A type I collagen reporter gene construct for protein engineering studies

Functional equivalence of transfected reporter COL1A1 and endogenous gene products during biosynthesis and in vitro extracellular matrix accumulation

Shireen R. LAMANDÉ and John F. BATEMAN*  
Department of Paediatrics, University of Melbourne, Royal Children's Hospital, Parkville, Victoria 3052, Australia

A type I collagen reporter gene construct, designed to facilitate detailed analysis of the consequences of introduced structural and regulatory mutations on collagen biosynthesis and participation in the extracellular matrix, was produced by site-directed mutagenesis of the mouse COL1A1 gene. The reporter construct, pWTCl-Ile*82, carried a single base change which converted the codon for amino acid 822 of the triple helix from methionine to isoleucine. This change allowed the reporter protein, [Ile*82]α1(I), to be distinguished from the wild-type α1(I), and quantified, by its altered CNBr cleavage pattern. In mouse Mov13 cells, which synthesize no endogenous proα1(I), reporter chains associated with endogenous proα2(I), formed pepsin-stable triple helices and were secreted efficiently from the cell.

The thermal stability of wild-type molecules and molecules containing the reporter [Ile*82]α1(I) chains was identical. The biosynthetic characteristics of wild-type and reporter chains were directly compared in stably transfected 3T6 cells. These cells did not make a distinction between reporter and endogenous α1(I) chains, which were secreted from the cells at the same rate and were processed and deposited into the 3T6 cell in vitro accumulated extracellular matrix with equal efficiency. These data demonstrate that the helical sequence alteration in the reporter protein is functionally neutral and that the reporter construct, pWTCl-Ile*82, is a suitable vector for the analysis of the biochemical effects of site-directed mutations in the putative COL1A1 functional domains.

INTRODUCTION

The molecular characterization of naturally occurring mutations in the fibrillar collagen genes has provided much information on the biochemistry of collagen structure and biosynthesis. The critical importance of the triple-helical domain for tissue integrity and function has been demonstrated in inherited connective tissue diseases such as osteogenesis imperfecta [1], Ehlers-Danlos syndrome [2] and the chondrodysplasias [3–5], where disruption of the strict repetitive helical sequence motif by glycine substitutions or larger gene mutations leads to gross connective tissue abnormalities. The bias inherent in these studies, i.e. that the mutation has resulted in a clinically significant phenotype, has meant that much less is known about the effect of more subtle helix sequence alterations or propeptide mutations. A better understanding of the role of these other collagen sequence domains in molecular assembly, secretion, propeptide processing, fibrillogenesis and modulation of macromolecular interactions is essential if we are to fully appreciate the complex biochemistry of extracellular matrix formation and regulation in normal development as well as in disease processes.

The development of cell culture and transgenic mouse models of collagen diseases [6-9] has made it possible to approach these questions by introducing specifically engineered alterations into the putative functional domains and then assessing the biochemical consequences. Crucial to the measurement of the biosynthetic parameters, and thus to an understanding of how these processes and interactions are disturbed by mutations, is the ability to experimentally discriminate between normal and mutant chains at all stages of collagen biosynthesis and matrix formation.

The mutations themselves cannot be utilized as markers unless they alter a readily assessable characteristic, such as the electrophoretic migration, of the structural helical domain of the molecule. This problem is most graphically demonstrated in the case of C- or N-propeptide mutations where, once the propeptides have been cleaved during the normal processes leading to matrix deposition, α-chains which once carried an altered propeptide domain will be indistinguishable from wild-type chains.

In order to overcome these experimental limitations and provide a versatile model system for the study of type I collagen mutations, a collagen protein reporter construct was produced by site-directed mutagenesis of the mouse α1(I) gene. The criteria for production of this marker were that the alteration must be silent at the protein level and readily detected experimentally. Amino acid 822 of the triple helix was changed from a methionine to an isoleucine so that the protein produced from the construct could be detected and quantified in processed collagen molecules by its altered CNBr cleavage pattern. In transfected cells this conservative substitution had no detectable effect on helix conformation, thermal stability, secretion or deposition into an in vitro accumulated extracellular matrix, but allowed the discrimination and detailed comparison of the introduced gene and endogenous gene products.

