Rearranged and Germline Immunoglobulin κ Genes: Different States of DNase I Sensitivity of Constant κ Genes in Immunocompetent and Nonimmune Cells†

Ursula Storb,* Ronald Wilson, Erik Selsing, and Alan Walfield

ABSTRACT: The rearrangement of a variable (V) and a constant (C) gene appears to be a necessary prerequisite for immunoglobulin gene expression. Multiple different rearranged κ genes were found in several mouse myelomas, although these cells produce only one type of κ chain [Wilson, R., Miller, J., & Storb, U. (1979) *Biochemistry 18*, 5013–5021]. It is therefore of interest to understand how only one allele within a lymphoid cell becomes expressed, while the other allele remains nonfunctional ("allelic exclusion"). We have studied the chromatin conformation of κ genes by making use of the preferential digestion of potentially active genes by DNase I described, for example, for globin genes [Weintraub, H., & Groudine, M. (1976) *Science (Washington, D.C.) 193*, 848–856]. The DNase I sensitivity of κ genes in myeloma tumors, in a B cell lymphoma, and in liver was determined by hybridization with DNA on Southern blots. It was found that rearranged C_κ genes are DNase I sensitive in myelomas in which several κ genes are rearranged, regardless of whether the rearranged genes code for the κ chains synthesized by the cell. Furthermore, the C_κ gene in germline configuration is also DNase I sensitive in a B cell lymphoma; i.e., it is in the same chromatin state as the rearranged C_κ gene which probably codes for the κ chains produced by the cell. The altered chromatin state appears to be localized: V_κ genes in germline context are not DNase I sensitive in myeloma or B lymphoma cells while C_κ genes present in a κ gene cluster on the same chromosomes are sensitive. When rearranged, however, the V_κ genes are as sensitive to DNase I as are rearranged C_κ genes. V_λ and C_λ genes are not DNase I sensitive in κ myelomas. Thus, commitment to κ gene expression is apparently correlated with a chromatin conformation which confers increased DNase I sensitivity to the DNA in the vicinity of all C_κ genes in the cell. "Allelic exclusion" does not operate on the level of chromation conformation which can be detected by altered DNase I sensitivity.

Immunoglobulins are coded by closely linked genes, V and C, for the variable and constant regions of the protein (Dreyer & Bennett, 1965). Studies of the structure of DNA have supported the hypothesis that before transcription particular V and C genes are rearranged so that they come to be in close proximity, although not contiguity (Brack et al., 1978; Seidman & Leder, 1978; Early et al., 1979; Wilson et al., 1979). The expression of a given class of immunoglobulin genes seems to obey "allelic exclusion"; i.e., a given plasma cell in a heterozygous individual produces only one of the two allotypes found in the serum (Pernis et al., 1965). It appeared possible that allelic exclusion was due to the exclusive rearrangement of one immunoglobulin gene on one of the two allelic chromosomes. This view had to be reconsidered when it was found by restriction mapping that many myeloma cells have multiple rearranged C_κ genes (Lenhard-Schuller et al., 1978; Wilson et al., 1979; Steinmetz & Zachau, 1980; Rabbits et al., 1980).

We have therefore begun to search for the level of C_κ gene expression at which allelic exclusion is manifest first. It has been observed in other systems that expressed genes have an altered chromatin conformation which renders them sensitive to digestion when nuclei are incubated for short times with low amounts of deoxyribonuclease I (Weintraub & Groudine, 1976). Thus, globin genes can be digested in nuclei of chicken erythrocytes (Weintraub & Groudine, 1976; Stalder et al., 1980) and mouse erythroleukemia and fetal liver cells (Miller et al., 1978) but not in nuclei of chicken fibroblasts and brain or mouse adult liver and hepatoma. Also active and inactive Rous sarcoma virus genes (Groudine et al., 1978), adenovirus genes (Flint & Weintraub, 1977), and ovalbumin genes (Garel & Axel, 1976; Palmiter et al., 1977) can be distinguished in this way.

It has, futhermore, been found that genes which once had been transcribed (Weintraub & Groudine, 1976; Shepherd et al., 1977) or which are in a state of extremely low transcriptional activity but inducible to high transcriptional activity (Miller et al., 1978) are also susceptible to digestion by DNase I.

DNase I sensitivity seems to be limited to active genes and closely neighboring DNA sequences (Flint & Weintraub, 1977). Mild DNase I digestion with subsequent analysis of the DNA by hybridization appeared therefore to provide a specific technique to probe for active or potentially active immunoglobulin genes. Recently, Wu et al. (1979) and Stalder et al. (1980) have extended the usefulness of the method by combining it with restriction enzyme–Southern blot analysis of the DNA. In the present study, we have used this modified method to determine the conformation of C_κ genes in cells which produce immunoglobulins (myelomas, B lymphoma) and in cells which do not (liver). We have analyzed DNA from mildly DNase I digested nuclei of myelomas, a B cell lymphoma, and liver have determined the states of DNase I sensitivity of germline and rearranged κ genes.

Experimental Procedures

Mice and Tumors. These were the same as in Wilson et al. (1979) except for B lymphoma WEH1 279 (Warner et al., 1975). The myelomas MOPC-41, MOPC-21, MOPC-167, and MOPC-321 and the B lymphoma WEH1 279 produce κ chains, and the myeloma S-178 produces λ chains but also contains a moderate level of κ RNA (Storb et al., 1977).

Preparation and DNase Digestion of Nuclei. A modification of the procedures of Weintraub & Groudine (1976) and Miller et al. (1978) was used. Tissues were excised into ice-

† From the Deparment of Microbiology and Immunology, University of Washington, Seattle, Washington 98195. *Received July 2, 1980.* Supported by National Institutes of Health Grants CA/AI 25754 and DE 02600 and National Science Foundation Grant PCM 78-13205.

