

## Biologie du collagène et maladies héréditaires de la matrice extracellulaire

Il existe au moins vingt gènes codant pour les différentes espèces de collagène dont les molécules s'associent en « fibrilles », constituants essentiels des matrices extracellulaires. Toute mutation perturbant la capacité des molécules de collagène à s'associer sous la forme d'une triple hélice caractéristique aura des conséquences sur les propriétés mécaniques des matrices extracellulaires, entraînant différents tableaux pathologiques dont le plus caractéristique est « l'ostéogénèse imparfaite » et sa grande fragilité osseuse.

### Michel van der Rest

Directeur de la section du collagène, unité de génétique, hôpital Shriners. Professeur agrégé, faculté de médecine, université McGill, Montréal.

### RÉFÉRENCES

1. Muir H. Macromolecular interactions and connective tissue metabolism. *Biochem Soc Trans* 1977 ; 5 : 397-402.
2. Mayne R, Burgeson RE. *Structure and Function of Collagen*. Orlando, FL : Academic Press, 1987 (sous presse).
3. van der Rest M. Collagen structure and biosynthesis. In : Cruess RL, ed. *The Musculoskeletal System*. New York : Churchill Livingstone, 1982 : 65-79.

### ADRESSE

M. van der Rest : unité de génétique, hôpital Shriners pour les enfants infirmes, 1529, avenue Cedar, Montréal, QC, Canada H3G 1A6.

La matrice extracellulaire remplit une grande variété de fonctions, à la fois biomécaniques comme élément de support des tissus ou organes, et biochimiques comme milieu de croissance des cellules. Loin d'être un élément statique et figé, elle est en évolution dynamique constante et en relation intime avec les cellules qui la synthétisent ou la modifient. A cette variété de fonctions, correspond une palette de macromolécules constitutives dont les proportions et les modifications post-traductionnelles confèrent aux diverses matrices les propriétés biomécaniques et biochimiques appropriées. Parmi ces macromolécules, le collagène constitue la classe de protéine la plus abondante. Sa fonction la plus évidente est de conférer aux tissus conjonctifs leur résistance à la traction. Ainsi les tendons, dont la fonction quasi exclusive est de transmettre une force de traction, sont composés presque uniquement de collagène. Dans les cartilages hyalins qui servent entre autres à amortir les chocs, la résistance à la traction des fibres de collagène s'oppose à la pression osmotique due à la forte densité de charges des protéoglycanes emprisonnées dans leur réseau [1]. Dans le matériau composite des tissus minéralisés, non seulement le col-

lagène contrebalance le caractère cassant du constituant minéral mais, de plus, il lui fournit, avec d'autres macromolécules, une trame sur laquelle il vient se déposer. Le collagène remplit également bien d'autres fonctions. A cette variété de fonctions correspond une variété de structures primaires et d'organisations moléculaires qui ont été classées à l'intérieur des différents types de collagène, dont onze ont été décrits jusqu'à présent [2]. Suivant l'usage accepté le plus communément, un collagène est une protéine structurale de la matrice extracellulaire dont une grande partie est formée par une triple hélice de trois chaînes polypeptidiques entrelacées [3]. Ces parties hélicoïdales ont une structure primaire répétitive (Gly-X-Y)<sub>n</sub> dans laquelle X est le plus souvent une proline et Y une hydroxyproline. Cette définition exclut des molécules comme Clq ou l'acétylcholinestérase qui ont en partie une structure protéique similaire au collagène, mais qui n'appartiennent pas à la matrice extracellulaire. Il s'agit là bien sûr d'une définition quelque peu arbitraire mais qui servira utilement à limiter le champ de ma discussion. De nombreuses maladies héréditaires sont caractérisées par des anomalies systémiques des tissus conjonctifs. Depuis des années, on

## RÉFÉRENCES

- McKusick VA. *Heritable Disorders of Connective Tissue*. Saint-Louis : The C.V. Mosby Company, 1972 : 33-46.
- Ramachandran GN, Ramakrishnan C. Molecular structure. In : Ramachandran GN, Reddi AH, eds. *Biochemistry of Collagen*. New York : Plenum Press, 1976 : 45-84.
- Berg RA, Prockop DJ. The thermal transition of a non-hydroxylated form of collagen. Evidence for a role for hydroxyproline in stabilizing the triple helix of collagen. *Biochem Biophys Res Commun* 1973 ; 52 : 115-20.
- Chapman JA, Hardcastle RA. The staining pattern of collagen fibrils. *Connect Tissue Res* 1974 ; 2 : 151-9.
- Eyre DR. Collagen : Molecular diversity in the body's protein scaffold. *Science* 1980 ; 207 : 1315-22.
- Chu M-L, de Wet W, Bernard M, et al. Human pro  $\alpha 1(I)$  collagen gene structure reveals evolutionary conservation of a pattern of introns and exons. *Nature* 1984 ; 310 : 337-40.
- Myers JC, Dickson LA, de Wet WJ, et al. Analysis of the 3' end of the human pro  $\alpha 2(I)$  collagen gene. *J Biol Chem* 1983 ; 258 : 10128-35.
- Cheah KSE, Stocker NG, Griffin JR, Grosveld FG, Solomon E. Identification and characterization of the human type II collagen gene (COL2A1). *Proc Natl Acad Sci USA* 1985 ; 82 : 2555-9.
- Chu M-L, Weil D, de Wet W, Bernard M, Sippola M, Ramirez F. Isolation of cDNA and genomic clones encoding human pro-alpha 1(III) collagen. *J Biol Chem* 1985 ; 260 : 4357-63.
- Myers JC, Loidl HR, Seyer JM, Dion AS. Complete primary structure of the human alpha-2 type V procollagen COOH-terminal propeptide. *J Biol Chem* 1985 ; 260 : 11216-22.
- Prockop DJ, Kivirikko K. Heritable diseases of collagen. *N Engl J Med* 1984 ; 311 : 376-86.
- Halila R, Peltonen L. Purification of human procollagen type III N-proteinase from placenta and preparation of antiserum. *Biochem J* 1986 ; 239 : 47-52.
- Broek DL, Madri J, Eikenberry EF, Brodsky B. Characterization of the tissue form of type V collagen from chick bone. *J Biol Chem* 1985 ; 260 : 555-62.
- Henkel W, Glanville RW. Covalent cross-linking between molecules of type I and type III collagen. *Eur J Biochem* 1982 ; 122 : 205-13.

a soupçonné que des mutations affectant la structure ou la biosynthèse du collagène sont à l'origine de plusieurs d'entre elles [4]. De telles mutations ont, de fait, été démontrées dans plusieurs cas ; elles sont toutes différentes les unes des autres, ce qui suggère une hétérogénéité biochimique encore plus grande que l'hétérogénéité clinique déjà bien décrite. Mon propos sera ici de montrer le nouvel éclairage que les découvertes récentes sur la biologie du collagène apportent à notre compréhension des maladies du tissu conjonctif. L'étude des maladies génétiques du tissu conjonctif, qui sont presque toujours de malheureuses « expériences de la nature », contribue par ailleurs de façon importante à élucider le fonctionnement normal de ces

assemblages complexes de macromolécules que sont les matrices extracellulaires.

