormonaux : hyperproduction de testostérone et carence en minéralocorticoïdes. Le gène de la 21-hydroxylase est localisé sur le chromosome 6, dans la région des gènes de classe III du complexe majeur d'histocompatibilité, au niveau d'un segment dupliqué d'ADN : il existe ainsi un gène actif et, à proximité immédiate, un pseudogène. Cette disposition explique la fréquence des recombinaisons inégales et des événements de conversion génique qui sont les mécanismes presque constants des déficits.

Étienne Mornet  
André Boué

### RÉFÉRENCES

1. Mornet E, Boué J, Raux MC, *et al.* First trimester prenatal diagnosis of 21-hydroxylase deficiency by linkage analysis to HLA class I and class II probes. *Hum Genet* 1986 ; 73 : 358-64.
2. White PC, New MI, Dupont B. Cloning and expression of cDNA encoding a bovine adrenal cytochrome P-450 specific for steroid 21-hydroxylation. *Proc Natl Acad Sci USA* 1984 ; 81 : 1986-90.

### ADRESSE

E. Mornet : docteur ès sciences. A. Boué : professeur de génétique. Inserm U.73, château de Longchamp, 75016 Paris, France.

**L**es progrès de la biologie moléculaire font apparaître les différents types de mutations responsables de maladies génétiques humaines. Le cas du déficit en 21-hydroxylase (21-OH), responsable de l'hyperplasie congénitale surrénalienne, est particulièrement intéressant parce que la plupart des mécanismes moléculaires concevables au niveau du gène sont effectivement trouvés dans cette maladie génétique métabolique : délétions complètes ou partielles, substitutions, insertions, microdélétions, mutations affectant l'épissage. La diversité de ces mutations est due à la disposition particulière des gènes de la 21-OH dont l'un est un pseudogène qui a accumulé ces mutations au cours de l'évolution.

### Le déficit en 21-hydroxylase

Le déficit en 21-OH obéit au modèle récessif autosomique ; sa fréquence

en Europe est de l'ordre de 1/12000 mais est sous-estimée du fait de l'existence de formes sans expression phénotypique franche. Le gène responsable est localisé sur le bras court du chromosome 6 dans la région des gènes de classe III du complexe majeur d'histocompatibilité de l'homme. Il s'agit du gène d'un cytochrome P450 (P450c21) responsable de la 21-hydroxylation de la progestérone en déoxycorticostérone et de la 17-hydroxyprogestérone (17-OHP) en 11-déoxycortisol (*figure 1*). Lorsque cette enzyme est déficiente, c'est-à-dire soit absente soit inactive, il s'ensuit un déficit en cortisol et en aldostérone et une accumulation de 17-OHP qui se convertit alors en testostérone. L'excès de synthèse de testostérone entraîne une virilisation des sujets atteints qui peut aller chez la fille jusqu'à l'ambiguïté sexuelle complète. De plus, 65 % des patients souffrent aussi de syndromes de perte de sel dus à une absence de rétention du sodium rénal liée au déficit en aldostérone. A côté de la



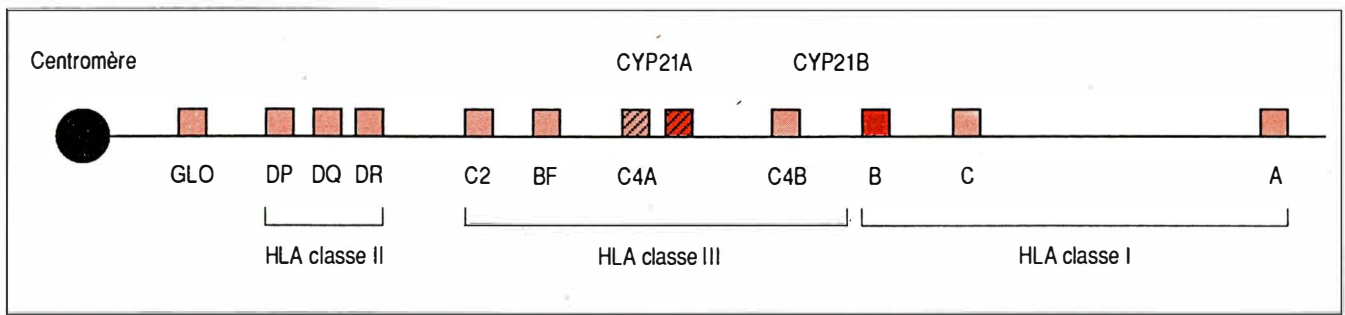


Figure 2. **Représentation schématique de la région du complexe majeur d'histocompatibilité sur le bras court du chromosome 6.** Les deux gènes de la 21-OH CYP21A et CYP21B sont situés à côté de chacun des deux gènes du C4. L'ensemble de cette région représente approximativement 1 800 kilobases, GLO = gène de la glyoxalase érythrocytaire.