0006-2960/81/0420-0990$01.00/0

cold $^1/_2$ × TKM buffer (50 mM Tris-HCl, pH 7.5, 50 mM KCl, and 15 mM $MgCl_2$ is 1 × TKM), chopped with scissors, and disrupted by two to five strokes with an A pestle in a Dounce homogenizer. Nuclei were pelleted onto a cushion of 60% sucrose–TKM by spinning at 2000 rpm for 10 min in an International centrifuge. The crude nuclei were mixed with the sucrose and resuspended in $^1/_2$ × TKM. Triton X-100 was added to 0.5%; the nuclei were pelleted at 1000 rpm for 10 min, washed in $^1/_2$ × TKM until the wash fluid was almost clear, and then washed once in reticulocyte standard buffer (RSB) (0.01 M Tris-HCl, pH 7.4, 0.01 M NaCl, and 3 mM $MgCl_2$). The washed nuclei were suspended in RSB as 2 × 10^8 nuclei/mL and treated for 1 min at room temperature with 0.25–16 μg/mL DNase I (Schwartz-Mann, RNase-free). There were considerable variations in DNase I susceptibility. The DNase activity was stopped by the addition of EDTA, pH 7.0, to 5 mM, $NaDodSO_4$ to 0.5%, and 100 μg/mL proteinase K.

Preparation of DNA. DNA was prepared from very mildly DNase I treated and untreated nuclei as described (Wilson et al., 1979).

Cloned κ DNA. The cDNA clone used as a hybridization probe was constructed by Philip Early in L. Hood's laboratory, California Institute of Technology (Joho et al., 1980), from MOPC-167 κ mRNA. It consists of the V and C regions plus the 3′ untranslated region. The total cloned DNA or the isolated C region DNA was nick translated with [^{32}P]NTPs (Rigby et al., 1977).

Cloned λ DNA. The $p\lambda_{I-1}$ clone was constructed by A. Bothwell in D. Baltimore's laboratory, Massachusetts Institute of Technology, from myeloma MOPC-104E. The insert (λ V + C) of about 800 base pairs is excisable from PBR322 by *Pst*I.

Restriction Endonuclease Digestion and Southern Blots. These procedures were performed as described previously (Southern, 1975; Jeffreys & Flavell, 1977; Wilson et al., 1979).

Scanning of Southern Blots. Southern blot films were scanned with a Helena Laboratories Quick Scan densitometer. The relative areas occupied by hybridization bands were determined by a Quick-Quant II integrator.

Results

In order to determine the DNase I sensitivity of specific κ genes, nuclei were digested with DNase I under extremely mild conditions so as to produce only one or a few breaks in a susceptible sequence. The DNAs purified from these nuclei were then treated with various restriction endonucleases and analyzed by Southern blots with a cloned cDNA from κ mRNA of myeloma MOPC-167 as a probe. The V region of this sequence, $V_{\kappa\text{-}167}$, is expressed only in the myeloma MOPC-167 and not in the other myelomas used in this study. The susceptibility of nuclei to DNase I varied between cell types; i.e., the same amount of DNase I resulted in differing degrees of DNA digestion. Therefore digestion stages indicated in the figures cannot be compared directly between different tumors. The V_{167} gene was used as an internal control as a germline gene in tumors where it is not expressed since its sensitivity to DNase I was related directly to the size of the bulk of the DNA (see below). The Southern blotting results are shown in Figures 1–4. The V and C designations indicated in the figures had been determined previously (Wilson et al., 1979). For orientation with respect to restriction maps of germline and rearranged genes, see Figure 6.

When nuclei from liver, a germline tissue, were DNase I digested, both $V_{\kappa\text{-}167}$ and C_κ germline genes disappeared at the same time (Figure 1A and Table I). The decrease in the intensity of the hybridization bands proportionately follows the degree of DNA digestion (Figure 5A). At digestion stage 4, where traces of the V and C bands are still visible (Figure 1A), the bulk of the DNA is smaller than 4 kb. On the other hand, when myeloma MOPC-41 nuclei were digested with DNase I and the DNA analyzed after digestion with *Bam*HI, the two rearranged C_κ genes were found to be considerably more sensitivity than the germline $V_{\kappa\text{-}167}$ gene which is not expressed in this tumor (Figure 1B). By the third stage of DNase I treatment (Figure 1B, part 3) the C genes are barely apparent on the blot whereas the $V_{\kappa\text{-}167}$ gene is still visible after digestion with 2 times as much DNase (Figure 1B, part 4). These same MOPC-41 DNA samples were analyzed in Southern blots by using a cloned λ cDNA as the hybridization probe (Figure 2A). After restriction with *Eco*RI, four λ bands are seen: a 8.6-kb $C_{\lambda\text{-}1}$, a 4.8-kb $V_{\lambda\text{-}2}$, a 3.5-kb $V_{\lambda\text{-}1}$, and a 2.7-kb $J_{\lambda\text{-}I}$ gene containing band (Brack et al., 1978; Sakano et al., 1979). All four were as insensitive to DNase I as the $V_{\kappa\text{-}167}$ gene—the four bands are last seen in digestion stage 4 (compare Figure 1B and Figure 2A). This tumor does not produce λ chains or λ mRNA (Storb et al., 1977). The results suggest that rearranged sequences, in this case both C_κ genes, are more sensitive to DNase I digestion than the $V_{\kappa\text{-}167}$ gene