EXPERIMENTAL

Site-directed mutagenesis of the COL1A1 gene

The reporter construct was prepared using derivatives of the mouse COL1A1 genomic clone 10D [10]. Figure 1 summarizes the mutagenesis protocol. DNA between a Thal site and the Apor site was replaced with synthetic oligonucleotides whose sequence differed from the wild-type by a single G to C base change (Figure 1a). The substitution altered the codon for amino acid 822 from Met to Ile and deleted an Ncol site. Mutagenesis

Abbreviations used: CB, peptide produced by CNBr treatment; DMEM, Dulbecco’s modified Eagle’s medium.

* To whom correspondence should be addressed.
was initially performed on DNA from a 5.5 kb EcoRI subclone of the gene in which an intronic KpnI site had been destroyed so that the remaining KpnI site would be unique [11]. The 1676 bp KpnI–XbaI fragment of this clone (Figure 1b) was purified and digested with Thal and Apal, and the 482 bp KpnI–Thal and 1174 bp Apal–XbaI fragments were purified. Equinuclor amounts of the synthetic oligonucleotides dissolved in 10 mM Tris/HCl, pH 7.5, 100 mM NaCl and 1 mM EDTA were mixed, heated to 90 °C, then slowly cooled to room temperature to allow them to anneal. The annealed oligonucleotides were ligated firstly to the KpnI–Thal fragment. The resulting larger KpnI–Apal fragment was then purified and ligated to the Apal–XbaI fragment, and the reconstituted KpnI–XbaI fragment was inserted back into the 5.5 kb EcoRI subclone. The molecular alteration in the plasmid was confirmed by restriction mapping with NcoI and sequencing (Sequenase; United States Biochemical Corp.) of a 651 bp KpnI–BamHI fragment subcloned in M13mp18. The 2942 bp Nhel–XhoI fragment from a selected clone was then used to replace the normal fragment of pWTCl, the full length genomic clone, producing the final reassigned gene construct pWTCl-Ile822. Plasmids were again screened by NcoI mapping and checked for possible rearrangements by digestion with HindIII.

Transfection

The reporter construct, pWTCl-Ile822, was co-transfected into Mov3 and 3T6 cells with pSV2neo [12] at a molar ratio of 7:1. pSV2neo contains the neomycin phosphotransferase gene which confers resistance to the drug G418. The plasmids were linearized prior to transfection; pWTCl-Ile822 by SalI digestion and pSV2neo by digestion with PvuI. Transfection was achieved using lipofectin reagent (Gibco-BRL) according to the manufacturer’s protocol. Stably transfected Mov3 cells were selected in growth medium containing 200 μg/ml G418 (Gibco-BRL), and 3T6 cells in medium with 300 μg/ml G418. Individual G418-resistant colonies were isolated and expanded into cell lines. G418 was removed from the culture medium after the fourth passage.

Cell culture

Mouse Mov3 [13] and 3T6 (American Type Culture Collection; CCL-96) cells transfected with the reporter construct were grown in culture as previously described for human skin fibroblasts [14]. To minimize variability in quantitative long-term culture experiments, cells were plated at the same density and fed with the same batch of medium. Approx. 1 × 10⁶ cells were seeded into 75 cm² dishes and grown to confluence in Dulbecco modified Eagle’s medium (DMEM) containing 10% (v/v) foetal calf serum. From confluence (day 0) the cultures were grown in DMEM containing 10% (v/v) foetal calf serum and supplemented daily with 0.25 mM ascorbic acid.

