High Levels of Expression of Full Length Human Pro-α2(V) Collagen cDNA in Pro-α2(V)-deficient Hamster Cells

(Received for publication, July 10, 1989)

Daniel S. Greenspan, Guy G. Hoffman, and Bum-Soo Lee

From the Department of Pathology, University of Wisconsin, Madison, Wisconsin 53706

A full length cDNA encoding human pro-α2(V) collagen was constructed. Partial sequencing of the cDNA and primer extension analysis of mRNA from fibroblasts found that pro-α2(V) mRNA differs from the mRNAs of other fibrillar collagens in the increased length of its 5'-untranslated region. The pro-α2(V) cDNA was placed downstream of the human cytomegalovirus immediate early promoter/regulatory sequences for expression studies in cultured Chinese hamster lung cells. These cells have been shown previously to synthesize large quantities of pro-α1(V) homotrimers as their only collagenous product. Transfection resulted in a number of clonal cell lines that express human α2(V) RNA at levels comparable to, and in some cases greater than, levels found in normal human skin fibroblasts. Pro-α2(V) chains produced in the majority of clonal lines were of sufficient quantity to complex all available endogenous pro-α1(V) chains. Chimeric heterotrimers, composed of hamster α1(V) and human α2(V) chains in a 2:1 ratio, were stable to pepsin digestion and were found predominantly associated with the cell layer. Surprisingly, pro-α2(V) chains, in excess to pro-α1(V) chains, were found in the extracellular matrix and, in much greater abundance, in media. These chains were pepsin sensitive, indicating that pro-α2(V) chains can be secreted as nonstable homotrimers or as free chains.

The 12 or more types of collagens identified thus far comprise the major structural components of the extracellular matrix and together represent approximately 30% of total body proteins in humans. Type I collagen, the major fibrous component of connective tissue, is the most abundant protein in the body. The more recently described type V collagen is widely distributed in many tissues as a pericellular component. Type V collagen is also frequently found closely associated with the cell layer. Surprisingly, pro-α2(V) chains, in excess to pro-α1(V) chains, were found in the extracellular matrix and, in much greater abundance, in media. These chains were pepsin sensitive, indicating that pro-α2(V) chains can be secreted as nonstable homotrimers or as free chains.

* This work was supported by National Institutes of Health Grant GM-38544 (to D. S. G.) and by a grant-in-aid from the American Heart Association (to D. S. G.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be sent: Dept. of Pathology, University of Wisconsin, 1900 University Ave., Madison WI 53706.

1 The abbreviations used are: HCMV-IE, human cytomegalovirus immediate early; CHL, Chinese hamster lung; SDS, sodium dodecyl sulfate; bp, base pair(s).
in subsequent expression studies. Oligonucleotides were synthesized at the University of Wisconsin Biotechnology Center.

**Cell Culture and Transfection—**CHL cells (clone HT1) have been described previously (6, 10–13). AH1F cells are normal neonatal foreskin fibroblasts (16). All cells were cultured in Dulbecco's modified Eagle's medium with 10% fetal calf serum. Cotransfection of CHL cells with pSVneo and pGZ4H3 was by calcium phosphate precipitation as described (16), except that cloning lines were selected in 0.8 mg/ml G418 (GIBCO).

**Isolation and Analysis of Nucleic Acids—**Cytoplasmic RNA was isolated by disruption of cells in isotonic lysis buffer in the presence of Nonidet P-40 (Shell Chemicals), with removal of nuclei by centrifugation as described (17). Polyadenylated RNA was selected from cytoplasmic RNA by chromatography on (dT)15-cellulose (type 2; Collaborative Research, Inc.) (17).

For primer extension, a 22-base oligomer complementary to sequences from 112 to 133 bases upstream of the translation initiation codon of pro-a2(V) mRNA (see Fig. 1A) was 5' end labeled with polynucleotide kinase and hybridized to 3 nCi of polyadenylated RNA from the cytoplasm of AH1F cells. RNA-DNA hybridization was overnight at 52 °C. Otherwise, hybridization conditions and extension with avian myeloblastosis virus reverse transcriptase (Life Science) were as described previously (18).

