Biochem. J. (1985) **229**, 287–303
Printed in Great Britain

REVIEW ARTICLE

Collagen genes and inherited connective tissue disease

Kathryn S. E. CHEAH
Department of Biochemistry, University of Hong Kong, Sassoon Road, Hong Kong

Introduction

Since it was established that defects in collagen metabolism were associated with inherited connective tissue disease (McKusick, 1972), considerable effort has been put into discovering the molecular basis of these genetic disorders. Up to 5 years ago a link between structural mutations in collagen genes and inherited connective tissue disease was considered likely, but there was little direct evidence to demonstrate the association (Sykes *et al.*, 1977; Prockop *et al.*, 1979). But remarkable progress has since been made in analysing the molecular basis of inherited collagen disease at the gene level (Shapiro & Rowe, 1983; Prockop & Kivirikko, 1984; Byers & Bonadio, 1984). This article reviews the recent progress in seeking out the molecular defects in collagen genes associated with inherited disease and discusses the impact this knowledge has on our understanding of the control of collagen biosynthesis and the pathology of the disorders. In addition, I consider present and future prospects for detecting other mutations with the ultimate aim of identifying patients at risk.

The collagen family

The collagens are major structural components of the extracellular matrix. In vertebrates they consist of a family of at least 10 or more genetically distinct protein types, for which a minimum of 18 genes must exist to code for their constituent α chains (Table 1; see Bornstein & Sage, 1980; Miller & Gay, 1982; Burgeson, 1982, for reviews). Collagenous molecules contain triple helical regions characterized by the presence of a basic repeating triplet unit of amino acids consisting of Gly-Xaa-Yaa. Although a variety of amino acids occupy the Xaa and Yaa positions, proline and hydroxyproline often occupy the Xaa and Yaa positions respectively. This triplet unit is essential for the formation of a semi-rigid triple-helical molecule composed of three chains. It is the variation in length, primary structure and the presence, extent and location of both noncollagenous domains and interchain disulphide-bonding which form the basis for categorizing the different collagen types.

The types, quantity, and occasionally the molecular composition (type V) of collagen expressed varies in different tissues (Table 1; Bornstein & Sage, 1980; Hollister *et al.*, 1983; Alitalo *et al.*, 1983; Shimokamaki *et al.*, 1980; von der Mark *et al.*, 1982; Odermatt *et al.*, 1983; Sage *et al.*, 1981, 1983; Haralson *et al.*, 1980, 1984; Niyibizi *et al.*, 1984). Types I–III have been classed as interstitial collagens, having a fibrillar structure. Of the nonfibrillar collagens, type IV, a major component of basement membranes, is best understood. The remaining types are not well characterized and are usually present in small amounts. While structure–function relationships of the different collagens remain to be defined, it is believed that the coordinate expression of these collagen genes is important in vertebrate development and differentiation (Adamson, 1982). There is also evidence indicating that the expression of type I collagen during development is essential for life. Its absence cannot be compensated for by other collagen types. This is illustrated by the finding that the integration of a retrovirus into the 5′ end of the α1(I) gene results in developmental arrest in early mouse embryos which is lethal (Schnieke *et al.*, 1983; Harbers *et al.*, 1984; Lohler *et al.*, 1984).

To date the gene structures for only a few collagen α chains have been determined. (Table 2). The complete structure of the chicken α2(I) gene was the first to be established (see Cheah & Grant, 1982; Boedtker *et al.*, 1983, for reviews). From these studies a remarkable picture emerges for collagen gene structure. Although a minimum of 5kb is required to code for a pro-α2(I) chain (1500 amino acids) it was found that the actual size of the chicken gene was approx. 40kb divided into 51 exons (Vogeli *et al.*, 1980; Wozney *et al.*, 1981). The genes for chicken type III collagen (Yamada *et al.*, 1983, 1984) and for human type II (Cheah *et al.*,

Abbreviations used: OI, osteogenesis imperfecta; EDS, Ehlers Danlos syndrome; pro-α1^{s}, shortened pro-α1(I) chain; type I^{ss}, type I collagen with two shortened α1(I) chains; CB, CNBr peptide; kb, kilobases; bp, base pairs; RFLP, restriction fragment length polymorphism.

Table 1. *The vertebrate collagen gene family: types, molecular composition and tissue distribution*
Types I–V collagen has been the best characterized but a number of other collagen types have been recently discovered [Furthmayr *et al.* (1983), Odermatt *et al.* (1983) (type VI); Bentz *et al.* (1983) (type VII); Sage *et al.* (1983) (type VIII); Mayne *et al.* (1984), Ninomiya & Olsen (1984), van der Rest *et al.* (1985) (type IX); Gibson *et al.* (1982, 1983), Schmid & Conrad (1982), Kielty *et al.* (1985) (type X)] and some await type assignments [Burgeson & Hollister (1979), Reese & Mayne (1981), Burgeson *et al.* (1982) (collagen 1α, 2α, 3α)]. The designation of 1α. 2α, and 3α as a new type (XI) or as species related to other collagen types (V and II) is controversial, hence the question mark in parentheses. Other names given to the types before their classification are: short chain/intima (type VI); long chain (VII); EC (type VIII); HMW, LMW, CPS-1, CPS-2, M collagen, H, and J (type IX or part of the intact molecule); G, cartilage short chain (type X).

Collagen type	Constituent α chains	Chain association	Major tissue distribution
I	α1(I) α2(I)	$[\alpha1(I)]_2\alpha2(I)$	Skin, tendon, bone, placenta, arteries
I trimer	α1(I)	$[\alpha1(I)]_3$	Skin, tendon, liver, dentine, cultured skin, fibroblasts, tumours
II	α1(II)	$[\alpha1(II)]_3$	Cartilage, vitreous humour, chondrosarcoma, intervertebral disc.
III	α1(III)	$[\alpha1(III)]_3$	Skin, lung, arteries, uterus, chorioamnion, liver, stroma and connective tissue of organs
IV	α1(IV) α2(IV)	$[\alpha1(IV)]_2\alpha2(IV)$	Basement membranes
V	α1(V) α2(V) α3(V)	$[\alpha1(V)]_2\alpha2(V)$ $[\alpha1(V),\alpha2(V),\alpha3(V)]$ $[\alpha1(V)]_3$	Placenta, skin, chorioamnion, pericellular/stromal in most connective tissue, rhabdomyosarcoma
VI	α1(VI) α2(VI) α3(VI)	Unknown	Blood vessels, uterus, ligament, skin, lung, kidney
VII	α1(VII)	$[\alpha1(VII)]_3$	Chorioamniotic membranes, skin, oesophagus
VIII	α1(VIII)	Uncertain; 'cassette' structure composed of collagenous polypeptides proposed	Culture medium from endothelial, astrocytoma and other cells from normal and malignant tissues
IX	α1(IX) α2(IX) α3(IX)	$[\alpha1(IX)\alpha2(IX)\alpha3(IX)]$	Cartilage, vitreous humour, intervertebral disc
X	α1(X)	$[\alpha1(X)]_3$	Cartilage
XI(?)	1α 2α 3α	$(1\alpha2\alpha3\alpha)$	Cartilage, vitreous humour, intervertebral disc

1985) are likewise large, approx. 38 to 30 kb respectively. However, the human αl(I) gene is 18 kb long. Although it too is divided into 51 exons, its relatively compressed structure is the result of smaller introns (Chu *et al.*, 1984).

Another distinctive feature of these genes is the presence of 45, 54, 99 and 108 bp exons which encode exact multiples of Gly-Xaa-Yaa triplets, in the regions which code for the triple-helical domains. Exons of 54 and 108 bp are most common in this domain and variations thereof could have originated by removing or adding the coding potential for a Gly-Xaa-Yaa triplet. It has therefore been suggested that collagen genes arose by duplication and recombination of an ancestral 54 bp primordial gene (Yamada *et al.*, 1980; reviewed by Solomon & Cheah, 1981; Boedtker *et al.*, 1983). This 54 bp exon size is common to all the fibrillar collagen genes isolated thus far (Yamada *et al.*, 1983; Sandell *et al.*, 1983, 1984; Chu *et al.*, 1984; Cheah *et al.*, 1985). But it is of note that 54 bp exons are not seen in the helical coding region of the genes for nonfibrillar collagens in two invertebrates, *Drosophila* and nematode (Monson *et al.*, 1982; Kramer *et al.*, 1982). Until the gene structures of vertebrate nonfibrillar collagens and of ancient collagens, e.g. spongin, are available for comparison the primordial origin of this large gene family remains uncertain.

Collagen biosynthesis

The synthesis of a collagen molecule is a complex process consisting of many post-transcriptional and post-translational steps. These processes have been most extensively studied for type I collagen and its synthesis is therefore used as

Table 2. *Vertebrate collagen gene probes isolated to date*

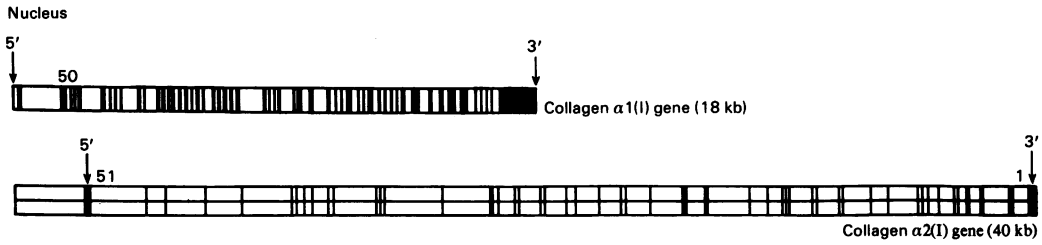
Collagen gene	Species	Type of DNA probe		Extent of gene covered	References
		cDNA	Genomic		
α1(I)	Chicken	+		Partial	Lehrach *et al.* (1979); Yamamoto *et al.* (1980)
	Mouse		+	Partial	Monson & McCarthy (1981)
	Human	+	+	Complete	Chu *et al.* (1982, 1984); Barsh *et al.* (1984)
	Rat	+			Genovese *et al.* (1984)
α2(I)	Chicken	+	+	Complete	Lehrach *et al.* (1978); Sobel *et al.* (1978); Vogeli *et al.* (1980); Wozney *et al.* (1981)
	Mouse		+	Partial	Schmidt *et al.* (1984)
	Sheep		+	Partial	Schafer *et al.* (1980)
	Human	+	+	Complete	Myers *et al.* (1981, 1983); Dalgleish *et al.* (1982); Tajima *et al.* (1984); Duvic *et al.* (1984)
	Rat	+			Genovese *et al.* (1984)
α1(II)	Chicken	+	+	Partial	Sandell *et al.* (1983); Lukens *et al.* (1983); Ninomiya *et al.* (1984); Young *et al.* (1984)
	Human	+	+	Complete	Strom & Upholt (1984); Cheah *et al.* (1985); Elima *et al.* (1985);
	Rat	+			Kohno *et al.* (1984)
α1(III)	Chicken		+	Almost complete	Yamada *et al.* (1983, 1984)
	Human	+	+	Partial	Loidl *et al.* (1984); Chu *et al.* (1985*b*); R. Dalgleish (personal communication)
	Mouse		+	Partial	Liau *et al.* (1985)
α1(IV)	Mouse	+		Partial	Kurkinen *et al.* (1983); dos Santos *et al.* (1984)
α2(IV)	Mouse	+		Partial	Wang & Gudas (1983)
α1(IX)	Chicken	+		Complete	Ninomiya & Olsen (1984)