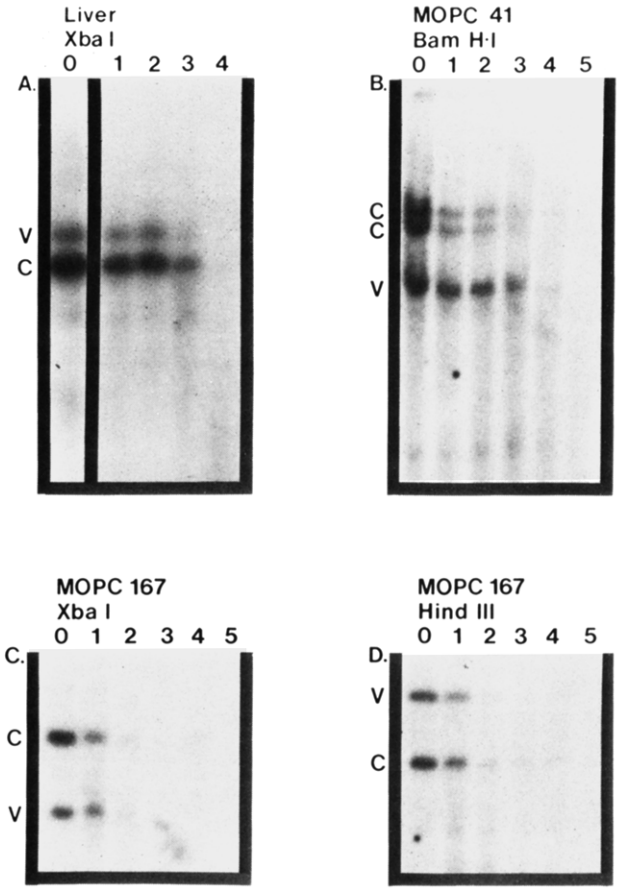

FIGURE 1: Souther blot analysis of DNA from DNase I digested and control nuclei. The DNAs were digested with the indicated enzymes. Hybridization probe: clone p167kRI: V + C sequences. (A) Liver: stage 0, undigested nuclei; stages 1–4, nuclei digested with 2, 4, 8, and 16 μg/mL DNase I. V is the fragment containing germline $V_{\kappa\text{-}167}$ gene (7.1 kb); C is the fragment containing germline C_κ gene (5.6 kb). (B) MOPC 41: stage 0, undigested nuclei; stages 1–5 nuclei digested with 0.5, 1, 2, 4, and 6 μg of DNase I/mL. C is the rearranged C gene containing fragments (12.9 and 11.1 kb). V is the $V_{\kappa\text{-}167}$ gene in germline arrangement (6.5 kb). (C and D) MOPC-167: stage 0, undigested nuclei; stages 1–5, nuclei digested with 1, 2, 3, 4, and 5 μg of DNase I/mL. *Xba*I: C is the rearranged C_κ gene (6.1 kb); V is the rearranged $V_{\kappa\text{-}167}$ gene (3.4 kb). *Hin*dIII: V is the rearranged $V_{\kappa\text{-}167}$ gene (7.3 kb); C is the rearranged C_κ gene (4.9 kb).

Table I: Relative Quantities of C_κ Genes after DNase I Digestion[a]

digestion stage	liver[d]	ratio C_κ ($V_{\kappa\text{-}167}$ + C_κ) MOPC-167[d]	MOPC-321[d]	MOPC-21[d]
0	0.78	0.60	0.72	0.51
1	0.74	0.59	0.43	NS[c]
2	0.69	0.67	0[b]	0.41
3	0.81	trace of C		0.38
4	no V or C	no V or C		0[b]

[a] The numbers represent the ratios of $C_\kappa/(C_\kappa + V_{\kappa\text{-}167})$ band intensities as determined by microdenitometry of Southern blot autoradiograms. For conditions of DNase I digestion, see Figures 1 and 3. The restriction enzyme used for liver, MOPC-321, and MOPC-21 was *Xba*I which produces a 7.1-kb V gene band, a 5.6-kb C gene band, and, in MOPC-321 only, an 8.1-kb C gene band. For MOPC-167, *Hin*dIII was used, producing V (7.3 kb) and C (4.9 kb) gene bands similar in size to those in the other cell types. [b] At this digestion stage, only a V gene band was visible. [c] NS, not scannable. [d] In liver and MOPC-167, the $C_\kappa/V_{\kappa\text{-}167}$ ratio remains approximately the same at first; in the last digestion stage where both V and C are visible, it rises slightly, probably due to the smaller size of the C band. In MOPC-321 and MOPC-21, the $C_\kappa/V_{\kappa\text{-}167}$ ratio falls continuously until only $V_{\kappa\text{-}167}$ is left, although this gene is present in a larger band than the C_κ genes of MOPC-21 and the C_κ genes in the smallest C_κ fragment of MOPC-321.

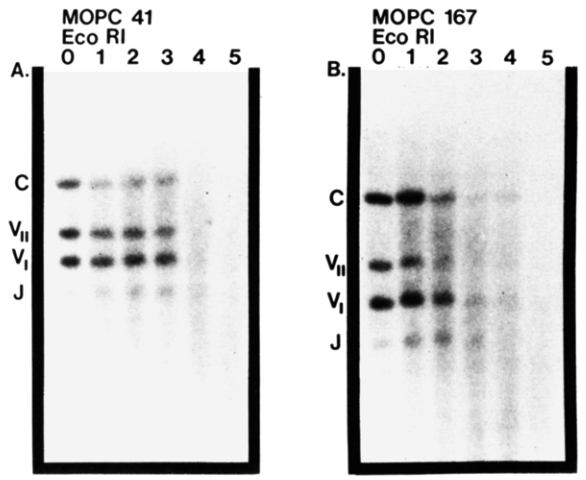

FIGURE 2: Southern blot analysis of DNA from DNase I digested and control nuclei. (A) MOPC-41: stage O, DNA from undigested nuclei; stages 1–5, DNA from nuclei digested with 0.25, 0.5, 1, 2, and 4 μg/mL DNase I. (B) MOPC-167: stage 0, undigested nuclei; stages 1–5, nuclei digested with 1, 2, 3, 4, and 5 μg of DNase I/mL. The hybridization probe was cloned DNA $p\lambda_{I\text{-}1}$ containing V_{λ_I} and C_{λ_I}. C is the C_{λ_I} gene (8.6 kb). V_{II} is the $V_{\lambda_{II}}$ (4.8 kb); this sequence cross-hybridizes with the V_{λ_I} probe but gives a weaker signal because of only partial homology (Brack et al., 1978). V_I is the V_{λ_I} gene (3.5 kb). The band marked "J" corresponds to a fragment containing J_{λ_I} sequences (Sakano et al., 1979).