des appariements non homologues entre ces gènes conduisant à des recombinaisons inégales et à des phénomènes de conversion génique (figures 4 et 5). Ces appariements sont d'autant plus stables que l'homologie est forte et c'est sans doute ce qui

explique leur grande fréquence dans cette région.

Une recombinaison inégale peut se produire à l'extérieur d'un gène (recombinaison intergénique), ou au contraire dans un gène CYP21 (recombinaison intragénique). Si la

recombinaison est intergénique, elle produit des délétions et des duplications complètes de ces gènes (figure 4). Lorsque les chromosomes recombinés ne contiennent plus que le pseudogène, ils ne portent plus d'information codant pour une acti-

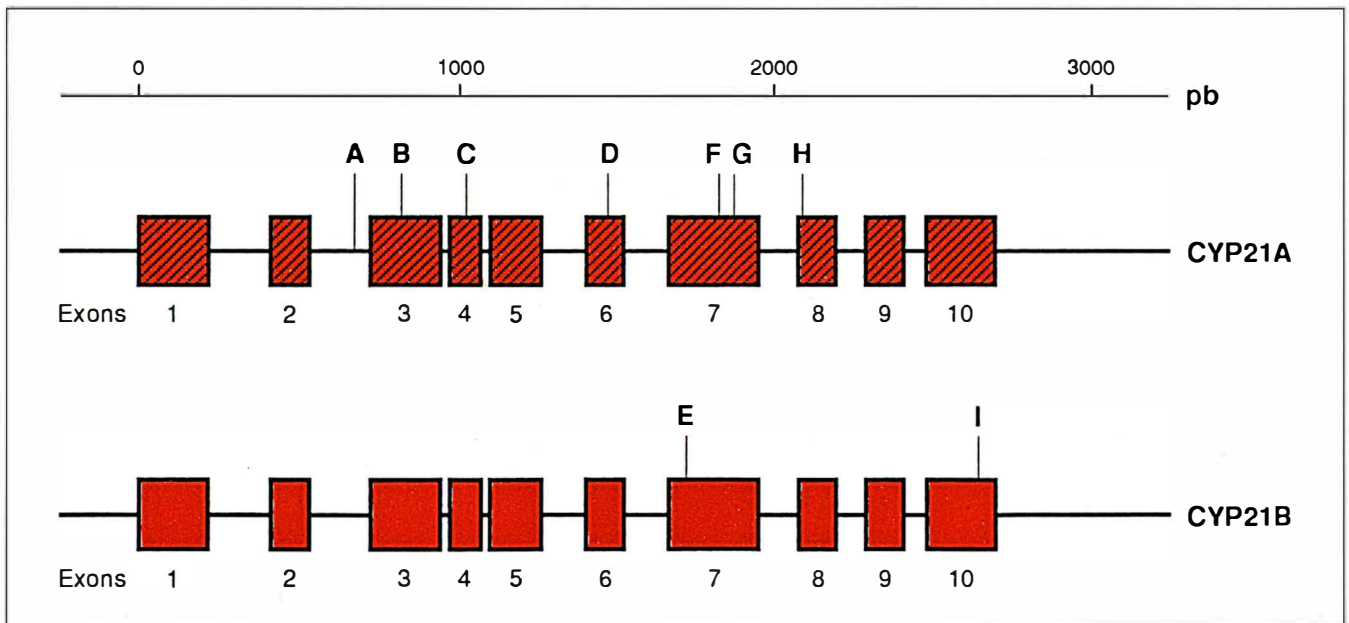


Figure 3. **Localisation des différentes mutations observées sur les gènes de la 21-OH.** Chacun de ces gènes est constitué de dix exons symbolisés par des rectangles. Le gène CYP21A est un pseudogène comportant sept mutations (notées de A à H) réparties dans les exons et dans le 2<sup>e</sup> intron. Le gène CYP21B est fonctionnel chez les individus normaux. La plupart des mutations responsables de la maladie existent également sur le pseudogène et ont été transférées par conversion génique ou recombinaison sur le gène CYP21B. Cependant deux substitutions (E et I) n'ont été décrites que sur le gène fonctionnel et sont donc probablement d'apparition récente sur ce gène. A = substitution C → G dans le 2<sup>e</sup> intron [9] ; B = délétion de huit paires de bases dans le 3<sup>e</sup> exon [11, 12, 14] ; C = substitution T → A dans le 4<sup>e</sup> exon [6] ; D = trois substitutions T → A dans le 6<sup>e</sup> exon [11, 12] ; E = substitution G → C dans le 7<sup>e</sup> exon [10] ; F = substitution G → T dans le 7<sup>e</sup> exon [13] ; G = insertion d'une T dans le 7<sup>e</sup> exon [11, 12, 14] ; H = substitution C → T dans le 8<sup>e</sup> exon [11, 12] ; I = substitution A → G dans le 10<sup>e</sup> exon [10].



