

Des facteurs de régulation spécifiques de la myogenèse

Les gènes de la famille MyoD1 codent pour des protéines nucléaires se fixant à l'ADN et capables d'entraîner la conversion de divers types de cellules en myoblastes et en myotubes. Ces protéines semblent actives sous la forme de dimères formés grâce aux interactions entre des régions des sous-unités adoptant une conformation « hélice-boucle-hélice ». La possible formation d'hétérodimères entre des sous-unités différentes de la famille MyoD1, ou entre une sous-unité de cette famille et une autre possédant le motif hélice-boucle-hélice mais non spécifique de la différenciation musculaire, est de nature à accroître la diversité et la flexibilité des processus de régulation auxquels participent ces protéines.

Serge Alonso

ADRESSE

S. Alonso : chargé de recherche au Cnrs, fait partie de l'unité de M. Buckingham à l'Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, France. Ce laboratoire est financé par l'AFM, l'ARC, le Cnrs, l'Inserm et l'Institut Pasteur.

m/s n° 7 vol. 6, septembre 90

Les quinze dernières années ont été marquées par un très grand bond en avant de la biologie, dû principalement au développement des techniques de génie génétique dans les laboratoires de recherche. Ceux qui s'intéressaient déjà à la différenciation musculaire (myogenèse) ont réussi alors à isoler un grand nombre de gènes codant pour des protéines exprimées spécifiquement dans les muscles (protéines contractiles, enzymes, récepteurs...) [1]. La diversité des isoformes musculaires résulte, soit de la transcription de gènes indépendants dispersés sur le génome, soit de la transcription d'un gène unique donnant naissance à plusieurs ARN messagers (ARNm) par épissage différentiel des exons, et c'est surtout au niveau transcriptionnel que s'opère le contrôle de

l'expression génétique. L'étape suivante a consisté à définir précisément les séquences qui jouent un rôle dans cette régulation. L'analyse fonctionnelle fondée sur les expériences de transfection dans les cellules musculaires a permis l'identification de régions d'ADN, dans le promoteur ou même à plus grande distance du site d'initiation de la transcription, qui sont essentielles à la fois pour l'activation transcriptionnelle et l'expression spatio-temporelle spécifique.

Les lignées myogéniques comme modèle de différenciation

Une fois ces motifs déterminés, il a été possible d'identifier des protéines qui se lient à ces séquences et confèrent la régulation. L'approche clas-

RÉFÉRENCES

1. Buckingham ME. Actin and myosin multigene families : their expression during the formation of skeletal muscle. In : *Essays in Biochemistry* 1985 ; 20 : 77-109.
 2. Reznikoff CA, Brankow DW, Heidelberg C. Establishment and characterization of a cloned line of C3H mouse embryo cells sensitive to postconfluence inhibition of division. *Cancer Res* 1973 ; 33 : 3231-8.
 3. Constantinides PG, Jones PA, Gevers W. Functional striated muscle cells from non-myoblast precursors following 5-azacytidine treatment. *Nature* 1977 ; 267 : 364-6.
 4. Taylor SM, Jones PA. Multiple new phenotypes induced in 10T1/2 and 3T3 cells treated with 5-azacytidine. *Cell* 1979 ; 17 : 771-9.
 5. Taylor SM, Jones PA. Changes in phenotypic expression in embryonic and adult cells treated with 5-azacytidine. *J Cell Physiol* 1982 ; 111 : 187-94.
 6. Konieczny SF, Emerson Jr CP. 5-azacytidine induction of stable mesodermal stem cell lineages from 10T1/2 cells : evidence for regulatory genes controlling determination. *Cell* 1984 ; 38 : 791-800.
 7. Lassar AB, Paterson BM, Weintraub H. Transfection of a DNA locus that mediates the conversion of 10T1/2 fibroblasts to myoblasts. *Cell* 1986 ; 47 : 649-56.
 8. Pinney DF, Pearson-White SH, Konieczny SF, Latham KE, Emerson Jr CP. Myogenic lineage determination and differentiation : evidence for a regulatory gene pathway. *Cell* 1988 ; 53 : 781-93.
 9. Davis RL, Weintraub H, Lassar AB. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* 1987 ; 51 : 987-1000.
 10. Weintraub H, Tapscott SJ, Davis RL, et al. Activation of muscle-specific genes in pigment, nerve, fat, liver and fibroblast cell lines by forced expression of MyoD. *Proc Natl Acad Sci USA* 1989 ; 86 : 5434-8.
 11. Tapscott SJ, Lassar AB, Davis RL, Weintraub H. 5-bromo-2'-deoxyuridine blocks myogenesis by extinguishing expression of *MyoD1*. *Science* 1989 ; 245 : 532-6.
- sique a consisté à utiliser les motifs régulateurs comme matrices pour piéger, dans des extraits nucléaires, les protéines capables de s'y fixer. Plusieurs facteurs transcriptionnels de type ubiquitaire ont ainsi été caractérisés et se sont révélés parfois agir aussi sur les promoteurs de gènes musculaires, mais le (ou les) facteur(s) spécifiquement myogénique(s) n'avaient pu être identifiés. En revanche, dans d'autres systèmes de différenciation cellulaire, comme la cellule hépatique par exemple, des facteurs spécifiques ont été identifiés par cette approche. Dans le cas du muscle, une tout autre démarche a donné des résultats étonnants. La stratégie utilisée concerne non seulement la question des facteurs de transcription spécifique mais aussi celle plus large des effecteurs de la différenciation cellulaire. La différenciation des cellules musculaires passe obligatoirement par la transition des myoblastes, cellules mononucléées se divisant, en myotubes multinucléés exprimant les différents marqueurs de la myogenèse. Depuis l'isolement, dans les années soixante-dix, de lignées cellulaires dérivées du muscle, il est possible de reproduire *in vitro* cette transition myoblaste/myotube et de l'étudier au niveau moléculaire. Ces lignées myogéniques, qui ont joué un rôle essentiel dans la mise en évidence des motifs d'ADN impliqués dans la transcription, n'ont pourtant pas permis, à l'origine, de découvrir un facteur spécifique et il a fallu faire appel à une lignée de... fibroblastes !

