Title
Tenascin-X: a novel extracellular matrix protein encoded by the human XB gene overlapping P450c21B.

Permalink
https://escholarship.org/uc/item/2ss0j25k

Journal
The Journal of cell biology, 122(1)

ISSN
0021-9525

Authors
Bristow, J
Tee, MK
Gitelman, SE
et al.

Publication Date
1993-07-01

DOI
10.1083/jcb.122.1.265

Peer reviewed
Tenascin-X: A Novel Extracellular Matrix Protein Encoded by the Human XB Gene Overlapping P450c21B

James Bristow,* Meng Kian Tee,* Stephen E. Gitelman,* Synthia H. Mellon,§ and Walter L. Miller*§

* Departments of Pediatrics, and of Obstetrics, Gynecology and Reproductive Sciences, and the § Metabolic Research Unit
University of California, San Francisco, California

Abstract. A human gene termed XB overlaps the P450c21B gene encoding steroid 21-hydroxylase and encodes a protein that closely resembles extracellular matrix proteins. Sequencing of phage and cosmid clones and of cDNA fragments shows that the XB gene spans 65 kb of DNA, consisting of 39 exons that encode a 12-kb mRNA. The predicted protein of over 400 kD consists of five distinct domains: a signal peptide, a hydrophobic domain containing three heptad repeats, a series of 18.5 EGF-like repeats, 29 fibronectin type III repeats, and a carboxy-terminal fibrinogen-like domain. Because the structure of the protein encoded by the XB gene closely resembles tenascin, we term this protein tenascin-X (TN-X), and propose a simplified nomenclature system for the family of tenascins. RNase protection experiments show that the TN-X transcript is expressed ubiquitously in human fetal tissues, with the greatest expression in the fetal testis and in fetal skeletal, cardiac, and smooth muscle. Two adrenal-specific transcripts, P450c21B (steroid 21-hydroxylase) and Y (an untranslated transcript) overlap the XB gene on the complementary strand of DNA, yielding a unique array of overlapping transcripts: a "polygene." In situ hybridization histochemistry experiments show that the TN-X transcript and the P450c21 and Y transcripts encoded on the complementary DNA strand are all expressed in the same cells of the human adrenal cortex. Genetic data suggest that TN-X may be essential for life.

Tenascin, a large glycoprotein of the extracellular matrix, has been studied extensively for about 10 years (for review see reference 11). Each of the six monomers that constitute tenasin exists in multiple size variants of 220–320 kD and contains four domains. The amino terminus comprises the hydrophobic "head group," which facilitates polymerization into the tenasin hexabrachion; this is followed in turn by a series of EGF-like repeats, a series of fibronectin type III (Fn III) repeats, and a carboxy-terminal fibrinogen-like domain (11, 16, 22, 41, 52, 58). Tenasin appears to mediate interactions between cells and the extracellular matrix (and possibly between cells) (7, 47), through an RGD-dependent (3, 29) or -independent (46, 52) receptor. Initial studies suggested a crucial role in embryonic development (6, 53, 58), although recent tenasin gene knockout experiments in transgenic mice suggest that tenasin serves no mandatory, irreplaceable role in development or in the adult (48). Tenasin appears to be one member of a family of related proteins, encoded by related genes. Each family member has the same general structure as tenasin: NH₂-head group-EGF domain-Fn III domain-fibrinogen domain-COOH. In addition to tenasin, a closely related protein termed restrictin (42) or J1-160/180 (13) has a similar structure. The structure of undulin, which contains a von Willebrand factor domain at the NH₂ terminus followed by Fn III repeats (23), indicates that there are variations on this theme.

We now describe the gene structure and tissue distribution of another member of this family termed tenasin X (TN-X), which is encoded by an unusual gene termed XB in the human major histocompatibility locus. This gene was discovered because it overlaps the gene encoding steroid 21-hydroxylase (39), and was later shown to encode a tenasin-like sequence (14, 60; Bristow, J., S. E. Gitelman, Y. Shi, and W. L. Miller, 1990. Pediatr. Res. 27:76a). Recent preliminary reports from another group have confirmed this (30) and extended the findings to show that this gene has the expected EGF-like domain (31). In addition to overlapping the P450c21B gene (formally designated CYP21B) (40) encoding steroid 21-hydroxylase, the human XB gene also extensively overlaps the YB gene (see Fig. 1), which encodes an untranslated adrenal-specific RNA (4). In this paper, we delineate the organization, sequence, and tissue distribution of expression of the XB gene. Furthermore, we propose a unified nomenclature system for these various tenascins and assign the name TN-X to the product of this new gene.
Materials and Methods

Genomic and cDNA Library Screening
Genomic libraries prepared from human leukocyte DNA in λ-DASH, and λ-Chang-4A were plated at a density of 3 × 10^6 pfu/plate. Duplicate lifts onto nitrocellulose filters were made and cross-linked by baking. A cosmid library in pWEIS (Stratagene, San Diego, CA) and a cDNA library in λ Zap 4 (λ) were plated at 5 × 10^5 colonies/plate, and duplicate replica filters were made and processed by standard procedures. All filters were probed in 50 mM Tris (pH 7.5)/50% formamide/5× SSC/Tı Denhardt's solution/0.1% NaDodSO4, at 42°C, with probes labeled to a specific activity of >5 × 10^6 cpm/μg using random primers (Pharmacia Fine Chemicals, Piscataway, NJ). Washing was carried out in 0.1% SSC/0.1% NaDodSO4 for 10 min at room temperature and for 30 min at 65°C. Positive plaques were picked and purified further by two additional rounds of plating and probing. Bacteriophage DNA was prepared by the method of Helms et al. (17). Cosmid DNA purified further by two additional rounds of plating and probing. Bacterio-

DNA Sequencing
Genomic and cDNA fragments for sequencing were subcloned into pBlue
script KS or SK vectors (Stratagene) and purified over a CsCl/ethidium bro-

Rapid Amplification of cDNA Ends
Rapid amplification of cDNA ends (RACE) was performed as described (12) using a reagent kit (Bethesda Research Laboratories); cDNA synthesis from 2 μg of total muscle RNA was primed with a gene-specific antisense primer and reverse transcribed at 42°C for 1 h. After removal of the initial primer and RNA, the cdNA was taild with dCTP and terminal transferase. Subsequent PCR used nested antisense primers and a 5’ adapter primer. PCR products were subcloned into pBlue
script (Stratagene) or pCRII (Invitrogen, San Diego, CA) for restriction mapping and DNA sequencing.