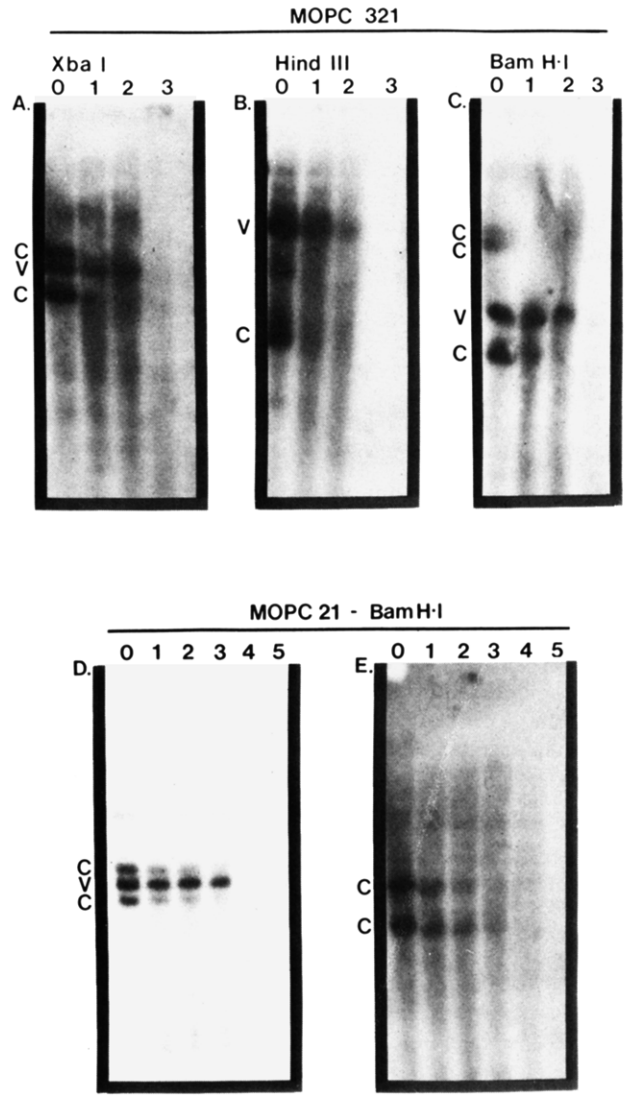

FIGURE 3: Southern blot analysis of DNA from DNase I digested and control nuclei. Hybridization probes: clone p167kRI containing V + C sequences (A–D); isolated C region from this clone (E). (A–C) MOPC-321: stage 0, undigested nuclei; stages 1–3, nuclei digested with 0.5, 1.0, and 2.0 μg/mL DNase I. *Xba*I: C is the rearranged C_κ genes (8.1 and 5.6 kb), and V is the germline $V_{\kappa\text{-}167}$ gene (7.1 kb). *Hin*dIII: V is the germline $V_{\kappa\text{-}167}$ gene (9.8 kb), and C is the rearranged C_κ genes (4.8 and 4.5 kb). Ba*m*HI: C is the "germline" C_κ gene (12.9 kb) and rearranged C_κ genes (11.4 and 4.6 kb), and V is the germline $V_{\kappa\text{-}167}$ gene (6.5 kb). (D and E) MOPC-21: stage 0, undigested nuclei; stages, 1–5, nuclei digested with 0.5, 1.0, 2.0, 4.0, and 8.0 μg/mL DNase I. C is the rearranged C_κ genes (6.9 and 5.2 kb); V is the germline $V_{\kappa\text{-}167}$ gene (6.5 kb).

and the λ genes, all of which are in the germline configuration. However, the almost 2-fold size difference between the $V_{\kappa\text{-}167}$ and C_κ fragments may have led to preferential destruction of the larger fragments due to a greater target size for the enzyme. We have therefore analyzed MOPC-321, another myeloma for which certain restriction enzymes produce V and C fragments of intermediate and very similar sizes.

Analysis of MOPC-321 DNA confirmed that the $V_{\kappa\text{-}167}$ germline gene is less sensitive to DNase I digestion than the rearranged C genes independent of fragment size (Figure 3A–C). MOPC-321 contains a germline C_κ gene (Lenhard-Schuller et al., 1978) which is seen on *Bam*HI blots as a 12.9-kb fragment (Wilson et al., 1979). This gene seems to be DNase I sensitive (Figure 3C). The situation is complicated, though, by the faintness of the fragment. We have therefore analyzed a B cell line which contains an equal complement of germline and rearranged C_κ genes (see below).

The difference between germline $V_{\kappa\text{-}167}$ and rearranged C_κ genes was further supported in the myeloma S178. S178 is a λ-producing tumor, which like other λ tumors (Stavnezer et al., 1974; Ono et al., 1977; Rabbits et al., 1977; Alt et al., 1979) also synthesizes moderate amounts of κ RNA (Storb et al., 1977). We have previously shown that this tumor has several rearranged but not germline C_κ genes (Wilson et al., 1979). The DNase I digestion of this tumor again showed that the $V_{\kappa\text{-}167}$ gene which is in germline configuration in this tumor was insensitive to DNase I and that the C_κ genes were sensitive (data not shown).

The above results are not due to inherent differences between V and C genes in general. We compared the DNase

I sensitivity of the expressed V_κ and C_κ genes in myeloma MOPC-167. We also compared the sensitivity of these expressed κ genes with unexpressed λ genes in order to further verify whether the former are relatively more DNase I sensitive. MOPC-167 produces κ chains with the V_{167} variable region and appears to be haploid or homozygous for the V_{167} and C_κ sequences (Wilson et al., 1979). It was found that $V_{\kappa\text{-}167}$ and C_κ genes were equally susceptible to mild DNase I digestion in this tumor (Figure 1C,D and Table I). This degree of susceptibility does indicate that κ genes are in the active configuration since when the same DNA samples are analyzed by a λ gene probe the λ genes prove to be relatively insensitive. The κ genes are almost absent in digestion stage 2 (Figure 1C,D) whereas the λ genes are still clearly present at this stage (Figure 2B). ($V_{\lambda\text{-}2}$ genes may be DNase sensitve, though this has not bee further investigated; the case of $V_{\lambda\text{-}2}$ is complicated by the fact that the probe is only partially homologous $V_{\lambda\text{-}1}$.)

It appeared in the results with MOPC-321 (Figure 3A–C) that there may exist a differential digestibility of the different C_κ genes of MOPC-321. However, the results may have been complicated by differences in fragment sizes. We have further analyzed the question of differential DNase I sensitivity of different rearranged C_κ genes in the myeloma MOPC-21. Here we know that the smallest, i.e., 5.2-kb C_κ fragment produced with *Bam*HI codes for the MOPC-21 κ chains and that the other rearranged C_κ gene in the 6.9-kb *Bam*HI fragment is not expressed on the protein level (Wilson et al., 1980; Storb et al., 1980) (note that the fragment sizes have been slightly revised due to the use of radiolabeled internal standards in the Southern blots). We found that again in this tumor the germline $V_{\kappa\text{-}167}$ gene was less sensitive to DNase I than the two rearranged C_κ genes (Figure 3D and Table I). For elimination of the problems with overlap by the $V_{\kappa\text{-}167}$ gene which migrates in a 6.5-kb fragment between the two C_κ genes, we also analyzed several blots with a C_κ probe alone. The conclusion from these data was that the DNase I sensitivity of the two rearranged C_κ genes does not differ significantly (Figure 3E and Table II).

