

La régulation des gènes de globine au cours du développement : plusieurs contrôles simultanés

La commutation ou *switch* des gènes de β -globine a été définie comme l'activation séquentielle des cinq gènes du *locus* selon un programme ordonné au cours du développement. L'intérêt persistant du sujet s'explique tant par la fonction hautement prioritaire de la molécule d'hémoglobine et la fréquence des variants naturels, pathologiques ou non, que par le caractère de modèle d'une famille multigénique soumise à une série de régulations différentes et simultanées (*m/s* n° 3 vol. 8, p. 255). On a pu démontrer que le contrôle de ce processus spécifique est majoritairement transcriptionnel, et que son mécanisme est l'interaction entre des éléments d'ADN situés en *cis* au voisinage de chacun des gènes et d'autres séquences situées beaucoup plus en amont, et qui constituent le LCR (*locus control region*); cette interaction est réalisée par l'intermédiaire de facteurs transactivateurs, en petit nombre et toujours les mêmes, ubiquitaires ou spécifiques des tissus érythroïdes. La fonction du LCR est à la fois d'activer les gènes de globine, assurant à chacun un niveau d'expression élevé, dépendant du nombre de copies intégrées et indépendant du site d'intégration, et de stabiliser l'ouverture de la chromatine, c'est-à-dire de maintenir les promoteurs libres de nucléosomes.

Le schéma trop simple élaboré ces dernières années ne répond cepen-

dant pas à toutes les questions que pose la régulation des gènes de globine au cours du développement. L'étude de modèles naturels, et surtout des expériences d'expression chez des souris transgéniques, ou en lignées permanentes, ont fait apparaître un modèle de commutation de plus en plus complexe, qui a pu être comparé à un nœud gordien. Si quelques données semblent établies, d'autres varient largement d'une expérience à l'autre, en particulier en ce qui concerne l'expression des gènes de γ globine, et donc de l'hémoglobine fœtale.

L'étude de modèles naturels a été considérée pendant longtemps comme une voie d'abord privilégiée de cette compréhension; une série de travaux a ainsi identifié la base moléculaire des différentes formes de persistance héréditaire d'hémoglobine fœtale (PHHF ou HPFH pour les Anglo-Saxons). Il s'agit, en fait, d'un groupe très hétérogène: au niveau morphologique, on a distingué des formes pancellulaires et des formes hétérocellulaires; au niveau biochimique, des hémoglobines fœtales où s'exprimait un seul des gènes de γ globine ou, au contraire, les deux gènes simultanément; sur le plan moléculaire enfin, des formes dues à une délétion de taille très variable, ou d'autres dans lesquelles n'était retrouvée qu'une mutation ponctuelle dans la région promotrice d'un gène. Accompa-

gnant ces différents travaux, des modèles explicatifs ont été fournis à la surexpression persistante d'hémoglobine fœtale. On a ainsi évoqué la délétion, ou non-délétion, d'une zone critique en amont du gène δ , le rapprochement post-délétionnel d'une séquence activatrice et des gènes γ , des modifications de la structure chromatinienne, la fixation entravée, ou au contraire amplifiée, de l'un ou l'autre facteur transactivateur par suite d'une mutation à proximité d'un gène de γ globine. Si les explications proposées étaient chaque fois séduisantes, elles n'étaient, cependant, applicables qu'à un nombre limité de cas. Aucune des explications proposées à ce jour ne se présente comme générale.