## De la structure à la fonction

La figure 1 donne une représentation schématique d'une section transversale d'une molécule de collagène [5]. On y voit trois chaînes polypeptidiques qui forment chacune une hélice gauche. Celles-ci s'enroulent entre elles en une superhélice droite appelée triple hélice. Tous les trois résidus, un acide aminé de chaque chaîne se retrouve au centre de la triple hélice. Seule la glycine, dont la chaîne latérale est réduite à un atome d'hydrogène, peut se trouver dans cette position sans provoquer d'encombrement stérique

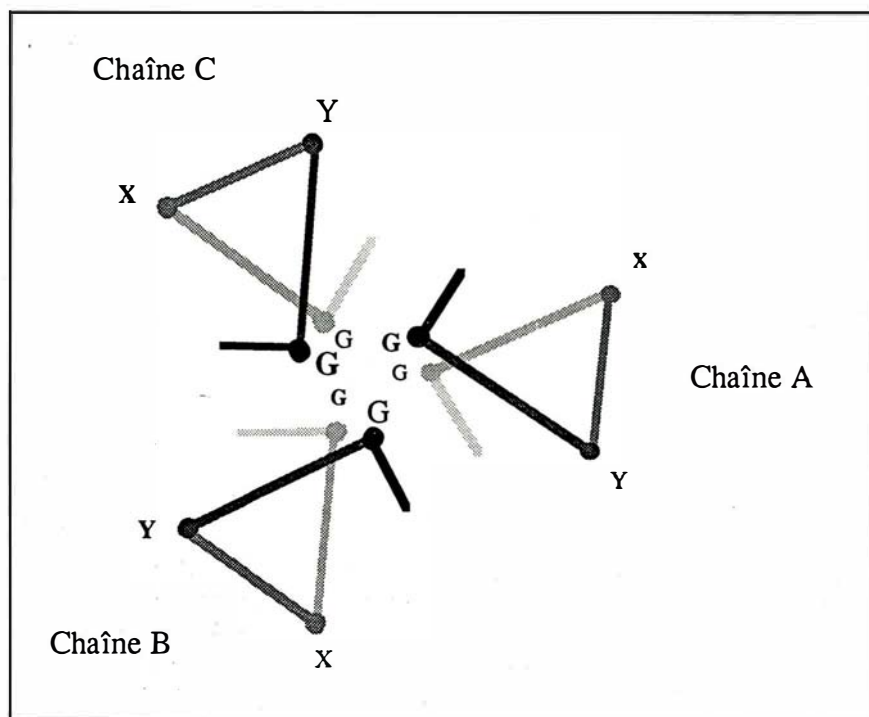


Figure 1. **Représentation schématique d'une coupe transversale de la triple hélice du collagène.** Trois chaînes polypeptidiques hélicoïdales (A, B et C) s'enroulent entre elles pour former la triple hélice. Seuls sont indiqués les carbones  $\alpha$  des acides aminés constituants. Les liens peptidiques sont symbolisés par les droites. Les résidus glycines (G) occupent le centre de l'hélice alors que les résidus X sont souvent de la proline et Y de l'hydroxyproline. Les tailles et l'intensité des traits et des lettres suggèrent l'éloignement par rapport au plan de la figure.

et rompre par conséquent la structure de l'hélice elle-même, d'où la structure primaire (Gly-X-Y)<sub>n</sub>. Dans environ un tiers des triplets, X est une proline et Y est une hydroxyproline. Ces acides iminés sont nécessaires au maintien des angles requis pour la formation de l'hélice. De plus, l'hydroxyproline stabilise l'hélice en établissant des ponts hydrogènes supplémentaires. Ainsi, la température de dénaturation d'un collagène non hydroxylé (l'hydroxyproline est produite par hydroxylation post-traductionnelle de résidus prolines) est inférieure de 15° C à celle d'un collagène normalement hydroxylé [6]. Les résidus en X et Y ont leurs chaînes latérales dirigées vers l'extérieur de la triple hélice. C'est donc eux qui seront responsables des interactions latérales des molécules de collagène entre elles

ou avec d'autres constituants de la matrice extracellulaire. Ces interactions confèrent aux molécules de collagène la propriété essentielle de s'agréger latéralement. Dans le cas des collagènes fibrillaires, on se rend compte facilement que ces interactions ont lieu tout le long de la triple hélice. La *figure 2* nous montre l'arrangement caractéristique des collagènes fibrillaires en lignes parallèles décalées d'une ligne à l'autre, d'une longueur de 66,8 nm (communément appelée D). Ce décalage périodique provient des interactions ioniques et hydrophobes des molécules adjacentes. On a ainsi des regroupements des charges ioniques en zones distinctes, visibles en microscopie électronique et apparaissant comme une striation transversale des fibrilles, bien connue des histologistes

(*figure 2*) [7]. L'alignement latéral permet, de plus, la réticulation des collagènes par la formation de liaisons covalentes à partir de résidus oxydés de lysine ou de dérivés de la lysine (hydroxylysine et hydroxylysine glycosylée). La réticulation des collagènes accroît de façon substantielle la résistance des fibres à la traction [8].

Il ressort de cette description de la structure générale des collagènes que des mutations perturbant la formation de l'hélice, l'alignement latéral ou le mécanisme de réticulation entraîneront une altération des propriétés biomécaniques de la matrice extracellulaire et seront donc morbides. Les sites qui sont critiques pour le maintien de la structure du collagène sont nombreux : par exemple les 666 glycines des chaînes  $\alpha 1(I)$  et  $\alpha 2(I)$  du collagène de type I sont