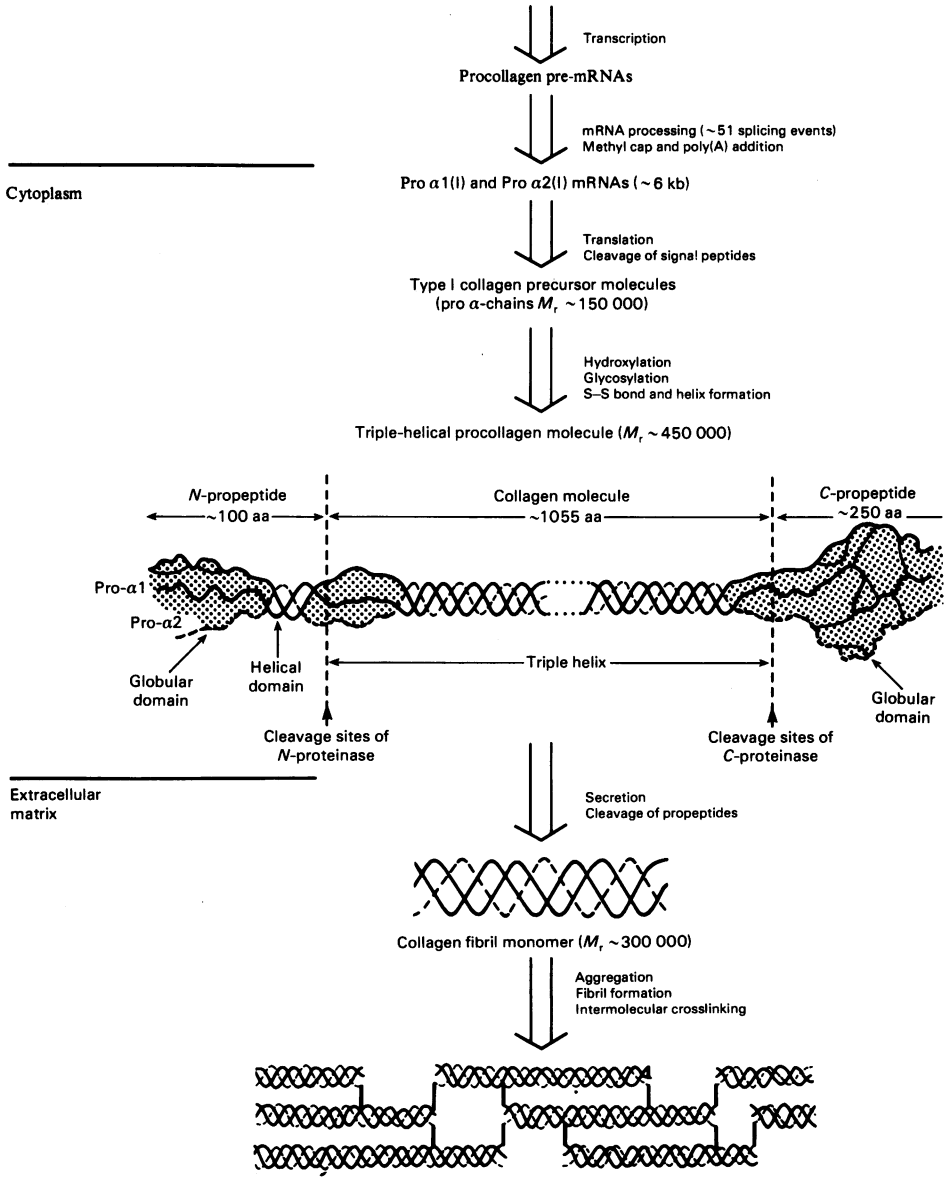
an illustration from which general principles may be applied for the other collagens (Fig. 1). Detailed descriptions of this process may be found in reviews by Grant & Jackson (1976), Fessler & Fessler (1978), Prockop *et al.* (1979) and Hollister *et al.* (1983).

Type I collagen is synthesized as a precursor molecule, procollagen, consisting of two pro-α1(I) chains and a pro-α2(I) chain which associate to form a triple helix with globular domains at the *N*- and *C*- termini, the *N*- and *C*- propeptides. As illustrated in Fig. 1, there are many steps in this complex process at which errors in biosynthesis may occur and lead to the production of a faulty collagen molecule, hence contributing to the pathology of disease.

Inherited collagen disease

Genetic disorders of collagen metabolism usually affect tissues in which the proper development and integrity of connective tissue is of paramount importance (for reviews see Hollister *et al.*, 1983; Pinnell & Murad, 1983; Krane, 1984). Hence the majority of these disorders affect bone (osteogenesis imperfecta; OI), skin and joints (Ehlers Danlos syndrome; EDS), arteries (Marfan syndrome) and cartilage development (chondrodystrophies). They form an extremely heterogeneous class within the wider category of heritable disorders with connective tissue involvement which include cutis laxa (loose, inelastic skin), Menkes syndrome (abnormal copper metabolism), epidermolysis bullosa (abnormal and repeated blister formation) and alkaptonuria (degenerative joint disease). However, not all these inherited diseases are associated with structural defects in collagen genes. The first true molecular defect in collagen metabolism to be identified was due to lysyl hydroxylase deficiency in EDS VI (Pinnell *et al.*, 1972) and in the majority of the other disorders


Nucleus
5′
3′
50
Collagen α1(I) gene (18 kb)
5′
3′
51
1
Collagen α2(I) gene (40 kb)


Transcription
Procollagen pre-mRNAs
mRNA processing (~51 splicing events)
Methyl cap and poly(A) addition
Cytoplasm
Pro α1(I) and Pro α2(I) mRNAs (~6 kb)
Translation
Cleavage of signal peptides
Type I collagen precursor molecules
(pro α-chains M_r ~150 000)
Hydroxylation
Glycosylation
S–S bond and helix formation
Triple-helical procollagen molecule (M_r ~450 000)
N-propeptide
~100 aa
Collagen molecule
~1055 aa
C-propeptide
~250 aa
Pro-α1
Pro-α2
Globular domain
Helical domain
Triple helix
Globular domain
Cleavage sites of N-proteinase
Cleavage sites of C-proteinase
Extracellular matrix
Secretion
Cleavage of propeptides
Collagen fibril monomer (M_r ~300 000)
Aggregation
Fibril formation
Intermolecular crosslinking

the biochemical basis remains unclear. While the biochemistry of the structure and biosynthesis of collagens I–V are relatively well understood, specific defects in a few collagen α chains and their phenotypic consequences have been shown in only a limited number of patients with inherited disease.

The clinical, genetic and biochemical hallmarks of the three best studied types of inherited diseases of collagen in which abnormalities of collagen genes and metabolism have been found (OI, EDS and Marfan syndrome) are summarized in Table 3.

The molecular basis of inherited collagen disease

Experimental strategies

The simplest method used to detect changes in collagen metabolism has been to compare the gross morphologies of the affected tissue, usually skin, bone or cartilage, by light and electron microscopy (Sillence, 1981; Holbrook & Byers, 1982; Hollister *et al.*, 1983, for reviews). Because of the relative ease of obtaining skin biopsies this tissue has been most extensively studied, whether the primary defect is in skin, bone or arteries (Holbrook & Byers, 1982). Abnormalities in general collagen fibre structure in skin are seen for a wide range of inherited disease e.g. EDS, OI, pseudoxanthoma elasticum, cutis laxa, Marfan syndrome, etc. But these studies also show that there are no unique disease-specific structural alterations in collagen fibril structure. In addition, there is difficulty in correlating morphological changes with biochemical changes. Intracellularly, engorged endoplasmic reticulum has also been observed in skin fibroblasts from OI and EDS patients (Byers *et al.*, 1979; Holbrook & Byers, 1981; Hollister *et al.*, 1983; Steinmann *et al.*, 1984), and there is evidence for the accumulation of type I procollagen in OI cells (Barsh & Byers, 1981). These morphological changes may be symptomatic of structural abnormalities within the molecule which result in reduced secretion.

Biochemical analyses probe deeper for the defect. Again, because of the heterogeneity of inherited collagen disease, it has been difficult to assign specific biochemical characteristics to each type or sub-class. However the primary collagen defect in a small number of patients has been successfully identified. Comparisons of amino acid compositions, levels of hydroxyproline and hydroxylysine and the extent of glycosylation in collagen isolated from patient and control tissue may indicate the primary cause of the defect (Peltonen *et al.*, 1980; Hollister, 1981; Byers *et al.*, 1981*a*; Hollister *et al.*, 1983). The relative amounts of the different collagens in the affected tissue may also indicate the nature of the abnormality. For example, altered type III:type I collagen ratios have been found in the skin from a number of OI patients (Sykes *et al.*, 1977; Barsh *et al.*, 1982) and have been shown to be due to reduced type I production. The absence or reduction of type III collagen synthesis in skin fibroblasts from patients with EDS IV has been shown (Pope *et al.*, 1975; Aumailley *et al.*, 1980; Hollister *et al.*, 1983), although there is an instance where normal amounts have been found (Sulh *et al.*, 1984). Inappropriate expression of collagen type may also play a role in the pathogenesis of disease. Two examples of perinatal lethal OI have been reported where type III collagen was found in bone in which it is not normally present. Also, the level of type V collagen in OI bone was elevated (Pope *et al.*, 1980).

