

## **Fonction des gènes Pax : synergie de liaison à l'ADN entre le domaine paired et l'homéodomaine**

Claude Desplan

Les gènes du développement *Pax*, codant pour des régulateurs de transcription, sont caractérisés par la présence d'un domaine paired (PD) capable de se lier à l'ADN. Celui-ci comporte deux motifs hélice-tour-hélice, appelés PAI et RED. Certaines protéines Pax possèdent également un troisième domaine de liaison à l'ADN, l'homéodomaine (HD). Des études *in vitro* et *in vivo* chez la drosophile pourraient expliquer, par une combinatoire de ces divers domaines de liaison à l'ADN, la très grande spécificité fonctionnelle des protéines Pax, alors que ces domaines reconnaissent des séquences d'ADN très semblables.

L'un des plus importants succès de la biologie du développement, qui vient d'être couronnée par l'attribution en 1995 du prix Nobel de Médecine et de Physiologie à trois généticiens de la drosophile, a été l'identification des gènes de développement chez la drosophile [1, 2]. La présence dans ces gènes de motifs fonctionnels très fortement conservés comme l'homéodomaine (HD) ou le domaine paired (PD), deux domaines de liaison à l'ADN, a aussi permis d'isoler les analogues vertébrés des gènes de développement de drosophile. Ces gènes ont souvent des fonctions comparables, et sont des outils puissants qui ont permis à la biologie du développement des vertébrés d'entrer dans une ère

nouvelle et de passer au niveau moléculaire [3]. La conservation entre la drosophile et la souris de la fonction et de l'organisation chromosomique des gènes *Hox*, qui contiennent un HD [4], et l'importance des propriétés parallèles observées entre les fonctions des gènes *Pax* [5, 6], caractérisés par la présence d'un PD [7, 8], représentent des illustrations spectaculaires de ces succès. Le sujet de cet article concerne une classe de gènes dont les produits contiennent à la fois un HD et un PD et sont donc appelés les gènes *Pax/HD*. Ceux-ci sont représentés par *Pax-3/Pax-7* et *Pax-6* chez les vertébrés, par *paired* et *eyeless* (*Pax-6*) chez la drosophile [9].

La présence d'un domaine conservé dans un gène permet l'identification

### ADRESSE

C. Desplan : *associate professor*. Howard Hughes Medical Institute, The Rockefeller University, 1230 York Avenue, New York NY 10021-6399, États-Unis.  
E-mail : desplan@rockvax.rockefeller.edu

## RÉFÉRENCES

1. Nüsslein-Volhard C, Wieschaus E. Mutations affecting segment number and polarity in *Drosophila*. *Nature* 1980; 287: 795-801.
2. Deutsch J, Lamour-Isnard C, Lepesant J. Le Prix Nobel 95 attribué à Ed Lewis, Christiane Nüsslein-Volhard et Eric Wieschaus : la reconnaissance de la génétique du développement. *Med Sci* 1995; 11: 1625-8.
3. Walther C, Guenet JL, Simon D, Deutsch U, Jostes B, Goulding MD, Plachov D, Balling R, Gruss P. Pax: a murine multigene family of paired box-containing genes. *Genomics* 1991; 11: 424-34.
4. Duboule D, Dolle P. The structural and functional organization of the murine *HOX* gene family resembles that of *Drosophila* homeotic genes. *EMBO J* 1989; 8: 1497-505.
5. Gruss P, Walther C. Pax in development. *Cell* 1992; 69: 719-22.
6. Noll M. Evolution and role of Pax genes. *Curr Op Genet Dev* 1993; 3: 595-605.
7. Bopp D, Burri M, Baumgartner S, Frigerio G, Noll M. Conservation of a large protein domain in the segmentation gene *paired* and in functionally related genes in *Drosophila*. *Cell* 1986; 47: 1033-49.
8. Frigerio G, Burri M, Bopp D, Baumgartner S, Noll M. Structure of the segmentation gene *paired* and the *Drosophila* PRD gene set as part of a gene network. *Cell* 1986; 47: 735-46.
9. Babinet C. Une famille de gènes du développement : les gènes Pax. *Med Sci* 1993; 9: 87-91.
10. Desplan C, Theis J, O'Farrell PH. The *Drosophila* developmental gene, *engrailed*, encodes a sequence-specific DNA binding activity. *Nature* 1985; 318: 630-5.
11. Treisman J, Harris E, Desplan C. The paired Box encodes a second DNA-binding domain in the paired homeo domain protein. *Genes Dev* 1991; 5: 594-604.
12. Treisman J, Gönczy P, Vashishtha M, Harris E, Desplan C. A single amino acid can determine the DNA binding specificity of homeodomain proteins. *Cell* 1989; 59: 553-62.
13. Treisman J, Harris E, Wilson D, Desplan C. The Homeodomain: a new face for the helix-turn-helix? *BioEssays* 1992; 14: 145-50.
14. Wilson D, Sheng G, Lecuit T, Dostatni N, Desplan C. Cooperative dimerization of paired class homeodomains on DNA. *Genes Dev* 1993; 7: 2120-34.

d'analogues. Elle donne aussi une indication de la fonction du produit de ces gènes. Le HD et le PD sont tous deux des domaines de liaison à l'ADN, permettant tous deux l'activation de la transcription dans des cellules en culture [10, 11], ce qui indique que les protéines qui les contiennent sont des facteurs de transcription multifonctionnels. Dans la mesure où existent des protéines qui contiennent seulement un HD, ou seulement un PD, la présence de ces deux domaines dans la même protéine avec des configurations très conservées d'une protéine à l'autre représente un sujet d'étude passionnant : ces protéines sont-elles capables d'agir différemment lorsqu'elles se lient à des promoteurs par l'intermédiaire de l'un ou de l'autre de leurs deux domaines; ou bien y a-t-il une interaction entre ces domaines, comme le suggèrent la configuration conservée et la présence de caractères particuliers aux HD coexistants avec un PD et, *vice versa*, des PD coexistants avec un HD? Par exemple, le déterminant le plus important de la spécificité des HD est le résidu à la position 50 [12]. Or, cet acide aminé est toujours une sérine lorsque un HD est associé à un PD, et réciproquement, si ce résidu est une sérine, cela signifie qu'un PD est présent en amont de cet HD (*figure 1*) [13]. En outre, ce type de HD appartient à une classe de protéines aux propriétés particulières (classe Prd/Pax) qui sont les seules à être capables de former des dimères homotypiques ou hétérotypiques avec d'autres membres de cette même classe [14].