Les études expérimentales sur des souris transgéniques maintiennent, elles aussi, des zones d'obscurité. La régulation au cours du développement, activation et extinction successives de chaque gène, a fait proposer deux modèles de contrôle moléculaire, une extinction autonome ou une compétition. Le gène ϵ fonctionne de façon autonome, c'est-à-dire de la même façon, qu'il ait été introduit chez l'animal à l'état isolé ou avec les autres gènes du *locus*. Les gènes adultes β et δ , quant à eux, fonctionnent de façon compétitive: ils sont exprimés à tous les stades du développement s'ils sont injectés seuls et en présence du

I.C.R ; leur régulation temporelle est rétablie dans un contexte de compétition avec les gènes fœtaux. Les résultats sont discordants en ce qui concerne ces derniers. Injectés sans I.C.R mais dans un fragment de 3,3 kb, leur régulation temporelle est correcte, quoique faible, ce qui prouve l'existence de séquences régulatrices au voisinage immédiat des gènes γ et β . L'adjonction du I.C.R réduit à ses séquences majeures, le μ I.C.R, fait perdre ce contrôle, et le gène γ , comme le gène β , est alors exprimé à tous les stades du développement. Des séquences supplémentaires dans les cellules adultes normales, absentes des deux types de transgènes présentés ci-dessus, sont donc indispensables à l'extinction normale des gènes γ . Deux séries d'expériences apparemment contradictoires ont cherché à déceler la nature de ces séquences.

D'une part, plusieurs équipes ont réalisé l'introduction des gènes γ en souris transgénique dans leur contexte chromosomique : dans un cosmide [1], ou mieux encore dans un YAC [2-4] (*m/s n° 11 vol. 9, p. 1282*), ce qui, dans tous les cas, rétablit une expression analogue à ce qui est observé *in vivo*. D'autres expériences ont pu montrer une régulation correcte en fonction des étapes du développement, en injectant à la souris un gène γ et une construction des quatre sites du I.C.R. Dans cette dernière expérience, les résultats n'étaient clairs que chez les animaux ayant incorporé une ou deux copies du transgène, les autres étaient éliminés sous prétexte d'artefact [5].

Un travail récent apporte des données complémentaires sur la structure des promoteurs γ , et permet peut-être une hypothèse unificatrice

d'interprétation valable tant pour les données phénotypiques normales ou pathologiques que pour les résultats expérimentaux. L'influence d'une série de délétions du promoteur sur l'expression d'un gène γ par une souris transgénique a été étudiée dans des constructions comportant les quatre sites du I.C.R, et le gène γ avec différents fragments de son promoteur, représentant de 141 à 730 pb en amont du gène (*figure 1*). Cela a permis d'identifier trois zones régulatrices, dont deux de régulation négative et une de régulation positive. La construction que nous appellerons -141 γ s'éteint spontanément dans les tissus adultes, son expression est faible, sans doute en relation avec l'absence d'une boîte CAC. On peut en inférer l'existence d'un élément de régulation négative spécifique entre -141 et le site Cap. L'addition du segment suivant, de -141 à -201, supprime l'extinction du gène γ au stade adulte. L'expression est, par ailleurs, importante, en rapport avec le nombre de copies intégrées, ce qui suggère qu'un élément de régulation positive, en interaction avec le I.C.R, est dominant par rapport à l'élément négatif vu précédemment, insuffisant alors pour assurer l'extinction du gène γ . C'est par l'addition du fragment allant de -382 à -730 que réapparaît une régulation négative des gènes γ au stade adulte, dominante sur la régulation positive située plus en aval. Mais une corrélation entre le nombre de copies intégrées et l'expression du gène n'est plus retrouvée, comme si celle-ci était modulée par la position d'intégration et par la chromatine environnante, ce qui correspond bien aux résultats expérimentaux de plusieurs auteurs [6].

L'ensemble de ces résultats permet sans doute de dire que l'arrêt d'expression des gènes γ est soumis au contrôle d'éléments multiples (*figure 2*). Dans une zone proche du gène lui-même, dans les 141 pb en amont du site de début de la transcription, il y aurait un site fixant un répresseur de transcription. Entre