tivité 21-hydroxylase puisqu'il n'y a plus de gène fonctionnel. La délétion de la région C4A CYP21A ou de la région CYP21A C4B n'entraîne naturellement pas de déficit puisque le gène *CYP21A* est un pseudogène et que la délétion d'un gène *C4* n'a pas (ou peu) de conséquence dans la mesure où les deux gènes *C4A* et *C4B* sont fonctionnels. Ils ne diffèrent que par quelques mutations silencieuses et par l'existence d'un intron supplémentaire dans le gène *C4A*.

Les chromosomes porteurs d'une duplication pourraient théoriquement avoir soit deux gènes *CYP21A* soit deux gènes *CYP21B*. Cependant la duplication du gène *CYP21B* n'a jamais été observée. La duplication du pseudogène a, en revanche, été fréquemment observée, notamment associée à la forme tardive de la maladie et à l'antigène HLA B14 [3, 4]. Ce phénomène a fait suggérer quelque temps l'existence dans cette forme clinique d'un effet de dose dû à une activité 21-OH résiduelle des pseudogènes : trois pseudogènes pourraient produire suffisamment d'activité 21-OH pour supprimer partiellement le bloc. Cette hypothèse doit maintenant être écartée car la structure du pseudogène *CYP21A* est telle qu'aucune activité même résiduelle ne peut être produite par ce gène. Comme nous le verrons plus loin, cette forme clinique est due à une mutation ponctuelle dans le gène actif entraînant la substitution d'un acide aminé conduisant à une protéine légèrement modifiée dans sa structure ou dans son activité. L'association entre la forme tardive de la maladie et la duplication du pseudogène est donc due à deux événements simultanés : une substitution nucléotidique et la duplication du pseudogène par recombinaison inégale intergénique. Dans les autres formes tardives de la maladie non associées à la duplication du pseudogène, d'autres mutations peuvent être impliquées.

Si la recombinaison est intragénique, elle conduit à des gènes « hybrides » constitués d'une partie 5' de l'un des gènes et d'une partie 3' de l'autre gène. De plus ces gènes « hybrides » sont associés à des délétions ou à des duplications des régions qui les bordent (figure 4). Ces gènes sont en fait inactifs, puisqu'ils contiennent une partie (5' ou 3') du pseudogène.