### Dimérisation des HD de la classe Prd/Pax

Depuis sa découverte en 1984, le HD a été l'un des motifs de liaison à l'ADN le mieux étudié. La structure tridimensionnelle obtenue par RMN et cristallographie aux rayons X a permis de décrire comment ce motif hélice-tour-hélice entre en contact avec des bases dans le sillon majeur de l'ADN, et comment un bras aminoterminal pénètre dans le sillon mineur [15, 16]. Ensemble, ces deux motifs reconnaissent le motif TAAT caractéristique des sites de liaison de la plupart des HD (*figure 2*) [13]. Le reste du site de liaison, c'est-à-dire les

deux bases qui suivent le TAAT, sont reconnues par le résidu 50 présent au milieu de l'hélice de reconnaissance (*figure 2*). Ce résidu n'établit pas en fait de contact direct avec les bases, et son rôle est resté obscur jusqu'à très récemment lorsqu'est apparu, dans la structure cristalline d'un complexe entre un HD de la classe Prd et son site de liaison palindromique, que ce contact est indirect et fait intervenir plusieurs molécules d'eau [17].

Cette reconnaissance par un monomère du HD est peu spécifique, puisqu'un grand nombre de HD fortement reliés par leur séquence (en particulier au niveau des acides aminés impliqués dans la reconnaissance de l'ADN) reconnaissent tous les mêmes 6 paires de bases. Le HD doit donc employer d'autres stratégies pour augmenter sa spécificité fonctionnelle et effectuer la fonction spécifique de la protéine dont il est le domaine essentiel. Les membres de la classe Prd/Pax de HD (*figure 1*) semblent avoir trouvé deux solutions pour résoudre le problème de spécificité : la première, une stratégie commune à de nombreux domaines de liaison à l'ADN, est de se lier à l'ADN comme dimères; l'autre est d'ajouter un second domaine de liaison à l'ADN (le PD) au HD. Tous les membres de la classe de HD Prd/Pax, que la protéine contient ou non un PD (c'est-à-dire quel que soit le résidu 50), reconnaissent des séquences palindromiques appelées P2 ou P3 faites de deux séquences TAAT inversées et séparées par deux ou trois paires de bases [14]. La nature de ces paires de bases varie avec le résidu présent à la position 50. La dimérisation ne se produit pas en solution. C'est, en fait, une liaison extrêmement coopérative qui fait intervenir des changements de conformation réciproques de l'ADN et du HD [17]. La *figure 3* décrit cette adaptation réciproque : en solution, l'ADN est en forme B, le HD est un monomère dont la partie aminoterminal est désordonnée. Lorsqu'un monomère du HD entre en contact avec l'ADN, sa partie aminoterminal acquiert une conformation particulière en pénétrant dans le sillon mineur, alors que l'ADN se tord d'un angle de 30° sous l'influence du HD. Ces deux phéno-