Biosynthetic labelling of collagen

Procollagens were biosynthetically labelled routinely at 1–2 days post-visual confluency and at other relevant times during long-term culture experiments. The medium was removed and replaced with 9.9 ml of DMEM containing 10% (v/v) foetal calf serum and 0.25 mM ascorbic acid. After 4 h, 0.1 ml of DMEM containing 50 μCi of L-[5-³H]proline (8.5 Ci/mmol; NEN Research Products) was added to the medium and the incubation was continued for a further 18 h. The final concentration of proline in the medium was 0.1 mM. In some experiments cells were labelled for 6, 9, or 24 h, with 5 μCi of L-[³H]proline (284.6 Ci/mmol; NEN Research Products), the labelling stopped by the addition of proline to a final concentration of 50 mM, and the radioactive collagen chased for up to 24 h. Following incubation, the cell layer and medium fractions were treated separately as previously described [14,15]. Briefly, after disruption of the cell layer by sonication, procollagens and collagens were precipitated from the cell and medium fractions with ammonium sulphate at 25% saturation. The precipitate was redissolved in 2 ml of 50 mM Tris/HCl, pH 7.5, containing 0.15 M NaCl and the proteinase inhibitors 5 mM EDTA, 10 mM N-ethylmaleimide and 1 mM phenylmethylsulphonyl fluoride. Aliquots of procollagens were precipitated with 75% (v/v) ethanol and subjected to limited
pepsin digestion (100 µg/ml pepsin in 0.5 M acetic acid, 4 °C, 16 h) to remove non-collagen sequences.

**Extraction of collagen from the extracellular matrix**

Cell layers from the long-term culture experiments were scraped into 2 × 5 ml of 50 mM Tris/HCl, pH 7.5, containing 0.15 M NaCl and proteinase inhibitors. The cell fractions were disrupted by sonication and extracted overnight at 4 °C [16]. Insoluble material was removed by centrifugation and the collagen in the supernatant (neutral-salt-soluble fraction) was precipitated with ammonium sulphate (25% saturation) and redissolved in 2 ml of 50 mM Tris/HCl, pH 7.5, 0.15 M NaCl and proteinase inhibitors. Insoluble collagen in the pellet was extracted by digestion with pepsin. Digestion was terminated by lyophilization.

**CNBr peptide mapping**

Freeze-dried samples of pepsin-digested collagen were resuspended in 100 mM ammonium bicarbonate and incubated at room temperature for 30 min to inactivate the pepsin. CNBr cleavage was performed in 70% (v/v) formic acid containing 50 mg/ml CNBr for 4 h at room temperature, as described by Scott and Veis [17]. After cleavage the samples were diluted with water and freeze-dried.

**Thermal stability of collagen**

The temperature at which the collagen helix was thermally denatured and thus became susceptible to proteolysis was determined by the method of Bruckner and Prockop [18]. [3H]Proline-labelled collagen samples dissolved in 0.1 M Tris/HCl, pH 7.4, containing 0.4 M NaCl were heated stepwise (1 °C/10 min) from 34 °C to 43 °C. At 1 °C intervals, samples were digested with a mixture of trypsin and chymotrypsin [18] and analysed by SDS/PAGE.

**SDS/PAGE analysis**

Collagen chains were resolved on 5% (w/v) polyacrylamide separating gels with a 3.5% (w/v) stacking gel. Collagen CNBr peptides were analysed either on 10% (w/v) polyacrylamide gels or by two-dimensional PAGE [19], which separates the peptides on the basis of both charge and size. Sample preparation, electrophoresis conditions, Coomassie Brilliant Blue staining and fluorography of radioactive gels have been described elsewhere [14,15]. Coomassie Blue-stained bands were quantified by densitometry in comparison with standard collagen samples loaded on each gel. The radioactivity in the collagen bands was quantified by excision and scintillation counting [20].

**RESULTS**

**Mutagenesis and confirmation of the introduced change**

A COL1A1 reporter gene construct was produced in which the codon for amino acid 822 of the triple-helical region was changed from the wild-type ATG (Met) to ATC (Ile). A synthetic Thal–Apal DNA fragment carrying the Met→Ile coding alteration, and as a DNA marker, a deleted NcoI site (Figure 1a), were introduced by a three-step ligation process into a 5.5 kb EcoRI genomic subclone (Figure 1b, see the Experimental section for details). Plasmids were screened for the deletion of the unique NcoI site and the modified region of the selected clone was sequenced to confirm that only the desired change had been made (results not shown). A 2.9 kb NheI–XhoI fragment of the cloned COL1A1 gene was replaced with the altered DNA. Clones were again screened for a deleted NcoI site (Figure 2a) and the construct checked for possible rearrangements by HindIII digestion. The seven bands generated by HindIII digestion of the Ile822 reporter construct were indistinguishable from those of the wild-type gene (Figure 2b). The final construct, pWTCl-Ile822, included the 3.7 kb promoter region, 17 kb of coding sequences and 3 kb of 3′-untranslated region.