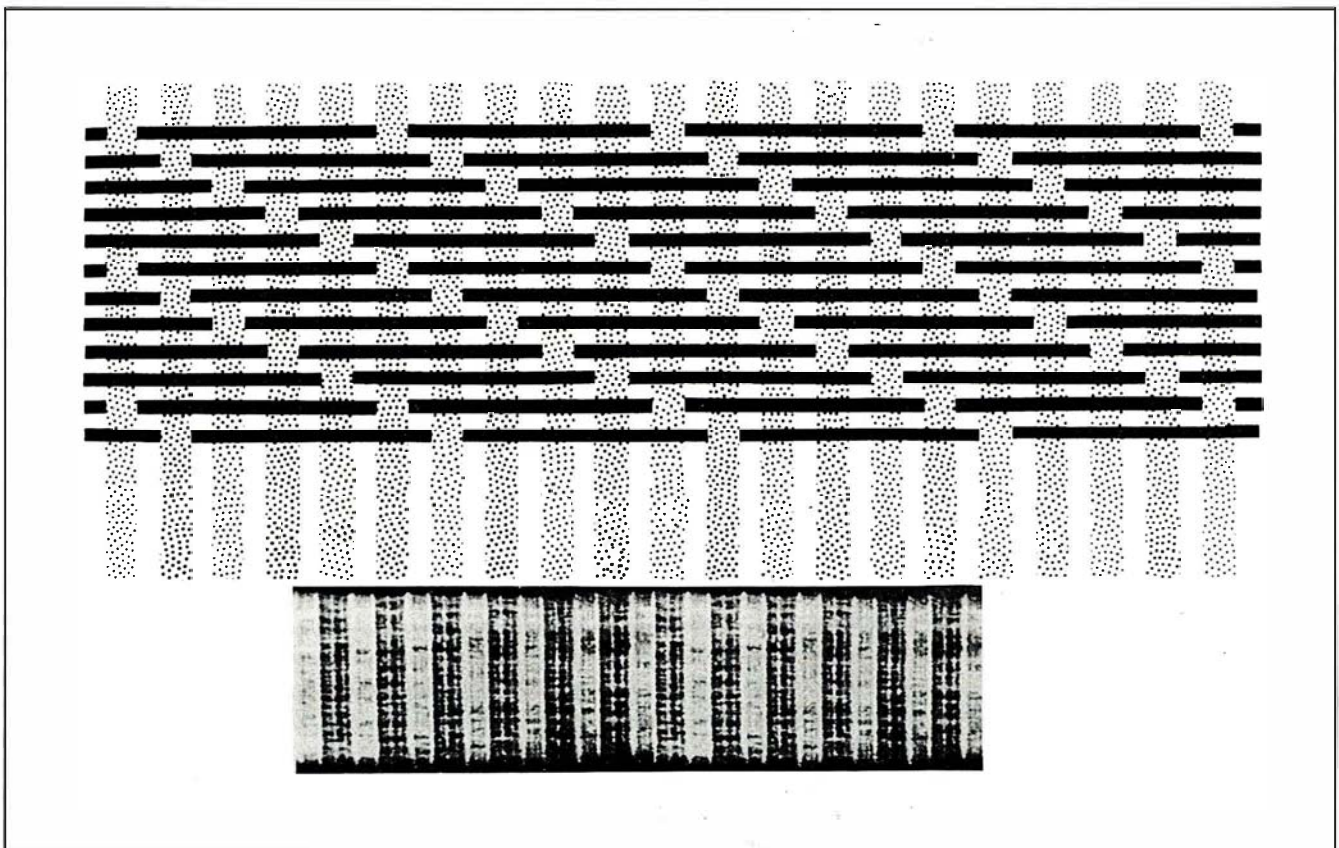


Figure 2. **Représentation plane de l'étagement des molécules de collagène de type I à l'intérieur de la fibrille** [5]. Les molécules de collagène sont représentées par des barres. En réalité, les molécules mesurent 294 nm de long pour un diamètre de 1,5 nm. Elles sont décalées d'une ligne à l'autre d'une longueur D de 66,8 nm. Les alignements des intervalles et des chevauchements apparaissent respectivement comme des bandes sombres et claires sur la micrographie électronique en coloration négative présentée en dessous (Figure tirée de [3]).

## RÉFÉRENCES

18. Linsenmayer TF, Fitch JM, Gross J, Mayne R. Are collagen fibrils in the developing avian cornea composed of two different collagen types? Evidence from monoclonal antibody study. *Ann NY Acad Sci* 1985; 460: 232-45.
19. Fleischmajer R, Perlish JS, Timpl R. Collagen fibrillogenesis in human skin. *Ann NY Acad Sci* 1985; 460: 246-57.
20. Brodsky B, Eikenberry EF. Supramolecular collagen assemblies. *Ann NY Acad Sci* 1985; 460: 73-84.
21. Broek DL. *Isolation and characterization of the intact forms of type V and 1 $\alpha$ , 2 $\alpha$ , 3 $\alpha$  collagen*. Ph. D. thesis. Piscataway: UMDNJ-Rutgers Medical School, 1984: 77-118.
22. Kühn K, Glanville RW, Babel W, et al. The structure of type IV collagen. *Ann NY Acad Sci* 1985; 460: 14-24.
23. Mayne R, van der Rest M, Ninomiya Y, Olsen BR. The structure of type IX collagen. *Ann NY Acad Sci* 1985; 460: 38-46.
24. Huber S, van der Rest M, Bruckner P, Rodriguez E, Winterhalter KH, Vaughan L. Identification of the type IX collagen polypeptide chains. *J Biol Chem* 1986; 261: 5965-8.
25. Yamada Y, Avvedimento VE, Mudryj M, et al. The collagen gene: Evidence for its evolutionary assembly by amplification of a DNA segment containing an exon of 54 bp. *Cell* 1980; 22: 887-92.
26. Sakurai Y, Sullivan M, Yamada Y. Alpha-1 type IV collagen gene evolved differently from fibrillar collagen genes. *J Biol Chem* 1986; 261: 6654-7.
27. Lozano G, Ninomiya Y, Thompson H, Olsen BR. A distinct class of vertebrate collagen genes encodes chicken type IX collagen polypeptides. *Proc Natl Acad Sci USA* 1985; 82: 4050-4.
28. Ninomiya Y, Gordon M, van der Rest M, Schmid T, Linsenmayer T, Olsen BR. The developmentally regulated type X collagen contains a long open reading frame without introns. *J Biol Chem* 1986; 261: 5041-50.
29. Timpl R, Wiedemann H, van Delden V, Furthmayr H, Kühn K. A network model for the organisation of type IV collagen molecules in basement membranes. *Eur J Biochem* 1981; 120: 203-11.
30. Engvall E, Hessel H, Klier G. Molecular assembly, secretion and matrix deposition of type VI collagen. *J Cell Biol* 1986; 102: 703-10.
31. Sakai LY, Keene DR, Morris NP, Burgesson RE. Type VII collagen is a major structural component of anchoring fibrils. *J Cell Biol* 1986; 103: 1577-86.

essentielles au maintien de la structure hélicoïdale.

### Collagènes fibrillaires et non fibrillaires

Depuis le début des années 1970, de nouveaux collagènes ont été découverts en nombre toujours croissant et classés suivant une nomenclature prêtant facilement à confusion. Les collagènes fibrillaires (*Tableau I*) se distinguent clairement des autres collagènes (*Tableau II*). Ils ont de nombreuses caractéristiques communes et dérivent d'un gène ancestral commun dont la structure génomique a été conservée de façon remarquable puisque la taille et la position des nombreux exons sont pratiquement identiques pour les cinq gènes étudiés jusqu'à présent, souvent dans plusieurs espèces [9-13]. Le collagène de type I en est l'archétype. Composé de deux chaînes polypeptidiques,  $\alpha 1(I)$  et  $\alpha 2(I)$ , assemblées généralement dans un rapport 2 : 1, [ $\alpha 1(I)$ ]<sub>2</sub>  $\alpha 2(I)$ , il est l'élément de base des fibres de collagène de la plupart des tissus conjonctifs. Ce collagène subit un métabolisme intense après la traduction (*figure 3*). Ce métabolisme a été très bien décrit [14] et ne sera pas discuté ici davantage. Associés au collagène de type I dans des proportions variées dans les tissus mous et à l'état de trace dans les tissus minéralisés et les tendons, on retrouve deux autres types de collagène qui participent aussi à la formation des fibrilles, à savoir les collagènes de type III et de type V. Ces deux types de collagène se distinguent du type I par leur sensibilité différente au clivage protéolytique du propeptide amino-terminal (N-propeptide). En effet, le N-propeptide du collagène de type I est presque complètement coupé par une enzyme spécifique, tandis que celui du collagène de type III l'est beaucoup plus lentement et par une enzyme distincte [15]. Quant au type V, sa forme tissulaire comprend une extension globulaire importante, sans doute le N-propeptide [16]. Il semble que ces trois collagènes ne