DNA sequences were obtained by diodeoxy chain termination (16).** Isolation of Radiolabeled Collagen and Procollagen—**Radiolabeling and isolation of collagen species for pepsin digestion were as described (16) except that precipitation from medium was achieved with 30% acetic acid (pH 2.0) for 6 h at 4 °C. Type V procollagen, which was not to be pepsinized (see Fig. 6), was extracted from the extracellular matrix associated with the cell layer by sequential extraction with 1% deoxycholate followed by 4% SDS (19). Cell layers were scraped from the Petri dish and centrifuged at 100,000 × g for 30 min. Deoxycholate-insoluble material was rinsed in 1% deoxycholate and then solubilized by boiling in 4% SDS gel buffer prior to electrophoresis.

**RESULTS**

**Sequence Analysis—**The full length cDNA cloned here contains 102 bp more 5'-untranslated sequence than reported previously (4) (Fig. 1A). In addition, DNA sequences at the 5' end of the full length pro-a2(V) cDNA differ from recently published a2(V) sequences derived from partial length cDNAs (3, 4). Of particular interest is a thymidine residue, not appearing in sequences reported previously. Sequences in brackets are from Ref. 4, otherwise, hybridization conditions and extension with avian myeloblastosis virus reverse transcriptase (Life Science) were as described previously (18).

**Expression of Human a2(V)-specific RNA in Transfected CHL Cells—**The full length pro-a2(V) cDNA was placed downstream of the immediate early promoter/regulatory region of human cytomegalovirus (Fig. 3). These HCMV-IE regulatory sequences are significantly more active in directing transcription of transfected genes in a variety of eukaryotic cells than the Rous sarcoma virus long terminal repeat (24), which had previously been used in this laboratory to direct transcription of pro-a2(I) collagen cDNA (16). In order to provide pro-a2(V) cDNA transcripts with the ability to splice and be polyadenylated, the expression vector was furnished with the SV40 small t splice site and early polyadenylation signal. The resultant recombinant (pGHH31) was cotransfected with the selectable marker pSV2neo (25) into CHL cells. Cytoplasmic RNA from clonal lines of transfected CHL cells, resistant to the neomycin analogue G418, was then

| Amino Acid | Position | Change |
|------------|----------|--------|
| Gly        | 204      | S       |
| Gly        | 218      | A       |
| Gly        | 254      | G       |

**Fig. 1. Sequence analysis.** A, the 5' end of the full length pro-a2(V) cDNA comprised of 102 bp of untranslated sequences not reported previously. The 5 bp that overlap but differ from the first 5 bases at the 5' end of pro-a2(V) sequences reported previously (4) are underlined, and the corresponding sequences from Ref. 4 are drawn beneath (parentheses). An arrow, drawn beneath 22 bases of the 5'-untranslated sequence (from 112 to 133 bases upstream of the a2(V) start codon), represents the complementary sequences that comprise an oligonucleotide, used in primer extension studies. B, the position of the thymidine residue (circled), reported here but absent in pro-a2(V) sequences reported previously (4), is shown 16 bases upstream of the start site for pro-a2(V) translation. Themino acid sequences of two tetrapeptides, potentially encoded by small open reading frames upstream of the 5' region of human a2(V) mRNA, are shown. Asterisks represent translation stop codons. A stem-loop structure potentially formed by highly conserved sequences containing the newly reported thymidine residue is also shown. C, amino acids and nucleotides that differ from sequences reported previously. Sequences in parentheses are from Ref. 4. Sequences in brackets are from Ref. 3. The numbering system for the amino acid sequences is from Ref. 4.

**Expression of Human a2(V)-specific RNA in Transfected CHL Cells—**The full length pro-a2(V) cDNA was placed downstream of the immediate early promoter/regulatory region of human cytomegalovirus (Fig. 3). These HCMV-IE regulatory sequences are significantly more active in directing transcription of transfected genes in a variety of eukaryotic cells than the Rous sarcoma virus long terminal repeat (24), which had previously been used in this laboratory to direct transcription of pro-a2(I) collagen cDNA (16). In order to provide pro-a2(V) cDNA transcripts with the ability to splice and be polyadenylated, the expression vector was furnished with the SV40 small t splice site and early polyadenylation signal. The resultant recombinant (pGHH31) was cotransfected with the selectable marker pSV2neo (25) into CHL cells. Cytoplasmic RNA from clonal lines of transfected CHL cells, resistant to the neomycin analogue G418, was then

**S. T. Lee and D. S. Greenspan, unpublished data.**
analyzed by the S1 nuclease protection assay for the presence of human α2(V)-specific transcripts (Fig. 4). Of the 28 lines tested, 17 of which are shown (Fig. 4), the RNAs of 23 were found to protect a 393-base DNA fragment diagnostic for human α2(V) sequences. Surprisingly, some transfected CHL lines (Fig. 4, lanes 9 and 17) were found to produce higher levels of human α2(V) RNA than do normal human diploid fibroblasts (AH1F, Fig. 4). As expected, no protected fragment resulted from probe hybridized to the cytoplasmic RNA of untransfected CHL cells (Fig. 4).