Conversion myogénique de la lignée fibroblastique 10T1/2 par les facteurs de régulation

Cette étude a eu comme point de départ l'existence d'une lignée cellulaire multipotentielle : la lignée embryonnaire C3H10T1/2 (10T1/2) de fibroblastes de souris, isolée il y a de nombreuses années [2] et susceptible de se différencier en cellules musculaires [3]. En effet, la croissance des fibroblastes 10T1/2 en présence de 5-azacytidine (analogue structural non méthylable de la cytidine) les conduit à se convertir en trois types cellulaires d'origine mésodermique : tissu adipeux, cartilage et

muscle (*figure 1*) [4]. Cette lignée permet donc d'étudier l'étape précoce de la myogenèse qui détermine une cellule mésodermique non différenciée à devenir un myoblaste. C'est cette étape que l'on a appelée la détermination. On savait déjà que l'hypométhylation des gènes est souvent associée à leur activation. Après incorporation de la 5-azacytidine (responsable de l'hypométhylation) dans l'ADN des cellules 10T1/2, on observe avec une fréquence élevée la formation de myoblastes (de 20 % à 50 %), ainsi que d'adipocytes (7 % à 28 %) et de chondrocytes (1 % à 7 %) [3-6]. La conversion des cellules 10T1/2 par la 5-azacytidine a fourni un modèle d'abord à Konieczny et Emerson [6] et plus récemment à Lassar *et al.* [7] pour étudier les gènes de régulation impliqués dans la détermination des cellules musculaires. Ces derniers auteurs ont montré que l'introduction de l'ADN génomique des azamyoblastes* dans les cellules d'origine les convertit en myoblastes stables, et cela même en absence de 5-azacytidine. C'est plus particulièrement la modification (probablement une déméthylation) d'un seul locus qui permet la conversion myogénique des cellules 10T1/2. L'ADN génomique des cellules non traitées à la 5-azacytidine est dépourvu de cette capacité. Un clone génomique Myd humain, dont la séquence reste encore inconnue, est capable à lui seul de convertir les cellules 10T1/2 en myoblastes [8].

En fait, la stratégie qui a permis d'isoler et de caractériser complètement le premier ADNc codant pour un facteur responsable de la conversion myogénique est l'hybridation différentielle. L'idée de départ était d'isoler une séquence responsable de la détermination des myoblastes. La technique consiste à hybrider la population des ADNc des azamyoblastes avec l'ARN polyA+** des cellules originelles 10T1/2 non

* Azamyoblastes : myoblastes obtenus après traitement de la lignée 10T1/2 par la 5-azacytidine.

** ARN polyA+ : fraction des ARN, les messagers inclus, comportant une suite d'acides adényliques en position 3'.

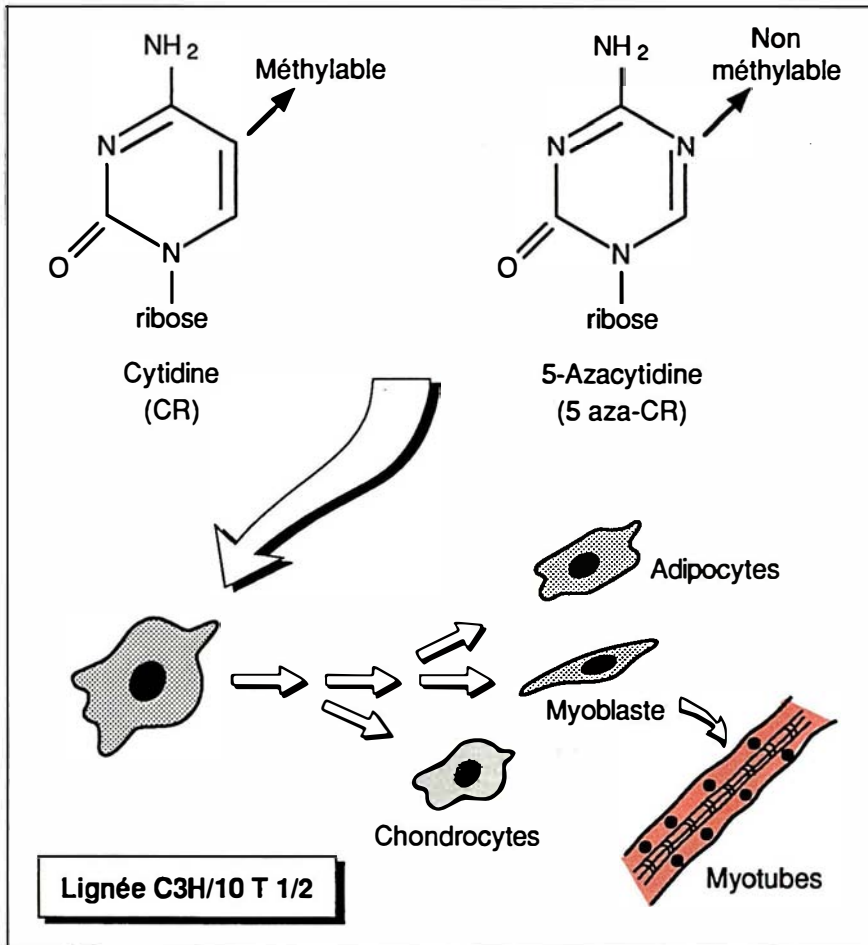


Figure 1. La croissance des fibroblastes 10T1/2 en présence de 5-azacytidine conduit à la conversion de cette lignée principalement en myoblastes mais aussi en adipocytes et chondrocytes.

traitées. Tous les ADNc qui n'hybrident pas à l'ARN 10T1/2 sont spécifiques du stade myoblaste. Plusieurs d'entre eux ont été utilisés comme sondes pour cribler une banque d'ADNc de myocytes, et parmi les clones obtenus, Davis *et al.* [9] ont effectivement montré que l'un d'eux était capable de convertir les cellules 10T1/2 en myoblastes stables. La séquence de cet ADNc révèle qu'il code pour une protéine, MyoD1, pour « *Myoblast Determination gene number 1* ». Le gène *MyoD1* est normalement exprimé *in vivo* dans le muscle squelettique et *in vitro* dans les myo-

blastes en division et les myotubes différenciés, mais non dans les cellules 10T1/2 non différenciées, ni dans celles des muscles cardiaque ou lisse. *MyoD1* est capable de convertir d'autres lignées d'origine mésodermique (adipocytes) et d'activer spécifiquement l'expression de gènes musculaires dans des lignées d'origine ectodermique (mélanome et neuroblastome) ou endodermique (hépatocytes) [10].

D'autres séquences régulatrices ont été recherchées dans les myoblastes déjà déterminés et sur le point de se différencier, mais avant que n'apparaissent les marqueurs spécifiques du muscle.

Dans ce cas, il s'agissait plutôt de rechercher une séquence de différenciation et c'est encore l'approche par criblage différentiel qui fut déterminante. C'est la lignée L6 de myoblastes de rat, donc déjà déterminée, qui a été utilisée. Cette lignée est incapable de se différencier en présence de 5-bromo-2'-désoxyuridine (BrdU), qui est un analogue de la thymidine [11], à part quelques colonies qui parviennent à surmonter ce blocage. Dans ce cas, l'hybridation différentielle est réalisée entre les ADNc de cellules sensibles non différenciées et ceux des cellules résistantes au BrdU capables de se différencier et supposées donc posséder une plus grande quantité de facteurs de différenciation [12]. Un clone d'ADNc appelé myogénine a ainsi pu être isolé. L'ARNm correspondant est présent à un haut niveau dans les myoblastes en train de fusionner et à un plus bas niveau dans les myotubes. Il est spécifique du muscle squelettique comme *MyoD1*. On constate que malgré un rôle supposé différent de celui de *MyoD1*, l'ADNc myogénine est lui aussi capable, comme *MyoD1*, de convertir la lignée 10T1/2 en cellules myogéniques. La caractérisation des séquences nucléotidiques des clones *MyoD1* et myogénine a révélé des domaines d'homologie entre elles et montré qu'il s'agissait en fait de deux gènes appartenant à la même famille multigénique.