forment pas de fibrilles distinctes mais que le diamètre des fibrilles, composées principalement de collagène de type I, soit modulé par ces collagènes fibrillaires mineurs [17-19]. Le collagène de type II est l'équivalent du collagène de type I dans certains tissus spécialisés, tels le cartilage hyalin ou l'humeur vitreuse. Cependant, il forme un assemblage légèrement moins dense que le collagène de type I, ce qui correspond sans doute à un besoin spécifique du cartilage hyalin [20]. Le collagène de type XI [2], encore souvent appelé 1 $\alpha$  2 $\alpha$  3 $\alpha$ , est sans doute le pendant du collagène de type V pour les tissus contenant le collagène de type II comme type principal. Comme le type V, sa forme tissulaire comprend une extension globulaire qui n'est pas éliminée en cours de biosynthèse [21].

Les autres collagènes (*Tableau II*) présentent une grande diversité de structures moléculaires et de ce fait leurs modes d'assemblage dans les tissus sont aussi fort variés. Plusieurs auteurs ont tenté de les regrouper suivant divers critères, en particulier leur taille ou la fréquence et la nature des interruptions de la triple hélice que l'on y retrouve. Ces regroupements me paraissent prématurés, même s'il est probable que certains de ces collagènes non fibrillaires existent dans divers tissus sous forme d'isotypes de structures similaires, mais produits de gènes différents. Ceux dont tout ou partie de la structure primaire a déjà été élucidée (types IV, IX et X) présentent des interruptions (six résidus et plus) ou des imperfections de la triple hélice (par exemple-Gly-X-Gly-, ou Gly-X-Y-X-Y-Gly) [12]. Certaines interruptions ont une incidence directe sur la structure de la molécule qui peut présenter des angles caractéristiques entre différentes régions hélicoïdales, comme cela a été observé pour le collagène de type IV [22] et le collagène de type IX [23]. Les modifications post-traductionnelles sont communes à tous les collagènes en ce qui concerne les hydroxylation

Tableau I  
LES COLLAGÈNES FIBRILLAIRES

A. Associés au type I			
Type	Chaînes	Molécules	Tissus représentatifs
I	$\alpha 1(I), \alpha 2(I)$	$[\alpha 1(I)]_2 \alpha 2(I)$	Peau, os, tendon, dentine, etc.
		$[\alpha 1(I)]_3$	Dentine, peau (forme mineure)
III	$\alpha 1(III)$	$[\alpha 1(III)]_3$	Peau, vaisseaux
V	$\alpha 1(V)^{(a)}, \alpha 2(V), \alpha 3(V)$	$[\alpha 1(V)]_3$	Culture de cellules de poumon de hamster
		$[\alpha 1(V)]_2 \alpha 2(V)$	Membranes fœtales, peau, os
		$\alpha 1(V) \alpha 2(V) \alpha 3(V)$	Placenta, membranes synoviales
B. Associés au type II			
Type	Chaînes	Molécules	Tissus représentatifs
II	$\alpha 1(II)$	$[\alpha 1(II)]_3$	Cartilage hyalin, humeur vitrée
XI <sup>(b)</sup>	$\alpha 1(XI), \alpha 2(XI), \alpha 3(XI)^{(c)}$	$\alpha 1(XI) \alpha 2(XI) \alpha 3(XI)^{(d)}$	Cartilage hyalin

<sup>(a)</sup> Une chaîne similaire dans sa triple hélice, mais différente dans ses propeptides a été décrite et appelée  $\alpha 1(V)'$  ou  $\alpha 4(V)$ .

<sup>(b)</sup> Souvent appelé  $1\alpha$   $2\alpha$   $3\alpha$ .

<sup>(c)</sup>  $\alpha 3(XI)$  est probablement identique à  $\alpha 1(III)$ , sauf en ce qui concerne les modifications post-traductionnelles.

<sup>(d)</sup> Cette stœchiométrie n'est pas absolument démontrée.

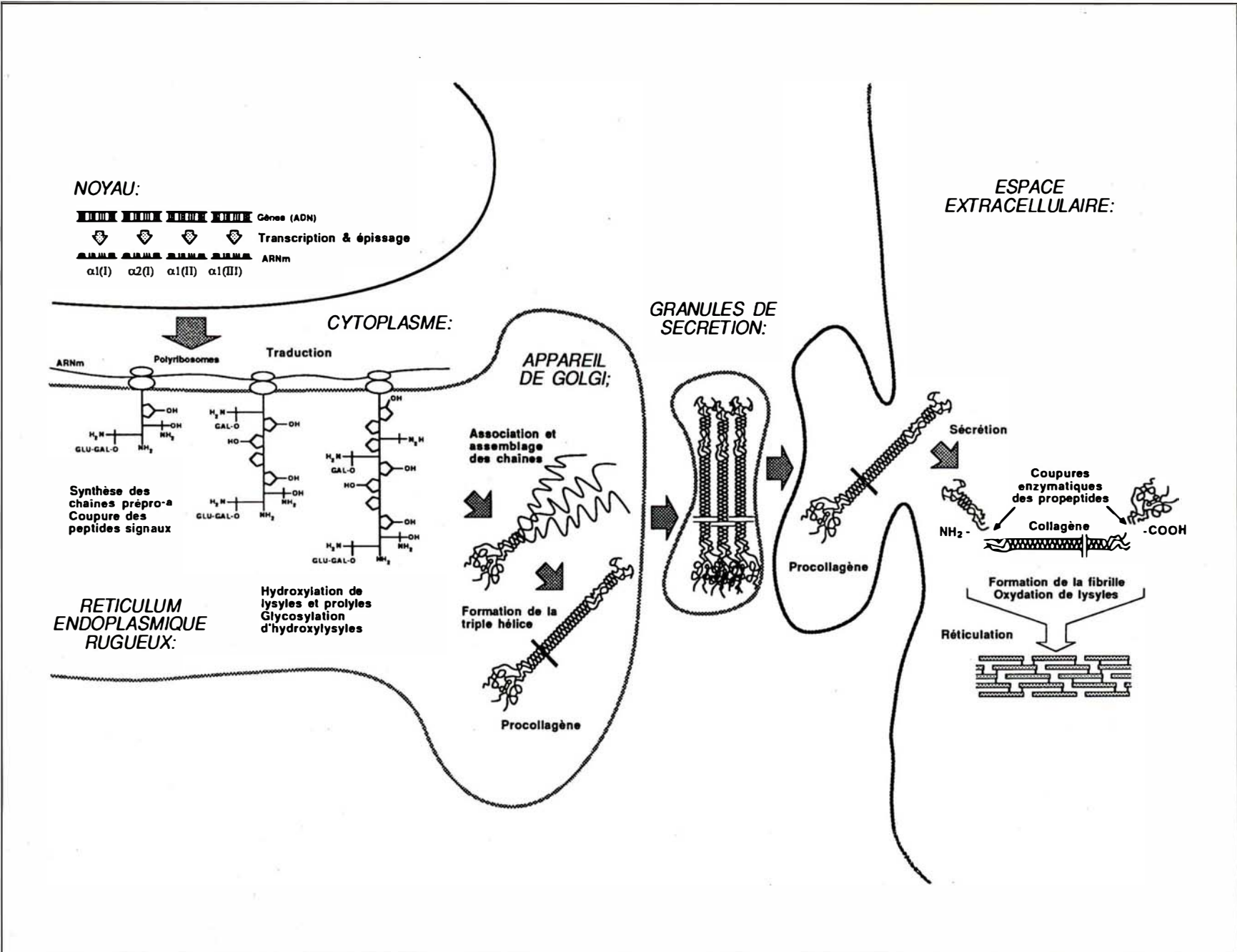
Tableau II  
LES COLLAGÈNES NON FIBRILLAIRES