RNase Protection
RNase protection experiments were done essentially as described (14). For RNase protection of XB gene transcripts, relevant genomic or cdNA frag-
ments were subcloned into pBlueScript and linearized. An antisense probe was synthesized with 32P-UTP using bacteriophage T7 RNA polymerase. After purification, the probe was hybridized to 15 μg of total RNA in 80% formamide/50 mM PIPES (pH 6.4)/0.4M NaCl/5 mM EDTA overnight at 42°C. Single-stranded RNA was digested with 5–8 μg of DNase-free RNase A (Boehringer-Manheim, Indianapolis, IN) and 50 U RNase T1 (Pharma-
cia, Piscataway, NJ) at 37°C for 1 h, and the protected fragments were sepa-
rated by electrophoresis on a 7% polyacrylamide gel containing 7 M urea.

In Situ Hybridization Histochemistry
Probes for in situ hybridizations were synthesized from linearized pBlue
script plasmids containing cDNAs for human P450c21 (36), TN-X, and Gene Y (4). Antisense riboprobes labeled with 35S-labeled UTP were synthesized with T3 or T7 RNA polymerase as described for the RNase protection probes. Human tissues (20–23-wk gestation) were obtained through approved protocols of the UCSF Department of Obstetrics, Gynecology and Repro-
ductive Sciences at the time of elective cervical dilatation and evacuation. Tissues were transported in medium 199 on ice, dried to remove excess moisture, and frozen in Tissue Tek OCT Compound (Miles, Inc., Elk-
hurst, IN) in a methanol/dry-ice bath. 10-μm sections were cut with a cryostat, collected onto microscope slides, and coated with a gelatin chrom alun mixture (0.4%/0.04% wt/vol). Frozen sections were fixed for 5 min in 4% paraformaldehyde in PBS, pH 7.4, and dehydrated in 70 and 100% ethanol. Dried sections were stored with desiccant at −70°C, then treated with protease K (2.5 μg/ml in 2× SSC) at 37°C for 15 min, rinsed in 2× SSC, and acetylated for 30 min as described (34). Probes were boiled for 5 min and applied to tissue sections at 100 μg/ml in prehybridization solu-
tion 50% formamide, 4× SSC, 1× Denhardt's solution, 140 μg/ml heparin, 25 μg/ml denatured Escherichia coli DNA, 2.5 μg/ml polyadenylic acid, and 100 μM DTT, and then hybridized overnight in a humidified environ-
ment at 37°C. Slides were washed twice for 15 min at 37°C in 2× SSC con-
taining 100 mM 2-mercaptoethanol, twice for 15 min at 37°C in 0.5× SSC/100 mM 2-mercaptoethanol, and once for 15 min in 2× SSC at 37°C. Slides were then treated with 20 μg/ml RNase A in 2× SSC at 37°C for 30 min, washed one final time in 2× SSC for 15 min at room temperature, and coated with Ilford K-5 nuclear emulsion diluted 1:1 with 2× SSC. Slides were stored at 4°C, developed after 2 wk of autoradiography, and counter-
stained with cresyl violet to define tissue architecture.

Results
Arrangement of the Gene Cluster
Tenascin X is encoded by the large XB gene in the class III region of the major histocompatibility locus. This gene is part of a complex array of genes encoding the fourth compo-
ent of serum complement (C4), the form of cytochrome P450 having steroid 21-hydroxylase activity (P450c21), and a transcript of unknown function operationally termed gene Y (4, 14, 36, 39). This locus is duplicated in the human ge-
ome, resulting in the complex array of genes shown in Fig. 1. We initially discovered the X genes overlapping the 3’ ends of the P450c21 genes by cloning a 2.7-kb cdNA fragment identified with a P450c21B cdNA probe (39). The 2.7-kb cdNA is encoded by the large XB gene; the XA gene in turn was truncated by the nonhomologous recombination event that led to the duplication of the human gene cluster, so that it does not appear to encode protein (14). By contrast, the XB gene appeared to encode a 12-kb RNA (14), which in turn encoded a protein resembling the extracellular matrix protein tenasin (14, 60).

Structure of the XB Gene
To clone the XB gene, we first screened a human genomic DNA library in A-Dash using a 1.8-kb BamHI/EcoRI fragment of the 2.7-kb XB cdNA (39). This probe, which was chosen because it excluded the regions overlapping P450c21, identified XBX-1 as an XB genomic clone. This clone hybridized to an 850-bp PstI/EcoRI fragment from the 5’ end of the 2.7-kb X cdNA and did not hybridize to a P450c21-specific cdNA fragment or a 5’ C4A cdNA (Fig. 1). Probing of a
Figure 1. Map of the XB gene-encoding TN-X. (A) Organization of the duplicated C4/P450c21/X/Y locus on the short arm of human chromosome 6; the centromere is to the left and the telomere to the right. The upper line shows the scale in kilobases. The diagram shows the TN-X and XA genes as hatched boxes, the P450c21 genes (21A and 21B) as black boxes, and the Y genes as open boxes. The arrows indicate transcriptional orientation. In B, the TN-X gene is enlarged to show its exon/intron structure (note the scale marker at the left). The upper lines show the extent of the genomic DNA in the various lambda phage (λ) and cosmid (c) clones discussed in Results. The exons of the TN-X gene are shown as boxes, and the Roman numerals below the line identify the Fn III repeats. cDNAs are indicated as horizontal bars over the relevant exons. Exon 1 encodes the leader peptide and heptad domain; exon 2 encodes the EGF-like repeats, exons 3-34 encode the 29 Fn III repeats, and exons 35-39 encode the fibrinogen-like domain. Note that Fn III repeats ii, xxvii, xxviii, and xxix are each encoded by a pair of "split" exons, while repeats viii and ix are encoded by a single "fused" exon. A map of the BamHI and EcoRI restriction sites is shown below. (C) Diagram of the structure of a TN-X monomer, using the symbols popularized by Spring et al. (52). The five "balloons" on the heptad, on Fn III repeat xxviii, and on the fibrinogen-like domain (FG) designate potential N-linked glycosylation sites.