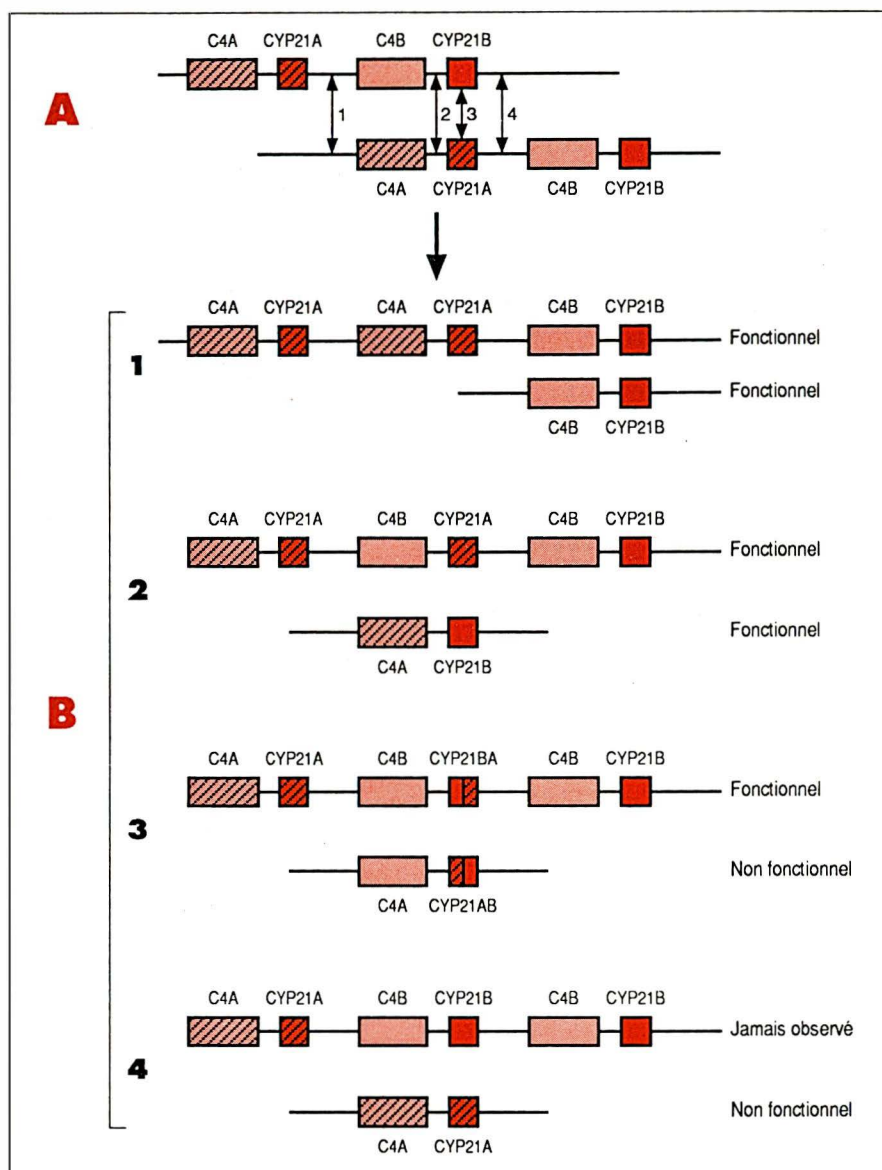


Figure 4. **A. Appariement non homologue des gènes C4 et 21-OH conduisant à des recombinaisons inégales à la méiose B. Produits de recombinaison obtenus.** Si la recombinaison est intergénique [1, 2, 4], elle conduit à des délétions et des duplications complètes des gènes C4 et 21-OH. Si elle est intragénique [3], ces délétions et duplications s'accompagnent de la formation de gènes « hybrides » CYP21AB ou CYP21BA constitués en partie du pseudogène et en partie du gène fonctionnel. Dans le cas où les deux gènes CYP21B originaux sont normaux, les produits 1 et 2 sont tous fonctionnels parce que chaque chromosome porte un gène CYP21B, la recombinaison 3 produit un chromosome fonctionnel et un chromosome porteur d'un gène « hybride » CYP21AB non fonctionnel. La recombinaison 4 produit un chromosome porteur de deux gènes fonctionnels CYP21B et un chromosome non fonctionnel avec une délétion du gène CYP21B.