Type	Chaînes	Molécules	Tissus représentatifs
IV	$\alpha 1(IV), \alpha 2(IV)$	$[\alpha 1(IV)]_2 \alpha 2(IV)$	Membranes basales
VI	$\alpha 1(VI), \alpha 2(VI), \alpha 3(VI)$	$\alpha 1(VI) \alpha 2(VI) \alpha 3(VI)$	Vaisseaux, peau, disque intervertébral
VII	$\alpha 1(VII)$	$[\alpha 1(VII)]_3$	Membranes fœtales
VIII	(?)	(?)	Cellules endothéliales, membrane de Descemet
IX	$\alpha 1(IX), \alpha 2(IX), \alpha 3(IX)$	$\alpha 1(IX) \alpha 2(IX) \alpha 3(IX)$	Cartilage hyalin
X	$\alpha 1(X)$	$[\alpha 1(X)]_3$	Plaque de croissance

des résidus lysines et prolines, les glycosylations des hydroxylysines et la désamination oxydative des lysines ou hydroxylysines préalable à la formation de ponts covalents. Cependant, la protéolyse spécifique des propeptides telle qu'elle existe pour les collagènes fibrillaires semble généralement absente dans les collagènes non fibrillaires et, lorsqu'elle existe, les propeptides clivés sont beaucoup plus petits. Enfin, on trouve dans certains des collagènes non fibrillaires des modifications qui sont absentes dans les collagènes fibrillaires telle l'addition d'une chaîne latérale de glycosaminoglycane à la chaîne  $\alpha 2(IX)$  du collagène de type IX [24].

La distinction entre collagènes fibrillaires et non fibrillaires se reflète dans la structure des gènes. Alors que les régions hélicoïdales des collagènes fibrillaires sont codées surtout par des exons de 54 paires de bases (ou des multiples de 54 paires de bases), correspondant à une séquence (Gly-X-Y)<sub>6</sub>, et que les quelques exceptions correspondent cependant toujours à des séquences (Gly-X-Y)<sub>n</sub> [25], la taille des exons des collagènes non fibrillaires déjà étudiés (les types IV, IX et X) s'écarte de cette règle [26-28]. Ainsi l'extrémité 5' de plusieurs exons commence par la deuxième base du codon glycine. Dans le cas du collagène de type IX, il semble y avoir eu, au cours de l'évolution, fusion entre un gène constitué d'exons de 54 paires de bases, codant pour une région de la molécule strictement triple hélicoïdale, et un gène à exons de tailles variables, codant pour une région présentant des imperfections dans la triple hélice (B.R. Olsen, communication personnelle).

Les collagènes non fibrillaires s'assemblent en superstructures d'une grande variété (figure 4). Les assemblages les mieux décrits sont ceux des types IV, VI et VII. Il est frappant de constater que les interactions latérales, qui sont favorisées par la structure hélicoïdale du collagène, jouent ici aussi un rôle important, comme



◀ **Figure 3. Biosynthèse du collagène de type I.** Près de dix enzymes sont impliquées dans la modification de la molécule après sa traduction. Les enzymes d'hydroxylation et de glycosylation localisées dans la lumière du réticulum endoplasmique n'agissent que sur les molécules non hélicoïdales. Lorsque la triple hélice est stabilisée par l'hydroxylation d'un nombre suffisant de résidus prolines en position Y, l'action de ces enzymes est inhibée. Par ailleurs, seules les pro-molécules hélicoïdales sont sécrétées. Elles sont empaquetées sous formes d'agrégats latéraux dans les granules de sécrétion et s'ajoutent probablement en bloc à la fibrille en voie d'assemblage. L'action des protéases des propeptides se fait à ce niveau, sans doute dans un microenvironnement créé par des invaginations de la membrane cellulaire.

pour l'arrangement des molécules fibrillaires. Ces interactions se font entre molécules antiparallèles dans la région de pontage du collagène de type IV (domaine 7 S) [29] et dans le collagène de type VI [30]. Ce dernier s'assemble linéairement par ses extrémités globulaires pour former les filaments perlés observés dans divers tissus conjonctifs. Dans le cas du collagène de type VII, les molécules sont alignées parallèlement pour former les structures connues sous le nom de « fibres d'ancrage » (*anchoring fibers*), rencontrées au niveau de la jonction de certaines membranes basales et du tissu conjonctif sous-jacent [31].

Le collagène-protéoglycane du cartilage (collagène de type IX) est lié de façon covalente à la fibrille de type II par des ponts dérivés de la lysine (M. van der Rest et R. Mayne, soumis pour publication). Il pourrait jouer un rôle de lien entre les fibrilles et d'autres éléments de la matrice [32]. Ici aussi on pense que le collagène de type IX interagit initialement avec la fibrille par le biais des interactions latérales de sa partie hélicoïdale. L'angle défini par un court domaine non hélicoïdal, site d'attachement du glycosaminoglycane [33], permettrait au bras hélicoïdal court (*figure 4*) d'exposer un domaine globulaire et basi-

que, qui constitue l'extrémité amino-terminale de la molécule, permettant ainsi des interactions avec d'autres éléments de la matrice extracellulaire. Des travaux récents du Dr M. Gordon, dans le laboratoire du Dr B.R. Olsen à Harvard, et du Dr B. Dublet, dans mon laboratoire, démontrent l'existence d'un nouveau collagène, le type XII, similaire au type IX mais produit d'un ou plusieurs gènes distincts dans certains tissus riches en collagène de type I. Il semble donc y avoir un ensemble de collagènes dont l'expression et la synthèse sont coordonnées et liées à la nature du tissu qui les contient.