Southern blot of multiple restriction enzyme digests of λXB-1 with the 1.8-kb cDNA probe revealed a 7-kb BamHI fragment and other RFLP markers predicted to be in the DNA 3' of the P450c21B gene (45). DNA sequencing could be initiated by a 20-mer corresponding to the 5' end of the 2.7-kb cDNA, further confirming that phage λXB-1 corresponded to the XB region, and that the 2.7-kb cDNA was encoded by XB. Mapping of λXB-1 showed it contained 11 kb of human DNA. Sequencing of this DNA showed the 2.7-kb cDNA was encoded by 15 exons spanning 6.6 kb (sequence included in Fig. 2). We found only two bona fide nucleotide differences between the gene and the cDNA. One of these differences was in the third position of a codon and does not result in an amino acid change, while the other results in a substitution of isoleucine for valine (amino acid number 3260 in Fig. 2). A sequencing error (bases 25,817-25,819 in Fig. 2) was also detected. The intron/exon junctions matched the GT/AG rule perfectly, and showed no evidence that the 2.7-kb cDNA corresponded to an alternately spliced mRNA.

λXB-1 also contained the 5' portion of the exon encoding the 5' end of the 2.7-kb cDNA, thus slightly extending the open reading frame. However, Northern blots suggested that the XB transcript might be substantially longer. To elucidate the sequence of XB, we then screened two additional human adrenal cDNA libraries, a human testis cDNA library and a human placenta cDNA library with 5' fragments of the 2.7-kb cDNA; all of these yielded 3' fragments of XB cDNA, but none was longer than the original 2.7-kb clone. Therefore, we extended our knowledge of XB by sequencing the XB gene.

 Sequencing of the remainder of the 5' DNA in λXB-1 re-
Figure 2.

The polyadenylation signal of TN-X is boxed, and that of P450c21B is indicated by a dashed overline. EMBL accession number as follows: X71923 = exons 1 and 2; X71924 = intron 2; X71925 = Fm repeat i; X71926 = Fm repeat ii; X71927 = Fm repeat iii; X71928 = Fm repeat iv; X71929 = Fm repeat v; X71930 = Fm repeat vi; X71931 = Fm repeat vii; X71932 = Fm repeat viii; X71933 = Fm repeat ix; X71934 = Fm repeat x; X71935 = Fm repeat xi; X71936 = Fm repeat xii; X71937 = Fm repeat xiii; X71938 = Fm repeat xiv; X71939 = Fm repeat xv; X71940 = Fm repeat xvii; X71941 = Fm repeat xviii-xxix; fibronectin domain, 3' untranslated region.

Bristow et al. Tenascin-X
revealed two additional exons encoding 100 amino acid repeats (xxi and xxi in Fig. 1) similar to those found in the 5' end of the X cDNA. Genomic DNA lying 5' of that in λXB-1 was obtained by identifying an overlapping phage (λXB-2) from bands, suggesting that additional repeats remained to be sequenced. X-B cosmids (cTNX: 2-1, 6-2, and 1A in Fig. 1) were Complete sequencing of this DNA (Fig. 2) identified repeats revealed two additional exons encoding 100 amino acid repeats DNA was followed by subcloning of overlapping BamHI and Probing of restriction digests of these cosmids with a genomic library in XCharon 4A by hybridization to a 1.4-kb BglII fragment (bases 19,651-21,088 in Fig. 2) from the 5' end of origin of the X cDNA. Genomic DNA lying 5' of that in XXB-1 was near the amino terminus of tenascin by using a PCR-amplified probe encompassing 1.1 kb of the EGF domain of EcoRI fragments. The ends of each subclone were sequenced, and the fragments were ordered by further mapping and sequencing. Repeats i, iii, v-vii, and xii-xiv contained, or were very near BamHI or EcoRI sites, and each was repeats xvi and xvii. Repeat iv was identified by RT/PCR. Repeat ii was mapped but not sequenced (31), and has not repeats vii and xii were identified by overlapping libraries. Sequence data were obtained near the 5' end of 3-kb BamHI fragments, which showed particularly intense hybridization to the genomic probe containing repeats xvi and xvii. Repeat iv was identified by RT/PCR. Repeat ii was mapped but not sequenced (31), and has not been verified by us. Each repeat had canonical splice junction sites at its boundaries, and maintained the same open reading frame as the original 2.7-kb cDNA. We searched for EGF-like repeats similar to those found near the amino terminus of tenascin by using a PCR-amplified probe encompassing 1.1 kb of the EGF domain of

Figure 2.
human tenasin. However, low-stringency hybridization detected no such sequences in cosmids cTNX: 1A, 6-2, or 21. Based on the partial sequence of a recently reported genomic DNA fragment of TN-X (31), we synthesized PCR primers corresponding to the apparent EGF-like domain (bases 903-922) and the apparent upstream heptad domain of TN-X (bases 250-269). These primers amplified a 327-bp fragment from adrenal cDNA and a 673-bp fragment from genomic DNA. This 327-bp 5’ cDNA was then used to identify our final cosmids cTNX: 5 (Fig. 1).