**Expression of the marker construct in Mov13 cells**

To investigate the properties of type I collagen molecules composed of α1(I) chains containing the Met→Ile change and wild-type α2(I) chains, the marker construct was transfected into Mov13 cells. Mov13 cells synthesize no endogenous collagen proα1(I) chains since the transcription of both COL1A1 genes is blocked by a retroviral insertion in the first intron [13,21]. Synthesis of proα2(I) is unaffected. Transfected colonies were isolated, expanded into cell lines (Mov13-Ile822–2 to Mov13-Ile822–12) and examined for expression of the transfected marker gene. Of the eleven cell lines tested, nine synthesized [Ile822]α1(I) chains. The marker chains associated with endogenous proα2(I) to form pepsin-stable molecules and were efficiently secreted into the culture medium (Figure 3).

Substitution of a methionine residue was selected to allow the reporter protein to be distinguished from the wild-type by an altered CNBr digestion pattern. Lack of CNBr cleavage at amino acid 822 generates a 42 kDa composite peptide, CB7/6, with an expected migration on polyacrylamide gels distinct from those of other collagen CNBr peptides (Figure 4a). The presence of the Met-Ile substitution was thus confirmed by CNBr digestion. As expected, type I collagen produced by Mov13 cells
Collagens were labelled with $[^3]H$proline and prepared as described in the Experimental section. Pepsin-digested samples from the cell layer (C) and medium (M) were analysed without reduction on 5% polyacrylamide gels. Collagens produced by the untransfected parental cell line Mov13 (lanes 1 and 2), and four clonal cell lines transfected with the replicator construct pWTCI-lle822 (lanes 3 to 10) are shown. The migration positions of the type I collagen $\alpha$1(I) and $\alpha$2(I) chains are indicated.

Radiolabelled CNBr peptides were prepared and analysed by two-dimensional electrophoresis as described in the Experimental section. Medium collagens produced by 3T6 cells (a), and Mov13 cells transfected with pWTCI-lle822 (b) were examined. The main wild-type $\alpha$1(I) chain CNBr peptides (CB8, CB7, CB6 and CB3) and the incompletely cleaved CNBr peptides (CB7/6, CB6/5, CB3/7) are marked in (a). The migration positions of the CB8, CB7 and CB3/7 peptides, which are absent from the type I collagen containing reporter [lle822]$\alpha$1(I) chains, and of the composite CB7/6 peptide are indicated in (b).

**Normal thermal stability of type I collagen containing [lle822]$\alpha$1(I) chains**

The integrity of the triple-helical region of type I collagen produced by transfected Mov13 cells was determined by measuring the resistance of the helices to trypsin/chymotrypsin digestion after exposure to increasing temperatures. The thermal stability of molecules containing [lle822]$\alpha$1(I) marker chains was identical that of molecules produced by Mov13 cells transfected with the wild-type gene (Figure 6). Both types of molecules had a $T_m$ of 39.4 °C, which is higher than the 38 °C previously reported for type I collagen in this cell line [11]. This difference most likely reflects slight variations in the experimental protocols.

**Expression and quantification of [lle822]$\alpha$1(I) marker chains in 3T6 cells**

The construct was transfected into 3T6 cells to determine if cells discriminated between endogenous $\alpha$1(I) and [lle822]$\alpha$1(I) chains. Individual G418-resistant colonies were expanded and analysed separately. Expression of chains containing lle822 was readily detected after CNBr cleavage of pepsin-digested collagen by the increased relative intensity of the CB7/6 peptide. Figure 7 (lanes 2–5) shows the CNBr cleavage patterns of collagen produced by four transfected 3T6 cell lines expressing different levels of the reporter protein.