### Maladies héréditaires liées au collagène

La structure de la triple hélice du collagène suggère qu'il existe un grand nombre de sites où une mutation aussi limitée qu'une substitution d'un acide aminé pourrait avoir un effet dramatique sur la stabilité de la molécule affectée et par conséquent sur les tissus où cette dernière est utilisée. De telles mutations ont été récemment démontrées dans le collagène de type I, et ont été associées dans la plupart des cas à l'une ou l'autre des formes d'ostéogenèse imparfaite (OI), maladie héréditaire caractérisée par une fragilité des os et de sévérité très variable [4].

Les mutations du collagène de type I observées jusqu'à présent peuvent être subdivisées en deux catégories suivant qu'elles entraînent un défaut strictement quantitatif, dû à une diminution de la synthèse d'un collagène apparemment normal, ou qualitatif. Dans ce cas, la sécrétion par les cellules d'un collagène structurellement anormal provoque l'affaiblissement de la texture des matrices dans lesquelles il est intégré. L'anomalie qualitative est ici souvent associée à une anomalie quantitative. Les défauts quantitatifs du collagène peuvent être liés à des anomalies situées à n'importe quel niveau entre la transcription du gène et la sécrétion du collagène.

Lorsque ces défauts affectent le collagène de type I, normalement composé de deux chaînes  $\alpha 1(I)$  identiques et d'une chaîne  $\alpha 2(I)$ , le résultat clinique est fort différent selon la chaîne qui est affectée. Lorsqu'un des allèles du pro  $\alpha 1(I)$  n'est pas exprimé ou que la chaîne produite par un des allèles n'est pas intégrée dans la promolécule, la sécrétion de collagène de type I est diminuée de 50 %. Une telle mutation entraîne la forme dominante classique de l'OI (sclérotique bleue, hérédité dominante et régression de la fréquence des fractures après la puberté) [34]. Lorsqu'un des allèles du pro  $\alpha 2(I)$  n'est pas exprimé ou que la chaîne produite par un des allèles n'est pas intégrée dans la promolécule, les porteurs de ces mutations sont cliniquement normaux.

Les anomalies qualitatives du collagène de type I décrites jusqu'à présent sont fort variées. La forme d'OI la plus étudiée est la forme létale [35] (type II dans la classification de Sillence *et al.* [36]). Byers *et al.* [37] ont suggéré que la cause la plus commune en est une substitution d'une glycine d'un triplet Gly-X-Y par un autre acide aminé, entraînant une altération de la structure hélicoïdale. Comme la formation de l'hélice est amorcée par la réunion des trois chaînes polypeptidiques par leurs propeptides carboxyterminaux et se propage ensuite tout au long de la molécule, la présence d'un acide aminé autre que la glycine dans l'axe central de la triple hélice (*figure 1*) entraîne une déstabilisation et un ralentissement dans la cinétique de la formation de cette dernière. Ceci se manifeste par un excès de modifications post-traductionnelles des résidus lysines [38]. En effet, les enzymes responsables de l'hydroxylation de certains résidus lysines et de la glycosylation de certains résidus hydroxylysines ne sont actives que sur les chaînes non assemblées en triple hélice. Un ralentissement dans la formation de l'hélice, secondaire à une mutation, entraîne par conséquent un excès des modifications post-traduc-

## RÉFÉRENCES

32. Müller-Glauser W, Humbel B, Glatt M, Sträuli P, Winterhalter K, Bruckner P. On the role of type IX collagen in the extracellular matrix of cartilage : type IX collagen is localized at intersections of collagen fibrils. *J Cell Biol* 1986 ; 102 : 1931-9.
33. McCormick D, van der Rest M, Goodship J, Lozano G, Ninomiya Y, Olsen BR. Structure of the glycosaminoglycan domain in the type IX collagen-proteoglycan. *Proc Natl Acad Sci USA* (sous presse).
34. Barsh GS, David KE, Byers PH. Type I Osteogenesis Imperfecta : a nonfunctional allele for pro  $\alpha 1(I)$  chains of type I procollagen. *Proc Natl Acad Sci USA* 1982 ; 79 : 3838-42.
35. Maroteaux P, Cohen-Solal L. L'Ostéogénèse Imparfaitée létale. *Ann Genet (Paris)* 1984 ; 27 : 11-5.
36. Silience DO, Senn AS, Danks DM. Genetic heterogeneity in Osteogenesis Imperfecta. *J Med Genet* 1979 ; 16 : 101-6.
37. Cohn DH, Byers PH, Steinmann B, Gelinas RE. Lethal Osteogenesis Imperfecta resulting from a single nucleotide change in one human pro  $\alpha 1(I)$  collagen allele. *Proc Natl Acad Sci USA* 1986 ; 83 : 6045-7.
38. Bonadio J, Holbrook KA, Gelinas RE, Jacob J, Byers PH. Altered triple helical structure of type I procollagen in lethal perinatal Osteogenesis Imperfecta. *J Biol Chem* 1985 ; 260 : 1734-42.
39. Pihlajaniemi T, Dickson LA, Pope M, et al. Osteogenesis Imperfecta : Cloning of a pro  $\alpha 2(I)$  collagen gene with a frameshift mutation. *J Biol Chem* 1984, 259 : 12941-4.
40. Deak SB, van der Rest M, Prockop DJ. Altered helical structure of a homotrimer of  $\alpha 1(I)$  chains synthesized by fibroblasts from a variant of Osteogenesis Imperfecta. *Coll Relat Res* 1985 ; 5 : 305-13.
41. Nicholls AC, Osse G, Schloon HG, et al. The clinical features of homozygous  $\alpha 2(I)$  collagen deficient Osteogenesis Imperfecta. *J Med Genet* 1984 ; 21 : 257-62.
42. van der Rest M, Hayes A, Marie P, Desbarats M, Kaplan P, Glorieux FH. Lethal Osteogenesis Imperfecta with amniotic band lesions : collagen studies. *Am J Med Genet* 1986 ; 24 : 433-446.
43. Falk CT, Schwartz RC, Ramirez F, Tsipouras P. Use of molecular haplotypes specific for the human pro  $\alpha 2(I)$  collagen gene in linkage analysis of the mild autosomal dominant forms of Osteogenesis Imperfecta. *Am J Hum Genet* 1986 ; 38 : 269-79.
44. Wenstrup RJ, Tsipouras P, Byers PH. Osteogenesis Imperfecta Type IV. Biochemical confirmation of genetic linkage to the pro  $\alpha 2(I)$  gene of type I collagen. *J Clin Invest* 1986 ; 78 : 1449-55.
- tionnelles au niveau de la partie de la molécule située entre la mutation et l'extrémité aminoterminal de la triple hélice (figure 5). Ce phénomène a été observé fréquemment même si les données établissant le lien entre ces modifications « épigénétiques » et les mutations ne sont encore que très fragmentaires. D'autres mutations, en particulier des délétions, ont également été démontrées dans quelques cas d'OI létale [14].
- Qu'advient-il des chaînes anormales qui sont synthétisées par les cellules des patients affectés de mutations entraînant un défaut qualitatif du collagène de type I ? Dans un cas exceptionnel, le processus a pu être pleinement élucidé. Ce patient atteint d'OI sévère (type III) et descendant de parents consanguins, est homozygote pour un gène anormal de la chaîne pro  $\alpha 2(I)$ . Ce gène a une délétion de quatre bases codantes, située près du codon stop. Il en résulte un décalage de phase de lecture\* et une substitution des 17 derniers acides aminés du pro  $\alpha 2(I)$  [39]. Ceci suffit à empêcher l'assemblage des propeptides et donc l'insertion de la chaîne pro  $\alpha 2(I)$  dans la molécule de collagène. Dans ce cas, une troisième chaîne pro  $\alpha 1(I)$  vient remplacer la chaîne manquante, produisant un homotrimère [pro  $\alpha 1(I)$ ]<sub>3</sub>. La chaîne pro  $\alpha 2(I)$  non utilisée n'est pas sécrétée mais dégradée à l'intérieur de la cellule. Quant à l'homotrimère [pro  $\alpha 1(I)$ ]<sub>3</sub>, il est transformé de façon similaire à la molécule normale en [ $\alpha 1(I)$ ]<sub>3</sub> et forme des fibrilles de qualité suffisante pour être compatible avec la vie. Cet homotrimère [ $\alpha 1(I)$ ]<sub>3</sub> est cependant moins stable que l'hétérotrimère normal [ $\alpha 1(I)$ ]<sub>2</sub>  $\alpha 2(I)$  [40] et l'OI est sévère chez ce patient [41]. Chez d'autres patients, en particulier ceux affectés par les formes les plus sévères de la maladie, on a pu montrer que les molécules anormales sont détruites à l'intérieur de la cellule par un mécanisme qui a été appelé le « suicide de protéine » [14]. Ce mécanisme n'est pas parfait et des molécules structurellement anormales peuvent être insérées dans la matrice extracellulaire et y causer de grands dommages [42]. Dans le cas de substitution d'une glycine par un autre acide aminé, on peut penser que la mutation interfère avec l'alignement parfait des molécules requis pour la formation de la fibrille, soit à la suite de l'hypermodification posttraductionnelle de certains acides aminés, décrite plus haut, soit en causant un changement de pas dans l'hélice, ce qui entraîne le défaut d'alignement d'une grande partie de la molécule.
- Notons enfin que des études de liaison génétique basées sur l'utilisation des polymorphismes de longueur de fragments de restriction (RFLP) ont démontré de façon claire l'association de la forme dominante de l'OI sans sclérotiques bleues (type IV) au gène de la chaîne pro  $\alpha 2(I)$  [43]. Cette liaison génétique a été confirmée dans un pedigree par la description d'une délétion d'environ 10 résidus d'acides aminés dans une chaîne pro  $\alpha 2(I)$  [44].