To identify additional upstream coding regions, we used the RACE procedure (12) using oligonucleotide sequences within the 327-bp 5’ cDNA fragment. Complementary DNA was synthesized from adrenal or muscle RNA using an oligonucleotide primer in the EGF domain (bases 831-850). This cDNA was 3’ tailed with dC, PCR amplified with poly(dG) linked to the adaptor primer and a second oligonucleotide in the EGF repeat (bases 806-826) (the 3’ primer). A second round of PCR was performed with the adaptor primer and an oligonucleotide in the heptad repeats (bases 263-282 in Fig. 2). The final product was cloned, and colonies were screened with the 327-bp 5’ cDNA. The sequences of positively hybridizing clones extended 220 bp into and beyond a region reported as the 5’-most intron by Matsumoto et al. (31), indicating that this was actually an exonic sequence. To demonstrate that this sequence was actually part of an exon, we performed an additional RT/PCR experiment (Fig. 3). PCR amplification of fetal muscle cDNA using specific primers (bases 60-80 and 806-826 in Fig. 2) yielded a 419-bp fragment, but PCR amplification of fetal muscle RNA without reverse transcription yielded no product, demonstrating that the PCR product arose from newly transcribed cDNA and not from contaminating cDNA or genomic DNA. Amplification of genomic DNA or of cosmids cTNX: 5 yielded a 765-bp product, differing from the 419-bp RT/PCR product by retention of the 346 bp in intron 1. Inclusion of this new DNA in exon 1 was also verified in fetal adrenal RNA by RNase protection assay (data not shown).

Determination of the correct intron/exon organization provides the sequence of the amino-terminus, including two ATG translational initiation codons, immediately preceded by a TGA stop codon. The algorithm of von Heijne (55) strongly predicted that the first 22 amino acids constitute a leader peptide for the secretory pathway, consistent with the hypothesis that TN-X is a secreted protein of the extracellular matrix. We mapped and sequenced a total of 37 exons encoding 3,626 amino acids of the XB gene (Fig. 2). In addition, Fig. 1 shows Fn III repeat ii, encoded by exons 4 and 5; this was mapped by Matsumoto et al. (30), but the sequence was not reported. Sequencing of three additional Fn III repeats is incomplete, although the lengths of missing sequences can be estimated from Fig. 4.4. Thus we estimate that the complete TN-X gene contains 39 exons encoding a protein of about 3,816 amino acids. Because fairly large regions of the gene remain unsequenced, additional exons encoding Fn III repeats may remain undiscovered.

**Structural Features of the XB gene and Encoded TN-X Protein**

The organization of the 39 exons of the XB gene is very similar to that of tenasin. Exon 1 contains a small amount of 5’-untranslated sequence and encodes the signal sequence and the “head piece,” which in tenasin and restrictin is responsible for polymerization of the protein monomers into multi-armed “brachion” structures. The carboxy-terminal half of the head group is similar in TN-X and tenasin, and contains three heptad repeats flanked by five conserved cysteine residues, suggesting that TN-X is also capable of forming multimers, as described for tenasin (52). Sequences upstream from these cysteine residues are unique in TN-X and lack the additional cysteine residues found in tenasin and restrictin which may permit assembly of higher-order brachion structures. Exon 2 encodes a series of 18.5 EGF-like repeats that are 55% similar to the 13.5 EGF-like repeats of tenasin. Exons 3-34 encode a series of 29 repeats resembling Fn III repeats (aligned and compared in Fig. 4). The second repeat (exons 5 and 6) and the last three repeats (exons 29-34) are each encoded by a pair of “split exons.” Splitting of Fn III repeats into two exons is also seen in the genes for chicken (52) and human tenasin (16). The maximum similarity between tenasin and TN-X Fn III repeats is in the 3’ split exons, but is only 40%. This region encodes a cell-binding domain (tenasin repeat 14), identified by Spring et al. (52), but the most similar Fn III repeat in TN-X (xxviii) is only 34% similar. Unlike the tenasin (or restrictin) genes, re-
tron/exon structure of the 3' end of the XB gene is very similar to the carboxy-terminal fibrinogen-like domain and the middle of TN-X arose more recently. Exons 35-39 encode the Fn II repeats suggests a series of evolutionary duplication events. Each of the other 24 repeats is encoded by a single exon. The encoded amino acid sequences of the Fn II repeats are conserved in evolution (2, 60), and the sequence and intron/exon structure of the 3' end of the XB gene are encoded by a single exon. The encoded amino acid sequences of the Fn II repeats viii and ix of the XB gene are encoded by a single exon. The carboxy-terminal fibrinogen-like domain is widely conserved in evolution (43). This domain is 54% similar in human tenascin (43). To our knowledge, such a "fused exon" has not been described in other genes encoding proteins with Fn repeats. Based on the alignment in A and using the Dayhoff cost matrix, the dendrogram shows that Fn II repeats at the tip of the dendrogram are more similar to each other than they are to the rest of the repeats. To determine whether these repeats are encoded by a single exon, we performed in situ hybridization histochemistry experiments using probes for P450c21, Y, XA, and XB. All of the overlapping transcripts are all expressed in the same cells, with greatest expression in fetal testis and in fetal smooth, striated, and cardiac muscle. Overlapping RNA transcripts are expressed in a single tissue, so that tenascin mRNA is not detected. As shown in Fig. 5, the XB gene encoding TN-X is expressed at varying levels in all tissues examined, with greatest expression in fetal testis and in fetal smooth, striated, and cardiac muscle.

**Cellular Expression of TN-X**

Expression of the XB gene encoding TN-X can be evaluated by highly specific RNase protection assays. We used a single-stranded probe from a region (Fig. 4B) that lies beyond the limit of duplication of this locus; hence, it will not hybridize to and be protected by XA mRNA species. This probe has <50% nucleotide sequence identity with the corresponding region of tenascin, so that tenascin mRNA is not detected. As shown in Fig. 5, the XB gene encoding TN-X is expressed at varying levels in all tissues examined, with greatest expression in fetal testis and in fetal smooth, striated, and cardiac muscle.