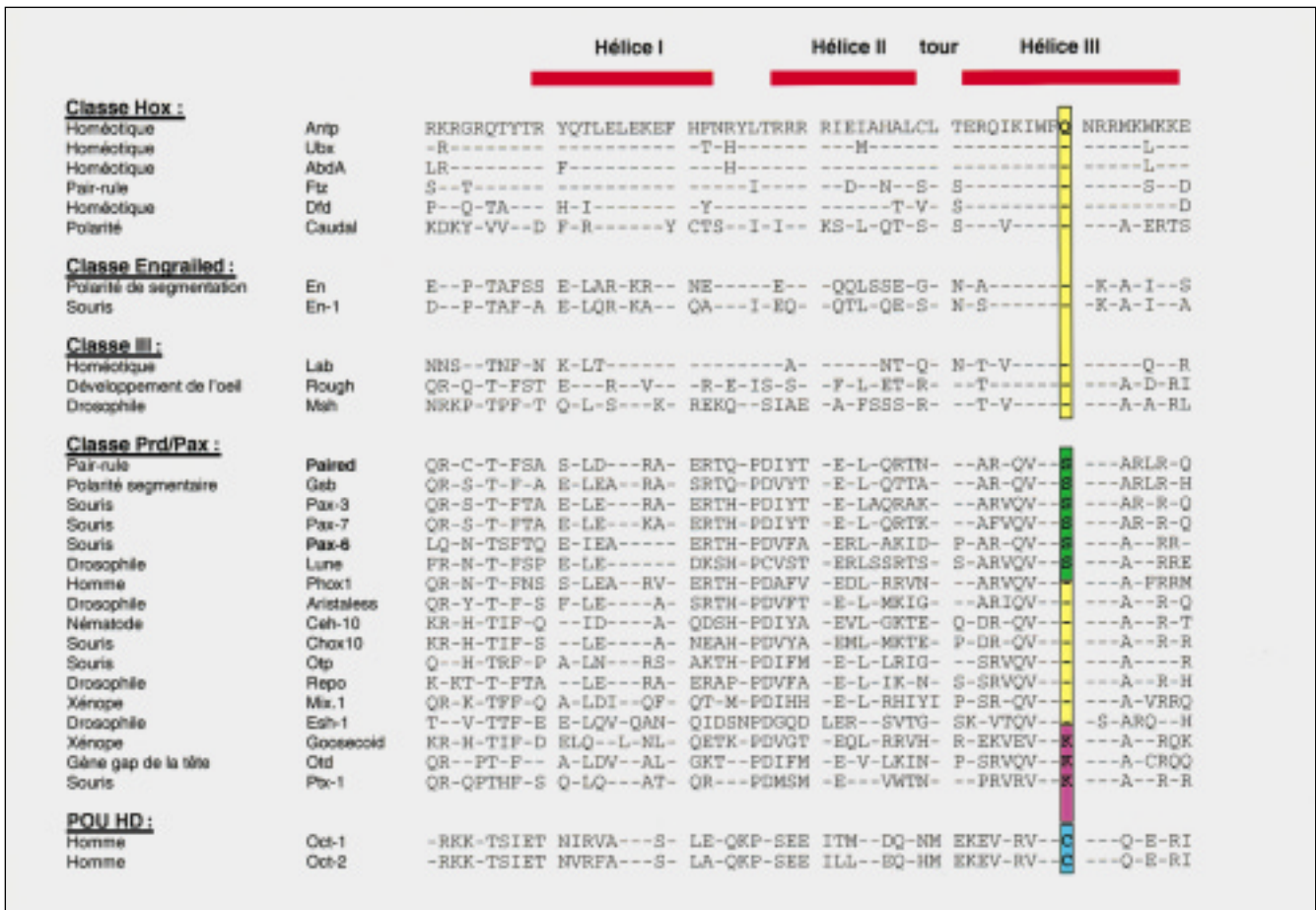


Figure 1. **Alignement de séquences représentatives de chaque groupe d'homéodomains.** Les séquences sont alignées par rapport à celle de Antennapedia, et seuls les acides aminés qui diffèrent sont indiqués. La position des hélices est indiquée au-dessus. Lorsque l'espèce n'est pas mentionnée, il s'agit de séquence de drosophile. La classe Prd/Pax est la seule qui présente trois résidus possibles à la position 50: S (Ser) (lorsque le HD est associé à un PD, dans les protéines Pax); Q, (Gln) dans la majorité des cas; K (Lys) dans un groupe de protéines possédant une spécificité de liaison à l'ADN particulière. Les références pour les séquences peuvent être trouvées sur le site world wide web « homeobox homepage » : <http://copan.bioz.unibas.ch/homeo.html>.

mènes peuvent en fait être observés avec tous les HD, quelle que soit leur classe [16, 17]. Cependant, ce n'est que dans le cas des HD de la classe Prd/Pax que ce changement de conformation permet à un deuxième HD d'entrer dans le complexe et d'établir un contact coopératif avec l'autre monomère. Cette coopérativité dépend aussi de la nature du résidu 50 [14]; elle peut être extrêmement élevée, multipliant l'affinité par environ 300. Enfin, le même résidu peut aussi agir sur l'espace-ment entre les deux TAAT, qui peut être réduit à deux paires de bases si

une sérine occupe cette position, comme c'est le cas des HD des protéines Pax qui contiennent aussi un PD [14].

### Reconnaissance d'un long site par un monomère du domaine paired

Si la dimérisation du HD apporte une spécificité importante aux protéines de la classe Prd/Pax, la présence d'un PD dans les protéines Pax accentue encore cette capacité de reconnaître de longues séquences.

Des études *in vitro* ont montré que le PD est, en fait, lui-même un domaine bipartite dont la partie aminotermi-nale (domaine PAI) joue un rôle fon-damental, alors que la partie carboxy-terminale (domaine RED) semble jouer un rôle moins important [18-20]. Cette situation rappelle celle du domaine POU pour lequel le HD est nécessaire à la liaison du domaine spécifique POU [21]. La structure cristalline d'un complexe entre le PD de Paired et son site de liaison de 15 paires de bases [22] a confirmé cette organisation bipartite : le PD est composé de deux sous-domaines

## RÉFÉRENCES

15. Qian YQ, Billeter M, Otting G, Muller M, Gehring WJ, Wuthrich K. The structure of the antennapedia homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors. *Cell* 1989; 59: 573-80.
16. Kissinger CR, Liu B, Martin-Blanco E, Kornberg TB, Pabo CO. Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Angstrom resolution: a framework for understanding homeodomain-DNA interactions. *Cell* 1990; 63: 579-90.
17. Wilson DS, Guenther B, Desplan C, Kuriyan J. High resolution crystal structure of a paired (Pax) class cooperative homeodomain dimer on DNA. *Cell* 1995; 82: 709-19.
18. Jun S, Desplan C. Cooperative interaction between homeodomain and paired domain. *Development* 1996; 122: 2639-50.
19. Czerny T, Schaffner G, Busslinger M. DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site. *Genes Dev* 1993; 7: 2048-61.
20. Epstein JA, Glaser T, Cai J, Jepeal L, Walton DS, Maas RL. Two independent and interactive DNA-binding subdomains of the Pax-6 paired domain are regulated by alternative splicing. *Genes Dev* 1994; 8: 2022-34.
21. Klemm JD, Rould MA, Aurora R, Herr W, Pabo CO. Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. *Cell* 1994; 77: 21-32.
22. Xu W, Rould M, Jun S, Desplan C, Pabo C. Crystal structure of a paired domain-DNA complex at 2.5 Angstrom resolution reveals structural basis for Pax developmental mutations. *Cell* 1995; 80: 639-50.
23. Hill R, VanHeyningen V. Mouse mutations and human disorders are paired. *Trends Genet* 1992; 8: 119-20.
24. Bertuccioli C, Fasano L, Jun S, Sheng G, Desplan C. *In vivo* requirement for the paired domain and the homeodomain of the paired segmentation gene product. *Development* 1996; 122: 2673-85.
25. Fujioka M, Miskiewicz P, Raj L, Gullledge A, Weir M, Goto T. *Drosophila* paired regulates late *even-skipped* expression through a composite binding site for the paired domain and the homeodomain. *Development* 1996; 122: 2697-707.