Comparison of the molecular sizes of the collagen and pro-α chains by gel electrophoresis or column chromatography may reveal structural alterations in these chains. But the extent of post-translational modification affects the relative mobilities of collagenous molecules in polyacrylamide gels (Cheah *et al.*, 1979). Thus, comparative studies of the procollagen chains synthesized in the presence or absence of inhibitors of hydroxylation often enable an apparent M_r change due to a structural defect in the molecule to be distinguished from one arising because of faulty processing enzymes (Barsh & Byers, 1981; Byers *et al.*, 1983*a*; Bonadio *et al.*, 1985).

The collagen molecule itself contains several landmarks which have proven useful for the determination of approximate locations of structural

Fig. 1. *Biosynthesis of procollagen I*

The pathway of synthesis of procollagen type I from transcription of the α1(I) and α2(I) genes to fibril formation is shown diagrammatically. The genes for pro-α1(I) and pro-α2(I) chains are transcribed and the 50 introns excised to yield the mature transcript. These procollagen mRNAs are translated on membrane-bound polysomes (Cheah & Grant, 1982, for review) and pre-pro chains containing signal peptides at their *N*-termini are synthesized. The signal peptides are cleaved off within the rough endoplasmic reticulum, glycosylation, hydroxylation, disulphide bonding and helix formation occurs and the mature procollagen molecule is secreted to the extracellular matrix. Outside the cell the propeptides are cleaved by specific *N*- and *C*- proteinases and the collagen monomers are assembled to form collagen fibres by aggregation and cross-linking. The drawing of the human α1(I) gene is reproduced from Chu *et al.* (1984) and is reprinted with the permission of the authors and of the publishers (Macmillan Journals Ltd.). The structure of the chicken α2(I) gene is reproduced from Vogeli *et al.* (1980) with the permission of the authors and the publishers (Cold Spring Harbour Laboratory). Gene structure: black box, exon; open box, intron.

Table 3. *Summary of clinical and biochemical features of inherited connective tissue diseases associated with a collagen gene defect*

See the text for reference sources. Abbreviations used: + −, heterozygosity; AD, autosomal dominant; AR, autosomal recessive; CB3/CB5, CNBr peptides. Note that the defective collagen chains described were found in only 13 patients, in whom each defect was unique. In the vast majority of patients studied no defined abnormality in the collagen polypeptides could be assigned.

Disease	Number of distinct types	Type (number of sub-types)	Inheritance	Characteristic symptoms	Biosynthetic abnormality	Collagen chain defect
Osteogenesis imperfecta (OI)	>6	I (3)	AD (mostly + −)	Mildest. Brittle bones; childhood onset of fractures, short stature, joint laxity, early hearing loss. Blue sclerae, dentinogenesis imperfecta present or absent.	Decreased type I collagen synthesis. Normal type III collagen.	Non-functional α1(I) chains (+ −). Cysteine present in α1(I) triple helix (+ −). Shortened α2(I) (+ −).
		II (3)	AR (+ −; compound) AD (+ −; new mutation)	Perinatal lethal. Multiple bone fractures, deficient ossification, tissue fragility, dwarfism, dark blue sclerae.	Decreased type I collagen synthesis and secretion. Intracellular accumulation, overmodification and increased degradation. Presence of type III in bone. Increased amount of type V in bone. Increased type III synthesis.	Lowered melting point of type I, unstable helix. Deletions in α1(I). Glycine → cysteine substitution in α1(I). Structural alterations in α1(I) or α2(I); delayed helix formation and overmodification. Possible point mutations. Insertion in α2(I). Non-functional α2(I) allele and a deletion of 18 amino acids (exon 25) in the other allele.
		III	AR	Progressive deformity, extreme bone fragility, frequent fractures, growth failures. Normal sclerae.	Presence of type I trimer, absence of normal type I. Decreased collagen production. Excess sugar residues in the *C*-propeptides of type I.	α2(I) chains synthesized abnormal because of a 4 bp deletion, causing a frame shift mutation and altered *C*-propeptide amino acid sequence.
		IV	AD	Similar to III, white sclerae. Bone fragility; dentinogenesis imperfecta may be present.	Mixture of normal and overmodified type I synthesized. Intracellular retention of overmodified type I.	Decreased T_m of overmodified molecules. Altered α2(I) chain affecting triple helix formation?
Non-assigned combinations			Heterogeneous	Combined clinical features of other collagen diseases e.g. OI/EDS; OI I/III.	Noncleavage of amino propeptide (+ −) of α2(I); persistence of pN collagen.	Shortened α2(I) resulting in alteration of *N*-proteinase recognition site? 54 bp (exon 42) deletion in α2(I) gene resulting in 18 amino acid shortening of α2(I) chain.

Ehlers Danlos syndrome (EDS)	10	IV	AD AR	Vascular variety. Thin, translucent and fragile skin. Visible veins, easy bruising. Bowel and arterial rupture, common uterine rupture near term in pregnancy. No joint hypermobility.	Decreased or absent synthesis of type III collagen. One example of normal type III synthesis. Decreased secretion of type III and increased intracellular retention.	Mostly not known. Apparent amino acid insertion in α1(III) (+ −).
		VII	AD AR	Extreme joint laxity, soft but non-fragile skin.	Persistence of *N*-propeptide in 50% of pro-α2(I).	Mutation affecting *N*-proteinase recognition site? (+ −). Insertion in α2 chain. CB4? new mutation.
Marfan syndrome	4		AD	Severe cardiovascular abnormalities, skeletal alterations (long, thin, limbs). Ocular lens dislocation.	Increased collagen solubility. Abnormal α2 chain synthesized.	Unknown. 20 amino acid insertion in α2(I) CB3/CB5 and 38 bp insertion in intron (same patient). Unknown effect.

defects. For example, each collagen α chain yields a characteristic set of CNBr peptides delineated by the position of methionine residues. Cysteines mark the characteristic position of interchain and intrachain disulphide bonding. Type I procollagen is disulphide-bonded only in the *C* propeptides and not in the helix, but there are disulphide bonds in the helix of type III procollagen (Fessler & Fessler, 1978). Mammalian collagenases cleave type I collagen once, specifically within the helix, to give two fragments, one one-quarter (*C*-terminal, B fragment) and the other three-quarters (*N*-terminal, A fragment) the length of the original molecule (Werb, 1982). In the extracellular matrix, *N*- and *C*-proteinases cleave off the *N*- and *C*-propeptides at specific sites (Heathcote & Grant, 1980).

Direct analysis of collagen gene structure has been the most recent experimental approach in the search for the cause of inherited collagen disease and in the study of their genetics. For this the availability of DNA probes for human collagen genes is essential and to date ones for types I–III have been isolated (Table 2). These probes may be used to detect insertions and deletions in the genes. Collagen gene probes may also be used to screen for polymorphic markers with which one may detect the defective locus or predict the occurrence of an inherited disease by linkage analysis. Apart from structural comparisons, biosynthetic studies of collagen mRNAs and the rates of synthesis and degradation of pro-α chains in normal and patient tissue provide important indications for the site of the primary defect.

In the vast majority of patients studied no abnormality in collagen at the primary level has been detected. But using a combination of all these experimental approaches described above the molecular basis of inherited collagen disease in a small number of patients has been determined.

Structural mutations in type-I collagen genes

The structural alterations found in type I procollagen in patients with OI, EDS and Marfan syndrome are illustrated and described in Fig. 2 and Table 3. These include deletions or insertions within the α1(I) and α2(I) genes and subtle changes resulting in overmodification or non-processing of the procollagen molecule.

Deletions, insertions and amino acid substitutions in α1(I)

One of the two best studied examples of a structural mutation in type I collagen is a patient with lethal OI in which collagen production was reduced in skin fibroblasts compared with normal cells (Penttinen *et al.*, 1975). Cultured skin fibroblasts from this patient secreted type I procollagen

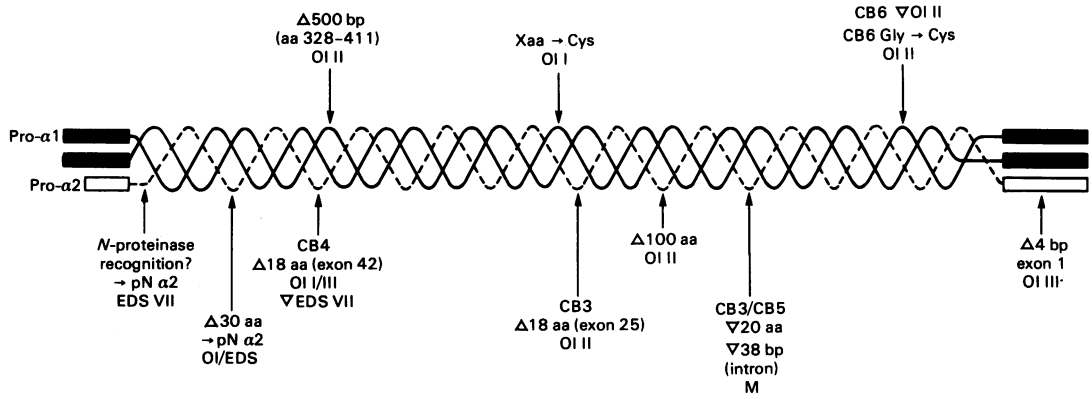


Fig. 2. *Structural alterations in type I procollagen found in patients with inherited collagen disorders*
Diagrammatic representation of type I procollagen showing the approximate locations of structural alterations found, some of which affect the biosynthesis of the pro-α1(I) and pro-α2(I) chains. Mutations affecting pro-α1(I) chains are indicated above the molecule. Those affecting pro-α2(I) are shown below the molecule. The absence of a size indication for an insertion or a deletion implies it is unknown. Symbols and abbreviations: △ represents a deletion; ▽ represents an insertion; OI, osteogenesis imperfecta; EDS, Ehlers Danlos syndrome; M, Marfan syndrome; Xaa, amino acid of unknown identity.