*An unusual feature of this genetic system is that all of these overlapping RNA transcripts are expressed in a single tissue, the adrenal cortex. The P450c21 mRNA and the four Y transcripts (YA-short, YA-long, YB-short, YB-long) (4) overlap the XA and XB transcripts. To determine whether these overlapping transcripts are all expressed in the same cells, we performed in situ hybridization histochemistry experiments using probes for P450c21, Y, and XA. All of the probes homogeneously and diffusely stain adult human adrenal cortex (Fig. 6). The P450c21 probe is a 0.7-kb KpnI/EcoRI fragment of our human P450c21 cDNA (32) that excludes Y. Fig. 6A shows that P450c21 is ubiquitously expressed in all cells of the adrenal cortex, but not in the adrenal capsule. The XB probe
pressed in most adrenocortical cells. In adrenocortical cells, all of these transcripts appear to be expressed throughout the adrenal cortex, but, as with P450c21, neither XA nor Y is expressed by the adrenal capsule. As each of these four probes hybridizes with RNA in over 90% of the cells of the adrenal cortex also contain XA RNA, and that there is substantial TN-X expression in the connective tissue of the capsule. The XA and Y probes are the opposite strands of a genomic fragment encompassing Fn III repeat xxiii. Note that the protected 180-bp fragment is seen in all samples; however, the level of expression varies widely among tissues, and is greatest in fetal testis, gut (small intestine, mostly smooth muscle), muscle (striated), and heart (whole).

As indicated by the RNase protection experiment in Fig. 2, there is a single-stranded copy of a 353-bp PstI fragment that encompasses exon 23 (Fn III repeat xxi, bases 19,006-19,368 in Fig. 2). This exon lies 3-kb upstream from the cap site of the XA transcript, and has only ∼50% nucleotide sequence identity with any portion of XA; hence this probe is specific for XB. Fig. 6 B shows that >90% of cells in the adrenal cortex also contain XB (TN-X) mRNA, and that there is substantial TN-X expression in the connective tissue of the capsule. The XA and Y probes are the opposite strands of a 238-bp SacI fragment extending from bases 2,390 to 2,628 in the XA gene sequence (Fig. 3 A in reference 14); because this sequence lies in the region corresponding to intron 26 in XB (bases 22,876-23,114 in Fig. 2), it is not found in XB (TN-X) mRNA, so that one strand of the probe is specific for XB and the other is specific for Y. Fig. 6 C shows that all cells of the adrenal cortex also contain XA RNA, and Panel D shows that Y RNA is also expressed homogeneously throughout the adrenal cortex, but, as with P450c21, neither XA nor Y is expressed by the adrenal capsule. As each of these four probes hybridizes with RNA in over 90% of the adrenocortical cells, all of these transcripts appear to be expressed in most adrenocortical cells.

As indicated by the RNase protection experiment in Fig. 5, TN-X is widely expressed in human fetal tissues, including the adrenal. In Fig. 7 A, a 400-bp probe from repeats xxvi and xxvii that detects transcripts from both the XB and XA genes diffusely hybridizes to RNA throughout the human fetal adrenal cortex. Similar diffuse hybridization is seen with probes for Y (Fig. 7 B), the cholesterol side-chain cleavage enzyme, P450sc (Fig. 7 C), and P450c21 (Fig. 7 D). Thus the pattern of expression of the P450c21/X/Y genes is similar in both the adult and fetal adrenal cortex.

Discussion

Overlapping Genes and Polygenes

Overlapping genes occur commonly in prokaryotes, bacteriophages, and viruses, but only a few examples have been described in higher eukaryotes. These include a cardiac-specific gene on the opposite strand of the gene for rat hypothalamic gonadotropin-releasing hormone (I) two Droso- phila loci (18, 51), a mouse locus of unknown function (59), a member of the c-erb-A gene family on the opposite strand from the gene encoding rat (28) and human (37) thyroid hormone receptor, and transcripts on the opposite strand of rat TGF-β (44) N-myc (27), and murine p53 (24). In prokaryotic systems, overlapping genes encoding "antisense" RNAs often are involved in gene regulation (for review see references 20 and 50); the roles of overlapping eukaryotic genes are now being studied.

All of these overlapping gene systems described to date consist of only two transcripts. By contrast, the TN-X/ P450c21/Y system appears to be the first eukaryotic example of three linked, overlapping genes: a "polygene." It is not wholly surprising that such a complex was first discovered in the HLA class III region. This region of the genome appears to be a hot spot for genetic recombination, as evidenced by the highly allelic nature of the HLA locus in general, and by the very high incidence of gene conversion events in the P450c21 and C4 loci (for reviews and references see references 35 and 38). This may have contributed to the very high density of genes in this region: as shown in Fig. 1, little intragenic DNA separates the various known transcriptional units in this locus.

Studies with antisense oligonucleotides, antisense RNAs generated in vitro, and transfected antisense genes (reviewed in reference 54) indicate that the abundance of an mRNA and its encoded protein can be decreased by introducing exogenous antisense sequences, presumably by a mechanism dependent on RNA/RNA hybridization (21, 25). However, endogenously produced antisense RNAs differ from these experimental systems in two important ways. First, heteroduplexes of mRNA and synthetic DNA expose the RNA component to rapid degradation by RNase H–like activities (9, 56); this would not occur with endogenously produced antisense RNA transcripts. Second, experimental inhibition of gene expression generally requires a large molar excess of the antisense sequence, probably because synthetic single-stranded oligonucleotides are unstable in vivo (25). However, antisense nucleotides are stabilized by integration into an open reading frame or by polyadenylation (10, 19). Hence endogenously produced P450c21/TN-X duplexes or TN- X/Y duplexes are likely to be much more stable than are corresponding synthetic experimental systems. Our in situ hy-
Figure 6. Darkfield photomicrographs of in situ hybridization histochemistry. A frozen surgical sample of human adult adrenal was sectioned and hybridized with the following probes: (A) P450c21 cDNA fragment that excludes XA and TN-X sequences; (B) a fragment of TN-X cDNA that does not overlap P450c21, XA, or Y; (C) a 5' fragment of XA cDNA that excludes TN-X, P450c21 and Y; and (D) a fragment of YA cDNA that hybridizes to both YA and YB, but not to P450c21, XA, or TN-X sequences. The nonhybridizing "holes" in each photograph are due to holes in the desiccated frozen surgical sample. An opposite strand TN-X sense probe is used as a control in panel E. Because of the overlapping nature of the other transcripts, the controls in E, G, and H are sections treated with RNase A before hybridization with cDNAs for P450c21 (E), XA (G), and Y (H). All photomicrographs were taken at the same magnification. Bar, 200 μm.