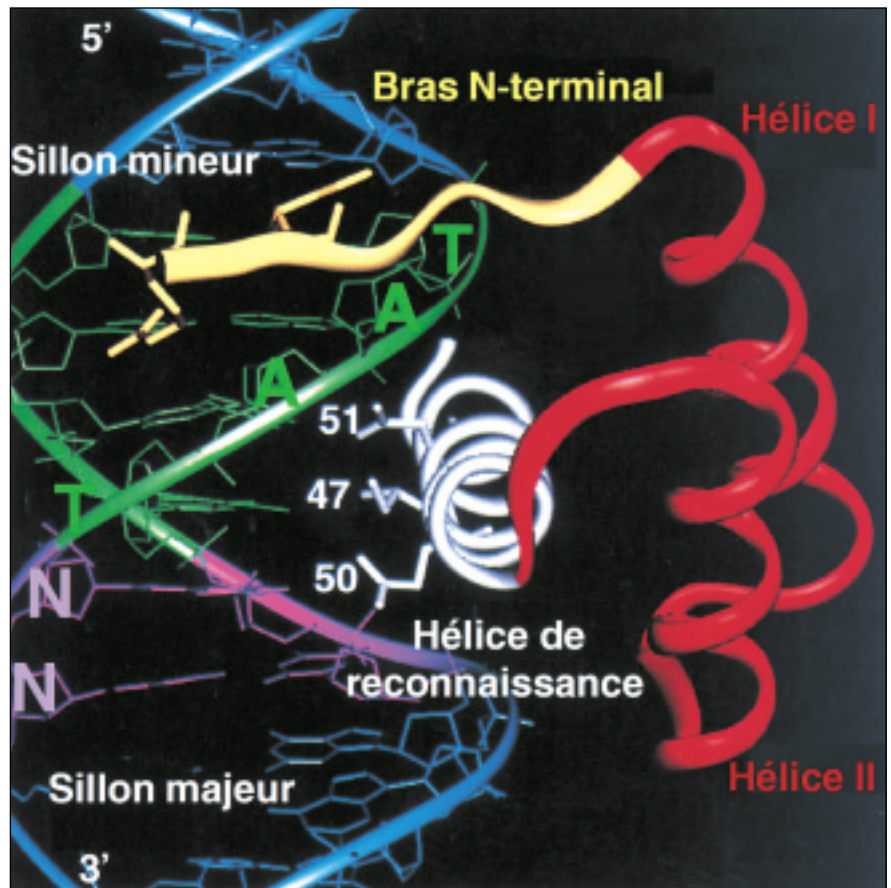


Figure 2. **Reconnaissance de l'ADN par l'homéodomaine.** Deux régions du HD contribuent à la reconnaissance des sites TAATNN. La séquence TAAT (en vert) est contactée dans le sillon majeur par deux résidus de l'hélice de reconnaissance qui sont fortement conservés chez tous les HD (positions 47 et 51), ainsi que par le bras aminoterminal, lui aussi très conservé. Le résidu 50 entre en contact avec les deux paires de bases qui suivent le TAAT (NN en violet) par l'intermédiaire de molécules d'eau. Le site total est de 6 paires de bases.

(PAI et RED) qui, tous deux, adoptent une conformation hélice-tour-hélice [18, 22]. Le PAI entre en contact avec l'ADN dans le sillon majeur par son hélice de reconnaissance et, dans le sillon mineur, par deux structures nouvelles: un tour  $\beta$ , ainsi qu'une longue séquence d'attachement entre les domaines PAI et RED; toutes deux rentrent profondément dans le sillon mineur et établissent de nombreux contacts spécifiques, qui expliquent la séquence étonnamment longue reconnue par un monomère de PD [22]. Le domaine RED n'entre pas en contact avec l'ADN dans cette structure, mais

il est fort probable qu'il le fasse par son motif hélice-tour-hélice dans d'autres PD.

L'une des caractéristiques les plus intéressantes des gènes *Pax* est le fait que de nombreuses mutations de développement affectent ces gènes. Les mutations dans *Pax-3* sont responsables du syndrome de Waardenburg chez l'homme ou du phénotype *Splotch* chez la souris (*m/s n°1*, vol. 11, p. 133); celles de *Pax-6* correspondent aux phénotypes *Aniridia* chez l'homme et *Small eye* chez la souris (*m/s n°2*, vol. 8, p. 181) (voir [9, 23] pour des synthèses). Ces mutations sont haplo-insuffisantes dans la

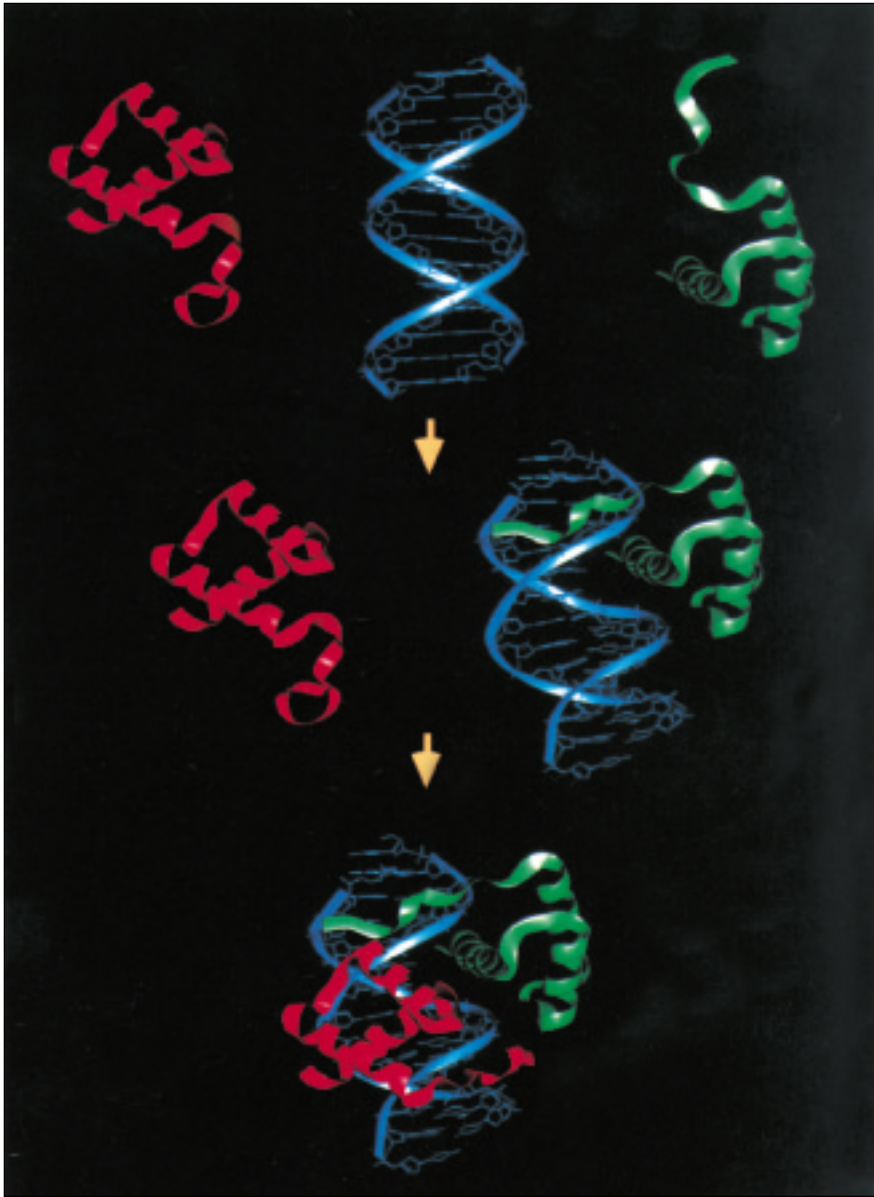


Figure 3. **Séquence des événements structuraux accompagnant la liaison d'un HD de la classe Prd/Pax à un site palindromique P3.** En solution, l'ADN est sous forme B alors que le HD est un jeu de trois hélices comprenant un motif hélice-tour-hélice. Le bras aminoterminal n'est pas structuré en solution. Au contact de l'ADN, ce bras acquiert une conformation particulière qui lui permet de rentrer dans le sillon mineur et de reconnaître les bases du motif TAAT. Cette liaison a aussi pour conséquence de tordre l'ADN avec un angle de 30°. Ces deux événements permettent au premier HD d'entrer en contact avec un deuxième qui se lie de manière très coopérative.