at half the rate of normal cells and there was increased intracellular accumulation of these molecules. Analyses of the pro-α1(I) chains synthesized by these cells in the presence or absence of α′α-dipyridyl, an inhibitor of hydroxylation, indicated that half the pro-α1(I) chains were altered in length, but not the pro-α2(I) chains (Barsh & Byers, 1981). It was proposed that the abnormal type I procollagen molecules contained either one or two shortened pro-α1(I) chains (pro-α1^{s}). The population of type I procollagen molecules therefore consisted of normal type I, ones with one short pro-α1(I) chain (type I^{s}) and ones with both pro-α1(I) chains shortened (type I^{ss}) in the ratio of 1:2:1 (Williams & Prockop, 1983). Pro-α1^{s} chains were able to associate with pro-α2(I) chains to form a triple helix although such molecules (type I^{s} and type I^{ss}) were less thermally stable and were not cleaved by procollagen *N*-proteinase. Because the melting temperature of the abnormal molecules was reduced below 37°C it was also unlikely that they could remain helical at body temperature. Molecules with one shortened α1(I) chain were also secreted more slowly (hence the observation of intracellular accumulation) and were rapidly degraded in the extracellular matrix. Those containing two shortened chains were either not secreted or degraded even more rapidly. Cell-free translation of mRNA from these cells yielded shortened and normal-sized pro-α1(I) chains, indicating that the defect lay in the gene (Chu *et al.*, 1983). Digestion of the abnormal trimers with mammalian collagenase indicated that the shortening of the pro-α1(I) chain lay *N*-terminal to the cleavage site. Because the entire gene for human α1(I) collagen has been isolated it was possible to locate the exact position of the missing portion of the chain. Using a combination of cDNA and genomic probes an internal deletion of approx. 500 bp was found in one allele (Chu *et al.*, 1983). DNA sequencing of the isolated defective allele reveals that the deletion effectively removed three exons (nos. 27–29) coding for 84 amino acids in the triple helix (amino acids 328–411) (Barsh *et al.*, 1985; Chu *et al.*, 1985*a*). The loss of 84 amino acids but the retention of the reading frame therefore resulted in the synthesis of shortened pro-α1(I) chains which could form a triple helix, albeit of reduced stability. Despite heterozygosity, the severity of the disease in this patient leading to perinatal death was probably because the molecular composition of type I collagen ensured that three-quarters of the molecules synthesized were abnormal and therefore ineffective in providing adequate tissue integrity. There was also some evidence for an elevation of type III collagen synthesis and gene expression (Williams & Prockop, 1983; Chu *et al.*, 1985*a*), perhaps in a compensatory attempt to overcome the type I deficiency. It is significant, however, that as for the developmentally lethal homozygous mutations in the mouse α1(I) gene (Schnieke *et al.*, 1983, Lohler *et al.*, 1984), the defect could not be overcome by increased synthesis of another collagen type.

Lengthening of α1(I) chains may also result in lethal disease. In two patients with lethal OI all pro-α chains synthesized were lengthened (Uitto *et al.*, 1983) and the pro-α chains were post-translationally overmodified, possibly because of a

reduced rate of helix formation. The location of the lengthening in the pro-α1(I) chain was tentatively assigned to CNBr peptide CB6 which is located beyond amino acid residue 820 in the helix. The exact effect of lengthening pro-α1(I) chains on helix formation and on the functional ability of the resulting type I collagen molecules remains to be defined. But it is likely that nonfunctional helices may be formed also with deleterious effect.

There is evidence that a mutation resulting in complete non-function in one α1(I) allele may cause disease but is less severe than one producing malfunctioning type I collagen molecules as just described. Barsh *et al.* (1982) have studied three patients with less severe, non-lethal type of OI (type I) in which reduced levels of type I collagen was found in bone and skin. Cultured skin fibroblasts from these patients synthesized one-half the normal level of type I collagen which was reflected in the amount of α1(I) chains, since pro-α1(I) and pro-α2(I) chains were in a 1:1 ratio instead of the normal 2:1. mRNA levels for pro-α1(I) were also reported to be reduced (D. Rowe, personal communication; Byers & Bonadio, 1984). A non-functioning allele for pro-α1(I) was suggested to be the cause for this defect. One may speculate that the lesser severity of this type of OI type I is because no abnormal α1(I) chains capable of being incorporated into a triple helix were synthesized. Hence all type I collagen molecules synthesized were normal, albeit at reduced level within the critical limit for life. This is a less severe situation than one in which three-quarters of the type I collagen molecules are abnormal, probably not helical at body temperature and rapidly degraded. The absence of triple-helical molecules composed of α2(I) chains confirms that they are unable to form triple helices and therefore unlikely to act as functional substitutes for α1(I).

The stability and formation of the collagen triple helix is dependent on the conservation of the Gly-Xaa-Yaa triplet. Destabilization of the type I collagen helix may result by substituting another amino acid for a glycine residue. Steinmann *et al.* (1984*b* and B. Steinmann, personal communication) have shown in skin fibroblasts from an OI type II patient that the substitution of a cysteine residue for glycine (residue 988) in CNBr peptide CB6 (towards the *C*-terminus) within one or both α1(I) chains of the triple helix results in reduced thermal stability of type I collagen, increased post-translational modification, increased intracellular degradation of collagen and decreased collagen production. In addition, collagen fibrils were found to be thinner than normal and abnormal intramolecular and interchain disulphide bonding was observed.

The molecular and functional effect of such a mutation remains to be proven. But this substitution probably delayed helix formation and would account for the increased post-translational modification, since hydroxylation and glycosylation do not occur once a helix is formed (Kivirikko & Myllyla, 1982). Again, despite heterozygosity, three-quarters of the type I collagen molecules synthesized would be unstable and/or nonfunctional and may be the major contributory cause of perinatal death. The molecular change in the α1(I) gene was the result of a single base change (GGT → TGT; B. Steinmann, unpublished work). Another example of the presence of cysteine in the α1(I) chains in skin fibroblasts and bone from an OI type I patient has been reported (Nicholls & Pope, 1984) although the location and effect on collagen biosynthesis has not yet been identified.

Structural alterations in α2(I)

The other well-characterized example of a structural mutation in type I collagen is in a patient with type III (moderately severe) OI, first noted by Nicholls *et al.* (1979) because of the apparent absence of pro-α2(I) chains in the collagen synthesized. All the collagen synthesized by the skin fibroblasts of the patient consisted of a trimer of pro-α1(I) chains. Subsequent study revealed that pro-α2(I) mRNA was present and pro-α2(I) chains were in fact synthesized both in cell-free translation and within the cells. But these pro-α2(I) chains could not associate with pro-α1(I) chains (Deak *et al.*, 1981, 1983). Because triple-helix formation begins at the *C*-terminus (Bachinger *et al.*, 1981; Bruckner *et al.*, 1981) and is dependent on the correct alignment of the *C*-propeptides it was likely that the primary defect lay within that region. Analyses of the mRNA and the isolated defective allele confirm this premise. Dickson *et al.* (1984) tested for mismatched regions between the mRNA and the coding regions of the gene by S1 nuclease digestion of RNA–DNA hybrids. These studies uncovered a region of mismatch within exon 1 of the α2(I) gene about 35 amino acids from the termination codon. This mismatch was found to be a 4 bp deletion in the gene which caused a frameshift mutation and created a new termination codon 3′ to the original site (Pihlajaniemie *et al.*, 1984). Consequently the amino acid sequence of the last 33 amino acids of the α2(I) chain was altered, changing the proportions of charged residues and creating a new cysteine residue. The phenotypic consequence of this mutation is that pro-α2(I) chains cannot form a triple helix with pro-α1(I) chains, confirming the critical role of the *C*-propeptide in helix formation. A secondary observation, as yet unexplained, was that pro-α2(I) chain synthesis was decreased. This example also implies that, unlike α1(I), the presence of α2(I)

chains in type I collagen is not essential for life, possibly because trimers of α1(I) chains are able to form a stable helix and provide partial functional compensation.

There is preliminary evidence that a deletion of approx. 100 amino acids from the central portion of the triple helical domain of α2(I) chains is the cause of another example of lethal OI (Byers *et al.*, 1983*b*; Byers & Bonadio, 1984). Both normal and shortened α2(I) chains were synthesized; however, helical molecules containing the mutant chains were unstable at 37°C and were degraded intracellularly. Other deletions in the α2(I) gene or shortening of pro-α2(I) chains have been reported. Interestingly, a number of these could correspond to a deletion of a 54 bp exon (18 amino acids or 6 Gly-Xaa-Yaa triplets). Fibroblasts from a type II OI patient synthesized only pro-α2(I) chains which were shortened by approx. 20 amino acids but could still form a helix (de Wet *et al.*, 1983). Digestion of the collagen with vertebrate collagenase and CNBr place the shortening in the peptide 2CB3/5^{A}, which lies in the middle of the helix (amino acids 361–775). A preliminary report of R-looping analysis of DNA–mRNA hybrids showed that exon 25, coding for 18 amino acids (i.e. a 54 bp exon), had been deleted (Sippola *et al.*, 1984*b*). Because neither of the proband's parents synthesized shortened α2(I) chains it was considered unlikely that both alleles could have mutated identically in the generation of a new mutation. But α2(I) mRNA levels and α2(I) chain synthesis were reduced in both father and patient. It was likely therefore that the shortened α2(I) chains originated from a mutated allele which was expressed while the other allele was not, or produced unstable transcripts. Similar-sized mutations are also probably involved in two other patients with a milder variant of OI intermediate between type I and III (Byers *et al.*, 1983*a*; Sippola *et al.*, 1984*a*). Both show a shortening by approx. 20 amino acids in the α2(I) CNBr peptide CB4. There is also a preliminary report in one patient that this is correlated, by R-looping, with the deletion of exon 42, a 54 bp exon (Sippola *et al.*, 1984*b*).

It is interesting that in one situation a deletion of 18 amino acids is associated with a lethal form of OI, and yet another similar-sized mutation is seen in the context of mild to moderate OI. In the former, it is possible that it is the combination of a null allele (de Wet *et al.*, 1983) together with one containing a deletion, so that all type I collagen molecules produced are abnormal, that results in lethality. In the latter, normal α2(I) chains were also synthesized and therefore a proportion of normal type I molecules (Byers *et al.*, 1983*a*) were present. It is also possible that the loss of exons coding for different parts of the protein molecule result in varying severity and may be tolerated depending on whether they occur in an important functional domain.