Some incompletely cleaved CB7/6 was always present in
Figure 6  Thermal denaturation of type I collagen

Pepsin-digested [3H]proline-labelled collagen from the medium of Mov13 cells transfected with either the wild-type COL1A1 gene or the reporter gene, pWTCI-Ile822, were warmed slowly from 34 °C to 43 °C. At 1 °C intervals, samples were digested with trypsin and chymotrypsin. α1-Chains resistant to proteolysis were resolved on 5% polyacrylamide gels and quantified as described in the Experimental section.

Figure 7  Electrophoresis of CNBr peptides of pepsin-digested collagens produced by 3T6 cells transfected with the reporter gene pWTCI-Ile822

[3H]Proline-labelled CNBr peptides were prepared from the cell culture medium and resolved on 12.5% polyacrylamide gels as described in the Experimental section. [3H]Proline incorporation into the CNBr peptides was determined by scintillation counting. Lane 1, collagen produced by the untransfected parental 3T6 cell line; lanes 2 to 5, collagen produced by four clonal 3T6 cell lines which had been transfected with the reporter construct pWTCI-Ile822. [Ile822]α1(I) chains represented 11%, 29%, 29% and 48% respectively of the total α1(I) chains synthesized by these cells. The migration positions of the α1(I) CB6, CB7 and CB7/6 peptides are indicated.

devolved that accounted for the low level of uncleaved CB7/6 contributed by the endogenous α1(I):

\[
[Ile^{822}]\alpha1(I) \text{ expression (\%) } = \\
\left(\frac{\text{[total CB7/6] } - \text{[endogenous CB7/6]}}{\text{[total CB7/6] } + \text{CB6 + CB7}}\right) \times 100
\]

where [endogenous CB7/6] = (CB6 + CB7) x 0.148.

Of the 17 transfected cell lines examined, 15 synthesized [Ile822]α1(I) chains, the level of expression ranging from 5% to 48% of total α1(I). There was no indication of slowly migrating CNBr peptides due to excess post-translational hydroxylation and glycosylation of lysine residues, which might be expected if the Met→Ile substitution interfered with the normal folding of the triple helix [23] (Figure 7).

[Ile^{822}]\alpha1(I) chains are secreted at the same rate as endogenous \alpha1(I) chains

The rates of secretion of endogenous α1(I) and [Ile^{822}]α1(I) chains were compared in two cell lines by pulse-labeling with [14C]proline for 6 h followed by a chase of up to 24 h in the presence of excess unlabelled proline. At each time point the proportion of endogenous and Ile822-substituted chains which had been secreted was the same (Figure 8). The observed apparent failure to secrete 100% of the pulse-labelled collagen into the culture medium most likely resulted from the experimental

control CNBr digests (Figure 4b, lane 1; Figure 5a; Figure 7, lane 1). This incomplete cleavage has been attributed to oxidation of methionine residues, steric blocking or interactions with neighbouring amino acids [17]. Attempts were made to increase the digestion efficiency by incubation for 24 h in the presence of β-mercaptoethanol prior to cleavage to ensure a reducing environment, and by extending the CNBr digestion to 24 h [22]. No increase in the digestion efficiency was achieved. The extent of cleavage by CNBr was, however, very consistent and showed the relationship CB7/6/(CB7 + CB6) = 0.148 ± 0.022 (mean ± S.D., n = 26). To calculate the level of expression of the reporter construct in transfected 3T6 cells, the following equation was

Figure 8  Secretion of endogenous \alpha1(I) and [Ile^{822}]\alpha1(I) reporter chains

Cells were pulse-labelled with [14C]proline for 6 h and the radiolabelled collagen chased for up to 24 h in the presence of excess unlabelled proline. At each point the pepsin-resistant cell layer and medium collagens were cleaved with CNBr and resolved on 10% polyacrylamide gels. [14C]Proline incorporation into the CNBr peptides was determined by scintillation counting. The proportions of endogenous α1(I) (○) and reporter [Ile^{822}]α1(I) (●) chains which had been secreted into the medium were calculated using the formula described in the Results section. (a) Secretion of α1(I) chains by 3T6-Ile822-21 cells; (b) secretion of α1(I) chains by 3T6-Ile822-23 cells.
limitations of being unable to discriminate between intracellular molecules and the small proportion of secreted molecules which remain associated with the pericellular matrix. This would not be expected to mask a significant secretion difference, particularly during the first few hours of chase time.