mesure où le phénotype se manifeste à l'état hétérozygote, ce qui a permis l'identification de très nombreux allèles. Or, la grande majorité des mutations ponctuelles (qui se comportent comme des pertes totales de fonction) touchent des résidus qui sont à l'interface entre le domaine

PAI et l'ADN [22] (quelques autres mutations affectent le HD). Cela signifie, d'une part, que le domaine PAI joue un rôle prépondérant dans la reconnaissance de l'ADN (plus que le RED qui n'est touché par aucune mutation) et que, dans les molécules qui ont à la fois un PD et

un HD, telles que Pax-3 et Pax-6, la perte d'activité de l'un des domaines de liaison à l'ADN se traduit par une perte de fonction totale de la molécule (bien que l'autre domaine soit toujours fonctionnel). Ces observations sont confirmées par les résultats obtenus *in vivo* avec la protéine Paired (*voir plus bas*) [24]. Cela suggère donc que ces molécules utilisent en même temps le PD et le HD pour la plupart de leurs fonctions. La synergie de liaison à l'ADN observée entre le PD et le HD permet d'expliquer ces résultats [11, 18].

### Coopérativité entre domaine paired et homéodomaine

Les premières observations faites sur la protéine Paired indiquaient que le PD et le HD étaient tous deux nécessaires pour la liaison à des sites de faible affinité [11]. Depuis, cette coopérativité des deux domaines pour la liaison à l'ADN a été confirmée et généralisée à d'autres molécules Pax/HD [18]. En utilisant les techniques de sélections successives suivies par des amplifications par PCR (de type SELEX) [14], un site optimal de liaison très long, reconnu à la fois par le PD et le HD, a été identifié à partir d'une bibliothèque d'oligonucléotides dont la séquence est aléatoire. Il s'agit de la juxtaposition du site de liaison du domaine PAI (11 paires de bases) et d'un site de liaison monomérique du HD (TAAT) en directions opposées, appelé PH0 (espacement de 0 paire de bases entre PD et HD) [18] (*figure 4*). L'ajout de deux paires de bases ou plus entre ces deux sites détruit complètement la coopérativité de la liaison, de même que l'inversion ou la mutation de l'un des sites. Ce long site (de près de 20 paires de bases) apparaît donc comme ayant une très forte affinité et une spécificité inégalée pour un site reconnu par un monomère. Lorsque l'on poursuit la sélection PCR, le site de liaison du HD devient palindromique, indiquant que deux molécules Paired peuvent se fixer avec une spécificité plus grande encore et reconnaissent un site de 35 paires de bases ! Récemment, un site qui montre une très grande similitude avec le site PH0 a été identifié dans le promoteur d'un

## RÉFÉRENCES

26. Kilchherr F, Baumgartner S, Bopp D, Frei E, Noll AM. Isolation of the *paired* gene of *Drosophila* and its spatial expression during early embryogenesis. *Nature* 1986; 321: 493-9.

27. Benedyk MJ, Mullen JR, DiNardo S. *Odd-paired*: a zinc finger pair-rule protein required for the timely activation of *engrailed* and *wingless* in *Drosophila* embryos. *Genes Dev* 1994; 8: 105-77.

28. DiNardo S, O'Farrell PH. Establishment and refinement of segmental pattern in the *Drosophila* embryo: spatial control of *engrailed* expression by pair rule genes. *Genes Dev* 1987; 1: 1212-25.

29. Ingham PW, Baker NE, Martinez-Arias AA. Regulation of segment polarity genes in the *Drosophila* blastoderm by Fushi Tarazu and Even Skipped. *Nature* 1988; 331: 73-5.

30. Mismar D, Rubin GM. Definition of cis-acting elements regulating expression of the *Drosophila melanogaster ninaE* opsin gene by oligonucleotide-directed mutagenesis. *Genetics* 1989; 121: 77-87.

31. Fortini ME, Rubin GM. Analysis of cis-acting requirements of the *Rh3* and *Rh4* genes reveals a bipartite organization to *rhodopsin* promoters in *Drosophila melanogaster*. *Genes Dev* 1990; 4: 444-63.

32. Halder G, Callaerts P, Gehring WJ. Induction of ectopic eyes by targeted expression of the *eyeless* gene in *Drosophila*. *Science* 1995; 267: 1788-92.

33. Salvini-Plawen LV, Mayr E. On the evolution of photoreceptors and eyes. *Evol Biol* 1977; 10: 207-63.

34. Zuker CS. On the evolution of eyes: would you like it simple or compound? *Science* 1994; 265: 742-3.

35. Grindley J, Davidson D, Hill R. The role of Pax-6 in eye and nasal development. *Development* 1995; 121: 1433-42.

36. Carrière C, Plaza S, Martin P, Quatanens B, Bailly M, Stehelin D, Saule S. Characterization of quail Pax-6 (Pax-QNR) proteins expressed in the neuroretina. *Mol Cell Biol* 1993; 13: 7257-66.

37. Plaza S, Dozier C, Martin P, Saule S. Coup d'œil sur PAX-6. *Med Sci* 1996; 12: 921-8.

38. Zhang Y, Emmons SW. Specification of sense-organ identity by a *Caenorhabditis elegans* Pax-6 homologue. *Nature* 1995; 376: 55-9.

39. Chisholm AD, Horvitz HR. Patterning of the *Caenorhabditis elegans* head region by the Pax-6 family member *vab-3*. *Nature* 1995; 376: 52-5.