One consequence of an altered primary structure may be non-cleavage of the propeptides because of a modification of the proteinase recognition sites or incorrect folding of the *N*-terminus. Two such examples have been reported. Steinmann *et al.* (1980) found that cultured skin fibroblasts from a patient with EDS type VII produced two classes of type I procollagen molecules. One was normal but the other was resistant to cleavage by procollagen *N*-proteinase. Hence partially processed pro-α2(I) chains [pN-α2(I)], in which the *N*-terminal peptide remained, were present. In the report on a patient with clinical characteristics of both milder OI and EDS, the shortening of the pro-α2(I) chain by approx. 30 amino acids in α2(I)-CB4 apparently also results in the persistence of partially processed pN-α2(I) chains and a proportion of thermally unstable molecules on a background of normal ones (Sippola & Prockop, 1983; Sippola *et al.*, 1984*a*). Insertions have also been implicated in inherited collagen disease. Byers *et al.* (1981*b*) reported the first incidence of an insertion of 20 amino acids in the α2(I) chain in a Marfan patient. Skin fibroblasts from the patient synthesized normal and lengthened α2(I) chains. The source of the extra material was located *N*-terminal to the vertebrate collagenase site in α2(I) CNBr peptides CB3 or CB5. The consequence of this insertion seems to be a reduction in crosslinking between type I collagen molecules, although the fibrils themselves were not altered morphologically. The molecular and biochemical basis for disease in this patient still remains obscure. An attempt has been made to isolate the defective α2(I) allele in this patient but so far only an insertion of 38 bp within an intron, has been found (Henke *et al.*, 1985). It is uncertain whether the additional material lay within the normal or altered allele, since it could not have caused a 20 amino acid deletion. The relationship of this 38 bp insertion to the disease therefore remains unclear and may reflect small polymorphisms in fragment length which may occur normally (discussed later). There has also been a preliminary report of an insertion in the α2(I) CB4 peptide in a patient with EDS type VII (Steinmann *et al.*, 1985). Both normal and abnormal α2(I) chains were observed, but again the role of this change in the aetiology of the disease is unknown.

Structural mutations in other collagen genes

The involvement of other collagen genes in the pathology of inherited connective tissue disease has not been well studied. The observed absence or reduction in level of type III collagen in EDS IV

patients (Pope *et al.*, 1975; Aumailley *et al.*, 1980; Byers *et al.*, 1981*a*) point towards a defect in the α1(III) collagen gene. Pyeritz *et al.* (1984) report on an EDS IV patient with markedly reduced levels of type III collagen content in skin. Skin from the patient synthesized normal amounts of pro-α1(III) mRNA but secreted pro-α1(III) chains of normal and abnormal electrophoretic mobility which were unusually sensitive to proteolytic degradation. The defect in the pro-α1(III) chain is probably an insertion of about 20 amino acids in CNBr peptide CB5 in the region of residues 555–775 (Pyeritz *et al.*, 1984; Stolle *et al.*, 1985).

Molecular mechanisms and disease phenotype

From the above studies it appears that the severity of an inherited collagen disease may depend on a number of factors. Small or large changes in type I procollagen structure may result in the same or different clinical phenotype depending on the part of the molecule affected and the proportion of abnormal trimers formed. Lethal OI has now been seen in several contexts. Alterations in α1(I) appear to be more severe in consequences probably because a defect in one allele will affect 75% of the procollagen I molecules. The consequence of changes in α2(I) vary in severity possibly because, in the heterozygous state, 50% normal type I collagen is synthesized if the α2(I) chain is capable of forming a helix with α1(I). If only type I trimer is made instead, all molecules would contain normal α1(I) chains, but this class of collagen obviously cannot fully replace type I collagen in function.

It is uncertain exactly what level of reduction in procollagen I production is incompatible with life. Furthermore, the precise effect of a mixture of abnormal and normal collagen molecules on the function of the tissue (e.g. bone for OI) and in causing the disease is still unknown. It is clear however that alterations in type I collagen probably play a major role in causing some forms of OI, EDS and Marfan syndrome. Since type I collagen is the main collagen component of skin, bone and arteries, its involvement with disease possibly explains the overlapping characteristics seen in OI and EDS and suggests that alteration in type I collagen can affect the extracellular matrix as a whole. Considering the data available, it has been suggested however that the type of mutation and its position in the molecule may affect the resulting clinical phenotype (Byers & Bonadio, 1984; Bonadio *et al.*, 1985). According to this hypothesis alterations in the *C*-terminal 400 residues of the triple helix of type I collagen would result in type II OI, whereas *N*-terminal modification would produce type III OI. Nevertheless, it is difficult to make similar predictions for the different diseases without more extensive evidence.

It is unlikely that lesions in type I collagen will account for all the inherited collagen disease. Other contributory factors are probably involved. Changes in the other collagen types normally present in smaller amounts have not been well studied. But there is some evidence for an insertion in pro-α1(III) chains in one form of EDS (IV) (Stolle *et al.*, 1985) and inappropriate expression of type III collagen and elevated type V levels in OI bone (Pope *et al.*, 1980). Within the extracellular matrix, collagen interacts with proteoglycans and glycoproteins (Lindahl & Hook, 1978) some of which are tissue specific. For example, some noncollagenous matrix proteins are specific to bone. These include osteonectin (Terminè *et al.*, 1981), bone proteoglycan (Fisher *et al.*, 1983*a*) and bone sialoprotein (Fisher *et al.*, 1983*b*). Osteonectin binds to hydroxyapatite and collagen and promotes mineral deposition onto type I collagen (Termine *et al.*, 1981). It is interesting that in bovine osteogenesis imperfecta, in addition to the reduction in collagen content often seen in the human context, there was a marked depletion of osteonectin and proteoglycans. Sialoprotein content was also reduced but to a lesser extent (Termine *et al.*, 1984). Arneson *et al.* (1980) have described a family with features of EDS I or II in which abnormalities in fibronectin have found in affected members. Also, in a Marfan patient while collagen amounts were normal, the amino acid, desmosine and isodesmosine compositions of elastin were abnormal, and levels were reduced (Abraham *et al.*, 1982). Elastin is a major matrix component of arteries, the affected tissue in Marfan patients. Altered elastic fibrils have been observed in skin from patients with cutis laxa and EDS (Holbrook & Byers, 1982). It is likely that changes in other collagenous and noncollagenous matrix components pass undetected because the majority of studies are on cultured skin fibroblasts, whereas the major affected tissue may be bone or artery or cartilage. But, at present, the precise effect of changes in matrix composition on tissue function remains uncertain.

Technical limitations and possible solutions

The molecular basis of inherited disease for at least two patients was discovered because of the application of recombinant DNA technology to systems where the collagen biochemistry was well understood. The isolation of DNA probes for entire human collagen genes has been essential for detecting mutations affecting any part of the collagen molecule. It is only within this past year that the human α1(I), α2(I) and α1(II) collagen genes have been completely isolated (Table 2).

However, there are limitations in the use of conventional methods for the detection of mutations in collagen at the protein and gene levels.

At the protein level, the limit of detection of an alteration in pro-α chain size is 15–20 amino acids. At the DNA level the lower limit for detecting insertions or deletions in collagen genes against a background of total genomic DNA without cloning the defective gene is approx. 100 bp. These methods therefore cannot detect small deletions and insertions or point mutations, single base changes and amino acid substitutions. Of course it is possible to clone the collagen genes from each patient and seek out changes over the entire DNA sequence, but is impractical because of the size of these genes with so many exons. There is therefore a necessity to develop other methods for detecting and locating small changes in procollagen. Nuclease treatment of hybrids formed between cloned single-stranded DNA and total genomic DNA have been recently used to detect small changes in globin genes (Cheboune *et al.*, 1984; Myers *et al.*, 1985) and may have similar applications for collagen genes.

In most of the examples of structural changes in procollagen, including a single amino acid substitution, all or half the pro-α chains were over-modified and were less thermally stable. Byers and coworkers have suggested that this change in physical property in pro-α chains may be a symptom of small changes e.g. single amino acid substitutions (Byers & Bonadio, 1984; Bonadio *et al.*, 1985).

They postulate that amino acid substitutions which involve glycine in the repeating triplet will destabilize the helix and hence slow down the rate of helix formation from that point onwards. Because post-translational modification of chains ceases once a helix is formed, a non-enzymic reason for increased hydroxylation and glycosylation could be delayed helix formation. The region at which increased post-translational modification begins may pinpoint the appropriate location of a small insertion, deletion or substitution. But this method has three limitations. It does not distinguish which α chain in the type I helix carries the defect. Also, modifications which occur towards the *N*-terminus [beyond α1(I)CB8] are difficult to localize. Finally, defects which alter other properties of the type I molecule not associated with post-translational modification cannot be detected. These findings are interesting, however, because once such a region is identified it may be possible to determine the exact molecular defect using the appropriate DNA probes, much in the same way that the frameshift mutation was identified in the type III OI patient (Pihlajaniemi *et al.*, 1984). In several cases of OI, mixed populations of normal and partially or wholly over-modified type I collagen were found in the absence of any other detectable difference in molecular size (Bateman *et al.*, 1984; Sayed *et al.*, 1984; Wenstrup *et al.*, 1984; Bonadio *et al.*, 1985) and may be symptomatic of changes which cause delayed helix formation.

DNA polymorphisms, linkage studies and prenatal diagnosis

Another important aim in studying the molecular basis of inherited collagen disease is the detection of patients at risk. In order to offer prenatal diagnosis it may not be always necessary to determine the biochemical basis for an inherited disease. Variations exist in the DNA base sequence within and around genes. These give rise to DNA fragments of varying lengths when cleaved by restriction endonucleases. The existence of restriction fragment length polymorphisms (RFLP) has been useful for the analysis of inherited disease. Some RFLPs may be linked with defective genes where these are known, e.g. those in the globin gene cluster associated with the thalassaemias (Orkin *et al.*, 1982). Others may act as linkage markers for an inherited disease without being directly involved in the causation of the disorder, e.g. Duchenne muscular dystrophy and Huntington's disease (Murray *et al.*, 1982; Gusella *et al.*, 1984). The detection of a collagen gene RFLP which segregates with an inherited connective disease is not only useful for prenatal diagnosis, but may also indicate which gene is defective.

Because of the size and number of collagen genes this approach may prove to be increasingly important for the determination of the genetic basis of inherited connective tissue disease. There are a number of prerequisites for such studies. The first is the availability of DNA probes specific for a particular collagen gene or its immediate surrounding area. Second, the range of polymorphisms within the normal population and different ethnic groups has to be determined. Third, in order to establish linkage and determine if the trait is autosomal dominant, recessive or a new mutation, a complete study of the family pedigree covering several generations is essential. Of these, the third requirement is often the hardest to fulfil. In many cases families are small or incomplete or other members unavailable for study. The chromosomal locations and copy number of the collagen genes is also of importance for linkage studies. The collagen genes mapped so far are widely dispersed in the genome. The α1(I) gene is on chromosome 17 (Huerre *et al.*, 1982; Solomon *et al.*, 1984*b*) and the α2(I) gene is on chromosome 7 (Junien *et al.*, 1982; Solomon *et al.*, 1983). Three others are likewise dispersed: the α1(II), α1(III) and α1(IV) genes are

on chromosomes 12, 2 and 13 respectively (Solomon *et al.*, 1984*a*, 1985). The type I and II collagen genes are present in single copies (Dalgleish *et al.*, 1982; Tsipouras *et al.*, 1983; R. Dalgleish, unpublished work).