Après ces travaux initiaux, d'autres laboratoires utilisèrent la séquence *MyoD1* comme sonde pour cribler différentes banques d'ADNc. L'homologue de la séquence myogénine chez la souris a déjà été isolé [13]. Par ailleurs, deux ADNc, *Myf-3* et *Myf-4*, respectivement homologues à *MyoD1* et myogénine, ont été isolés à partir d'une banque d'ADNc de muscle squelettique de fœtus humain [14]. L'ADNc *Myf-5* issu de la même banque code, lui, pour un facteur myogénique de régulation (le troisième) différent de *MyoD1* et de la myogénine et capable comme eux de convertir la lignée 10T1/2 [15]. Un quatrième facteur, *MRF4*, provient d'une banque d'ADNc de muscle squelettique de rat adulte [16] et équivaut à *Myf-6* chez l'homme [17] et à l'herculine

RÉFÉRENCES

12. Wright WE, Sassoon DA, Lin VK. *Myogenin*, a factor regulating myogenesis, has a domain homologous to *MyoD*. *Cell* 1989 ; 56 : 607-17.
13. Edmondson DG, Olson EN. A gene with homology to the myc similarity region of *MyoD1* is expressed during myogenesis and is sufficient to activate the muscle differentiation program. *Genes Dev* 1989 ; 3 : 628-40.
14. Braun T, Bober E, Buschhausen-Denker G, Kotz S, Grzeschik K-H, Arnold HH. Differential expression of myogenic determination genes in muscle cells: possible autoactivation by the *Myf* gene products. *EMBO J* 1989 ; 8 : 3617-25.
15. Braun T, Buschhausen-Denker G, Bober E, Tannich E, Arnold HH. A novel human muscle factor related to but distinct from *MyoD1* induces myogenic conversion in 10T1/2 fibroblasts. *EMBO J* 1989 ; 8 : 701-9.
16. Rhodes SJ, Konieczny SF. Identification of MRF4: a new member of the muscle regulatory factor gene family. *Genes Dev* 1989 ; 3 : 2050-61.
17. Braun T, Bober E, Winter B, Rosenthal N, Arnold HH. *Myf-6*, a new member of the human gene family of myogenic determination factors: evidence for a gene cluster on chromosome 12. *EMBO J* 1990 ; 9 : 821-31.
18. Miner JH, Wold B. Herculin, a fourth member of the *MyoD* family of myogenic regulatory genes. *Proc Natl Acad Sci USA* 1990 ; 87 : 1089-93.
19. Lin Z-Y, Dechesne CA, Eldridge J, Paterson BM. An avian muscle factor related to *MyoD1* activates muscle-specific promoters in nonmuscle cells of different germ-layer origin and in BrdU-treated myoblasts. *Genes Dev* 1989 ; 3 : 986-96.
20. De La Brousse F, Emerson Jr CP. Localized expression of a myogenic regulatory gene, *gmf1*, in the somite dermatome of avian embryos. *Genes Dev* 1990 ; 4 : 567-81.
21. Hopwood ND, Pluck A, Gurdon JB. *MyoD* expression in the forming somites is an early response to mesoderm induction in *Xenopus* embryos. *EMBO J* 1989 ; 8 : 3409-17.
22. Scales JB, Olson EN, Perry M. Two distinct *Xenopus* genes with homology to *MyoD1* are expressed before somite formation in early embryogenesis. *Mol Cell Biol* 1990 ; 10 : 1516-24.
- chez la souris [18]. Ce quatrième facteur présente les mêmes propriétés que les ADNc *MyoD1*, myogénine et *Myf-5*, mis à part qu'on ne le détecte dans aucune des lignées myogéniques testées et qu'il est prédominant dans le muscle squelettique adulte de chacune des espèces. Les auteurs n'ayant pas cru bon de se mettre d'accord sur une nomenclature commune, il est conseillé de se reporter au *Tableau I* qui indique les différents noms des quatre facteurs en fonction de l'espèce.
- On remarque que la conversion des cellules 10T1/2 en myoblastes est obtenue par n'importe lequel des quatre ADNc, à la condition que ce dernier soit sous le contrôle d'un promoteur fort (type LTR viral) mais qu'en revanche, avec l'ADN génomique des azamyoblastes, les promoteurs naturels suffisent. La famille multigénique comprenant *MyoD1* et *myogénine* contient également *Myf-5* et *MRF4* mais pas le gène *myd* qui doit appartenir à une classe différente de gènes régulateurs [14]. Si les facteurs de régulation ont été à l'origine mis en évidence chez les mammifères (souris, rat et homme), la famille multigénique qui leur a donné naissance est conservée au cours de l'évolution puisqu'on retrouve des séquences homologues à *MyoD1* chez le poulet [19], la caille [20], le xénope [21, 22] et la drosophile [23].

Expression des gènes de régulation dans les cellules musculaires

L'expression constitutive du gène *MyoD1* dans les azamyoblastes 10T1/2 en division précède celle du gène *myogénine* plutôt caractéristique du myoblaste terminal sur le point de se différencier. Les deux transcrits continuent d'être présents dans les myotubes différenciés, mais dans la lignée L6, par exemple, le niveau des transcrits myogénine est nettement plus faible que dans les myoblastes. Alors que toutes les lignées myogéniques différenciées expriment le gène *myogénine*, certaines d'entre elles (les lignées BC3H, L6 et L8) restent négatives pour *MyoD1* mais positives pour *Myf-5* aux stades myoblaste et myotube. Dans d'autres lignées (P2, MM14), alors que *MyoD1* est abon-

dant, *Myf-5* n'a pas été détecté par les techniques utilisées [14, 24, 25]. Le facteur *Myf-5* est-il capable de jouer le rôle de *MyoD1* dans les lignées négatives pour ce dernier (et réciproquement) ou bien les lignées ne reproduisent-elles pas plutôt des « morceaux choisis » différents du programme embryonnaire ?

L'isolement d'une lignée C2 inducible de cellules de souris permet de mieux appréhender les relations entre les différents facteurs. Sans inducteur, ces cellules, qui sont pourtant des myoblastes déterminés, ne se différencient que très peu et n'expriment pas *MyoD1* [26]. En présence de l'inducteur, elles se différencient en myotubes et expriment *MyoD1* au cours de la différenciation terminale des cellules [26]. Les transcrits myogénine ne sont trouvés ni dans les myoblastes C2 classiques ni dans les myoblastes inductibles et n'apparaissent qu'au stade myotube. Les facteurs *MyoD1* et myogénine ne sont donc pas nécessaires pour le maintien du phénotype myoblastique. *MyoD1* ne serait pas le premier gène de régulation entrant en action, comme cela avait été supposé, mais agirait plutôt en augmentant la probabilité de différenciation des myoblastes [voir aussi 27]. Dans les cellules C2 inductibles qui n'expriment ni le gène *MyoD1* ni le gène *myogénine* au stade myoblaste, on détecte les transcrits *Myf-5* dont le niveau décroît au cours de la différenciation des cellules (D. Montarras et C. Pinset, communication personnelle). Ces résultats suggèrent que le gène *Myf-5* est exprimé plus tôt dans les myoblastes présomptifs et pourrait agir en amont de *MyoD1* et de *myogénine* dans le processus de détermination. Le gène *MyoD1* est exprimé plus tardivement dans les myoblastes et participerait avec le gène *myogénine* à la différenciation des myoblastes C2 en myotubes. L'expression des transcrits *MRF4* n'est détectée dans aucune des lignées myogéniques qui ont été testées [16].