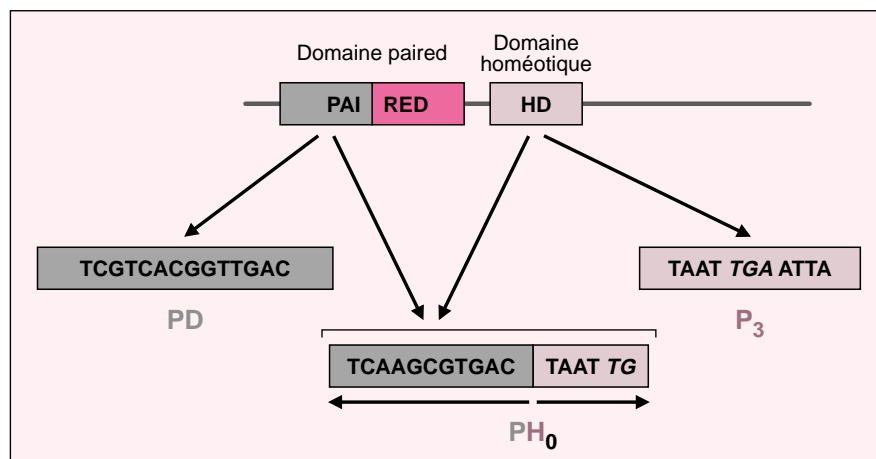


Figure 4. **Possibilités de liaison à l'ADN des protéines de la classe Pax/HD.** Des protéines comme *Paired* ou *Pax-6* contiennent à la fois un PD et un HD. Chacun de ces domaines peut se lier à des sites très spécifiques (sites PD ou P3). Le PD et le HD peuvent aussi coopérer pour se lier aux sites PH0 qui représentent la juxtaposition sans espacement d'un site PD et d'un motif TAAT en orientations inversées. La coopérativité est de l'ordre de 10 fois.

gène cible de *Paired*, *even-skipped*, et il a été démontré que l'intégralité de ce site était nécessaire pour la régulation de ce gène par *Paired* [25]. Il est à noter que, même si le site PH0 est extrêmement spécifique, toutes les molécules Pax/HD reconnaissent pratiquement la même séquence, puisqu'elles possèdent des HD pratiquement identiques et des PD extrêmement semblables (sauf pour *Pax-6* qui présente une différence significative avec les autres PD dans la reconnaissance de l'ADN). Cette spécificité n'est donc pas le fondement des différentes fonctions de ces protéines.

### Dissection in vivo de la fonction de la protéine Pax/HD *Paired*

Toutes les observations décrites ci-dessus représentent une analyse *in vitro* des propriétés de liaison à l'ADN des motifs PAI, RED et HD. La présence de ces trois motifs présents dans des configurations tout à fait semblables dans les produits d'une classe de gènes importants nécessite que l'analyse soit poursuivie pour une compréhension des mécanismes d'action *in vivo*. Le choix pour cette analyse s'est porté sur le produit du gène *paired*, qui est très bien caractérisé génétiquement. *paired* est un gène de segmentation *pair-rule* de la

*Drosophila* [26] dont le phénotype *prd*<sup>-/-</sup> est une perte alternée de segments (d'où son nom de *pair-rule*) (figure 5A). Il se situe en aval des gènes *gap* et est exprimé en bandes alternées dans les segments impairs (figure 5A). Il agit en amont des gènes de polarité segmentaire dont il règle l'expression. Le phénotype cuticulaire de perte de segments est le résultat de l'absence d'expression, dans ces segments, des gènes cibles *engrailed*, *wingless* et *gooseberry* [27-29] (figure 5B). Il est donc possible d'analyser le phénotype *prd*<sup>-/-</sup> par ses conséquences sur la cuticule de la larve mais aussi, plus directement, sur le profil d'expression de ses gènes cibles qui, au moins pour certains, sont probablement des cibles transcriptionnelles directes.

Pour cette analyse, un système de sauvetage du phénotype *prd*<sup>-/-</sup> a été mis en place [24]: il est possible d'introduire dans une mouche un transgène contenant toute l'information pour reproduire le profil d'expression normal de *prd*, et d'attacher cette longue séquence (de plus de 10 kb) à une cassette contenant la séquence codante de *prd* contenant des modifications introduites par mutagenèse dirigée dans l'un ou l'autre des domaines fonctionnels de la molécule [24]. Si la séquence introduite est normale, le sauvetage est complet et les cibles de *prd* sont exprimées

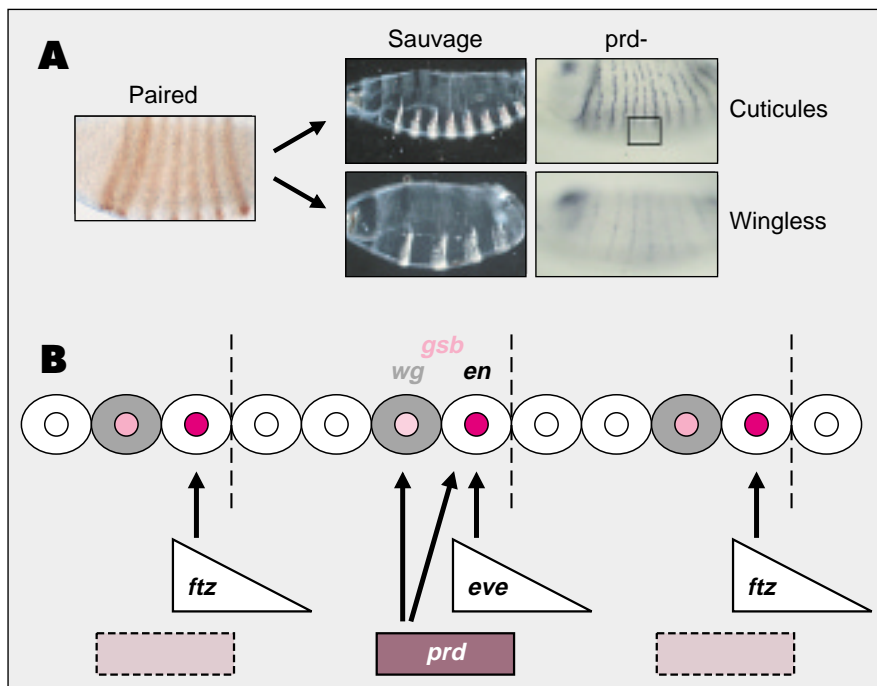


Figure 5. **A.** *paired* est un gène pair-rule. La protéine est synthétisée dans l'embryon de drosophile dans chaque segment alterné. Lorsque le gène est sauvage, la cuticule et les gènes cibles de *Paired* sont normaux. *Wingless*, par exemple, est synthétisée en bandes dans chacun des segments. En l'absence de la fonction *paired*, la cuticule n'a que la moitié du nombre de segments, qui est la conséquence de la perte d'expression de *wingless* dans les segments manquants. **B.** Diagramme de la fonction de *paired*. La région entourée d'un carré en A est représentée en B par 4 cellules dans chaque segment. *paired* est exprimé à haut niveau dans deux cellules des segments impairs où il contrôle l'expression de *engrailed* (dans la cellule la plus postérieure du segment), *wingless* (dans la cellule immédiatement antérieure) et *gooseberry* (dans les deux cellules). L'expression de *engrailed* nécessite la présence d'un autre gène pair-rule, *even-skipped*. Trois cibles de *Paired* sont donc réglées de manières différentes.