The greater DNase I sensitivity of C_κ genes in myelomas compared with C_κ genes in liver is also apparent when the digestibility of this sequence is compared in myelomas and liver relative to the general DNA fragment sizes. The C_κ genes are only visible in MOPC-21 up to digestion stage 3 (Figure 3D) where the DNA is essentially intact (Figure 5B, part 3). On the other hand, although liver DNA was already significantly degraded at stage 2 (Figure 5A, part 2), C_κ genes in liver were still clearly visible at stage 3 after digestion with twice as much DNase (Figure 1A, part 3).

The question of DNase I sensitivity of the germline C_κ gene in an immune B cell was analyzed in the lymphoma WEH1 279. It was found that both the rearranged and the germline C_κ genes were DNase sensitive compared with the germline $V_{\kappa\text{-}167}$ gene (Figure 4 and Table III). At digestion stage 5 where the C_κ genes are completely eliminated, the DNA is only slightly reduced in size. It is of the same size (not shown) as the DNase I treated liver DNA at stage 3 in Figure 5A. At this digestion stage, the germline C_κ gene of liver is still hybridizable (Figure 1A and Table I). Furthermore, in the myeloma PC3741, the germline C_κ gene is also DNase sensitive (not shown). Therefore, the germline C_κ gene is in an altered chromatin state in cells committed to κ gene expression.

Discussion

We presume in the following discussion that increased DNase I sensitivity of a gene indicates that it has a chromatin

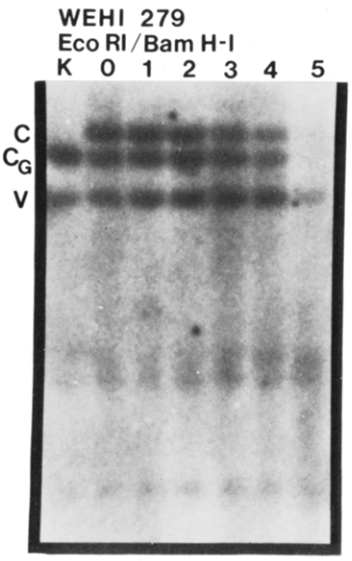

FIGURE 4: Southern blot analysis of DNA from DNase I digested and control nuclei of WEH1 279. Hybridization probe: clone p167kRI containing V and C sequences. K, kidney DNA. stage 0, undigested WEH1 279 nuclei; stages 1–5 WEH1 279 nuclei digested with 0.3, 0.6, 1, 1.5, and 3 μg/mL DNase I. C is the rearranged C_κ gene 8.2 kb. Cg is the germline C_κ gene 7.1 kb. V is the germline $V_{\kappa\text{-}167}$ gene 6.5 kb.

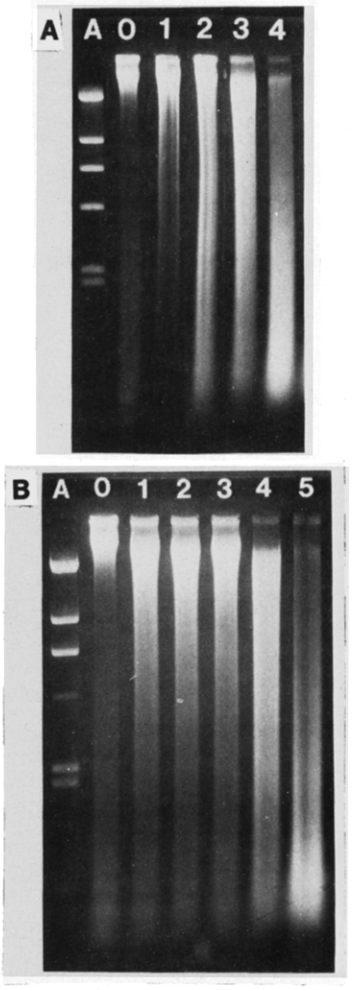

FIGURE 5: Electrophoresis in 0.6% agarose gel of DNA from liver (A) and MOPC-21 (B) nuclei digested with DNase I. Pattern of ethidium bromide staining. The DNAs are the same as shown in Figures 1A and 3D,E but were not treated with restriction enzymes. Slot A, *Hin*dIII cleaved phage λ DNA; fragment sizes are from top to bottom: 23.7 kb, 9.5 kb, 6.7 kb, 4.3 kb, 2.3 kb, 2.0 kb, and 0.59 kb.