## RÉFÉRENCES

3. Garlepp MJ, Wilton AN, Dawkins RL, White PC. Rearrangements of 21-hydroxylase genes in disease-associated MHC supratypes. *Immunogenetics* 1986 ; 23 : 100-5.
4. Mornet E, Couillin P, Kutten F, *et al.* Association between restriction fragment length polymorphism detected with a probe for human 21-hydroxylase (21-OH) and two clinical forms of 21-hydroxylase deficiency. *Hum Genet* 1986 ; 74 : 402-8.
5. Kourilsky P. Molecular mechanisms for gene conversion in higher cells. *TIG* 1986 ; 2 : 60-3.
6. Amor M, Parker KL, Globerman H, New MI, White PC. Mutation in the CYP21B gene (Ile → Asn) causes steroid 21-hydroxylase deficiency. *Proc Natl Acad Sci USA* 1988 ; 85 : 1600-4.
7. Steinmetz M, Uematsu Y, Lindhal KF. Hotspots of homologous recombination in mammalian genomes. *TIG* 1987 ; 3 : 7-10.
8. Collier S, Sinnott PJ, Dyer PA, Price DA, Rodney H, Strachan T. Pulse field gel electrophoresis identifies a high degree of variability in the number of tandem 21-hydroxylase and complement C4 genes repeats in 21-hydroxylase deficiency haplotypes. *EMBO J* 1989 ; 8(5) : 1393-402.
9. Higashi Y, Tanai A, Inoue H, Hiro-masa T, Fujii-Kuriyama Y. Aberrant splicing and missense mutation causing steroid 21-hydroxylase (P-450c21) deficiency in humans : possible gene conversion products. *Proc Natl Acad Sci USA* 1988 ; 85 : 7486-90.
10. Rodrigues NR, Dunham I, Yu CY, Carrol MC, Porter RR, Campbell RD. Molecular characterisation of the HLA-linked steroid 21-hydroxylase B gene from an individual with congenital adrenal hyperplasia. *EMBO J* 1987 ; 6 : 1653-61.
11. Higashi Y, Yoshioka H, Yamana M, Gotoh O, Fujii-Kuriyama Y. Complete nucleotide sequence of two steroid 21-hydroxylase genes tandemly arranged in human chromosome : a pseudogene and a genuine gene. *Proc Natl Acad Sci USA* 1986 ; 83 : 2841-5.
12. White PC, New MI, Dupont B. Structure of human steroid 21-hydroxylase genes. *Proc Natl Acad Sci USA* 1986 ; 83 : 5111-5.
13. Speiser P, New MI, White PC. Molecular analysis of non classical steroid 21-hydroxylase deficiency associated with HLA-B\*4, DR1. *N Eng J Med* 1988 ; 1 : 19-23.
14. Globerman H, Amor M, Parker KL, New MI, White PC. Nonsense mutation causing steroid 21-hydroxylase deficiency. *J Clin Invest* 1988 ; 82 : 139-44.
15. Lindberg RLP, Negeshi M. Alteration of mouse cytochrome P450c11 substrate specificity by mutation of a single aminoacid residue. *Nature* 1989 ; 339 : 632-4.
16. Mornet E, Dupont J, Vitek A, White PC. Characterization of two genes encoding human steroid 11-B hydroxylase (P450c11). *J Biol Chem* 1989 ; 264 : 20961-7.

La conversion génique est le remplacement d'une certaine quantité d'ADN (d'une à plusieurs centaines de paires de bases) par une autre qui lui est fortement homologue sans qu'il y ait réciprocité [5]. C'est cette absence de réciprocité qui différencie la conversion génique de la recombinaison inégale. Le gène donneur subit une réparation *ad integrum*, tandis que le gène receveur n'est pas réparé et conserve donc la séquence donnée. Ces gènes peuvent se trouver sur deux chromatides différentes ou sur une même chromatide (figure 5). Dans ce dernier cas, c'est une boucle du chromosome qui met face à face les deux séquences entre lesquelles va se réaliser la conversion génique\*. L'existence dans les gènes CYP21A et CYP21B de six séquences « GCTGGGG » dites Chi (*cross-over hotspots instigators*) [7] pourrait favoriser la conversion génique et la recombinaison inégale [6]. Dans le cas de la 21-OH, la conversion génique conduit à des gènes « hybrides » constitués en partie par des séquences du pseudogène et en partie par des séquences du gène fonctionnel. Seule la conversion génique de CYP21A vers CYP21B (c'est-à-dire la transformation d'une partie du gène actif B en la séquence du pseudogène A) conduit à une modification du génotype. En effet la conversion d'un gène CYP21B vers un gène CYP21A ne peut probablement pas conduire à un gène fonctionnel car trop de sites sont mutés sur ce pseudogène. Il faut enfin souligner que dans l'espèce humaine comme chez tous les eucaryotes supérieurs, le généticien ne peut observer que le gène receveur, c'est-à-dire une moitié du phénomène, et qu'il est souvent difficile de distinguer un produit de recombinaison inégale d'un produit de conversion génique. Cependant l'électrophorèse en champs pulsés permet maintenant de distinguer plus facilement ces deux phénomènes [8]. Le séquençage des gènes CYP21B d'individus malades a montré que la plupart des mutations observées sur ces gènes et responsables du déficit en 21-hydroxylase existait aussi sur le