Efficient Incorporation of [Ile$^{822}$]x1(I) chains into an in vitro accumulated collagen matrix

When cultured for several weeks in the continuous presence of ascorbic acid, 3T6 cells deposit an extensive collagenous extracellular matrix. As the matrix matures, much of the collagen can only be extracted by pepsin digestion and contains a substantial amount of cross-linked material [24]. To compare the patterns of deposition and maturation in the matrix of endogenous and the reporter Ile$^{822}$-substituted chains, 3T6-Ile$^{822}$-21 cells were grown for up to 22 days post-confluency. The total matrix deposited by the cells and the proportion sequentially extracted by neutral salt and then pepsin digestion is shown in Figure 9(a). The extracellular matrix accumulated until day 8, with an increasing proportion of the collagen being found in the mature, cross-linked, pepsin-extracted fraction. While the cells continued to synthesize collagen and efficiently incorporate the newly synthesized molecules into the matrix (results not shown), between day 8 and day 22 approx. 50% of the accumulated collagen was degraded (Figure 9(a)). The contribution of Ile$^{822}$-substituted chains to the total x1(I) chains in each matrix extract from day 3 to day 22 is shown in Figure 9(b). There was insufficient collagen present at day 0 for accurate quantification. Throughout the experiment Ile$^{822}$-substituted chains represented approx. 30–40% of the total x1(I) chains in both the neutral-salt fraction and the more mature, more extensively cross-linked, pepsin-extracted matrix. The constant proportion of Ile$^{822}$-substituted chains during the period of matrix degradation also indicated that the chains were not selectively degraded.

The rates at which endogenous and Ile$^{822}$-substituted chains were incorporated into the matrix were compared by analysing the newly synthesized $[^3H]$proline-labelled collagen. As the collagenous matrix accumulated during culture, an increasing proportion of the newly synthesized collagen was associated with the cell layer (results not shown). Throughout the culture period the reporter [Ile$^{822}$]x1(I) chains contributed equally to the total collagen pool secreted into the medium and to the collagen fractions extracted with neutral salt buffer or pepsin (Figure 10). These data clearly demonstrated that procollagen processing and deposition into the matrix occurred with equal efficiency for the endogenous type I collagen and for collagen containing the reporter protein.

There was, however, a steady decline in the synthesis of Ile$^{822}$-substituted chains relative to endogenous x1(I) chains during prolonged culture of the transfected cells. Gradual loss of copies of the transfected gene during growth of the cells may have led to decreased relative synthesis of the transfected gene product, but this may also have been due to differential matrix-dependent regulation of endogenous and transfected gene expression. To distinguish between these possibilities, the synthesis of Ile$^{822}$-substituted chains from passages 3 to 16 in 3T6-Ile$^{822}$-21 and two other transfected cell lines was examined. The cells were passaged on alternate days and only exposed to ascorbate during the preincubation and labelling periods, and so lacked an extensive extracellular matrix. In all three cell lines there was a decrease in the relative level of expression of the transfected gene with increasing passage number. The reduced expression of the

---

**Figure 9** Accumulation of type I collagen in the extracellular matrix of 3T6-Ile$^{822}$-21 cells

Cells were grown from confluence (Day 0) in the presence of 0.25 mM ascorbic acid. Collagen deposited into the neutral-salt and pepsin-soluble fractions was analysed on polyacrylamide gels, stained with Coomassie Brilliant Blue and quantified by densitometry as described in the Experimental section. Each value represents the average of determinations on three separate culture dishes. To indicate the reproducibility of the data, the highest value determined at each point is shown by the vertical lines. (a) Accumulated type I collagen in the extracellular matrix expressed as µg of collagen extracted/µg of DNA. (b) Contribution of [Ile$^{822}$]x1(I) chains to the accumulated matrix, calculated using the formula described in the Results section and expressed as a percentage of the total x1(I) chains extracted.