In a search for RLFPs in collagen genes, an *Eco*RI and a *Msp*I site polymorphism has been found in the α2(I) gene (Tsipouras *et al.*, 1983, 1984). The *Eco*RI polymorphism, which lies in the 5′ end of the gene, occurs at a high allelic frequency and has been found to segregate with mild autosomal dominant types of OI (types I and IV) in two families. This suggests that the clinical phenotype in this family may be linked to a defect in the α2(I) gene. However in three other families with autosomal dominant OI there was no linkage with the disease (Tsipouras *et al.*, 1984) and illustrates the heterogeneity of these disorders. By similar analyses of *Hin*dIII and *Pvu*II site polymorphisms in the type II collagen gene, the involvement of this locus in three families with autosomal dominant OI type I has been excluded (Sykes *et al.*, 1985*a*). Using the *Hin*dIII polymorphism, type II collagen has also been excluded as the defective locus in one family with EDS type II (Wordsworth *et al.*, 1985).

The polymorphic sites reported for the type II collagen gene and its 3′ flanking region should also prove useful for the genetic analysis of the chondrodysplasias. The *Hin*dIII site polymorphism (Driesel *et al.*, 1982; Sykes, 1983) occurs with high frequency and has a neutral ethnic distribution. There is also a variation in an *Eco*RI fragment length which was first observed in several families with a history of perinatal lethal OI (Pope *et al.*, 1984) and now appears to occur in West Indians and Asian Indians at higher frequencies than in Caucasians (Sykes & Ogilvie, 1984; Sykes *et al.*, 1985*b*; F. M. Pope, unpublished work). The *Eco*RI fragment length variation lies in the region immediately 3′ to the type II collagen gene (Pope *et al.*, 1984; Cheah *et al.*, 1985). This is the first length polymorphism within a single fragment close to a collagen gene to be reported which also appears to have a frequency which is ethnic-group-associated. Two variants of this gene have been isolated (K. S. E. Cheah, unpublished work). From this and other studies the variation in length of the *Eco*RI fragment appears to range from deletions of 30 bp to 300 bp, to an insertion of 100 bp when compared with the commonest fragment size (Sykes *et al.*, 1985*b*; F. M. Pope, unpublished work).

Instead of RFLPs, diminished α1(I) mRNA levels have been used as a marker for dominantly inherited OI. Reduced levels of α1(I) mRNA segregated with affected individuals within a three-generation family (D. Rowe, personal communication). This is unlikely to be a popular means for prenatal diagnosis because of the length of time and expense required to establish primary skin fibroblasts from many individuals. Measurements of mRNA levels are also subject to experimental artefacts associated with possible increased degradation and instability of the mRNA during extraction.

At present techniques for prenatal diagnosis of OI and achondroplasia are by ultrasound, radiography and the analysis of the procollagen synthesis of amniotic fluid cells (Shapiro *et al.*, 1982; Aylsworth *et al.*, 1984). These methods are disadvantaged by the lateness at which the disease may be detected (19–25 weeks of gestation). Also only type I and IV collagen synthesis may be monitored in the amnion (Crouch & Bornstein, 1978, 1979). Of the inherited diseases those involving the type II collagen gene may be the most difficult to analyse for several reasons. Skin fibroblasts do not synthesize type II collagen and cartilage biopsies are both difficult to obtain and to maintain phenotypically and biochemically in long-term culture. Biosynthetic studies are therefore severely handicapped. The use of RFLP linkage studies may therefore be of particular importance for the diagnosis and detection of a type II collagen gene defect with subsequent gene cloning as a means of determining the molecular basis of the disease.

Prenatal diagnosis is not possible for parents of lethal OI patients where the biochemical basis is that of a new mutation, as has been found in several cases (de Wet *et al.*, 1983; Steinmann *et al.*, 1984*b*; Bonadio *et al.*, 1985).

Future prospects

It is certain that linkage studies of polymorphic restriction enzyme sites will be used increasingly as a tool both to provide a pointer to the molecular defect and for prenatal diagnosis. For disorders involving types I and III collagen the future is promising for detecting gross structural alterations in the genes and the protein. Accompanying such discoveries will be a greater understanding of the functional importance of the different collagen types and their intramolecular domains. It will be harder to detect fine changes, e.g. point mutations or amino acid substitutions, but it is likely that new techniques for detecting these will be devised. For example, heterokaryons formed by cell fusion have been used in an attempt to detect mutations in type I collagen genes (Byers & Bonadio, 1984). Although the results are as yet inconclusive this may prove to be a useful strategy to seek out recessive gene mutations. The future for discovering mutations in other collagen genes is less clear and certainly will prove to be an even more demanding

task. Likewise, linking the clinical and biochemical heterogeneity of the inherited collagen diseases will take time. At any rate there is enough work to keep a great many researchers busy for a long time to come.

I am extremely grateful to Peter Byers, Raymond Dalgleish, Mike Grant, Elizabeth Henke, Sheldon Pinnell, Michael Pope, Darwin Prockop, Francesco Ramirez, David Rowe, Eero Vuorio, Ellen Solomon, Beat Steinmann and Bryan Sykes for communicating papers and data prior to publication to enable this review to be as up to date as possible. I also thank Mike Grant, Ellen Solomon, Neil Stoker and Bryan Sykes for their critical comments and helpful discussion. The patience and assistance of Irene Wu and Helena Tai in the preparation of this manuscript is also gratefully acknowledged.

References

Abraham, P. A., Perejda, A. J., Carnes, W. H. & Uitto, J. (1982) *J. Clin. Invest.* **70**, 1245–1252

Adamson, E. D. (1982) in *Collagen in Health and Disease* (Weiss, J. B. & Jayson, M. I. V., eds.), pp. 218–243, Churchill Livingstone, London

Arneson, M. A., Hammerschmidt, D. E., Furcht, L. T. & King, R. A. (1980) *J. Am. Med. Assoc.* **244**, 144–147

Alitalo, K. A., Bornstein, P., Vaheri, A. & Sage, H. (1983) *J. Biol. Chem.* **258**, 2653–2661

Aumailley, M., Krieg, T., Dessaus, W., Muller, P. K., Timpl, R. & Bricand, H. (1980) *Arch. Dermatol. Res.* **269**, 169–177

Aylsworth, A. S., Seeds, J. W., Guilford, W. B., Burns, C. B. & Washburn, D. D. (1984) *Am. J. Med. Genet.* **19**, 707–714

Bachinger, H. P., Fessler, L. I., Timpl, R. & Fessler, J. H. (1981) *J. Biol. Chem.* **256**, 13193–13199

Barsh, G. S. & Byers, P. H. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 5142–5146

Barsh, G. S., David, K. E. & Byers, P. H. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3838–3842

Barsh, G. S., Roush, C. L. L. & Gelinas, R. E. (1984) *J. Biol. Chem.* **249**, 14906–14913

Barsh, G. S., Roush, C. L., Bonadio, J., Byers, P. H. & Gelinas, R. E. (1985) *Proc. Natl. Acad. Sci. U.S.A.*, in the press

Bateman, J. F., Mascara, T., Chan, D. & Cole, W. G. (1984) *Biochem. J.* **215**, 103–116

Bentz, H., Morris, N. P., Murray, L. W., Sakai, L. Y., Hollister, D. W. & Burgeson, R. E. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 8168–8172

Boedtker, H., Fuller, F. & Tate, V. (1983) *Int. Rev. Connect. Tissue Res.* **10**, 1–63

Bonadio, J., Holbrook, K. A., Gelinas, R. E., Jacob, J. & Byers, P. H. (1985) *J. Biol. Chem.*, **260**, 1734–1742

Bornstein, P. & Sage, H. (1980) *Annu. Rev. Biochem.* **49**, 957–1003

Bruckner, P., Eikenberry, E. F. & Prockop, D. J. (1981) *Eur. J. Biochem.* **118**, 607–613

Burgeson, R. E. (1982) *J. Invest. Dermatol.* **79** *Suppl.* **1**, 25S–30S

Burgeson, R. E. & Hollister, D. W. (1979) *Biochem. Biophys. Res. Commun.* **87**, 1124–1131

Burgeson, R. E., Hebda, P. A., Morris, N. P. & Hollister, D. W. (1982) *J. Biol. Chem.* **257**, 7852–7856

Byers, P. H. & Bonadio, J. F. (1984) in *Butterworths International Medical Reviews, Paediatrics 5: Metabolic and Genetic Disease* (Schier, C. R. & Lloyd, J., eds.), Butterworths, London, in the press

Byers, P. H., Holbrook, K. A., McGillivray, B., MacLeod, P. M. & Lowery, R. B. (1979) *Hum. Genet.* **47**, 141–150

Byers, P. H., Barsh, G. S. & Holbrook, K. A. (1981*a*) *Collagen Relat. Res.* **5**, 475–489

Byers, P. H., Siegel, R. C., Peterson, K. E., Rowe, D. W., Holbrook, K. A., Smith, L. T., Chang, Y. H. & Fu, J. C. C. (1981*b*) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7745–7749

Byers, P. H., Shapiro, J. R., Rowe, D. W., David, K. E. & Holbrook, K. A. (1983*a*) *J. Clin. Invest.* **71**, 689–697

Byers, P. H., Bonadio, J. F., Steinmann, B., Barsh, G. S., Holbrook, K. S., Greenberg, C., Rowe, D. W. & Gelinas, R. (1983*b*) *Am. J. Hum. Genet.* **35**, 39A (abstr.)