pseudogène. La probabilité qu'un tel événement soit dû au hasard, c'est-à-dire qu'un gène CYP21B subisse la même mutation qu'un gène CYP21A au même endroit est quasiment nulle. Par conséquent une telle mutation provient presque certainement du pseudogène par conversion génique. Neuf mutations ponctuelles différentes ont été décrites dans les gènes de la 21-OH (figure 3). Trois d'entre elles conduisent à un codon de terminaison de la traduction, une autre entraîne un épissage aberrant par substitution nucléotidique dans un intron et les cinq dernières sont des substitutions conduisant à une modification de la structure primaire de la protéine et de son activité. Le gène CYP21A est non fonctionnel parce qu'il contient les trois mutations qui conduisent à une terminaison de la traduction : une délétion de huit paires de bases (pb) dans le 3<sup>e</sup> exon, une insertion d'une pb dans le 7<sup>e</sup> exon et une substitution d'une pb dans le 8<sup>e</sup> exon. De plus, il existe dans ce gène d'autres différences avec le gène fonctionnel qui conduisent à des changements d'acides aminés et donc à une modification de la structure et (ou) de l'activité de la protéine.

**Épissage aberrant.** Une substitution C → G a été décrite par Higashi *et al.* [9] dans le 2<sup>e</sup> intron à 13 pb du 3<sup>e</sup> exon. Il en résulte, à cet endroit, l'apparition d'une séquence AG (au lieu de AC) qui fait apparaître un site accepteur cryptique d'épissage conduisant à un messager plus long de 13 pb et par conséquent à un décalage du cadre de lecture qui entraîne l'apparition d'un codon de terminaison au début de l'exon 3. Cette mutation également décrite par Rodrigues *et al.* [10] semble relativement fréquente.

**Une délétion de 8 pb** dans le troisième exon décrite par Higashi *et al.* [11] et White *et al.* [12] entraîne l'apparition d'un codon de terminaison par décalage du cadre de lecture.

**Substitutions entraînant une modification de l'activité et (ou) de la structure de la protéine.** Sept mutations de ce type ont été décrites dont la plupart existent dans le pseudogène et ont été transférées au gène fonctionnel par conversion génique. Certaines d'entre elles ont des con-

\* Voir m/s n° 2, vol. 4, p. 123.

séquences importantes. C'est le cas par exemple d'une substitution T → A (Ile → Asn) dans le 4<sup>e</sup> exon décrite par Amor *et al.* [6]. En effet cette mutation est retrouvée chez les patients atteints de la forme sévère de la maladie. En revanche, ces patients n'ont pas de syndrome de perte de sel, ce qui signifie que la voie de l'aldostérone n'est pas complètement abolie. D'autre part ce codon Ile est retrouvé conservé dans différentes espèces comme la souris, le bœuf ainsi que dans les cytochromes P450 de la même famille (17-hydroxylase, phénobarbital, dioxine), ce qui indique qu'il joue probablement un rôle important dans le maintien de la conformation de l'enzyme.

Une autre substitution décrite par Speiser *et al.* [13] transforme une valine en leucine et est retrouvée chez des individus atteints de la forme tardive de la maladie. Cette substitution doit affecter la protéine

dans une moindre mesure puisque son activité résiduelle évite une forme très sévère de la maladie. Cependant le codon normal valine est également conservé dans les autres espèces indiquant que toute modification à ce niveau (même mineure) est mal tolérée et contre-sélectionnée.

Une substitution Gln → Stop dans le 8<sup>e</sup> exon introduit un codon de terminaison au lieu d'une glycine et conduit à une protéine tronquée et par conséquent inactive [11, 12, 14]. L'insertion d'une base dans le 7<sup>e</sup> exon entraîne l'apparition d'un codon de terminaison par décalage du cadre de lecture [11, 12].

### Variabilité clinique et polymorphisme génétique

On peut constater la grande diversité des mécanismes moléculaires responsables du déficit en 21-OH. Si une

telle diversité est retrouvée dans d'autres maladies métaboliques héréditaires, il est néanmoins certain que la disposition particulière des gènes et pseudogène de la 21-OH est responsable de cette diversité.