Bridging histochemistry experiments show that P450c21B, XB, and Y are all expressed in the same fetal adrenocortical cells. Thus, formation of such RNA duplexes in vivo is possible.

Some such endogenously produced antisense systems have been studied. N-myc and its antisense transcript form double-stranded hybrids in vivo, which influence N-myc mRNA processing (29). In contrast, the antisense transcript of p53 prevented normal translocation of p53 mRNA to the cytoplasm in murine erythroleukemia cells (24). Kimelman and Kirschner (26) described a small transcript arising from the opposite strand of the gene encoding Xenopus basic FGF.

This antisense transcript encoded a protein, but also directly interacted with the sense transcript converting half of the hybridized adenine residues to inosine, thus changing the informational content of the sense transcript. However, we have found no evidence of such RNA editing in the P450c21/TN-X/Y system.

Tenascins

Tenascin was first described as an extracellular matrix protein in tendons and muscle and later in developing chick brain. It has been variously termed GMEM, tenascin,
peptide capable of forming alpha helices. For TN-C and TN-X, the heptads appear to allow three polypeptides to form a tri-}

Figure 7. Darkfield photomicrographs of in situ hybridization histochemistry of human fetal adrenal sections. The probes used are: (A) a 350-bp HincII/PstI probe that detects transcripts from either the XB or XA genes; (B) the same Y-specific probe used in Fig. 6 D; (C) a 1.2-kb fragment of human P450ccc cDNA (8); (D) the same P450c21 probe used in Fig. 6 A. The controls for the fetal tissues are the same as in Fig. 6, and are not shown. All photomicrographs were taken at the same magnification. Bar, 100 μm.

cytotactin, JI, hexabrachion, and myotendinous antigen (reviewed in reference 11). Elucidation of its gene, mRNA, and protein structure revealed that the protein is a polymer of six chains, each comprising a carboxy-terminal fibrinogen-like domain, a series of Fn \[\] repeats, a series of EGF-like repeats, and an amino-terminal "head group" that permits the chains to aggregate into the characteristic "hexabrachion" seen in electron micrographs (52). The closely related protein restrictin (42), also termed JI-160 (13), is a trimer composed of similar but considerably smaller units: a "tribrachion." The XB gene described in this report encodes a protein that is structurally similar to TN-X, or TN-X. Although there is no direct evidence to date, we predict that TN-X will also polymerize into a "brachion" comprising an unknown number of arms. We suggest that this category of proteins be termed "brachions" or "tenascins." We favor the latter. Tenascin, the first-described member, which is also widely termed "cytotactin," would be termed tenascin-C or TN-C to designate tenascin-cytotactin; restrictin would be termed tenascin-R or TN-R; and the product of the XB gene described in this paper would be tenascin-X or TN-X. This system would emphasize the relatedness among the monomeric units of these proteins and would, to the extent possible, incorporate terminologies and letterings favored by various groups. It seems unlikely that the number of tenascins will exceed the confines of the alphabet.

The Journal of Cell Biology, Volume 122, 1993 276

TN-X protein appears to have well-conserved heptad repeats capable of forming alpha helices. For TN-C and TN-X, the heptads appear to allow three polypeptides to form a tri-ple coiled-coil which is stabilized by the flanking cysteine residues (46, 52). One interesting difference between TN-X and other members of this gene family is the sequence divergence at the extreme amino terminus of the mature protein. This region appears to mediate dimerization of or TN-C trimers joined in the heptad region to form the final hexabrachion. Cysteine residues are notably absent from this region of TN-X, and hence we predict that TN-X may form pairs or trimers, but not higher order polymers. While the available electron microscopic data show no heterogeneity in the lengths of the arms of TN-C and TN-X, it might be possible for monomers of TN-C, -R, and -X to form heteropolymers.

Several other features of TN-X merit comment. Unlike TN-C and TN-R, which are heavily glycosylated, TN-X has only five Asn-X-Ser (Thr) sites for N-linked glycosylation (Fig. 2). Whether this difference in the potential for glycosylation will confer unique biological properties to TN-X is not known. In addition, both TN-C and TN-R mRNAs undergo alternate splicing to yield different sized mRNAs and proteins. Because most of our structural data come from genomic DNA rather than from cDNA copies of mRNA molecules, we have no data to confirm or deny alternate splicing of TN-X. Finally, unlike human and chicken TN-C, TN-X does not contain an RGD sequence, and hence cannot bind to an RGD-dependent receptor. Recent evidence suggests that TN-C binding is not limited to the domain carrying the RGD sequence (46, 52). RGD-independent binding is consistent with the lack of an RGD sequence in mouse TN-C (57).

Overall, there is greater sequence similarity among the 29 Fn III repeats of TN-X (Table I) than among the 15 Fn III repeats of TN-C (15) or the 9 Fn III repeats of TN-R (42). Our phylogenetic analysis of Fn III repeats in TN-X and Gulcher's similar analysis of TN-C (15) suggest a recent reduplication of central Fn III repeats in these genes that has not occurred in TN-R. Table I shows that 15 of the 20 central Fn III repeats of TN-X (iv-xxiii) are most similar to repeat V of TN-R. This suggests that the central repeats of TN-X arose by duplication of the exon encoding the homolog of TN-R repeat V in the ancestral progenitor gene. These findings support a model in which TN-C, TN-R, and TN-X arose from a smaller primordial gene resembling TN-R, and that subsequent internal reduplications produced the central repeats of TN-C and TN-X after their divergence from one another. For TN-X, frequent duplication might be expected given the high incidence of genetic crossover and homologous recombination events in the MHC region of chromosome 6 (reviewed in reference 38). These events have led to homogenization of duplicated sequences in this region such as C4A and B, P450c21A and B, and the various Fn III repeats of TN-X.

No functional studies have been done to date on TN-X. However, previous studies of TN-C using polyclonal antibodies may have identified mixed populations of TN-C and TN-X. Our RNase protection experiments unambiguously establish that the TN-X mRNA is expressed in nearly all tissues, and is especially widely expressed in developing fetal tissues. It is conceivable that antibodies raised against TN-C might cross-react with TN-X (which might not be detected in Western blotting studies because of its size). Thus, previous studies of the functional role of tenascin based on immunologic reagents may need to be re-examined in the light
of the newly established existence of another, very similar, and widely expressed protein.