L'étude de l'expression de ces facteurs *in vivo* démontre qu'ils sont spécifiques du muscle squelettique. *MyoD1*, *myogénine* et *Myf-5* sont exprimés à un plus haut niveau chez le fœtus que chez l'adulte [9, 12, 14], alors que les transcrits *MRF4* appa-

Tableau I			
NOM DES QUATRE FACTEURS MYOGÉNIQUES SELON L'ESPÈCE CONSIDÉRÉE			
	souris	rat	homme
1	MyoD1		Myf-3
2	myogénine	myogénine	Myf-4
3	Myf-5		Myf-5
4	herculine	MRF4	Myf-6

raissent dans le fœtus tardif et sont prédominants chez l'adulte [16]. Il est évidemment important de savoir à quel moment ils apparaissent au cours de l'embryogenèse et si leur expression précède bien la formation des muscles. Les expériences d'hybridation *in situ* sur des coupes d'embryons de souris démontrent que les premiers transcrits sont décelables dans les somites dont le compartiment myotomal contient les cellules précurseurs des muscles squelettiques [28]. Certains gènes de protéines contractiles comme les actines sont également exprimés très tôt dans les somites alors que d'autres (gènes des chaînes lourdes de la myosine) sont transcrits plus tardivement. Au cours de l'embryogenèse, les gènes *MyoD1*, *myogénine* et *Myf-5* présentent des cinétiques d'expression différentes qui sont détaillées dans l'article de M.O. Ott, B. Robert et M. Buckingham de ce numéro.

Propriétés des facteurs myogéniques

• Autorégulation positive

Comme cela a déjà été observé pour certains gènes de la drosophile (*fushi tarazu*, *ultrabithorax* ou *deformed*) participant à la détermination de l'« architecture » de l'insecte, le gène *MyoD1* est sujet à une autorégulation positive. La transfection des fibroblastes 10T1/2 par un vecteur d'expression contenant soit une séquence *MyoD1*, soit une séquence myogénine sous contrôle d'un promoteur fort active la transcription du gène endogène *MyoD1* dans ces cellules [29]. De plus la transfection par un ADNc *MyoD1*

conduit à l'activation du gène *myogénine* endogène [29]. L'expression constitutive de MRF4 dans les cellules 10 T1/2 déclenche la conversion myogénique ainsi que l'expression des gènes *MyoD1* et *myogénine* [16]. Au contraire, le gène *MRF4* endogène n'est pas activé par la transfection des cellules par une séquence *MyoD1* ou myogénine [16]. Dans la cellule, l'autorégulation positive pourrait présenter au moins deux avantages : tout d'abord maintenir en activité de manière autonome et stable des gènes régulateurs essentiels et aussi amplifier la quantité de facteurs jusqu'à un seuil critique requis pour le déclenchement de la différenciation terminale.

• Inhibition de *MyoD1* par les oncogènes ou les facteurs de croissance

Les proto-oncogènes sont connus aujourd'hui pour jouer un rôle au cours du cycle cellulaire et de la différenciation des cellules [30]. L'expression d'oncogènes transformants est capable d'inhiber la différenciation de lignées cellulaires. Par exemple, l'expression de l'oncogène activé *ras* empêche la différenciation terminale des cellules musculaires [31]. Des myoblastes déterminés restent toujours capables de proliférer alors que leur différenciation terminale en myotubes et leur prolifération sont deux phénomènes qui s'excluent mutuellement. En effet, la différenciation des myoblastes s'accompagne obligatoirement de l'arrêt en phase G1 du cycle cellulaire, et c'est le moment où l'on observe, par exemple, une chute des transcrits du proto-oncogène *c-*

myc [32]. L'expression de l'oncogène *ras* activé dans les azamyoblastes 10T1/2 ou C2 inhibe l'expression du gène *MyoD1* dans les myoblastes, leur différenciation en myotubes, ainsi que l'expression de la plupart des marqueurs spécifiques du programme myogénique [33, 34]. Cette inhibition est levée par l'expression, dans les myoblastes transformés par *ras* (ou *fos*), d'un rétrovirus portant la séquence *MyoD1*, suffisante pour réactiver le programme myogénique. On peut en quelque sorte titrer *MyoD1* par un oncogène tel que *ras*. Au vu de l'homologie des facteurs de la famille *MyoD1* avec le proto-oncogène *c-myc* (voir section suivante), on peut proposer que le produit d'un proto-oncogène entre en compétition avec un ou plusieurs de ces facteurs myogéniques pour se lier à un autre effecteur clé (inconnu) de la différenciation myogénique. La différenciation des myoblastes pourrait alors être contrôlée par les niveaux relatifs des transcrits respectifs des deux compétiteurs. Cette hypothèse permettrait de comprendre pourquoi on doit forcer l'expression des ADNc du type *MyoD1* au moyen d'un promoteur fort afin d'obtenir la conversion des lignées. En effet, le niveau élevé de facteurs myogéniques ainsi produits permettrait de contrebalancer l'effet antagoniste des oncogènes que l'on peut supposer abondants dans les lignées cellulaires transformées, lesquels oncogènes maintiennent les cellules en division permanente.

L'addition du facteur de croissance FGF (*fibroblastic growth factor* ou TGF- β (*transforming growth factor* β)) dans des cultures de myoblastes produit une inhibition de la différenciation myogénique en réduisant le niveau des transcrits *MyoD1* dans ces cellules [35]. A la différence de l'inhibition produite par l'oncogène *ras*, l'expression constitutive d'une séquence *MyoD1* est incapable de surmonter ce blocage produit par FGF ou TGF- β .