**Figure 10** Distribution of newly synthesized [Ile$^{822}$]x1(I) chains in the matrix produced by 3T6-Ile$^{822}$-21 cells

Cells were grown from confluence (Day 0) in the presence of 0.25 mM ascorbic acid. At each point the cells were labelled with $[^3H]$proline for 18 h. Radiolabelled collagens released into the medium or digested into the neutral-salt and pepsin-extracted fractions were digested with CNBr, analysed on 10% polyacrylamide gels and detected by fluorography as described in the Experimental section. $[^3H]$Proline incorporation into the CNBr peptides was determined by scintillation counting. The percentage of total newly synthesized x1(I) chains in each fraction comprising reporter [Ile$^{822}$]x1(I) chains was calculated using the formula described in the Results section. Each value represents the average of determinations on three separate culture dishes. To indicate the reproducibility of the data, the highest value determined at each point is shown by the vertical lines.
transfected gene was, therefore, unlikely to be regulated by the matrix, since cells cultured under conditions which promoted very different levels of extracellular matrix synthesis and deposition showed a similar effect. Reduced synthesis of $[\text{Ile}^{\text{II}}]2(1)$ chains correlated with the time in ascorbate-supplemented culture or the number of times the cells had been passage, and thus appeared to be due to loss or inactivation of the gene during growth and cell division. Southern blots of DNA prepared from cells at passages 6, 11 and 16 indicated that the three cell lines contained multiple copies of the transfected gene in tandem arrays (results not shown). Because of the high copy number, estimated to be around 100 copies per cell in the 3T6-Ile$^{\text{II}}$-21 cell line, it was not possible to determine if any had been lost during growth.

**DISCUSSION**

This paper describes the development of an $\alpha 1(1)$ protein reporter construct for use as a universal vector in site-directed mutagenesis studies of type I procollagen synthesis and function in the extracellular matrix. The primary consideration in selecting a suitable marker alteration for introduction into the reporter gene construct was that the resultant protein could be easily detected and distinguished from the corresponding endogenous gene product, and would permit quantification. The sequence alteration needed to be in the structural triple-helical region of the molecule so that it would be present at all stages of collagen biosynthesis and fibrillogenesis and, importantly, it needed to be functionally neutral. The helical Met→Ile$^{\text{II}}$ substitution produces an easily detected CNBr cleavage pattern change and mimics an interspecies variation at residue 929 of $\alpha 1(1)$, which in the chick is a methionine residue but in ox, man and mouse is an isoleucine [25–28]. It is a chemically conservative change, maintaining charge and a similar hydrophobicity, and the sequence motif Gly-Pro-Ile-Gly-Pro generated by the change is also found in two other places in the published mouse $\alpha 1(1)$ chain sequence [28–30].

The type I collagen reporter protein construct was produced by site-directed mutagenesis of the cloned mouse COL1A1 gene and characterized in stably transfected mouse Mov13 and 3T6 cells. Detailed biochemical analysis of the reporter proα1(1) chains produced by Mov13 cells, which do not synthesize endogenous proα1(1), showed that reporter proα1(1) chains were able to associate with endogenous proα2(1) chains, formed pepsin-stable triple helices and were secreted efficiently from the cell. The collagen triple-helical structure was not disturbed by the Met→Ile alteration, since the thermal stability of molecules containing either wild-type $\alpha 1(1)$ or $[\text{Ile}^{\text{II}}]2(1)$ reporter chains was identical.

The biosynthetic characteristics of endogenous and $[\text{Ile}^{\text{II}}]2(1)$ chains were directly compared in 3T6 cells which had been stably transfected with the reporter construct. Examination of the CNBr-digested collagen produced by the transfected 3T6 cells revealed only normally migrating collagen peptides, indicating that the substitution did not alter the folding kinetics of the chains and led to abnormally high levels of modifications to lysine residues. Slow electrophoretic migration of type I collagen $\alpha$-chains and CNBr peptides is a common finding in osteogenesis imperfecta and is caused by underlying structural mutations delaying triple-helix formation and thus allowing excess posttranslational lysine hydroxylation and glycosylation [14,15,31].