## Mutants et variants

La seule étude du collagène de type I a révélé un très grand nombre de mutations différentes, menant à un éventail clinique allant d'une forme létale à une quasi-normalité. Et je n'ai pu mentionner en détail certaines maladies dominantes associées à d'autres mutations du collagène de type I (par exemple, certaines formes du syndrome d'Ehlers-Danlos de type VII). Dans la majorité des cas, il est frappant de constater que la structure de la molécule peut être sévèrement altérée sans que le développement fœtal ne soit interrompu. L'intégrité du collagène ne semble donc pas indispensable au développement de l'organisme, mais bien à sa résistance. D'autre part, le nombre de sites où une altération structurale du collagène aura un effet sur la qualité de la matrice extracellulaire, est très élevé, com-

\* Voir lexique m/s n° 5, vol. 2, p. 276.



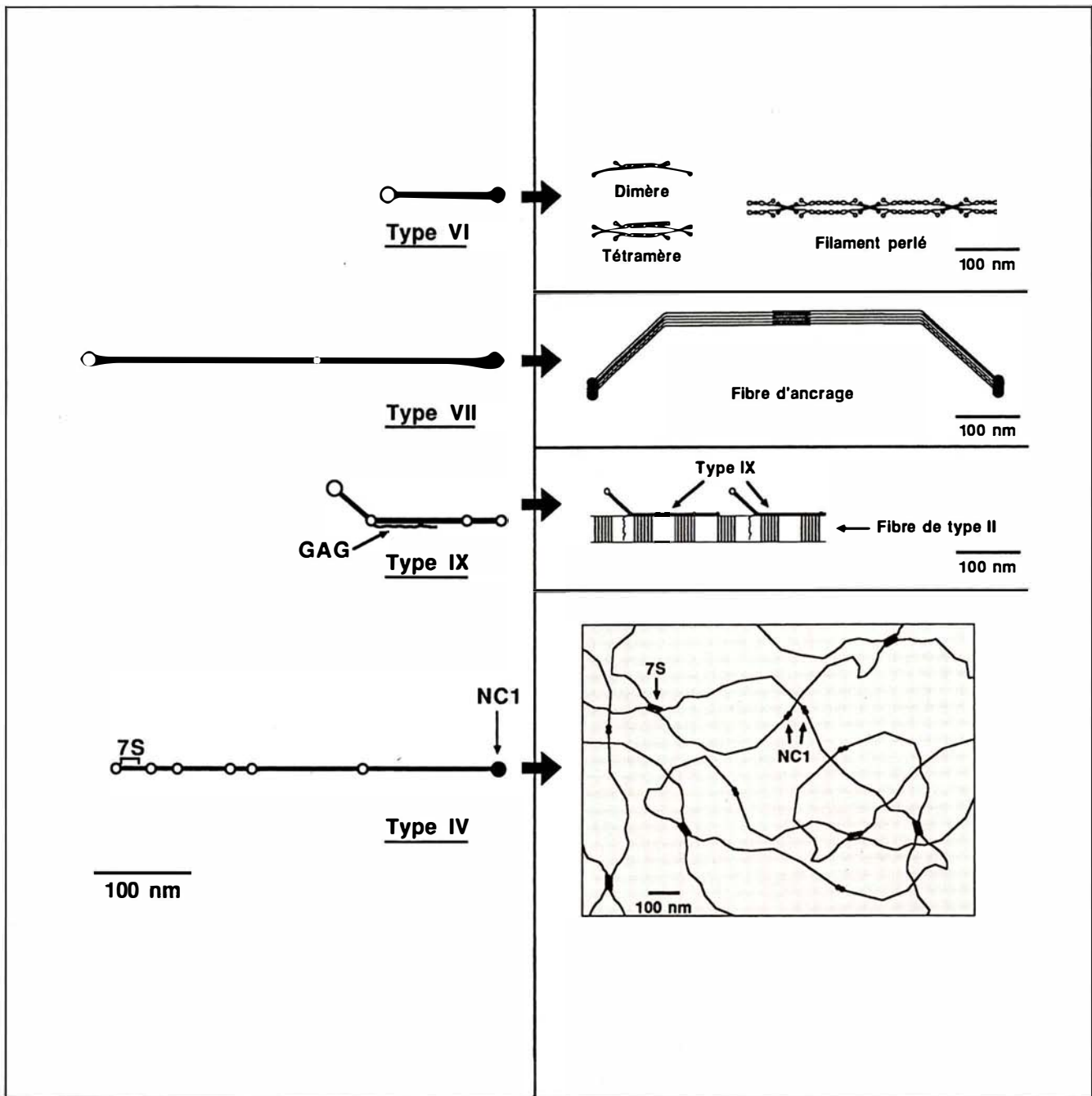


Figure 4. **Modes d'agrégation de quelques collagènes non fibrillaires.** Les molécules constitutives sont schématisées à gauche ; (—) : région hélicoïdale ; (○) : interruptions de la triple hélice (plus de 5 résidus) ; (○) : régions globulaires amino-terminales ; (●) : régions globulaires carboxy-terminales ; GAG = glycosaminoglycane. Les petites imperfections que l'on a démontrées dans les types IV et IX ne sont pas indiquées. Dans la partie droite sont indiqués les modes d'agrégation de ces collagènes. Le type VI forme d'abord des dimères de deux molécules antiparallèles, puis se forment des tétramères. Enfin, ces derniers s'accrochent par leurs domaines globulaires pour former les filaments perlés. Le type VII forme d'abord un dimère (non représenté) par chevauchement de l'extrémité N-terminale de deux molécules antiparallèles, puis un agrégat latéral non étagé pour former les fibres d'ancrage qui se lient à la membrane basale, probablement par les domaines globulaires C-terminaux. Le type IX se fixe sur la fibre de type II. La manière dont il servirait de lien avec d'autres éléments de la matrice du cartilage est encore hypothétique. Le collagène de type IV forme un tétramère (« l'araignée ») par l'interaction des régions 7 S amino-terminales de quatre molécules disposées tête-bêche. Les extrémités globulaires carboxy-terminales (NC 1) s'assemblent en dimères pour former un réseau irrégulier où viennent se fixer les autres constituants de la membrane basale.

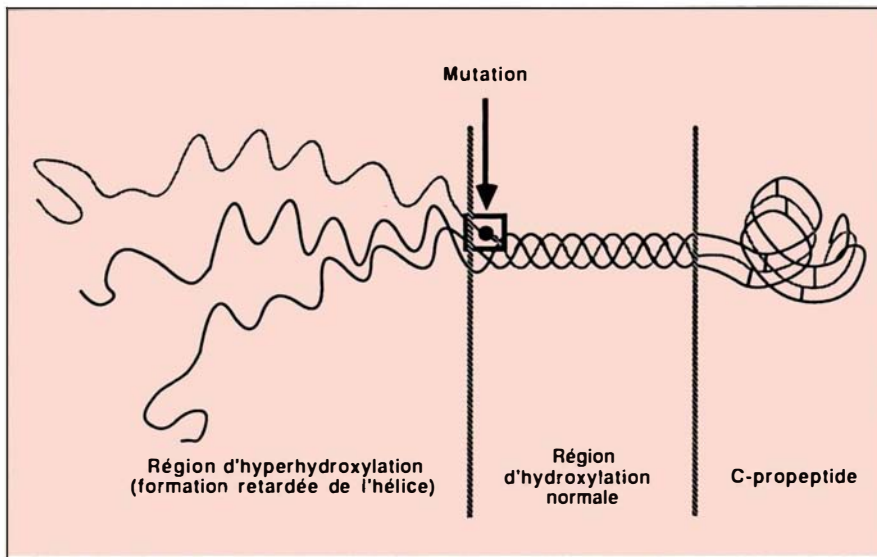


Figure 5. **Effet d'une substitution de glycine en position centrale de la triple hélice.** L'assemblage des propeptides C-terminaux se fait normalement et l'hélice s'amorce jusqu'à la mutation qui retarde la propagation de l'hélice vers l'extrémité amino-terminale, entraînant une hypermodification des résidus lysines et prolines dans cette région. Ce type de mutation semble fréquent dans l'ostéogenèse imparfaite létale.