normalement, la cuticule est normale, et les larves éclosent pour donner naissance à des adultes viables, cela en dépit de l'absence du gène endogène. En revanche, lorsque le sauvetage est effectué avec des phases codantes contenant une mutation ponctuelle dans le PD ou dans le HD, aucun sauvetage n'est observé, à aucun niveau [24]. De façon plus intéressante encore, lorsque deux constructions contenant l'une une mutation qui inactive le PD, l'autre une mutation qui inactive le HD sont exprimées conjointement dans la même mouche, aucun signe de sauvetage n'est détecté, impliquant que les deux domaines doivent être présents dans la même molécule pour que celle-ci soit active. Cela signifie que toutes les cibles connues de *prd*

doivent contenir dans leur promoteur des séquences de type PH0 définies ci-dessus [24]. Ces séquences n'ont pas encore été trouvées dans les promoteurs de *engrailed*, *wingless* et *gooseberry* mais, comme indiqué plus haut, un site de ce type est essentiel pour l'expression tardive du gène *even-skipped*, une cible directe de *Paired* [25].

Cette étude a aussi permis d'analyser le rôle du domaine RED qui n'entre pas en contact avec l'ADN dans la structure cristalline de la protéine *Paired* attachée à son site de liaison. Un gène *prd* dont la séquence qui code pour le domaine RED a été éliminée est toujours capable de remplir sa fonction normalement, ce qui suggère que ce domaine n'est pas essentiel à la fonction de *Paired*. Il

est pourtant certain que ce même domaine joue un rôle important pour la fonction des produits d'autres gènes *Pax* tels que ceux de *Pax-5* (qui ne contient qu'un PD et pas de HD) et de *Pax-6* qui contient à la fois un PD et un HD [19, 20].

### Le domaine *paired* et l'homéodomain fonctionnent-ils indépendamment ?

La question qui reste donc à résoudre à ce point est celle de savoir s'il existe des situations dans lesquelles des protéines comme *Paired* ou *Pax-6* agissent en utilisant seulement le PD, ou le HD (figure 6). Il a été démontré qu'une protéine *Paired* dont le HD a été muté est encore capable d'activer *in vivo* la transcription d'un gène rapporteur artificiel contenant des cibles du PD oligomérisées, démontrant que cette molécule peut agir dans l'embryon, même s'il s'agit de constructions artificielles [24]. Cependant, l'absence de cibles moléculaires caractérisées pour *Paired* (sauf le site PH0 décrit plus haut dans le promoteur de *even-skipped* [25], *Pax-3* ou *Pax-6*, n'a pas permis l'identification de gènes qui pourraient être activés par l'un des deux domaines seulement. La seule possibilité qui existe à ce jour est en fait extrêmement séduisante : il a été trouvé dans le promoteur de tous les gènes de rhodopsine et d'autres gènes impliqués dans la photoréception chez les invertébrés (et aussi dans certains gènes de rhodopsine de vertébrés) une séquence qui correspond précisément au site de liaison palindromique du HD de *Pax-6* (P3) [17, 30, 31]. Or, il a été suggéré que la conservation extraordinaire au sein du règne animal du rôle de *Pax-6* dans la formation de l'œil [32] (cela en dépit de la grande divergence évolutive de l'œil chez ces espèces (*m/s* n° 5, vol. 11, p. 776) [33]), pourrait être due à un rôle primordial de *Pax-6* comme régulateur de l'expression des gènes de photoréception dans un photorécepteur primitif [34]. Il est tentant de suggérer que ce rôle primordial ne faisait intervenir que le HD de *Pax-6* et que les sites P3 trouvés dans les promoteurs de rhodopsine sont en fait des cibles du HD de *Pax-6*. Bien que *Pax-*