Cheah, K. S. E. & Grant, M. E. (1982) in *Collagen in Health and Disease* (Weiss, J. B. & Jayson, M. I. V., eds.), pp. 73–100, Churchill Livingstone, London

Cheah, K. S. E., Grant, M. E. & Jackson, D. J. (1979) *Biochem. Biophys. Res. Commun.* **91**, 1025–1031

Cheah, K. S. E., Stoker, N. G., Griffin, J. G., Grosveld, F. G. & Solomon, E. (1985) *Proc. Natl. Acad. Sci. U.S.A.*, **82**, 2555–2559

Chebloune, Y., Trabuchet, G., Poncet, G., Cohen-Solal, M., Fauri, C., Verdier, G. & Nigon, V. M. (1984) *Eur. J. Biochem.* **142**, 473–480

Chu, M. L., Myers, J. C., Bernard, M. P., Ding, J. F. & Ramirez, F. (1982) *Nucleic Acids Res.* **10**, 5925–5934

Chu, M. L., Williams, C. J., Pepe, G., Hirsch, J. L., Prockop, D. J. & Ramirez, F. (1983) *Nature* (*London*) **304**, 78–80

Chu, M. L., de Wet, W., Bernard, M., Ding, J. F., Morabito, M., Myers, J., Williams, C. & Ramirez, F. (1984) *Nature* (*London*) **310**, 337–340

Chu, M. L., Gargiulo, V., Williams, C. & Ramirez, F. (1985*a*) *J. Biol. Chem.* **260**, 691–694

Chu, M. L., Weil, D., de Wet, W., Bernard, M., Sippola, M. & Ramirez, F. (1985*b*) *J. Biol. Chem.*, in the press

Crouch, E. & Bornstein, P. (1978) *Biochemistry* **17**, 5499–5509

Crouch, E. & Bornstein, P. (1979) *J. Biol. Chem.* **254**, 4197–4204

Dalgleish, R., Trapnell, B. C., Crystal, R. G. & Tolstoshev, P. (1982) *J. Biol. Chem.* **257**, 13816–13822

Deak, S. B., Chu, M. L., Nicholls, A. C., Pope, F. M., Rowe, D. & Prockop, D. J. (1981) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **41**, 852

Deak, S. B., Nicholls, A., Pope, F. M. & Prockop, D. J. (1983) *J. Biol. Chem.* **258**, 15192–15197

de Wet, W. J., Pihlajaniemi, T., Myers, J., Kelly, T. E. & Prockop, D. J. (1983) *J. Biol. Chem.* **258**, 7721–7728

Dickson, L. A., Pihlajaniemi, T., Deak, S., Pope, F. M., Nicholls, A. C., Prockop, D. J. & Myers, J. C. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4524–4528

dos Santos, C. L. S., Villa, L. L., Sonohara, S. & Brentani, R. R. (1984) *Nucleic Acids Res.* **12**, 2035–2046

Driesel, A. J., Schumacher, A. M. & Flavell, R. A. (1982) *Hum. Genet.* **62**, 175–176
Duvic, M., Pinnell, P. & Kaufman, R. (1984) *J. Invest. Dermatol.* **32**, 440 (abstr.)
Elima, K., Makela, J. K., Vuorio, T., Kauppinen S., Knowles, J. & Vuorio, E. (1985) *Biochem. J.*, **229**, 183–188
Fessler, J. H. & Fessler, L. I. (1978) *Annu. Rev. Biochem.* **47**, 129–162
Fisher, L. W., Termine, J. D., Dijter, S. W., Whitson, S. W., Yanagishita, M., Kimura, H. J., Hascall, V. C., Hassell, J. R. & Nilsson, B. (1983*a*) *J. Biol. Chem.* **258**, 6588–6594
Fisher, L. W., Whitson, S. W., Avioli, L. V. & Termine, J. D. (1983*b*) *J. Biol. Chem.* **258**, 12723–12727
Furthmayr, H., Weiderman, H., Timpl, R., Odermatt, E. & Engel, J. (1983) *Biochem. J.* **211**, 303–311
Genovese, C., Rowe, D. & Kream, B. (1984) *Biochemistry* **23**, 6210–6216
Gibson, G. J., Schor, S. L. & Grant, M. E. (1982) *J. Cell Biol.* **93**, 767–774
Gibson, G. J., Kielty, C. M., Garner, C., Schor, S. L. & Grant, M. E. (1983) *Biochem. J.* **211**, 417–426
Grant, M. E. & Jackson, D. S. (1976) *Essays Biochem.* **12**, 77
Gusella, J. F., Tanzi, R. E., Anderson, M. A., Hobbs, W., Gibbons, K., Raschtchian, R., Gilliam, T. C., Wallace, M. R., Wexler, N. S. & Conneally, P. M. (1984) *Science* **225**, 1320–1326
Haralson, M. A., Mitchell, W. M., Rhodes, R. K., Kresina, T. F., Gay, R. & Miller, E. J. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5206–5210
Haralson, M. A., Mitchell, W. M., Rhodes, R. K. & Miller, E. J. (1984) *Arch. Biochem. Biophys.* **229**, 509–518
Harbers, K., Kuehn, M., Delius, H. & Jaenisch, R. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1504–1508
Heathcote, J. G. & Grant, M. E. (1980) in *The Enzymology of Post-translational Modifications of Proteins* (Freeman, R. B. & Hawkins, H. C., eds.), pp. 457–506, Academic Press, London
Henke, E., Leader, M., Tajima, S., Pinnell, S. & Kaufman, R. (1985) *J. Cell. Biochem.* **27**, 169–174
Holbrook, K. A. & Byers, P. H. (1981) *Lab. Invest.* **44**, 342–350
Holbrook, K. A. & Byers, P. H. (1982) *J. Invest. Dermatol.* **79**, 7S–16S
Hollister, D. W. (1981) *Collagen Relat. Res.* **1**, 228–234
Hollister, D. W., Byers, P. H. & Holbrook, K. A. (1983) *Adv. Hum. Genet.* **13**, 1–87
Huerre, C., Junien, C., Weil, D., Chu, M.-L., Morabito, M., Van Cong, N., Myers, J. C., Foubert, C., Gross, M.-S., Prockop, D. J., Boue, A., Kaplan, J.-C., de la Chapelle, A. & Ramirez, F. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6627–6630
Junien, C., Weil, D., Myers, J. C., Van Cong, N., Chu, M.-L., Foubert, C., Gross, M.-S., Prockop, D. J., Kaplan, J.-C. & Ramirez, F. (1982) *Am. J. Hum. Genet.* **34**, 381–387
Kielty, C. M., Kwan, A. P. L., Holmes, D. F., Schor, S. L. & Grant, M. E. (1985) *Biochem. J.* **227**, 545–554
Kivirikko, K. I. & Myllyla, R. (1982) in *Collagen in Health and Disease* (Weiss, J. B. & Jayson, M. I. V., eds.), pp. 101–120, Churchill Livingstone, Edinburgh
Kohno, K., Martin, G. R. & Yamada, Y. (1984) *J. Biol. Chem.* **259**, 13668–13673
Kramer, J. M., Cox, G. N. & Hirsch, D. (1982) *Cell* **30**, 599–606
Krane, S. M. (1984) in *Extracellular Matrix Biochemistry* (Piez, K. A. & Reddi, A. H., eds.), pp. 413–463, Elsevier, New York
Kurkinen, M., Barlow, D. P., Hofman, D. M., Williams, J. G. & Hogan, B. L. M. (1983) *Nucleic Acids Res.* **11**, 6199–6209
Lehrach, H., Frischauf, A. M., Hanahan, D., Wozney, J., Fuller, F., Crkvenjakov, R., Boedtker, H. & Doty, P. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 5417–5421
Lehrach, H., Frischauf, A. M., Hanahan, D., Wozny, J., Fuller, F. & Boedtker, H. (1979) *Biochemistry* **18**, 3146–3152
Liau, G., Yamada, Y. & de Cronbrugghe, B. (1985) *J. Biol. Chem.* **260**, 531–536
Lindahl, U. & Hook, M. (1978) *Annu. Rev. Biochem.* **47**, 385–417
Lohler, J., Timpl, R. & Jaenisch, R. (1984) *Cell* **38**, 597–607
Loidl, H. R., Brinker, J. M., May, M., Pihlajaniemi, T., Morrow, S., Rosenbloom, J. & Myers, J. C. (1984) *Nucleic Acids Res.* **12**, 9383–9394
Lukens, L. N., Frischauf, A. M., Pawlowski, P. J., Brierley, G. T. & Lehrach, H. (1983) *Nucleic Acids Res.* **11**, 6021–6039
McKusick, V. A. (1972) *Heritable Disorders of Connective Tissue*, 4th edn., C. V. Mosby, St. Louis
Mayne, R., Wiedermann, H., Irwin, M. H., Sanderson, R. D., Fitch, Y. M., Linsenmayer, T. T. & Kuhn, K. (1984) *J. Cell Biol.* **98**, 1637–1644
Miller, E. J. & Gay, S. (1982) *Methods Enzymol.* **82**, 3–32
Monson, J. M. & McCarthy, B. J. (1981) *DNA* **1**, 59–69
Monson, J. M., Natzle, J., Friedman, J. & McCarthy, B. J. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 1761–1765
Murray, J. M., Davies, K. E., Harper, P. S., Meredith, L., Mueller, C. R. & Williamson, R. (1982) *Nature (London)* **300**, 69–71
Myers, J. C., Chu, M., Faro, S. H., Clark, W. J., Prockop, D. J. & Ramirez, F. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 3515–3520
Myers, J. C., Dickson, L. A., de Wet, W., Bernard, M. P., Chu, M. L., DiLiberto, M., Pepe, G., Sangiorgi, F. O. & Ramirez, F. (1983) *J. Biol. Chem.* **258**, 10128–10135
Myers, R. M., Lumelsky, N., Lerman, L. S. & Maniatis, T. (1985) *Nature (London)* **313**, 495–498
Nicholls, A. C. & Pope, F. M. (1984) *Br. Med. J.* **288**, 112–113
Nicholls, A. C., Pope, F. M. & Schloen, H. (1979) *Lancet* **i**, 1193
Ninomiya, Y., Showalter, A. M., van der Rest, M., Seidah, N. G., Chretien, M. & Olsen, B. R. (1984) *Biochemistry* **23**, 617–624
Ninomyia, Y. & Olsen, B. R. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3014–3018
Niyibizi, C., Fietzek, P. P. & van der Rest, M. (1984) *J. Biol. Chem.* **259**, 14170–14174