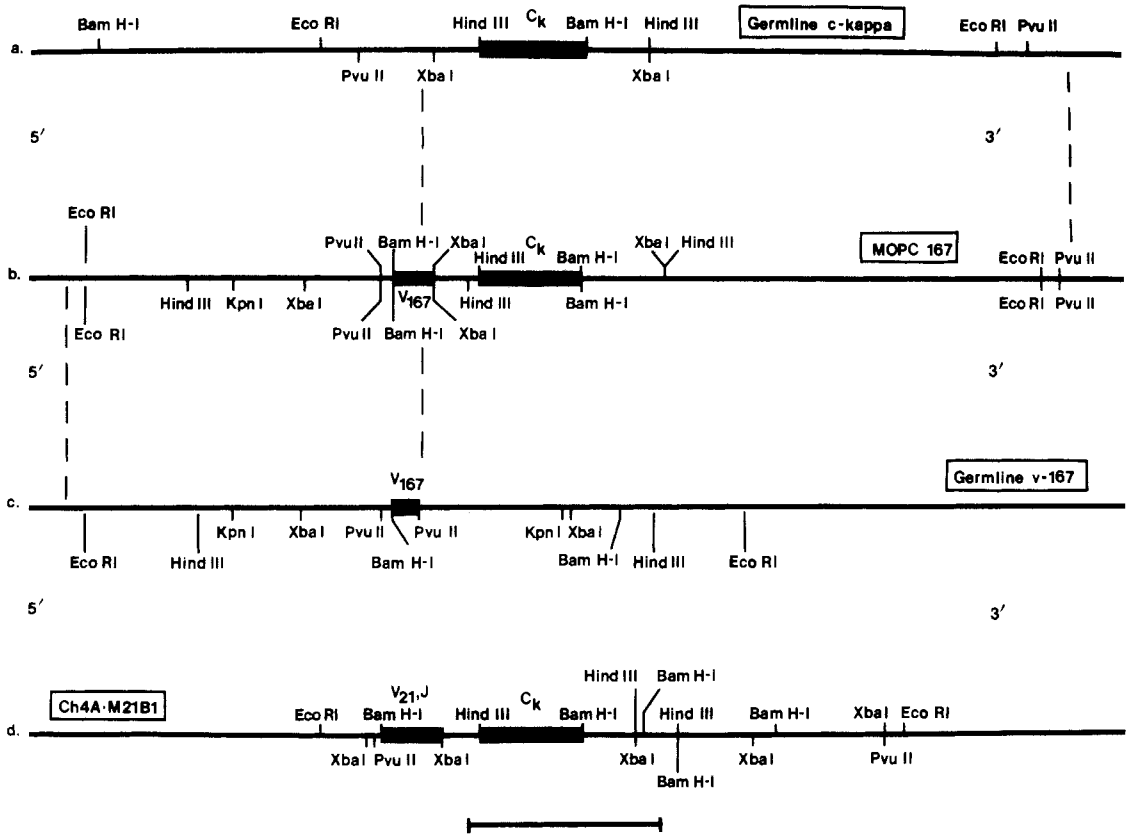

FIGURE 6: Comparison of germline MOPC-167 and MOPC-21 restriction maps. Maps a–c are from Wilson et al. (1979), slightly modified. (a) Germline C_κ map as obtained from liver and Krebs ascites cells. (b) Rearranged $V_{\kappa\text{-}167}$, C_κ map of MOPC-167. Note that the 5′ side is essentially identical with the restriction map flanking the V_{167} germline gene at one end (left side of part c) and that the 3′ side is essentially identical with the 3′ side of the germline C_κ map up to the *Xba*I site 5′ of the C_κ fragment (a). The figure shows slight discrepancies in restriction sites of comparable regions of DNA (e.g., compare the *Eco*RI and *Pvu*II sites at the 3′ ends of maps a and b). These maps present actual data; we assume that the small discrepancies are due to experimental error in molecular weight measurements. In map c, the *Eco*RI and *Hin*dIII sites cannot be positioned relative to the internal sites or to each other because fragments produced by these enzymes do not overlap with other fragments. (c) Germline $V_{\kappa\text{-}167}$ map as obtained from liver, Krebs ascites cells, and the myelomas MOPC-41, MOPC-321, S178, and MOPC-21. (d) Restriction map of a Charon 4A clone containing the rearranged $V_{\kappa\text{-}21}$ and C_κ genes of MOPC-21 (A. Walfield, U. Storb, E. Selsing, and P. Zentgraf, unpublished results). The MOPC-21 sequence is inserted at the *Eco*RI sites flanking the short and long arms of Charon 4A (Blattner et al., 1978). The *Bam*HI *Xba*I fragment indicated to contain V_{21},J reacts with a $V_{\kappa\text{-}21}$ probe exclusive of J. The cloned sequence is homologous to a germline C_κ clone from the 3′ end up to about 200 base pairs 3′ of the *Bam* site in the V_{21},J *Bam–Xba* fragment as determined by heteroduplexing, thus indicating that V_{21} is inserted next to the J_2 sequence as previously described by Sakano et al. (1979). The *Bam* fragment comprising both V_{21} and C_κ is 5.2-kb long and corresponds to the 5.2-kb expressed C_κ gene seen on Southern blots of MOPC-21 genomic DNA (Figure 3D). V (C) is the smallest restriction fragment containing a $V_\kappa(C_\kappa)$ region. Bar = 5 kb.

Table II: Relative Quantities of the Two C_K Genes in MOPC-21 after DNase I Digestion[a]

digestion stage	relative amount of 5.2-kb C gene[b]		
	expt I	expt II	expt III
0	0.54	0.52	0.56
1	0.60	0.59	0.53
2	0.67	0.58	0.45
3	0.69	0.55	0.45
4	no C	no C	no C

[a] For conditions of DNase I digestion, see MOPC-21, Figure 5. [b] The numbers represent the ratios of $C_{5.2}/C_{total}$ band intensities as determined by microdensitometry of three different Southern blot autoradiograms. The ratio between the two C_K genes remains approximately the same until both are completely digested.

conformation which differs from that of the bulk of the DNA. The studies reported here show that the vicinity of C_κ genes is DNase I sensitive, i.e., in a potentially "active" chromatin structure regardless of expression on the protein level. This conclusion is supported by the following data obtained by Southern blot analysis: (1) C_κ genes are more sensitive to DNase I than germline V_{167} genes in myeloma tumors and a B lymphoma; (2) C_κ genes are more sensitive than germline V_λ and C_λ genes in myelomas which do not produce λ mRNA, while the λ genes have the same DNase I insensitivity as germline $V_{\kappa\text{-}167}$ genes; (3) C_κ genes are more DNase I sensitive in myelomas and a B lymphoma than are germline C_κ genes in liver, on the basis of the comparison of the digestion stage at which the C_κ gene carrying fragments disappear from the Southern blots; (4) rearranged $V_{\kappa\text{-}167}$ genes are as DNase I sensitive as rearranged C_κ genes in myeloma MOPC-167, the myeloma which expresses $V_{\kappa\text{-}167}$.

No difference in DNase I sensitivity could be detected between the two rearranged C_κ genes of MOPC-21. In this myeloma, the C_κ gene contained in a 5.2-kb *Bam*HI fragment is the only C_κ gene which codes for κ chains (Wilson et al., 1980; Storb et al., 1980). In several tissue culture sublines, each of which had lost one of the rearranged C_κ genes, it is found that production of the MOPC-21 κ chains was correlated with the presence of this "5.2-kb" gene. Loss of this "expressed" gene eliminated the synthesis of the κ chains; loss of the other rearranged gene did not. The "5.2-kb" gene contains the $V_{\kappa\text{-}21}$ sequence (Figure 6d). Since both C_κ genes in MOPC-21 are equally sensitive to DNase I, the rearranged unexpressed gene must possess the signal for a change in chromatin conformation. This signal must also be present in the germline C_κ gene of committed B cells, as the germline

Table III: Relative Quantities of the C_K Genes of WEH1 279 after DNase I Digestion[a]

digestion stage	relative amount of C_K genes	
	germline[b]	rearranged[c]
0	0.61	0.60
1	0.56	0.57
2	0.53	0.52
3	0.52	0.50
4	0.44	0.39
5	0[d]	0[d]

[a] For conditions of DNase I digestion see Figure 4. The numbers are the ratios of germline C_K gene/(germline C_K plus $V_{K\text{-}167}$ genes)[b] and rearranged C_K/(rearranged C_K + $V_{K\text{-}167}$ genes).[c] [d] At this digestion stage, only the V_{167} gene was visible. Relative to the germline $V_{K\text{-}167}$ gene, the two C_K genes decrease at approximately the same rate with DNase I digestion.