**Association of Human α2(V) Chains with Hamster α1(V) Chains in Cell Layers**—The six clonal lines found by S1 analysis to contain the highest levels of human α2(V)-specific RNA were metabolically labeled with [3H]proline and their media and cell layers analyzed separately for the presence of collagen species. As reported previously (6, 13), untransfected CHL cells produce only α1(V) homotrimers that are localized to the cell layer (Fig. 5A, lane 1). In contrast to untransfected CHL cells, the cell layers of clonal lines that had been found previously to produce high levels of human α2(V)-specific RNA are shown to contain both pepsin-resistant α1(V) and α2(V) chains (Fig. 5A, lanes 2–7). Densitometric scanning of autoradiograms, exposed for varying lengths of time, gave ratios of α1(V) to α2(V) very closely approximating 2:1 for five of the six lanes containing the cell layer-associated collagens of clonal lines (Fig. 5A, lanes 2–6). The ratio of α1(V) to α2(V) chains in the seventh lane was 5:2:1.

The medium of untransfected CHL cells is shown to contain low levels of a pepsin-resistant species (Fig. 5B, lane 1) that has been reported previously to be an approximately 85-kDa proteolytic cleavage product of α1(V) chains (10). In contrast, the media of clonal lines that produce human α2(V) chains contained small amounts of full length α1(V) chains (Fig. 5B, lanes 2–7). A second pepsin-resistant band found in the media of these cells (Fig. 5B, lanes 2–7) appears to be slightly larger than the putative α1(V) proteolytic fragment found in media of untransfected CHL cells (Fig. 5B, lane 1) and to comigrate with α2(V) chains. Thus, it is likely that the pepsin-resistant species in the media of clonal lines includes small quantities of heterotrimers. However, the lower band is rather broad and is more abundant than the α1(V) band. It therefore probably contains both α2(V) chains and α1(V) proteolytic fragments that have not been resolved during electrophoresis.

**Analysis of Unpepsinized Collagen Species**—In order to analyze procollagen species, unpepsinized samples from the extracellular matrix and media of a clonal line, designated A1, which had been found previously to produce the highest levels
of α2(V) RNA (Fig. 4, lane 17) and protein (Fig. 5A, lane 2), were analyzed by SDS-polyacrylamide gel electrophoresis (Fig. 6). Unexpectedly, the media of these cells contained abundant pro-α2(V) chains that had not been incorporated into matrix. Based on the results of Fig. 5, the vast majority of these media pro-α2(V) chains are not stable to limited pepsin digestion. The extracellular matrix, isolated by sequential extraction of the cell layer in deoxycholate and SDS (19), was found to contain pro-α1(V) and pro-α2(V) chains in the ratio 1:1.7. This implies that in addition to pepsin-resistant heterotrimers found in the cell layer (Fig. 5A, lane 2), a small proportion of pepsin-sensitive pro-α2(V) chains, not bound to pro-α1(V) chains, has been incorporated or trapped in the extracellular matrix. The absence of processed type V collagen chains in the extracellular matrix of Fig. 6 is consistent with the results of Fessler et al. (13) which showed CHL cells to lack specific procollagen peptidase activity. In contrast to other reports in which varying amounts of pro-α2(V) have been found disulfide linked to pro-α1(V) chains (1), unreduced gels did not reveal disulfide links between pro-α2(V) and pro-α1(V) chains in the present study (data not shown).

Discussion

In this study, we have constructed a full length cDNA encoding the human pro-α2(V) collagen chain in its entirety. Analysis of pro-α2(V) cDNA sequences indicates pro-α2(V) mRNA to contain a highly conserved inverted repeat and two short open reading frames around the translation initiation site similar to what has been reported for the mRNAs of other fibrillar collagen chains (20). However, pro-α2(V) mRNA is shown, by primer extension, to differ from the mRNAs of other fibrillar collagens in the greater length of its 5’-untranslated region.