Crucial insights into function can come from deletion mutants—either created in the laboratory or spontaneously occurring in nature. The C4/P450c21/Gene X locus has been the subject of intensive genetic study because of its association with human disease, and especially with congenital adrenal hyperplasia (35, 36, 38). Among 811 informatively studied human chromosomes bearing mutated alleles causing congenital adrenal hyperplasia, 156 had large gene deletions and another 46 had other genetic rearrangements, yet none of these chromosomes has a deletion extending into the XB gene encoding TN-X (38), suggesting that TN-X may be necessary for survival.

Examination of the duplication junctions of the C4/P450c21/Gene X locus clearly shows that the duplicated human, bovine, and murine loci have different boundaries, suggesting gene duplications and genetic rearrangements that postdate mammalian speciation (14). Therefore, it is difficult to extrapolate from mouse genetics to human genetics—neither created in the laboratory or spontaneously occurring in nature. The C4/P450c21/Gene X locus has been the subject of intensive genetic study because of its association with human disease, and especially with congenital adrenal hyperplasia (35, 36, 38). Among 811 informatively studied human chromosomes bearing mutated alleles causing congenital adrenal hyperplasia, 156 had large gene deletions and another 46 had other genetic rearrangements, yet none of these chromosomes has a deletion extending into the XB gene encoding TN-X (38), suggesting that TN-X may be necessary for survival.

Examination of the duplication junctions of the C4/P450c21/Gene X locus clearly shows that the duplicated human, bovine, and murine loci have different boundaries, suggesting gene duplications and genetic rearrangements that postdate mammalian speciation (14). Therefore, it is difficult to extrapolate from mouse genetics to human genetics—neither created in the laboratory or spontaneously occurring in nature. The C4/P450c21/Gene X locus has been the subject of intensive genetic study because of its association with human disease, and especially with congenital adrenal hyperplasia (35, 36, 38). Among 811 informatively studied human chromosomes bearing mutated alleles causing congenital adrenal hyperplasia, 156 had large gene deletions and another 46 had other genetic rearrangements, yet none of these chromosomes has a deletion extending into the XB gene encoding TN-X (38), suggesting that TN-X may be necessary for survival.

Examination of the duplication junctions of the C4/P450c21/Gene X locus clearly shows that the duplicated human, bovine, and murine loci have different boundaries, suggesting gene duplications and genetic rearrangements that postdate mammalian speciation (14). Therefore, it is difficult to extrapolate from mouse genetics to human genetics—neither created in the laboratory or spontaneously occurring in nature. The C4/P450c21/Gene X locus has been the subject of intensive genetic study because of its association with human disease, and especially with congenital adrenal hyperplasia (35, 36, 38). Among 811 informatively studied human chromosomes bearing mutated alleles causing congenital adrenal hyperplasia, 156 had large gene deletions and another 46 had other genetic rearrangements, yet none of these chromosomes has a deletion extending into the XB gene encoding TN-X (38), suggesting that TN-X may be necessary for survival.

of the newly established existence of another, very similar, and widely expressed protein.

Crucial insights into function can come from deletion mutants—either created in the laboratory or spontaneously occurring in nature. The C4/P450c21/Gene X locus has been the subject of intensive genetic study because of its association with human disease, and especially with congenital adrenal hyperplasia (35, 36, 38). Among 811 informatively studied human chromosomes bearing mutated alleles causing congenital adrenal hyperplasia, 156 had large gene deletions and another 46 had other genetic rearrangements, yet none of these chromosomes has a deletion extending into the XB gene encoding TN-X (38), suggesting that TN-X may be necessary for survival.

Examination of the duplication junctions of the C4/P450c21/Gene X locus clearly shows that the duplicated human, bovine, and murine loci have different boundaries, suggesting gene duplications and genetic rearrangements that postdate mammalian speciation (14). Therefore, it is difficult to extrapolate from mouse genetics to human genetics—neither created in the laboratory or spontaneously occurring in nature. The C4/P450c21/Gene X locus has been the subject of intensive genetic study because of its association with human disease, and especially with congenital adrenal hyperplasia (35, 36, 38). Among 811 informatively studied human chromosomes bearing mutated alleles causing congenital adrenal hyperplasia, 156 had large gene deletions and another 46 had other genetic rearrangements, yet none of these chromosomes has a deletion extending into the XB gene encoding TN-X (38), suggesting that TN-X may be necessary for survival.

Examination of the duplication junctions of the C4/P450c21/Gene X locus clearly shows that the duplicated human, bovine, and murine loci have different boundaries, suggesting gene duplications and genetic rearrangements that postdate mammalian speciation (14). Therefore, it is difficult to extrapolate from mouse genetics to human genetics—neither created in the laboratory or spontaneously occurring in nature. The C4/P450c21/Gene X locus has been the subject of intensive genetic study because of its association with human disease, and especially with congenital adrenal hyperplasia (35, 36, 38). Among 811 informatively studied human chromosomes bearing mutated alleles causing congenital adrenal hyperplasia, 156 had large gene deletions and another 46 had other genetic rearrangements, yet none of these chromosomes has a deletion extending into the XB gene encoding TN-X (38), suggesting that TN-X may be necessary for survival.

Examination of the duplication junctions of the C4/P450c21/Gene X locus clearly shows that the duplicated human, bovine, and murine loci have different boundaries, suggesting gene duplications and genetic rearrangements that postdate mammalian speciation (14). Therefore, it is difficult to extrapolate from mouse genetics to human genetics—neither created in the laboratory or spontaneously occurring in nature. The C4/P450c21/Gene X locus has been the subject of intensive genetic study because of its association with human disease, and especially with congenital adrenal hyperplasia (35, 36, 38). Among 811 informatively studied human chromosomes bearing mutated alleles causing congenital adrenal hyperplasia, 156 had large gene deletions and another 46 had other genetic rearrangements, yet none of these chromosomes has a deletion extending into the XB gene encoding TN-X (38), suggesting that TN-X may be necessary for survival.