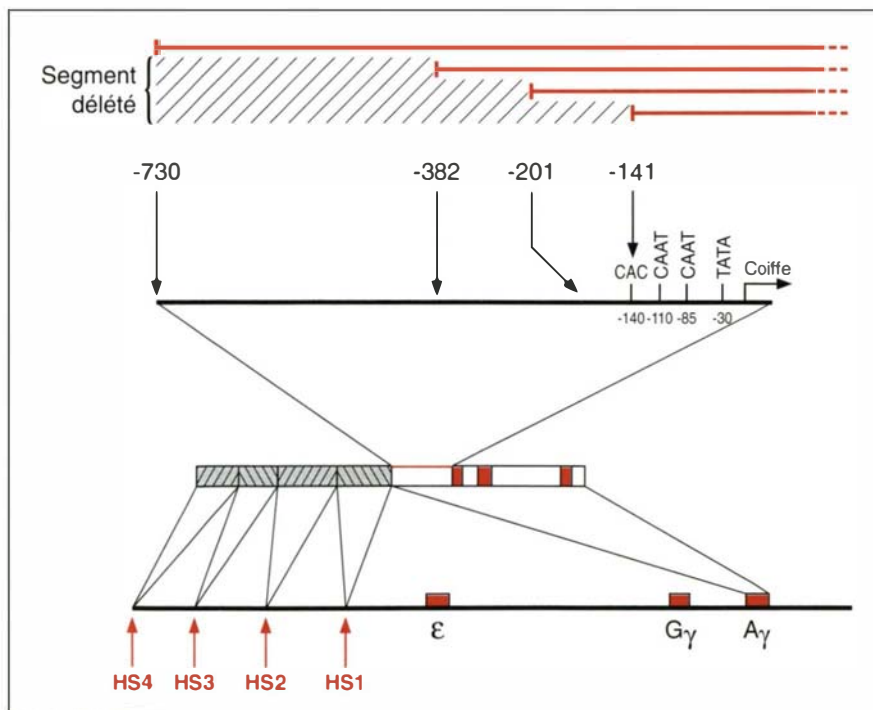


Figure 1. **Schéma des constructions qui ont été utilisées et exprimées chez des souris transgéniques.** La construction type comporte un μ -LCR fait des quatre sites hypersensibles et un gène $A\gamma$. La région promotrice, en 5' de la coiffe (capping site) est de longueur variable, une série de délétions ayant amputé des morceaux plus ou moins grands.

-730 et -382, se situe un élément dont l'action est spécifique du stade adulte, et dont on peut penser qu'il empêche l'interaction du LCR et des gènes foétaux. A ces deux éléments, nécessaires mais non suffisants, s'ajoute la contribution compétitive du gène β (*m/s n° 3 vol. 8, p. 255*). Si l'action combinée de ces trois facteurs est nécessaire pour une régulation parfaite, on pourra expliquer que l'un ou l'autre peut se trouver en défaut et entraîner une synthèse des gènes γ persistante chez l'adulte. On a pu ainsi trouver, à l'origine d'une PHHF, une mutation du promoteur minimal en -117 ou des délétions plus ou moins importantes supprimant la compétition du gène β . Les mutations situées plus en amont, dans la partie distale du promoteur, en -175 ou autour de -200, agiraient sans doute de façon indirecte. Si l'on suit la proposition des auteurs, les modifications induites dans la liaison de facteurs transactivateurs seraient susceptibles de renforcer une régulation positive et d'entraver de ce fait le caractère dominant normal du *silencer* adulte. On peut évidemment souhaiter qu'une mutation trouvée au niveau de cette séquence *silencer* vienne fournir la confirmation directe de cette interprétation.

Tout modèle proposé reste une hypothèse qui tente d'expliquer des données expérimentales. Il n'est pas définitif et appelle naturellement une vérification. Celui que les auteurs formulent ici est, dans l'état actuel des connaissances, particulièrement séduisant parce que d'application très large. La famille des gènes de globine est un modèle qui peut aider à repenser l'extraordinaire variété phénotypique de presque toutes les maladies génétiques, même dites monogéniques. Les schémas simples formulés initialement ne sont sans doute que le point de départ qui permettra d'aborder la complexité des processus biologiques.

D.L.