Pro-α2(V) cDNA, driven by HCMV-IE promoter/regulatory sequences, is shown to express at high levels upon transfection into CHL cells. Levels of human α2(V)-specific RNA in the cytoplasm of transfected lines of CHL cells were comparable to, and in some cases greater than, levels of α2(V) RNA found in normal human diploid fibroblasts. Although production of endogenous pro-α1(V) homotrimers has been shown previously to represent from 20 to 30% of the total capacity of these cells for protein synthesis (11, 12), levels of human pro-α2(V) chains in most clonal lines examined were sufficient to have complexed all endogenous pro-α1(V) chains.

The 2:1 ratio of pepsin-resistant hamster α1(V) chains to human α2(V) chains found in the cell layers of a majority of clonal lines is most consistent with the interpretation that these chains form chimeric hamster/human ((α1 V),(α2 V)) heterotrimers. Pepsin resistance indicates that these heterotrimers are in stable triple helical form. Moreover, the heterotrimers are predominantly incorporated into the extracellular matrix of the cell layer. This is characteristic of the close association reported for normal type V collagen and a variety of cultured cell types (13).

The 2:1 ratio of α1(V) and α2(V) chains found in pepsinized material from cell layers is also consistent with in vitro renaturation studies (9) and indicates that pro-α2(V) chains do not form pepsin-resistant triple helical pro-α2(V) heterotrimers in vivo. Surprisingly, however, media samples were found to contain abundant amounts of pro-α2(V) chains that were not resistant to limited pepsin digestion at 4°C. This suggests that pro-α2(V) chains can be secreted in a form that does not have a compact triple helical configuration. This is in contrast to numerous studies with type I collagen which suggest that pro-α chains must be in a stable triple helical conformation for productive secretion to occur (26). It remains to be determined whether pepsin-sensitive pro-α2(V) chains in media exist as nonstable homotrimers or as individual chains. In this regard, it is of interest that a line of transformed Syrian hamster cells that do not synthesize pro-α1(I) chains has been shown recently to secrete homotrimers of pro-α2(I) chains (27). These pro-α2(I) homotrimers were relatively unstable and were pepsin sensitive at 15°C. However, unlike the pro-α2(V) chains reported here, the pro-α2(I) homotrimers were triple helical and stable to pepsin digestion at 4°C.

The finding that small quantities of triple helical heterotrimers appear in the media (Fig. 5B) whereas intact triple helical pro-α1(V) homotrimers are found exclusively in the cell layer suggests that pro-α1(V) homotrimers may have greater affinity for the cell layer than heterotrimers containing the pro-α2(V) chain. The finding that pepsin-sensitive pro-α2(V) chains are predominantly secreted into media rather than retained in matrix suggests that pro-α2(V) chains, uncomplexed to pro-α1(V) chains, have poor affinity for matrix or that pro-α chains that are not in a stable triple helical configuration are in general poorly incorporated into matrix.

The high levels of expression of pro-α2(V) cDNA achieved here are in contrast to lower levels achieved previously with pro-α2(I) cDNA in the W8 line of rat cells (16). One element contributing to this difference appears to be use of the HCMV-IE promoter/enhancer sequences instead of the Rous sarcoma virus long terminal repeat used in the pro-α2(I) study. In transient expression assays in which pro-α2(I) cDNA is driven by either the Rous sarcoma virus long terminal repeat or HCMV-IE promoter/enhancer in otherwise identical expression vectors, the HCMV-IE sequences yield higher levels of expression. Recently, it has been reported that the HCMV-IE promoter/regulatory sequences can obviate the need for introns in the efficient expression of immunoglobulin cDNA (28). Preliminary results suggest that in constructs containing either pro-α2(I) or pro-α2(V) cDNA, the downstream SV40 small t splice site is used inefficiently, and only a small percentage of transcripts is spliced. Therefore, the positive effect of HCMV-IE sequences on expression may represent, to some degree, its ability to compensate for inefficient splicing in these constructs.

The high level functional expression of pro-α2(V) cDNA achieved in CHL cells provides a unique system for study of type V collagen. Site-directed mutagenesis of the cDNA, prior to transfection, may allow mapping of domains important in the biosynthesis and cellular functions of type V collagen. The enzymatic processing of type V procollagen into more
mature forms, which remains somewhat obscure (1), might also be addressed in this system, provided that (i) specific procollagen peptidase activities shown to be defective in CHL cells (13) are supplied from culture media of normal fibroblasts; and (ii) the properties of the human/hamster heterotrimer are sufficiently similar to those of naturally occurring type V heterotrimers. Alternatively, transfection of pGGH31 into normal human fibroblasts should yield levels of expression which would be high relative to levels of endogenous type V collagen. This should allow study of the effects of specific mutations on the expression which would be high relative to levels of endogenous type V collagen and perhaps reveal a role for aberrant type V collagen in some inherited disorders of connective tissue.