Odermatt, E., Risteli, J., van Delden, V. & Timpl, R. (1983) *Biochem. J.* **211**, 295–302
Orkin, S. H., Kazazian, H. H., Antonarakis, S. E., Goff, S. C., Boehm, C. D., Sexton, J. P., Waber, P. G. & Giardina, P. J. V. (1982) *Nature (London)* **296**, 627–631
Peltonen, L., Palotie, A. & Prockop, D. J. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6179–6183
Penttinen, R. P., Lichtenstein, J. R., Martin, G. R. & McKusick, V. A. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 586–589
Pihlajaniemi, T., Dickson, L. A., Pope, F. M., Korhonen, V. R., Nicholls, A., Prockop, D. J. & Myers, J. C. (1984) *J. Biol. Chem.* **259**, 12941–12944
Pinnell, S. R. & Murad, S. (1983) in *The Metabolic Basis of Inherited Disease*, 5th edn. (Stanbury, J. B., Wyngaarden, J. B., Fredrickson, D. S., Goldstein, J. C. & Brown, M. S., eds.), pp. 1425–1429, McGraw-Hill, New York
Pinnell, S. R., Krane, S. M., Kenzora, J. E. & Glimcher, M. J. (1972) *N. Engl. J. Med.* **866**, 12013–12020
Pope, F. M., Martin, G. R., Lichtenstein, J. R., Penttinen, R. P., Gerson, G., Rowe, D. W. & McKusick, V. A. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 1314–1316
Pope, F. M., Nicholls, A. C., Eggleton, C., Narcissi, P., Hey, E. N. & Parkin, J. M. (1980) *J. Clin. Pathol.* **33**, 534–538
Pope, F. M., Cheah, K. S. A., Nicholls, A. C., Price, A. B. & Grosveld, F. G. (1984) *Br. Med. J.* **288**, 431–434
Prockop, D. J. & Kivirikko, K. I. (1984) *N. Engl. J. Med.* **311**, 376–385
Prockop, D. J., Kivirikko, K. T., Tuderman, L. & Guzman, N. A. (1979) *N. Engl. J. Med.* **301**, 12–31, 77-85
Pyeritz, R. E., Stolle, C. A., Parfrey, N. A. & Myers, J. C. (1984) *Am. J. Med. Genet.* **19**, 607–622
Reese, C. A. & Mayne, R. (1981) *Biochemistry* **20**, 5443–5448
Sage, H., Pritzl, P. & Bornstein, P. (1981) *Biochemistry* **20**, 3778–3784
Sage, H., Trueb, B. & Bornstein (1983) *J. Biol. Chem.* **258**, 13391–13401
Sayed, A. K., Hollbrook, K. A., Philips, J. A., III & Byers, P. H. (1984) *Lab. Invest.* **50**, P11 (abstr.)
Sandell, L. J., Yamada, Y., Dorfman, A. & Upholt, W. B. (1983) *J. Biol. Chem.* **258**, 11617–11621
Sandell, L. J., Prentice, H. L., Kravis, D. & Upholt, W. B. (1984) *J. Biol. Chem.* **259**, 7826–7834
Schafer, M. P., Boyd, C. P., Tolstoshev, P. & Crystal, R. G. (1980) *Nucleic Acids Res.* **8**, 2241–2253
Schmid, T. M. & Conrad, H. E. (1982) *J. Biol. Chem.* **257**, 12444–12450
Schmidt, A., Yamada, Y. & de Crombrugghe, B. (1984) *J. Biol. Chem.* **259**, 7411–7415
Schnieke, A., Harbers, K. & Jaenisch, R. (1983) *Nature (London)* **304**, 315–320
Shapiro, J. E., Phillips, J. A., III, Byers, P. H., Saunders, R., Holbrook, K. A., Levin, L. S., Dorst, T., Barsh, G. S., Peterson, M. S. & Goldstein, P. (1982) *J. Paediatr.* **100**, 127–133
Shapiro, J. R. & Rowe, D. W. (1983) *Ann. Int. Med.* **99**, 700–704
Shimokomaki, M., Duance, V. C. & Bailey, A. J. (1980) *FEBS Lett.* **121**, 51–54
Sillence, D. O. (1981) *Clin. Orthop. Relat. Res.* **159**, 11–25
Sippolla, M. & Prockop, D. J. (1983) *J. Cell Biol.* **97**, 229a (abstr.)
Sippola, M., Kaffe, S. & Prockop, D. J. (1984*a*) *J. Biol. Chem.* **259**, 14094–14100
Sippola, M., Tromp, G., Prockop, D. J., de Wet, W. & Ramirez, F. (1984*b*) *Am. J. Hum. Genet.* **36**, 153S (abstr.)
Sobel, M. E., Yamamoto, T., Adams, S. L., DiLaurio, R., Avvedimento, V., de Crombrugghe, B. & Pastan, I. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 5846–5850
Solomon, E. & Cheah, K. S. E. (1981) *Nature (London)* **291**, 450–451
Solomon, E., Hiorns, L., Dalgleish, R., Tolstoshev, P., Crystal, R. & Sykes, B. (1983) *Cytogenet. Cell. Genet.* **35**, 64–66
Solomon, E., Hiorns, L. R., Cheah, K. S. E., Parker, M., Weiss, E. & Flavell, R. A. (1984*a*) *Cytogenet. Cell. Genet.* **37**, 588–589
Solomon, E., Hiorns, L., Sheer, D. & Rowe, D. (1984*b*) *Ann. Hum. Genet.* **48**, 39–42
Solomon, E., Hiorns, L. R., Spurr, N., Kurkinen, M., Hogan, B. L. M. & Dalgleish, R. (1985) *Proc. Natl. Acad. Sci. U.S.A.*, in the press
Steinmann, B., Tuderman, L., Peltonen, L., Martin, G. R., McKusick, V. A. & Prockop, D. J. (1980) *J. Biol. Chem.* **255**, 8887–8893
Steinmann, B., Rao, V. H. & Gitzelmann, R. (1985) *Ann. N. Y. Acad. Sci.*, in the press
Steinmann, B., Rao, V. H., Vogel, A., Bruckner, P., Gitzelmann, R. & Byers, P. H. (1984) *J. Biol. Chem.* **259**, 11129–11138
Stolle, C. A., Pyeritz, R. E., Myers, J. C. & Prockop, D. J. (1985) *J. Biol. Chem.*, **260**, 1937–1944
Strom, C. M. & Upholt, W. B. (1984) *Nucleic Acids Res.* **12**, 1025–2038
Sulh, H. M. B., Steinmann, B., Rao., V. H., Dudin, G., Zeid, J. A., Slein, M. & Der Kaloustian, V. M. (1984) *Clin. Genet.* **25**, 278–287
Sykes, B. (1983) *Dis. Markers* **1**, 141–146
Sykes, B. & Ogilvie, D. (1984) *Br. Med. J.* **288**, 1980
Sykes, B., Francis, M. J. O. & Smith, R. (1977) *N. Engl. J. Med.* **296**, 1200–1203
Sykes, B., Ogilvie, D. J. & Wordsworth, B. P. (1985*a*) *Hum. Genet.*, in the press
Sykes, B., Smith, R., Vipond, S., Paterson, C., Cheah, K. & Solomon, E. (1985*b*) *J. Med. Genet.*, in the press
Tajima, S., Ting, J. P. Y., Pinnell, S. R. & Kaufman, R. E. (1984) *J. Invest. Dermatol.* **82**, 265–269
Termine, J. D., Kleinman, H. K., Whitson, S. W., Conn, K. M., McGravy, M. L. & Martin, G. R. (1981) *Cell* **26**, 99–105
Termine, J. D., Robey, P. G., Fisher, L. W., Shimokawa, A., Drum, M. A., Conn, K. M., Hawkins, G. R., Cruz, J. B. & Thomson, K. G. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2213–2217
Tsipouras, P., Myers, J. C., Ramirez, F. & Prockop, D. J. (1983) *J. Clin. Invest.* **72**, 1262–1267
Tsipouras, P., Borresoen, A.-L., Dickson, L. A., Berg, K., Prockop, D. J. & Ramirez, F. (1984) *Am. J. Hum. Genet.*, **36**, 1172–1179

Uitto, J., Tan, E. M. L., Muller, P., Krieg, T. & Prockop, D. J. (1983) *Clin. Res.* **31**, 468A (abstr.)
van der Rest, M., Mayne, R., Ninomiya, Y., Seidah, N. G., Chretien, M. & Olsen, B. R. (1985) *J. Biol. Chem.* **260**, 220–225
Vogeli, G., Ohkubo, H., Avvedimento, V. E., Sullivan, M., Yamada, Y., Mudryj, M., Pastan, I. & de Crombrugghe, B. (1980) *Cold Spring Harbour Symp. Quant. Biol.* **45**, 777–783
von der Mark, K., van Menxel, M. & Wiedermann, H. (1982) *Eur. J. Biochem.* **124**, 57–62
Wang, S. Y. & Gudas, L. J. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 5880–5884
Wenstrup, R. J., Tsipouras, P. & Byers, P. H. (1984) *Am. J. Hum. Genet.* **36**, 23S (abstr.)
Werb, Z. (1982) in *Collagen in Health and Disease* (Weiss, J. B. & Jayson, M. I. V., eds.), pp. 121–134, Churchill Livingstone, London
Williams, C. J. & Prockop, D. J. (1983) *J. Biol. Chem.* **258**, 5915–5921
Wordsworth, P., Ogilvie, D., Smith, R. & Sykes, B. (1985) *Ann. Rheum. Dis.*, in the press
Wozney, J., Hanahan, D., Tate, V., Boedtker, H. & Doty, P. (1981) *Nature (London)* **294**, 129–135
Yamada, Y., Avvedimento, V. E., Mudryj, M., Ohkkubo, H., Vogeli, G., Irani, M., Pastan, I. & de Crombrugghe, B. (1980) *Cell* **22**, 887–892
Yamada, Y., Mudryj, M., Sullivan, M. & de Crombrugghe, B. (1983) *J. Biol. Chem.* **258**, 2758–2761
Yamada, Y., Liau, G., Mudryj, M., Obici, S. & de Crombrugghe, B. (1984) *Nature (London)* **310**, 333–337
Yamamoto, T., Sobel, M. E., Adams, S. L., Avvedimento, V. E., Di Laurio, R., Pastan, I., de Crombrugghe, B., Showalter, A., Pesciotta, D., Fietzek, P. & Olsen, B. (1980) *J. Biol. Chem.* **255**, 2612–2615
Young, M. F., Vogeli, G., Nunez, A. M., Fernandez, M. P., Sullivan, M. & Sobel, M. E. (1984) *Nucleic Acids Res.* **12**, 4207–4227