• Les différents domaines de *MyoD1*

L'unique phase de lecture de l'ADNc *MyoD1* code pour une protéine de 318 acides aminés présentant plusieurs caractéristiques des protéines se liant à l'ADN [9]. Les différents

RÉFÉRENCES

23. Michelson AM, Abmayr SM, Maniatis T. DrMyoD, a *Drosophila* homolog of the mammalian myogenic regulatory gene family. Abstract for EMBO workshop on cellular and molecular biology of muscle development, septembre 1989.
24. Mueller PR, Wold B. *In vivo* footprinting of a muscle specific enhancer by ligation mediated PCR. *Science* 1989 ; 246 : 780-6.
25. Brennan TJ, Edmondson DG, Olson EN. Aberrant regulation of *MyoD1* contributes to the partially defective myogenic phenotype of BC3H1 cells. *J Cell Biol* 1990 ; 110 : 929-37.
26. Montarras D, Pinset C, Chelly J, Kahn A, Gros F. Expression of *MyoD1* coincides with terminal differentiation in determined but inducible muscle cells. *EMBO J* 1989 ; 8 : 2203-7.
27. Hiti AL, Bogenmann E, Gonzales F, Jones PA. Expression of the *MyoD1* muscle determination gene defines differentiation capability but not tumorigenicity of human *Rhabdomyosarcomas*. *Mol Cell Biol* 1989 ; 9 : 4722-30.
28. Sassoon D, Lyons G, Wright WE, et al. Expression of two myogenic regulatory factors *myogenin* and *MyoD1* during mouse embryogenesis. *Nature* 1989 ; 341 : 303-7.
29. Thayer MJ, Tapscott SJ, Davis RL, Wright WE, Lassar AB, Weintraub H. Positive autoregulation of the myogenic determination gene *MyoD1*. *Cell* 1989 ; 58 : 241-8.
30. Kelly K, Siebenlist U. The regulation and expression of *c-myc* in normal and malignant cells. *Annu Rev Immunol* 1986 ; 4 : 327-38.
31. Olson EN, Spizz G, Tainsky MA. The oncogenic forms of *N-ras* or *H-ras* prevent skeletal myoblast differentiation. *Mol Cell Biol* 1987 ; 7 : 2104-11.
32. Endo T, Nadal-Ginard B. Transcriptional and posttranscriptional control of *c-myc* during myogenesis : its mRNA remains inducible in differentiated cells and does not suppress the differentiated phenotype. *Mol Cell Biol* 1986 ; 6 : 1412-21.
33. Konieczny SF, Drobos BL, Menke SL, Taparowsky EJ. Inhibition of myogenic differentiation by the *H-ras* oncogene is associated with the down regulation of the *MyoD1* gene. *Oncogene* 1989 ; 4 : 473-81.

domaines de la protéine MyoD1 à partir de l'extrémité NH₂-terminale sont : une région acide (résidus 1-61), un domaine dit en « doigts à zinc » (résidus 62-101) riche en résidus cystéine et histidine, ce domaine présentant de l'affinité pour l'ADN [36], une région fortement basique (résidus 102-104) et près de celle-ci une région de 22 acides aminés (résidus 141-162) appelée région *myc* parce que très homologue à un domaine conservé des protéines transformantes codées par la famille des gènes cellulaires *myc* (*N-myc*, *L-myc* et *C-myc*) (figure 2). La protéine MyoD1 est phosphorylée. La présence des régions basique et *myc* est nécessaire et suffisante pour obtenir la conversion des fibroblastes 10T1/2 en myoblastes stables. La région basique suffit, quant à elle, à assurer la localisation nucléaire de la protéine, aussi bien dans les myoblastes que dans les myotubes [37]. Cette localisation particulière de *MyoD1* suggère qu'il s'agit d'une protéine se liant à l'ADN. Bien que cela reste encore à démontrer, la région acide de la partie N-terminale de *MyoD1* est un domaine pouvant jouer le rôle d'activation dans la transcription des gènes auxquels se fixerait *MyoD1*. En effet, ce type de séquence est connu

pour activer la transcription de gènes sous le contrôle de promoteurs naturels ou hétérologues [voir 38].

La super-famille des protéines *myc*

Les quatre ADNc du type MyoD1 codent pour des protéines différentes mais qui partagent entre elles un domaine conservé d'environ 60 acides aminés comprenant à la fois la région basique, servant à la fixation à l'ADN, et la région *myc*. Ce domaine conservé est justement celui qui est essentiel à la conversion des fibroblastes 10T1/2 [37], ce qui suggère qu'il a probablement une fonction commune dans les trois facteurs. Il a été également retrouvé dans deux protéines humaines, E12 et E47, qui se lient au *enhancer* de la chaîne κ des immunoglobulines [39], ainsi que dans quatre protéines régulatrices de la drosophile : *achaete-scute* (As-C), impliquée dans la détermination neuronale [40], *daughterless* (da), essentielle pour la détermination du sexe et du système nerveux périphérique [41], *twist*, qui intervient au moment de la gastrulation [42], et *enhancer of split*, exprimé au cours de la neurogenèse [43] (figure 3). A l'intérieur de cette super-famille, on

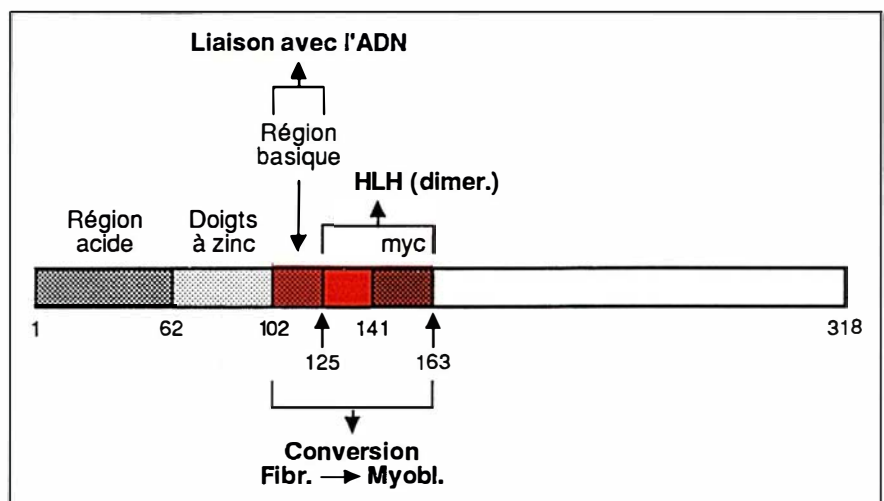


Figure 2. Les différents domaines de la protéine MyoD1 et leurs fonctions. HLH = hélice-boucle-hélice pour hélice-boucle-hélice. Le premier résidu de chaque domaine est indiqué par sa position dans la chaîne.