C_κ gene in WEH1 279 and MOPC-3741 cells was DNase I sensitive. It is possible that alteration of the chromatin state precedes rearrangement and may be a prerequisite for rearrangement. The exact location of the DNase I sensitive areas with respect to the C_κ gene is currently being studied in our laboratory in germline and different rearranged C_κ genes. There is some indication that DNase I hypersensitive hotspots may exist (see the smaller digestion products in Figure 4), but this question has not been dealt with systematically yet. The chromatin alteration must be relatively localized since the $V_{K\text{-}167}$ gene, which is in germline context in the myelomas MOPC-21, MOPC-41, MOPC-321, and S178 and in the B lymphoma WEH1 279 (see Figure 6 for map information), is DNase I insensitive in these cells although it is most probably present in the same κ gene cluster on the same chromosome as the rearranged C_κ gene (Hood et al., 1973).

What prevents the synthesis of κ chain proteins from the "nonfunctional" rearranged and the germline C_κ genes? The block may exist at the level of transcription (lack of a promoter, etc.), processing of precursor RNA (lack of proper RNA splicing signals, etc.), translation (wrong reading frame because of misalignment during gene rearrangement or lack of a ribosome binding site in functional relationship to an initiation codon of the mRNA) or posttranslational stability (structural instability of the polypeptide chain, etc.). In the case of germline C_κ genes of myelomas, Perry et al. (1980) have reported that the genes can be transcribed into a large precursor RNA. In the case of the "nonfunctional" C_κ gene of MOPC-21, we have found that it is transcribed, and the RNA is apparently processed and transported into the cytoplasm but not translated into stable κ chains (Walfield et al., 1980). Presumably there will be other examples found which may correspond to other possible blocks.

Obviously there must have been a strong selective pressure to develop a system which ensures allelic exclusion; otherwise the simultaneous synthesis of two different antibodies by a given cell would lead to scrambled molecules of lower affinity. Thus, in a lymphoid cell, often only one of the κ alleles is rearranged (Wilson et al., 1980; Joho & Weissman, 1980; E. Selsing and U. Storb, unpublished results; and Figure 4). In many myeloma cells and probably in some normal plasma cells (Wilson et al., 1980), more than one κ gene is rearranged, although generally these cells produce κ chains coded for by only one gene (Potter, 1967; and see Nisonoff et al., 1975). We hypothesize that "allelic exclusion" of immunoglobulin κ genes generally oprates by the functional rearrangement of one κ gene; other rearrangements are relatively frequent, but they are generally nonfunctional, thus preventing the production of scrambled antibodies. It remains to be determined whether a specific mechanism exists to create functional in contrast to nonfunctional rearrangements in lymphocyte precursors.

Added in Proof

The defect in the nonfunctional gene of MOPC-21 has now been found to be a misalignment of V,J which leads to a shift in translational reading frame (A. Walfield, E. Selsing, B. Arp, and U. Storb, unpublished experiments).

Acknowledgments

We thank P. Early and A. Bothwell for cDNA clones, F. Blattner for Charon 4 A phage and information about its use, M. Potter and M. Cohn for mouse tumors, C. Sibley for WEH1 279 (which she originally received from N. Warner), J. Beard and the National Cancer Institute for reverse transcriptase, and P. Bornstein and E. Click for the use of a scanner. We are also indebted to H. Weintraub for many suggestions concerning this work, to B. Arp for excellent technical help, and to T. E. Martin for comments on the manuscript.

References

Alt, F. W., Enea, V., Bothwell, A. L. M., & Baltimore, D. (1979) *ICN-UCLA Symp. Mol. Cell. Biol. 14*, 407–419.

Blattner, F. R., Blechl, A. E., Denniston-Thompson, K., Faber, H. E., Richards, J. E., Slightom, J. L., Tucker, P. W., & Smithies, O. (1978) *Science (Washington, D.C.) 202*, 1279–1284.

Brack, C., Hirama, M., Lenhard-Schuller, R., & Tonegawa, S. (1978) *Cell 15*, 1–14.

Dreyer, W. J., & Bennett, J. C. (1965) *Proc. Natl. Acad. Sci. U.S.A. 54*, 864–868.

Early, P. W., Davies, M. M., Kaback, D. B., Davidson, N., & Hood, L. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 857–861.

Flint, S. J., & Weintraub, H. M. (1977) *Cell 12*, 783–794.

Garel, A., & Axel, R. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 3966–3970.

Groudine, M., Das, S., Neiman, P., & Weintraub, H. (1978) *Cell 14*, 865–878.

Hood, L., McKean, D., Farnsworth, V., & Potter, M. (1973) *Biochemistry 12*, 741–749.

Jeffreys, A. J., & Flavell, R. A. (1977) *Cell 12*, 429–439.

Joho, R., & Weissman, I. L. (1980) *Nature (London) 284*, 179–181.

Joho, R., Weissman, I. L., Early, P., Cole, J., & Hood, L. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 1106–1110.

Lenhard-Schuller, R., Hohn, B., Brack, C., Hirama, M., & Tonegawa, S. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 4709–4713.

Miller, D. M., Turner P., Nienhuis, A. W., Axelrod, D. E., & Gopalakrishnan, T. V. (1978) *Cell 14*, 511–521.

Nisonoff, A., Hopper, J. E., & Spring, S. B. (1975) in *The Antibody Molecule*, Academic Press, New York.

Ono, M., Kawakami, M., Kataoka, T., & Honjo, T. (1977) *Biochem. Biophys. Res. Commun. 74*, 796–802.

Palmiter, R., McKnight, S., Mulvihill, E., & Senear, A. (1977) *Cold Spring Harbor Symp. Quant. Biol. 42*, 773–778.

Pernis, B., Chiappino, G., Kelus, A. S., & Gell, P. G. H. (1965) *J. Exp. Med. 122*, 853–876.

Perry, R. P., Kelley, D. E., Coleclough, C., Seidman, J. G., Leder, P., Tonegawa, S., Matthyssens, G., & Weigert, M. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 1937–1941.