Unassembled and misfolded proteins are selectively retained intracellularly in the endoplasmic reticulum [32] and protein secretion is thus a sensitive measure of the fidelity of protein folding and oligomerization. Collagen secretion is reduced by mutations which have an adverse effect on the initial association of the component chains or the folding of the triple-helical domain [14,15,31,33]. There was no difference in the rates at which reporter and endogenous $\alpha 1(1)$ chains were secreted from the cells into the extracellular space, again indicating that molecules containing $[\text{Ile}^{\text{II}}]2(1)$ chains have efficiently folded into their normal three-dimensional structure and are not discriminated by the cell.

Collagen molecules containing the reporter $\alpha 1(1)$ chains were deposited into the extracellular matrix accumulated by 3T6-Ile$^{\text{II}}$-21 cells in vitro. Ultrastructural analysis by electron microscopy demonstrated similar fibril organization in the matrices produced by the transfected and parental 3T6 cell lines (results not shown). During the 22-day ascorbate-supplemented culture period the dynamics of matrix formation changed from an initial period where only a small proportion of the collagen synthesized was processed and incorporated into the matrix, through to a period of efficient deposition. During the later stages of the culture, significant degradation of the accumulated collagenous matrix also occurred. Marker chains were deposited at the same rate as endogenous $\alpha 1(1)$ chains, and contributed to the accumulated matrix in proportion to the level of synthesis, throughout the 22-day culture and its changing pattern of matrix production, clearly demonstrating the functional equivalence of $[\text{Ile}^{\text{II}}]2(1)$ and endogenous chains in this extracellular matrix.

Site-directed mutagenesis has not yet been used to study the roles of the C- and N-propeptide domains, largely because of the inability to directly experimentally detect the mutant protein. This same restriction has limited the kinds of helical and telopeptide mutations to those which can be assessed by altered electrophoretic migration of the mutant protein. In most cases this was possible because the mutation introduced additional disulphide bonds, or changed the charge, size or extent of posttranslational modifications of the $\alpha$-chains [6,7,9,34]. The Met→Ile$^{\text{II}}$ reporter protein construct eliminates this need for the mutation itself to be detectable and allows normal and mutant proteins to be readily discriminated and quantified. This will facilitate detailed studies of poorly understood intracellular biosynthetic events such as the regulation of chain selection and subunit assembly and the targeting of abnormal proteins for degradation. The ability to express protein-engineered mutant collagen in the 3T6 cell culture system and monitor extracellular matrix formation and composition greatly extends the biosynthetic information that can be obtained from in vitro experiments, allowing questions about the regulation of propeptide processing and fibrillogenesis and the roles of other matrix molecules in these events to be addressed. It will be possible to biochemically screen mutations in this system before selecting suitable examples for phenotypic expression in transgenic mice.

The reporter construct will also be valuable for assessing the effects of mutations in collagen regulatory regions and in the transcription factors which interact with them. The deleted Neo site will allow reporter and wild-type mRNA transcripts to be quantified by methods which combine PCR amplification and differential restriction enzyme digestion [35]. Experimental approaches to collagen gene regulation have most commonly involved measurement of the responses of regulatory elements attached to reporting genes such as chloramphenicol acetyltransferase [36–38], human growth hormone [37] or mini-gene constructs [7]. These suffer from the potential problem that the regulatory elements are not necessarily within an appropriate sequence context that will allow natural interactions of the regulatory regions with other, possibly distant, gene domains [39]. The Met→Ile$^{\text{II}}$ reporter construct largely overcomes these.
problems by allowing the examination of regulatory mutations within the context of the entire \textit{COLIA1} gene.

We thank Jacinta White for electron microscope analysis of the Mov13 and 3T6 cell matrices. This work was supported by grants from the National Health and Medical Research Council of Australia and the Royal Children's Hospital Research Foundation.

REFERENCES


Received 24 November 1992/11 February 1993; accepted 17 February 1993