Acknowledgments—We are grateful to Dr. Michael A. Haralson (Dept. of Pathology, Vanderbilt University) for supplying us with the CHL cell line and for useful advice. We thank Seung-Taek Lee for helpful conversations. Dr. Burton Goldberg for critical reading of the manuscript, and Ellen Pluim for excellent technical assistance.

REFERENCES

1. Fessler, J. H., and Fessler, L. I. (1987) in Structure and Function of Collagen Types (Mayne, R., and Burgess, R. E., eds) pp. 81–109, Academic Press, Orlando, FL.
2. Myers, J. C., Loidl, H. R., Seyer, J. M., and Dion, A. S. (1985) J. Biol. Chem. 260, 11216–11222.
3. Weil, D., Bernard, M., Gargano, S., and Ramirez, F. (1987) Nucleic Acids Res. 15, 181–197.
4. Woodbury, D., Benson-Chanda, V., and Ramirez, F. (1989) J. Biol. Chem. 264, 2735–2738.
5. Miller, E. J. (1984) in Extracellular Matrix Biochemistry (Piez, K. A., and Reddi, A. H., eds) pp. 41–81, Elsevier Science Publishing Co., New York.
6. Haralson, M. A., Mitchell, W. M., Rhodes, R. K., Kresina, T. F., Gay, R., and Miller, E. J. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 5206–5210.
7. Rhodes, R. K., and Miller, E. J. (1978) Biochemistry 17, 3442–3448.
8. Deyl, Z., Macek, K., and Adam, M. (1979) Biochem. Biophys. Res. Commun. 89, 627–634.
9. Bentz, H., Bachinger, H. P., Glanville, R., and Kuhn, K. (1978) Eur. J. Biochem. 92, 563–567.
10. Haralson, M. A., Mitchell, W. M., Rhodes, R. K., and Miller, E. J. (1984) Arch. Biochem. Biophys. 229, 509–518.
11. Haralson, M. A., Frey, K. L., and Mitchell, W. M. (1978) Biochemistry 17, 864–868.
12. Haralson, M. A., Sonneborn, J. H., and Mitchell, W. M. (1978) J. Biol. Chem. 253, 5536–5542.
13. Fessler, L. I., Robinson, W. J., and Fessler, J. H. (1981) J. Biol. Chem. 256, 9646–9651.
14. Wolf, D., and Rotter, V. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 780–794.
15. Stinski, M. F., and Roehr, T. J. (1985) J. Virol. 55, 431–441.
16. Lee, S.-T., Smith, B. D., and Greenspan, D. S. (1988) J. Biol. Chem. 263, 13414–13418.
17. Greenspan, D. S., and Weissman, S. M. (1985) Mol. Cell. Biol. 5, 1894–1900.
18. Ghosh, P. K., Reddy, V. B., Piatak, M., Lebowitz, P., and Weissman, S. M. (1986) Methods Enzymol. 65, 580–595.
19. McKeown-Longo, P. J., and Mosher, D. F. (1984) J. Biol. Chem. 259, 12210–12215.
20. Yamada, Y., Mudryj, M., and de Crombrugghe, B. (1983) J. Biol. Chem. 258, 14914–14919.
21. De Crombrugghe, B., Schmidt, A., Liu, G., Setoyama, C., Mudryj, M., Yamada, Y., and McKeon, C. (1985) Ann. N. Y. Acad. Sci. 460, 154–162.
22. Chu, M.-L., de Wet, W., Bernard, M., and Ramirez, F. (1985) J. Biol. Chem. 260, 2315–2320.
23. de Wet, W., Bernard, M., Benson-Chanda, V., Chu, M.-L., Dickson, L., Weil, D., and Ramirez, F. (1987) J. Biol. Chem. 262, 16032–16036.
24. Cullen, B. R. (1986) Cell 46, 973–982.
25. Southern, P. J., and Berg, P. (1982) J. Mol. Appl. Genet. 1, 327–341.
26. Dziadek, M., Timpl, R., and Jaenicke, R. (1987) Biochem. J. 244, 375–379.
27. Majmudar, G., Schalk, E., Batenzaan, J., and Peterkofsky, B. (1988) J. Biol. Chem. 263, 5555–5559.
28. Neuberger, M. S., and Williams, G. T. (1988) Nucleic Acids Res. 16, 6713–6724.