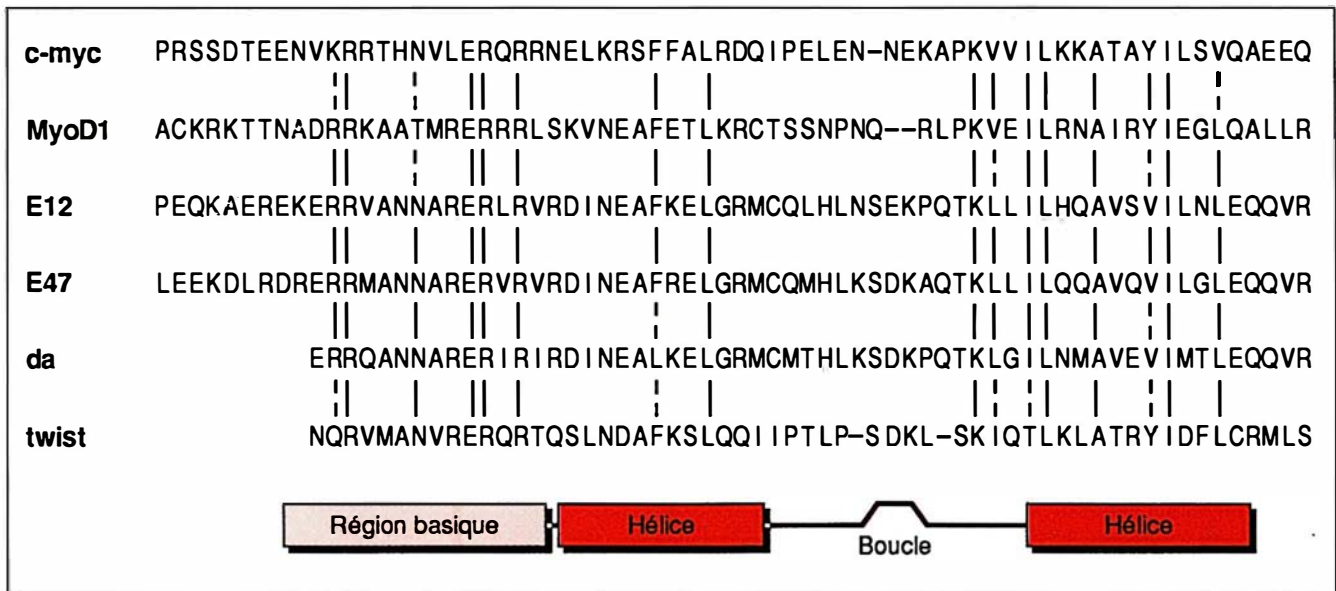


Figure 3. **Alignement de six protéines appartenant à la super-famille des protéines myc, au niveau des régions basique et HLH.** L'identité entre acides aminés est notée par un trait et la similarité par des pointillés, seulement lorsque les six protéines sont concernées.

distingue les facteurs ubiquitaires (E12, E47 et *daughterless*), les facteurs à spécificité tissulaire (*MyoD1*, myogénine, *Myf-5*, *MRF-4* et *achaete-scute*) et les protéines myc. Cette super-famille regroupe des facteurs issus d'espèces différentes qui sont essentiels à la différenciation de grandes voies ontogéniques comme la myogenèse ou la neurogenèse.

L'alignement des séquences en acides aminés des facteurs appartenant à la super-famille des protéines myc montre une répartition similaire de certains résidus hydrophiles et hydrophobes dans les régions basique et myc (figure 3) [39]. La région basique est immédiatement suivie d'un domaine incluant la région myc, lequel est susceptible de prendre dans l'espace une conformation en « hélice-boucle-hélice » (*helix-loop-helix* = HLH) (figure 2). Les deux hélices α sont reliées entre elles par un feuillet β en forme de boucle [44]. On retrouve cette structure HLH dans certains facteurs transcriptionnels qui se lient à l'ADN. Cette représentation hypothétique montre que les acides aminés hydrophobes sont concentrés d'un seul côté de chaque hélice

et s'offrent ainsi à d'éventuelles interactions avec d'autres protéines (voir section sur la dimérisation). Selon ce modèle, les résidus lysine et arginine de la région basique sont au contact des groupements phosphates acides de l'ADN et contribuent ainsi à la fixation du facteur.

MyoD1 se lie par sa région basique au enhancer amont du gène de la créatine kinase musculaire

La première démonstration que *MyoD1* se lie directement à l'ADN et peut agir comme facteur transcriptionnel sur un gène musculaire vient des résultats obtenus avec le gène de la créatine kinase musculaire (CKM). Ce gène, qui est exprimé au cours de la transition des myoblastes en myotubes, contient deux éléments activateurs (*enhancers*), un en amont du promoteur, l'autre dans le premier intron du gène. Ces deux éléments sont essentiels pour une expression spécifique du gène au niveau des muscles. C'est l'activateur amont qui est le plus important pour l'expres-

sion spécifique du gène *CKM* et sur lequel ont porté les expériences de liaison avec *MyoD1*. Le domaine minimal de *MyoD1* requis pour l'interaction spécifique comprend les régions basique et *myc*, c'est-à-dire celles mêmes qui permettent la conversion des fibroblastes. L'analyse fonctionnelle, fondée sur les expériences *in vitro* de retardement sur gel, de compétition et d'interférence à la méthylation, a montré que *MyoD1* est capable d'interagir directement avec deux sites distincts situés dans l'activateur amont [45]. Ces deux sites G (gauche) et D (droit) contiennent tous les deux une séquence CANNTG* et apparaissent ne fixer que des facteurs à spécificité musculaire. En effet, dans les myotubes de deux lignées myogéniques (C2C12 et MM14), Buskin et Hauschka ont identifié un facteur protéique musculaire, MEF-1, qui se fixe au site D de l'activateur amont ainsi qu'à l'activateur du premier intron [46]. La liaison avec MEF-1 n'est pas détectable dans les

* C : acide cytidylique ; A : acide adénylique ; T : acide thymidylique ; G : acide guanylique ; N : un nucléotide quelconque A, C, G ou T.