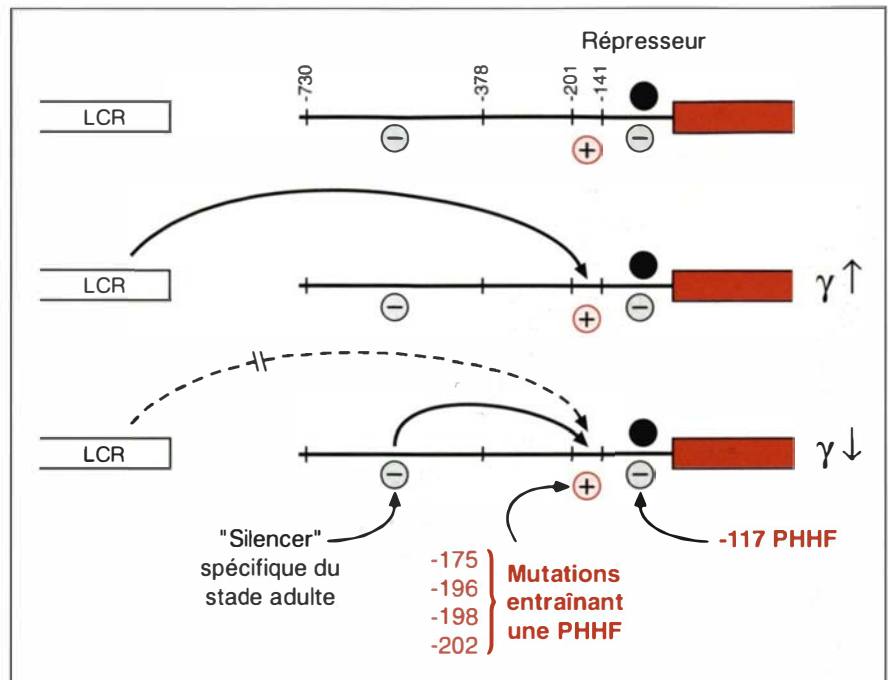


Figure 2. **Schéma théorique proposant le rôle fonctionnel des différents fragments d'un promoteur γ .** Deux éléments de régulation négative encadrent un élément de régulation positive. Le plus près du gène lui-même, on trouve la séquence de fixation d'un répresseur spécifique. C'est à ce niveau qu'on expliquerait l'action d'une mutation en -117. Plus en amont, entre -141 et -201, un élément de régulation positive, dont l'action est dominante et qui peut expliquer la persistance d'expression des gènes foétaux, quand il n'y a pas de compétition chez la souris au stade adulte. Entre -382 et -730, enfin, se fixerait un facteur de régulation négative puissant, spécifique du stade adulte, qui empêche alors l'interaction du LCR et du gène γ . Cette inhibition serait elle-même entravée par plusieurs mutations de la zone activatrice.

1. Strouboulis J, Dillon N, Grosveld F. Developmental regulation of a complete 70-kb human β -globin locus in transgenic mice. *Genes Dev* 1992; 6: 1857-64.
2. 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 β -globin locus display proper developmental control of human globin genes. *Proc Natl Acad Sci USA* 1993; 90: 7593-7.
3. Peterson KR, Zitnik G, Huxley C, Lowrey CH, Gnirke A, Leppig KA, Papayannopoulou T, Stamatoyannopoulos G. Use of yeast artificial chromosomes (YACs) for studying control of gene expression: correct regulation of the genes of a human β -globin locus fol-

lowing transfer to mouse erythroleukemia cell lines. *Proc Natl Acad Sci USA* 1993; 90: 11207-11.

4. Karin M, Gaensler L, Kitamura M, Kan YW. Germ-line transmission and developmental regulation of a 150-kb yeast artificial chromosome containing the human β -globin locus in transgenic mice. *Proc Natl Acad Sci USA* 1993; 90: 11381-5.
5. Dillon N, Grosveld F. Human γ -globin genes silenced independently of other genes in the β -globin locus. *Nature* 1991; 350: 252-4.
6. Stamatoyannopoulos G, Josephson B, Zhang JW, Li Q. Developmental regulation of human γ -globin genes in transgenic mice. *Mol Cell Biol* 1993; 13: 7636-44.