me l'indiquent la grande variété des mutations déjà démontrée ainsi que la grande conservation de la structure primaire et du polymorphisme du collagène entre les espèces. Ceci suggère qu'il puisse exister chez l'homme des collagènes mutants dont la présence, sans entraîner de maladie clairement détectable chez tous les porteurs, affaiblirait la résistance des tissus conjonctifs de façon apparente uniquement dans des conditions de stress exceptionnelles. On sait par exemple que l'absence totale ou quasi totale du collagène de type III a été associée au syndrome d'Ehlers-Danlos de type IV, caractérisé par la fragilité des vaisseaux sanguins. Quel sera l'effet de mutations du collagène de type III, similaires à celles observées dans l'OI pour le collagène de type I, sur la résistance des vaisseaux sanguins ?

### Conclusion

Au moins vingt gènes différents codent pour les chaînes de collagène. La grande diversité des mutations observées au niveau du

collagène de type I s'étendra sans doute aux autres collagènes. On peut s'attendre à ce que les études biochimiques qui viendront dévoiler les défauts sous-jacents à de nombreuses affections héréditaires des tissus conjonctifs, rendent, dans une première phase, la classification de ces maladies encore plus complexe que la classification clinique existante. Cependant, et cela se produit déjà pour les maladies liées au collagène de type I, émergera alors une nouvelle classification qui permettra, par exemple, de distinguer les mutations quantitatives des mutations qualitatives. Parmi ces dernières, on peut maintenant distinguer celles entraînant une perturbation de la formation de la triple hélice, celles affectant la chaîne  $\alpha 2(I)$ , etc. A long terme, ces études devraient rendre possible un affinement du diagnostic et l'établissement d'une corrélation entre la nature du défaut biochimique et ses conséquences cliniques. Ainsi, une meilleure adéquation entre la nature réelle de la maladie et la thérapeutique envisagée sera-t-elle peut-être possible ■

### Summary

The extracellular matrix serves a number of different functions and this is manifested in the polymorphism of its major structural component, collagen. At least 11 collagen types, encoded in more than 20 different genes, are present in the various connective tissues. Because of its helicoidal structure, which requires the strict maintenance of a glycine at every three amino-acid residues and because of the lateral interactions between collagen molecules themselves and between collagen and other matrix components, interactions which involve many other amino-acid residues, mutations affecting the collagen structure are likely to be numerous and diversified in their consequences. This is already exemplified in mutations affecting type I collagen, many of which cause the various forms of Osteogenesis Imperfecta. The clinical manifestations of described mutations range from normality to the lethal perinatal form. It is suggested that mutations affecting the various collagen types might be expressed not only in the form of clearly identifiable heritable disorders, but also in the propensity of individuals carrying variant collagen genes to develop, in an adverse environment, certain pathological conditions (e.g. aneurisms or osteoarthritis) that might result from a weakened extracellular matrix.

### TIRÉS A PART

M. van der Rest : unité de génétique, hôpital Shriners pour les enfants infirmes, 1529, avenue Cedar, Montréal, QC, Canada H3G 1A6.