Une question importante et non encore résolue est de savoir si la grande variabilité clinique observée dans cette maladie peut être corrélée à la diversité des mutations mises en cause. Il est très probable par exemple que la substitution G → T, dans le 7<sup>e</sup> exon du gène *CYP21B*, que l'on trouve associée à une duplication du gène *CYP21A* et à l'antigène HLA B14, est responsable d'une grande partie des formes tardives de la maladie (environ 60 %). Le criblage de grandes populations d'individus à l'aide de sondes oligonucléotidiques spécifiques de ces mutations devrait confirmer cette hypothèse. Un autre cas relativement simple est celui de la délétion totale du gène *CYP21B*

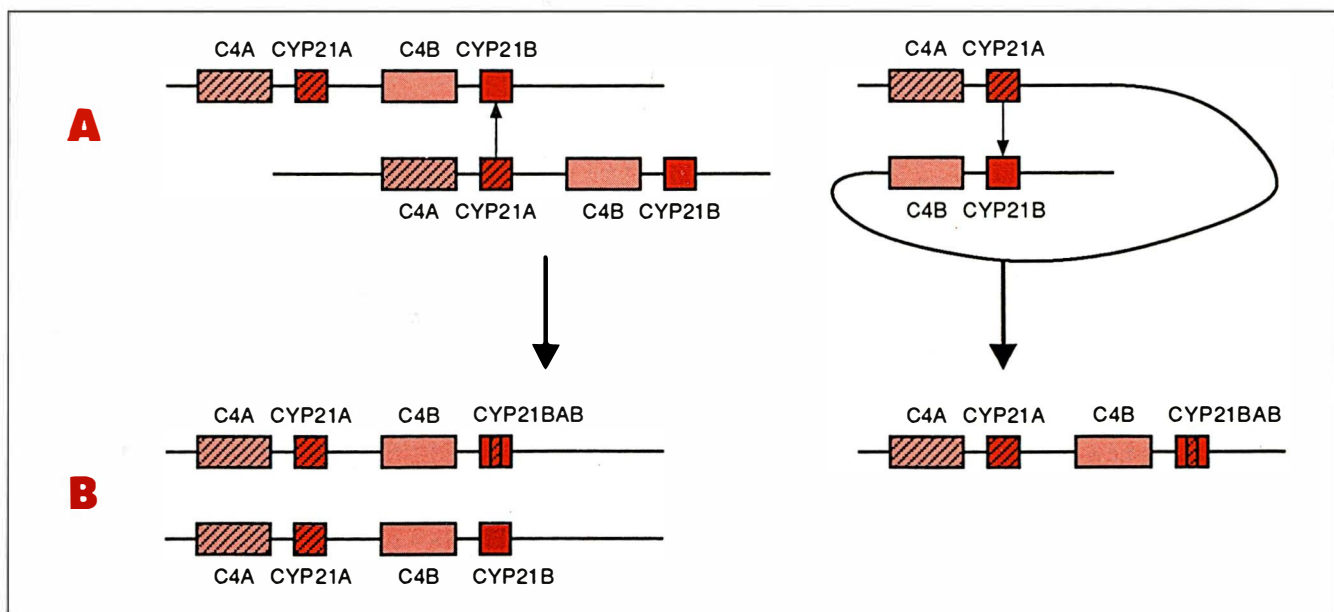


Figure 5. **Conversion génique du pseudogène *CYP21A* vers le gène fonctionnel *CYP21B*.** A. L'appariement des deux gènes peut se faire entre deux chromatides différentes ou au contraire sur une même chromatide par formation d'une boucle qui place face à face les gènes C4A et C4B, d'une part, et les gènes CYP21A et CYP21B, d'autre part. B. Produits de la conversion génique. Le gène donneur n'est pas modifié, tandis que le gène receveur est constitué d'une partie du pseudogène et d'une partie du gène fonctionnel.