RÉFÉRENCES

34. Lassar AB, Thayer MJ, Overell RW, Weintraub H. Transformation by activated *ras* or *fos* prevents myogenesis by inhibiting expression of *MyoD1*. *Cell* 1989 ; 58 : 659-67.
35. Vaidya TB, Rhodes SJ, Taparowsky EJ, Konieczny SF. Fibroblast growth factor and transforming growth factor β repress transcription of the myogenic regulatory gene *MyoD1*. *Mol Cell Biol* 1989 ; 9 : 3576-9.
36. Helbecque N, Hénichart JP. Les « doigts à zinc », éléments de reconnaissance de l'ADN. *médecine-sciences* 1988 ; 4 : 624-8.
37. Tapscott SJ, Davis RL, Thayer MJ, Cheng P-F, Weintraub H, Lassar AB. *MyoD1* : a nuclear phosphoprotein requiring a *myc* homology region to convert fibroblasts to myoblasts. *Science* 1988 ; 242 : 405-11.
38. Struhl K. Helix-turn-helix, zinc-finger, and leucine-zipper motifs for eukaryotic transcriptional regulatory proteins. *TIBS* 1989 ; 14 : 137-40.
39. Murre C, Schonleber McCaw P, Baltimore D. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, *daughterless*, *MyoD* and *myc* proteins. *Cell* 1989 ; 56 : 777-83.
40. Villares R, Cabrera CV. The *achaete-scute* gene complex of *D. melanogaster* : conserved domains in a subset of genes required for neurogenesis and their homology to *myc*. *Cell* 1987 ; 50 : 415-24.
41. Caudy M, Väassin H, Brand M, Tuma R, Jan LY, Jan YN. *Daughterless* a *Drosophila* gene essential for both neurogenesis and sex determination, has sequence similarities to *myc* and the *achaete-scute* complex. *Cell* 1988 ; 55 : 1061-7.
42. Thisse B, Stoetzel C, Gorostiza-Thisse C, Perrin-Schmitt F. Sequence of the *twist* gene and nuclear localization of its protein in endomesodermal cells of early *Drosophila* embryos. *EMBO J* 1988 ; 7 : 2175-83.
43. Klaembt C, Knust E, Tietze K, Campos-Ortega JA. Closely related transcripts encoded by the neurogenic gene complex *Enhancer of split of Drosophila melanogaster*. *EMBO J* 1989 ; 8 : 203-11.
44. Leszczynski JF, Rose GD. Loops in globular proteins : a novel category of secondary structure. *Science* 1986 : 234 : 849-55.
45. Lassar AB, Buskin JN, Lockshon D, et al. *MyoD* is a sequence-specific DNA binding protein requiring a region of *myc* homology to bind to the muscle creatine kinase enhancer. *Cell* 1989 ; 58 : 823-31.
46. Buskin JN, Hauschka SD. Identification of a myocyte nuclear factor that binds to the muscle-specific enhancer of the mouse muscle creatine kinase gene. *Mol Cell Biol* 1989 ; 9 : 2627-40.
47. Buskin JB, Lassar AB, Davis RL, Weintraub H, Hauschka SD. Is the myogenic determination factor *MyoD* identical to the myocyte-specific DNA binding factor *MEF 1*? *J Cell Biol* 1988 ; 107 : 98a.
48. Donoghue M, Ernst H, Wentworth B, Nadal-Ginard B, Rosenthal N. A muscle-specific enhancer is located at the 3' end of the myosin light-chain 1/3 gene locus. *Genes Dev* 1988 ; 2 : 1779-90.
49. Piette J, Bessereau J-L, Huchet M, Changeux J-P. Two adjacent MyoD1-binding sites regulate the expression of the acetylcholine receptor α -subunit gene. *Nature* 1990 ; 345 : 353-5.
50. Davis RL, Cheng P-F, Lassar AB, Weintraub H. The *MyoD* DNA binding domain contains a recognition code for muscle-specific gene activation. *Cell* 1990 ; 60 : 733-46.
51. Murre C, Schonleber McCaw P, Vaessin H, et al. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* 1989 ; 58 : 537-44.
52. Brennan TJ, Olson EN. *Myogenin* resides in the nucleus and acquires high affinity for a conserved enhancer element on heterodimerization. *Genes Dev* 1990 ; 4 : 582-95.
53. Benezra R, Davis RL, Lockshon D, Turner DL, Weintraub H. The protein *Id* : a negative regulator of helix-loop-helix DNA binding proteins. *Cell* 1990 ; 61 : 49-59.
54. Jones N. Transcriptional regulation by dimerization : two sides to an incestuous relationship. *Cell* 1990 ; 61 : 9-11.
55. Gros F, Montarras D, Pinset C, Mouly V. Hétérogénéité myoblastique et filiation myogénique. *médecine-sciences* 1990 ; 6 : 245-51.

myoblastes. Il a été montré que MEF-1 et MyoD1 ont des déterminants antigéniques communs, mais on ne peut pas encore conclure qu'il s'agit de la même protéine [45, 47]. Le fait que les deux sites G et D soient nécessaires à l'expression du gène *CKM* ne prouve pas que *MyoD1* joue le rôle de facteur transcriptionnel *in vivo*, d'autant plus que l'on sait que le gène *MyoD1* peut être exprimé de manière constitutive dans les myoblastes en absence des transcrits *CKM* et de ceux d'autres gènes musculaires (actines, myosines, etc.).

Le problème se pose de savoir comment *MyoD1* agit différemment dans les myoblastes et les myotubes et s'il se fixe *in vivo* aux promoteurs de gènes musculaires. Les empreintes à l'ADN des différents facteurs se liant à l'activateur amont du gène *CKM* ont été détectées *in vivo* par utilisation de la technique nouvelle de la PCR (*polymerase chain reaction*) [24]. Les interactions ADN/protéines ne sont détectées par cette technique que dans les myotubes où le gène *CKM* est activement transcrit. Dans les myoblastes, où ils sont présents, les facteurs spécifiques ou ubiquitaires restent incapables de se fixer à l'activateur, soit par inaccessibilité au niveau de la chromatine, soit qu'eux-mêmes doivent subir des modifications post-traductionnelles pour pouvoir interagir.

Puisque la région d'homologie entre les protéines *myc* contient la séquence basique de liaison à l'ADN, il était logique de penser qu'elles puissent reconnaître des sites communs. C'est le cas des protéines MyoD1, myogénine, E12 et E47 qui sont capables de se lier à un même site de l'ADN [45]. Plusieurs gènes exprimés spécifiquement au niveau du muscle présentent des séquences consensus de type *MyoD1* et les exemples suivants sont issus de publications aussi bien que de communications personnelles aimablement données par les chercheurs eux-mêmes : le *enhancer* 3' du gène des chaînes légères de la myosine de rat [48], les promoteurs de plusieurs gènes : la chaîne légère de la myosine embryonnaire de souris (F. Catala), les tropomyosines α et β de poulet (M.Y. Fiszman), la chaîne lourde de la myosine IIB de souris

(S. Takeda) et le *enhancer* 5' du gène de la sous-unité α du récepteur de l'acétylcholine de poulet [49]. Sauf dans le cas de ce dernier et du gène *CKM*, on ne sait pas encore si la fixation de *MyoD1* est requise pour l'activation de tels gènes, mais il est probable que certains cas ressemblent à la situation du gène *CKM*.

La substitution du domaine HLH de *MyoD* par la séquence analogue de la protéine *achaete-scute* de la drosophile n'a pas d'effet *in vitro* ni sur la liaison à l'ADN ni sur l'expression des gènes musculaires dans la lignée C3H10T1/2 [50]. En revanche, si le remplacement de la région basique de *MyoD* par celle d'une autre protéine *myc* n'empêche pas la fixation à l'ADN, il bloque le déclenchement du programme myogénique. Cela démontre que les domaines basique et HLH fonctionnent de manière indépendante et suggère également que la région basique doit contenir un sous-domaine nécessaire à l'activation des gènes musculaires [50].