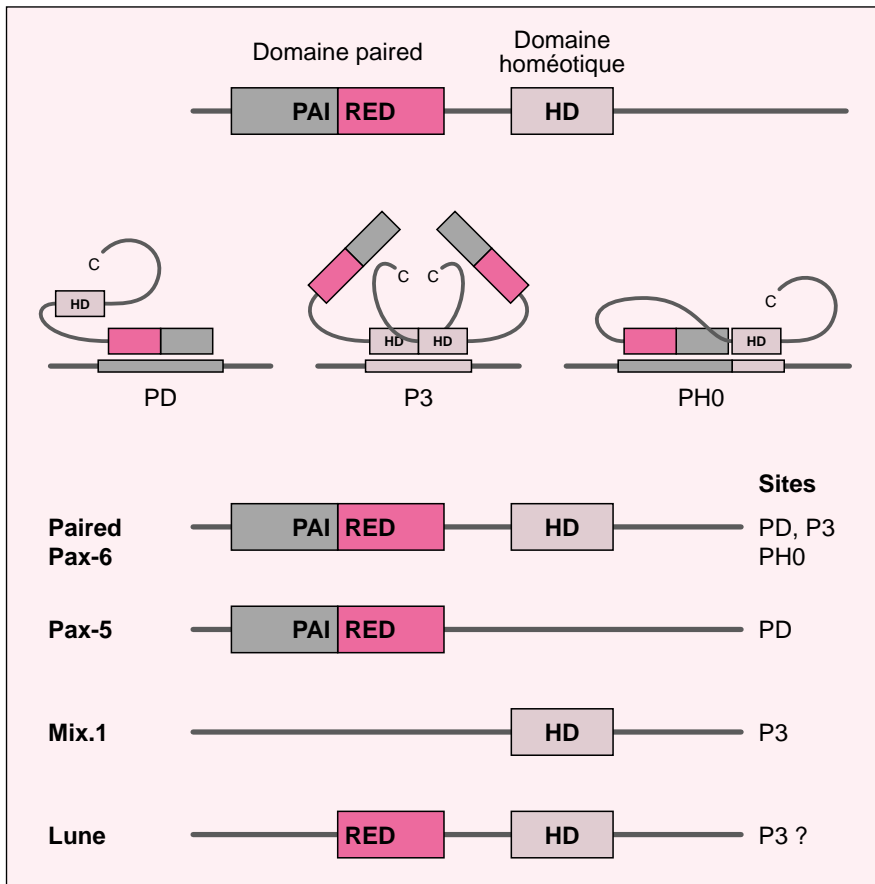


Figure 6. **Modalités de liaison à l'ADN de protéines à PD et/ou HD.** Les protéines Paired et Pax-6 contiennent à la fois un PD et un HD. Ces molécules peuvent se fixer à l'ADN, soit par leur domaine PAI (site PD), soit par un dimère du HD sur un site P3, soit par une interaction entre le PD et le HD sur des sites PHO. Des protéines, comme Pax-5, qui n'ont qu'un PD, utilisent à la fois leur domaine PAI et leur domaine RED. Des protéines contenant un HD de la classe Prd/Pax s'associent en dimères de manière coopérative pour se lier au site palindromique P3. Enfin, une nouvelle classe de protéines Pax, représentée par le produit du gène Lune [18], qui n'a qu'un domaine RED et pas de PAI, se lie aux sites P3 et peut-être aussi à d'autres sites auxquels le domaine RED se lie de manière coopérative avec le HD.

6 soit produit dans la rétine adulte de souris, il ne semble pas être abondant dans les cellules photoréceptrices [35]. Il se pourrait que des isoformes ne possédant pas le PD (qui ont été identifiées chez la caille [36]) soient synthétisées mais n'y aient pas encore été détectées [37]. De même, *C. elegans* (qui n'a pas d'œil) possède un gène *Pax-6* qui correspond à deux fonctions génétiques: l'une est sensorielle (*mab18*) et correspond à une séquence codante qui ne contient que le HD comme site fonctionnel [38]. Cette fonction pourrait correspondre au rôle de Pax-6 dans la régulation des gènes de rhodopsine.

L'autre (*vab3*) affecte la morphogénèse antérieure du système nerveux et nécessite à la fois le PD et le HD [39]. Il pourrait correspondre au rôle de Pax-6 dans le système nerveux central de la drosophile ou des vertébrés. Il convient de noter que, chez la caille, des isoformes contenant ou non le PD ont été identifiées.

## Conclusion

La famille de protéines qui vient d'être décrite a un potentiel énorme pour utiliser diverses combinaisons de domaines afin d'effectuer des fonctions très spécifiques (figure 6).

Ce nombre de combinaisons est restreint dans le cas des protéines Pax qui ne possèdent pas de HD et doivent donc utiliser uniquement le PD (certainement à la fois les domaines PAI et RED), ou lorsque des homéoprotéines de la famille Paired n'ont pas de PD et doivent agir uniquement par dimérisation de leur HD. Il est cependant probable que la famille de protéines Pax/HD a évolué à partir de molécules contenant seulement un HD et qui ont acquis ultérieurement un PD (ou *vice versa*); elle a donc conservé certaines des fonctions individuelles des molécules initiales, surtout pour des fonctions ancestrales. Le cas de Pax-6 représente un excellent exemple de diversification de fonction assumée par l'utilisation différentielle du PD et du HD ■

## Remerciements

Je voudrais remercier les membres du laboratoire qui ont contribué au travail scientifique présenté dans cet article et ont su créer l'ambiance très positive qui a toujours régné parmi nous: David Wilson, Susie Jun, Claudio Bertuccioli et Guojun Sheng ne sont pas des auteurs de cet article uniquement parce qu'ils ne pouvaient contribuer à son écriture. Ali Tahayato, Philippe Beauvils, Jean-René Huyn et Valérie Schaeffer ont aidé considérablement à cette écriture en corrigeant de multiple versions du manuscrit et je leur en suis très reconnaissant.

## Accès à la base de données internationale en Immunogénétique : IMGT

La base de données internationale ImMunoGeneTics, IMGT, initiée et coordonnée par Marie-Paule Lefranc, Montpellier, France, est accessible sur le serveur WWW du CNUC (<http://imgt.cnusc.fr:8104>). IMGT comprend deux bases de données: LIGM-DB (Immunoglobulines et Récepteurs T) et MHC/HLA-DB. IMGT/LIGM-DB contient à ce jour plus de 19 000 séquences (13 798 séquences d'Immunoglobulines et 5 850 séquences de récepteurs T) de 78 espèces différentes. Les fichiers à plat sont accessibles sur le serveur ftp anonyme d'EMBL-EBI (<ftp.ebi.ac.uk/pub/databases/imgt>).

## Contact :

Prof. Marie-Paule Lefranc, Coordinateur de IMGT  
Tel. : +33 (0)4 67 61 36 34 - Fax : +33 (0)4 67 04 02 31/45  
E-mail : [lefranc@ligm.crbm.cnrs-mop.fr](mailto:lefranc@ligm.crbm.cnrs-mop.fr)



## Summary

### Synergism in DNA binding of paired and homeo domains in Pax protein

The Pax genes encode a set of transcriptional regulators involved in several developmental processes. They are characterized by the presence of the paired domain (PD), a DNA binding domain composed of two Helix-Turn-Helix (HTH) motifs, the PAI and RED domains. Some Pax proteins contain a third HTH DNA binding motif, the homeodomain (HD). Since all PDs recognize highly related DNA sites, and since all HDs recognize a common TAAT sequence, it has been difficult to understand how these proteins achieve their functional specificity. Here, we describe how different Pax proteins use multiple combinations of their DNA binding motifs to target different promoters. *In vitro*, the *Drosophila* paired protein can bind either through its PAI domain, through its HD *via* cooperative dimerization, or through both domains. *In vivo*, using a transgenic rescue assay, we show that *prd* function requires the synergistic action of both the PAI domain and the HD on abutted PD and HD sites. Surprisingly, the RED domain, although conserved, can be deleted without loss of viability. This is in contrast to other Pax proteins that require both PAI and RED domains. Furthermore, specific isoforms of Pax6 as well as a new Pax protein, Lune, may rely on the RED domain alone. Finally, we propose that Pax6 may also act through its HD alone on a series of highly conserved palindromic TAAT sites found in all rhodopsin promoters. This may represent the ancient function of this master regulator of eye development before it acquired a PD.

## TIRÉS À PART

C. Desplan.

*m/s* n° 2, vol. 13, février 97