Potter, M. (1967) *Methods Cancer Res. 2*, 105–157.

Rabbits, T. H., Forster, A., Smith, M., & Gellam, S. (1977) *Eur. J. Immunol. 7*, 43–48.
Rabbits, T. H., Forster, A., Dunnick, W., & Bently, D. L. (1980) *Nature (London) 283*, 351–356.
Rigby, P. W. J., Dieckmann, M., Rhodes, C., & Berg, P. (1977) *J. Mol. Biol. 113*, 237–251.
Sakano, H., Hüppi, K., Heinrich, G., & Tonegawa, S. (1979) *Nature (London) 280*, 288–294.
Seidman, L. G., & Leder, P. (1978) *Nature (London) 276*, 790–795.
Shepherd, J., Mulvihill, E., & Palmiter, R. (1977) *J. Cell Biol. 75*, 353a.
Southern, E. M. (1975) *J. Mol. Biol. 98*, 503–517.
Stalder, J., Groudine, M., Dogson, J. B., Engel, J. D., & Weintraub, H. (1980) *Cell 19*, 973–980.
Stavnezer, J., Huang, R. C., Stavnezer, E., & Bishop, J. M. (1974) *J. Mol. Biol. 88*, 43–63.
Steinmetz, M., & Zachau, H. G. (1980) *Nucleic Acids Res. 8*, 1693–1706.
Storb, U., Hager, L., Wilson, R., & Putnam, D. (1977) *Biochemistry 16*, 5432–5438.
Storb, U., Arp, B., & Wilson, R. (1980) *Nucleic Acids Res. 8*, 4681–4687.
Walfield, A., Storb, U., Selsing, E., & Zentgraft, H. (1980) *Nucleic Acids Res. 8*, 4689–4707.
Warner, N., Harris, A., & Gutman, G. (1975) in *Membrane Receptors of Lymphocytes* (Seligman, M., Prud'homme, J. L., & Kourilsky, I. M., Eds.) pp 203–220, Elsevier, New York.
Weintraub, H., & Groudine, M. (1976) *Science (Washington, D.C.) 193*, 848–856.
Wilson, R., Miller, J., & Storb, U. (1979) *Biochemistry 18*, 5013–5021.
Wilson, R., Storb, U., & Arp, B. (1980) *J. Immunol. 124*, 2071–2076.
Wu, C., Bingham, P. M., Livak, K. J., Holmgren, R., & Elgin, S. C. R. (1979) *Cell 16*, 797–806.

Sequence Determination and Analysis of the 3′ Region of Chicken Pro-α1(I) and Pro-α2(I) Collagen Messenger Ribonucleic Acids Including the Carboxy-Terminal Propeptide Sequences†

Forrest Fuller‡ and Helga Boedtker*

ABSTRACT: Three pro-α1 collagen cDNA clones, pCg1, pCg26, and pCg54, and two pro-α2 collagen cDNA clones, pCg13 and pCg45, were subjected to extensive DNA sequence determination. The combined sequences specified the amino acid sequences for chicken pro-α1 and pro-α2 type I collagens starting at residue 814 in the collagen triple-helical region and continuing to the procollagen C-termini as determined by the first in-phase termination codon. Thus, the sequences of 272 pro-α1 C-terminal, 260 pro-α2 C-terminal, 201 pro-α1 helical, and 201 pro-α2 helical amino acids were established. In addition, the sequences of several hundred nucleotides corresponding to noncoding regions of both procollagen mRNAs were determined. In total, 1589 pro-α1 base pairs and 1691 pro-α2 base pairs were sequenced, corresponding to approximately one-third of the total length of each mRNA. Both procollagen mRNA sequences have a high G+C content. The pro-α1 mRNA is 75% G+C in the helical coding region sequenced and 61% G+C in the C-terminal coding region while the pro-α2 mRNA is 60% and 48% G+C, respectively, in these regions. The dinucleotide sequence pCG occurs at a higher frequence in both sequences than is normally found in vertebrate DNAs and is approximately 5 times more frequent in the pro-α1 sequence than in the pro-α2 sequence. Nucleotide homology in the helical coding regions is very limited given that these sequences code for the repeating Gly-X-Y tripeptide in a region where X and Y residues are 50% conserved. These differences are clearly reflected in the preferred codon usages of the two mRNAs.

Collagen is a fibrillar, structural protein responsible for the physical integrity of organs, tissues, and gross skeletal structures of vertebrates and of at least some invertebrates as well (Adams, 1978). The structures and functions of collagens have been recently reviewed (Fessler & Fessler, 1978; Prockop, et al., 1979a; Bornstein & Byers, 1980; Eyre, 1980; Bornstein & Sage, 1980; Olsen & Berg, 1979). Analysis of polypeptide chains implies that higher animals produce at least five different collagens from at least seven genes. Various cells synthesize different collagens and modify them to produce structural matrices specific to the cell type. Consequently, collagen expression is carefully regulated during differentiation and embryogenesis (von der Mark et al., 1976; Linsenmayer & Toole, 1977; Bornstein & Sage, 1980). Failure to correctly express or modify specific collagens has been implicated as the causative factor in several diseases of man and other animals (Prockop et al., 1979b; Bornstein & Byers, 1980). In order to understand how the cell regulates expression of these different collagens, it is necessary to determine the levels and rates of synthesis of collagen mRNA[1] and pre-mRNAs at various stages during development and in both normal and abnormal cells. It is also important to define the organization of the collagen sequences in the genome to ascertain how this

† From the Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138. *Received June 27, 1980.* Supported by Grants from the National Institutes of Health (HD-01229) and the Muscular Dystrophy Association, Inc.

* To whom reprint requests should be addressed.
‡ To whom requests for sequence data and computer programs should be addressed.

[1] Abbreviations used: cDNA, complementary DNA; mRNA, messenger RNA; poly(A), poly(adenylic acid); mC, 5-methyldeoxycytidine; I, inosine; helical, a peptide (or nucleotides coding for peptide) characterized by the repeating sequence $(Gly\text{-}X\text{-}Y)_n$; C-terminal, amino acids (or nucleotides coding for them) which occur in the carboxyl ends of procollagens and do not contain the repeating sequence, $(Gly\text{-}X\text{-}Y)_n$.

0006-2960/81/0420-0996$01.00/0