Les facteurs spécifiques et ubiquitaires se lient au site *MyoD1* sous forme hétérodimérique

La disposition des résidus hydrophobes d'un seul côté des hélices α du domaine HLH indiquait déjà que les protéines HLH pouvaient se fixer à l'ADN sous forme de dimère en associant leurs parties hydrophobes respectives (figure 4). Les régions HLH des facteurs sont responsables de la dimérisation. Des expériences réalisées *in vitro* ont montré que la formation de tels dimères a effectivement lieu. Les facteurs ubiquitaires E12, E47 et *daughterless* (da) peuvent se combiner aux facteurs spécifiques *MyoD1* et *achaete-scute* (As-C) pour engendrer des hétérodimères (E12/*MyoD1*, E47/*MyoD1*, E12/As-C, E47/As-C, As-C/da) qui se lient spécifiquement à la même séquence κ E2 [51]. Celle-ci est localisée dans le *enhancer* du gène de la chaîne κ des immunoglobulines. Ces hétérodimères présentent — en général — une meilleure affinité pour le site κ E2 que les homodimères (*MyoD1/MyoD1*, E47/E47). Alors que certaines combinaisons d'hétérodimères (ex. : myogénine/E12) présentent une grande

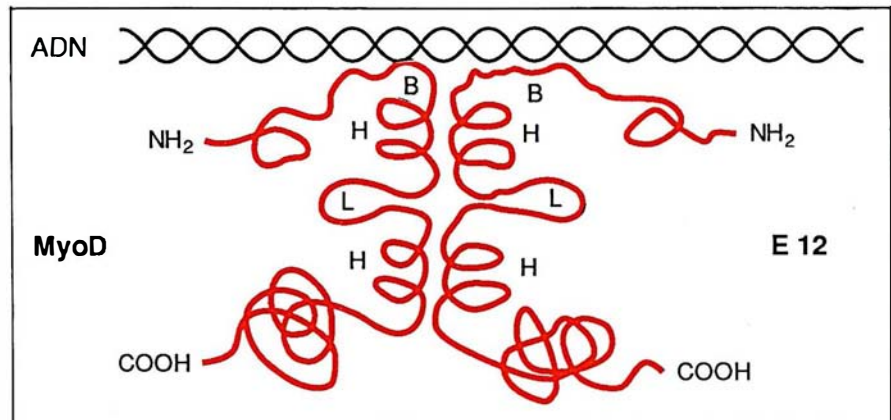


Figure 4. Schéma hypothétique de l'interaction entre l'hétérodimère *MyoD1/E12* et le « *enhancer* » amont du gène de la créatine kinase musculaire (*CKM*). La disposition des domaines HLH permet leur interaction au niveau des résidus hydrophobes. La fixation à l'ADN se fait au niveau des régions basiques (B).

affinité pour l'ADN [52], d'autres restent incapables de former un complexe avec l'ADN comme les hétérodimères utilisant la protéine N-myc. On ne sait pas encore si un hétérodimère qui se lie *in vitro* à l'ADN est actif *in vivo*, mais cela devient bien sûr une base de travail.

Récemment, un nouveau facteur ubiquitaire, Id (*inhibitor of DNA binding*), a été isolé sous forme d'ADNc [53]. Ce facteur, qui possède le domaine HLH, est capable de s'associer avec *MyoD* ou E12, mais du fait qu'il lui manque la région basique, l'hétérodimère ainsi formé reste incapable de se fixer à l'ADN. Le facteur Id (*inhi-*

bitor of DNA binding) est un facteur de régulation négative de la myogenèse qui inhibe en *trans* l'activation du gène *CKM*.

Conclusions

La nature hétérodimérique du complexe protéique se liant à l'ADN offre à la régulation de l'expression génétique une combinatoire bien plus riche que si le rapport était toujours d'un facteur par site de liaison. C'est l'association particulière de deux facteurs dans des conditions physiologiques données qui déterminerait la spécificité transcriptionnelle, le com-

MyoD1 :	souris/xénope	73 %
	souris/poulet	80 %
	souris/homme	90 %
	poulet/caille	94 %
Myogénine :	souris/rat	92 %
	rat/homme	85 %
MRF4/Myf-6/Herculine :	souris/rat	99 %
	souris/homme	95 %
	rat/homme	95 %

plexe pouvant être formé d'un facteur spécifique associé à un facteur ubiquitaire. Cette relation binaire entre les facteurs pourrait rendre compte à la fois de la spécificité tissulaire et de l'économie de la machinerie cellulaire, un même facteur ubiquitaire jouant des rôles différents dans des tissus distincts. D'autres familles de protéines nucléaires (Jun, Fos, ATF/CREB) se lient à l'ADN uniquement sous forme d'homodimère ou d'hétérodimère et montrent ainsi que la dimérisation des facteurs transcriptionnels joue un rôle essentiel dans la régulation de l'expression génétique [54].

On a vu qu'une courte portion conservée d'acides aminés (60), comprenant les régions basique et myc, est douée d'un grand nombre de propriétés : la conversion des fibroblastes en myoblastes, la fixation à l'ADN et une structure HLH permettant la formation d'hétérodimères. A quelques exceptions près, l'homologie se borne à ce domaine particulier lorsque l'on compare entre elles les séquences protéiques des quatre facteurs connus. Pourtant la conservation est bien plus étendue si on compare deux gènes homologues dans différentes espèces (*Tableau II, p. précédente*). On constate que le pourcentage d'homologie est presque toujours supérieur à 80 % et que même pour deux espèces (souris et xénope) ayant divergé depuis très longtemps, leurs séquences MyoD1 restent homologues à 73 %. Il n'y a donc pas seulement les régions basique et myc qui sont importantes dans la fonction des facteurs et les autres domaines conservés entre espèces pourraient éventuellement intervenir au niveau des interactions protéine/protéine, puisqu'ils ont été sélectionnés par l'évolution.

L'ordre dans lequel se fait l'activation des quatre gènes régulateurs du type MyoD1, leur importance hiérarchique au cours des événements conduisant à la formation des fibres musculaires ainsi que leur rôle exact respectif commencent à peine d'être entrevus. Il n'est pas exclu qu'à l'intérieur de la population de myoblastes déterminés mais non différenciés, il existe plusieurs classes différentes de cellules susceptibles d'exprimer des facteurs de régulation distincts [55] ■

Summary

Specific myogenic regulation factors

A differential cloning strategy based on the unique properties of the fibroblast C3H10T1/2 cell line, resulted in 1987 in the isolation and the characterization of the first cDNA clone coding for the myogenic factor MyoD1. Since this date, three additional myogenic regulatory sequences have been isolated as cDNAs and characterized : myogenin, Myf-5 and MRF4. Each sequence is able to convert the fibroblast cell line to myoblasts which can differentiate into myotubes. These regulatory proteins share a common domain of approximately 60 amino acids including both a basic region, allowing fixation to the DNA, and a helix-loop-helix domain found in other regulatory proteins such as myc. Experiments demonstrate that these proteins can act as transcriptional regulators of muscle expression. Their expression is specific to skeletal muscle cells and they have different patterns of expression.

TIRÉS A PART

S. Alonso.

Le Prix International de la Fondation SPA a été décerné au Professeur Pierre Corvol, Professeur au Collège de France, et à son équipe (les Docteurs François Alhenc-Gelas et Florent Soubrier) pour leur recherche intitulée « Biochimie et biologie moléculaire de la rénine et de l'enzyme de conversion de l'angiotensine I ». Le jury de ce prix est composé de scientifiques belges se réunissant sous l'égide du « Fonds National de la Recherche Scientifique ». Il délivre des prix dont le domaine de la recherche en cardiologie, rhumatologie et troubles rénaux.