retrouvée chez environ 20 % des patients et associée généralement à l'antigène HLA B47. L'activité 21-OH est totalement abolie et on peut donc supposer que ce type de mutation conduit à la forme la plus sévère de la maladie. C'est sans doute également le cas des mutations conduisant à un codon de terminaison, puisque la protéine tronquée qui en résulte est inactive. Dans les autres cas, il est difficile de relier directement la mutation avec la forme clinique du patient qui est souvent double hétérozygote du fait du nombre de mutations possibles. De plus, il est rare de trouver une seule mutation sur un gène *CYP21B*. Ainsi Rodrigues *et al.* [10] ont séquencé complètement le gène d'un patient atteint de la forme classique avec perte de sel. Ces auteurs ont relevé onze différences dans la séquence, dont cinq sont retrouvées sur le pseudogène. Deux de ces mutations se trouvent dans la région 5' flanquante, deux dans les introns, une dans la région 3' flanquante et quatre dans les exons. Parmi les quatre différences dans les exons, deux sont des substitutions silencieuses qui ne modifient pas la structure primaire de la protéine et deux entraînent des substitutions d'acides aminés dans les exons 7 et 10 respectivement. De plus, une des mutations dans l'intron 2 est responsable d'un mauvais épissage. Par conséquent, au moins trois mutations sont délétères dans ce gène et il est difficile de connaître leurs rôles respectifs dans la maladie. Ces trois mutations peuvent d'ailleurs avoir des rôles additifs à l'origine d'un phénotype particulier. C'est le cas du cytochrome P450c21 mitochondrial de la souris où trois substitutions éloignées les unes des autres entraînent une perte complète de l'activité de cette enzyme, alors que la présence d'une seule de ces mutations permet la production de la moitié de l'activité enzymatique [15]. Les gènes *21-OH* comportant plusieurs mutations ne sont probablement pas rares, du fait des conversions géniques et recombinaison, favorisées, comme nous l'avons déjà souligné par la disposition et la grande homologie des gènes et pseudogènes *21-OH*. La grande diversité des mutations qui en résulte est proba-

bablement la cause de la variabilité clinique de cette maladie. Une analyse systématique de l'ADN à l'aide de sondes oligonucléotidiques, effectuée sur une large population, devrait permettre de déterminer les rôles respectifs de ces mutations.

La très grande majorité des hyperplasies surrénaliennes (98 %) est due soit à un déficit en 21-OH soit à un déficit en 11 $\beta$ -OH, tandis que les autres cas (soit 2 %) sont dus à un déficit en 17-hydroxylase ou en desmolase. On sait que les gènes de la 11 $\beta$ -OH comme ceux de la 21-OH sont dupliqués et très homologues et que l'un d'entre eux n'est apparemment pas fonctionnel [16]. Les gènes de la 17-hydroxylase ou de la desmolase sont au contraire uniques. Par conséquent, la fréquence des mutations pourrait être particulièrement élevée dans des gènes liés et homologues. Il faut enfin souligner qu'outre les anomalies moléculaires du gène du cytochrome P450c21, des mutations affectant d'éventuels gènes de régulation de ce dernier pourraient également jouer un rôle dans l'expression phénotypique.

### 21-OH et génétique des populations

La fréquence élevée des mutations du gène *21-OH* fonctionnel, et par voie de conséquence de la maladie, peuvent s'expliquer par un effet sélectif favorable d'un allèle C4 associé sur le même chromosome. En effet, en raison de la liaison physique importante entre les *loci* C4 et 21-OH, toute mutation favorable ou défavorable à l'un des deux *loci* pourra entraîner l'augmentation ou la diminution de la fréquence de l'allèle présent à l'autre *locus*. On peut aussi expliquer cette fréquence élevée des mutations par le fait qu'à chaque génération, c'est-à-dire de manière récurrente, le pseudogène apporte par conversion génique, au gène fonctionnel, les mutations qui se sont accumulées facilement puisque sans conséquences sélectives. C'est ainsi que l'on peut expliquer la pénétration du déficit en 21-OH dans les populations alors que cette maladie était encore très récemment considérée comme létale au sens reproductif du terme ■

## Summary

### Molecular basis of congenital adrenal hyperplasia

Congenital adrenal hyperplasia is an autosomal recessive disorder due to a deficit of the 21-hydroxylase enzyme. Molecular analysis of the cytochrome P450c21 gene responsible for the disease shows multiple types of mutations including partial or complete deletion, substitutions, insertions. The high diversity of mutations is due to the particular disposition of the functional gene and of a pseudogene tandemly arranged in the class III region of the human major histocompatibility complex. This disposition associated with the high degree of homology between these genes favours the appearance of mutations by unequal recombination and gene conversion.

## TIRÉS À PART

E. Mornet.