Examination of the duplication junctions of the C4/P450c21/Gene X locus clearly shows that the duplicated human, bovine, and murine loci have different boundaries, suggesting gene duplications and genetic rearrangements that postdate mammalian speciation (14). Therefore, it is difficult to extrapolate from mouse genetics to human genetics—neither created in the laboratory or spontaneously occurring in nature. The C4/P450c21/Gene X locus has been the subject of intensive genetic study because of its association with human disease, and especially with congenital adrenal hyperplasia (35, 36, 38). Among 811 informatively studied human chromosomes bearing mutated alleles causing congenital adrenal hyperplasia, 156 had large gene deletions and another 46 had other genetic rearrangements, yet none of these chromosomes has a deletion extending into the XB gene encoding TN-X (38), suggesting that TN-X may be necessary for survival.

Examination of the duplication junctions of the C4/P450c21/Gene X locus clearly shows that the duplicated human, bovine, and murine loci have different boundaries, suggesting gene duplications and genetic rearrangements that postdate mammalian speciation (14). Therefore, it is difficult to extrapolate from mouse genetics to human genetics—neither created in the laboratory or spontaneously occurring in nature. The C4/P450c21/Gene X locus has been the subject of intensive genetic study because of its association with human disease, and especially with congenital adrenal hyperplasia (35, 36, 38). Among 811 informatively studied human chromosomes bearing mutated alleles causing congenital adrenal hyperplasia, 156 had large gene deletions and another 46 had other genetic rearrangements, yet none of these chromosomes has a deletion extending into the XB gene encoding TN-X (38), suggesting that TN-X may be necessary for survival.

Examination of the duplication junctions of the C4/P450c21/Gene X locus clearly shows that the duplicated human, bovine, and murine loci have different boundaries, suggesting gene duplications and genetic rearrangements that postdate mammalian speciation (14). Therefore, it is difficult to extrapolate from mouse genetics to human genetics—neither created in the laboratory or spontaneously occurring in nature. The C4/P450c21/Gene X locus has been the subject of intensive genetic study because of its association with human disease, and especially with congenital adrenal hyperplasia (35, 36, 38). Among 811 informatively studied human chromosomes bearing mutated alleles causing congenital adrenal hyperplasia, 156 had large gene deletions and another 46 had other genetic rearrangements, yet none of these chromosomes has a deletion extending into the XB gene encoding TN-X (38), suggesting that TN-X may be necessary for survival.

Examination of the duplication junctions of the C4/P450c21/Gene X locus clearly shows that the duplicated human, bovine, and murine loci have different boundaries, suggesting gene duplications and genetic rearrangements that postdate mammalian speciation (14). Therefore, it is difficult to extrapolate from mouse genetics to human genetics—neither created in the laboratory or spontaneously occurring in nature. The C4/P450c21/Gene X locus has been the subject of intensive genetic study because of its association with human disease, and especially with congenital adrenal hyperplasia (35, 36, 38). Among 811 informatively studied human chromosomes bearing mutated alleles causing congenital adrenal hyperplasia, 156 had large gene deletions and another 46 had other genetic rearrangements, yet none of these chromosomes has a deletion extending into the XB gene encoding TN-X (38), suggesting that TN-X may be necessary for survival.
29. Leahy, D. J., W. A. Hendrickson, I. Aukhil, and H. P. Erickson. 1992.
31. Matsumoto, K., N. Ishihara, A. Ando, H. lnoko, and T. Ikemura. 1992.
The Journal of Cell Biology, Volume 122, 1993 278
28. Lazar, M. A., R. A. Hodin, D. S. Darling, and W. W. Chin. 1989. A novel
25. Kim, S. K., and B. J. Wold. 1985. Stable reduction of thymidine kinase
23. Just, M., H. l-Herbst, M. Hummel, H. Dfirkop, D. Tripier, H. Stein, and
20. Inouye, M. 1988. Antisense RNA: its functions and applications in gene
19. Holt, J. T., T. V. Gopal, A. D. Moulton, and A. W. Nienhnis. 1986. Induc-
17. Helms, C., M. Y. Graham, J. E. Dutehik, and M. V. Olson. 1985. A new
16. Gulcher, J. R., D. E. Nies, M. J. Alexakos, N. A. Ravikant, M. E. Sturgill,
15. Guleber, J. R., M. J. Alexakos, M. M. Le Beau, R. S. Lemons, and K.
1. Erickson, H. P., and M. A. Bourdon. 1989. Tenascin: an extracelhilar ma-
9. Dash, P. I., I. Lotan, M. Knapp, E. R. Kandel, and P. Goellet. 1987. Selec-
7. Chiquet-Ehrismarm, R., P. Kalla, C. A. Pearson, K. Beck, and M. Chi-
8. Chung, B., K. J. Matteson, R. Voutilainen, T. K. Mohandas, and W. L.
12. 485--491.
19:1128-1136.
9:129-138.
258: 987-991.
129: 275-281.
56. Walder, R. Y., and J. A. Walder. 1988. Role of RNase H in hybrid-arrested
52. Spring, J., S. W. Gunther, and R. Chiquet-Ehrismann. 1989. Two contrary func-
51. Spencer, C. A., R. D. Giriz, and R. B. Hodgetts. 1986. Overlapping tran-
39. Morel, Y., J. Bristow, S. E. Gitelman, and W. L. Miller. 1992. Analysis
38. Miller, W. L., S. E. Gitelman, J. Bristow, and Y. Morel. 1990. Chromosomal
37. Miyajima, N., R. Horiuchi, Y. Shibuya, S. Fukushige, K. Matsubara, K.
36. Miller, W. L., and Y. Morel. 1989. The molecular genetics of 21-
35. Miller, W. L., and Y. Morel. 1989. The molecular genetics of 21-
34. Miller, W. L., S. E. Gitelman, J. Bristow, and Y. Morel. 1992. Analysis
33. Mellon, S. H., and W. L. Miller. 1989. Extracellular steroid 21-
32. Mellon, S. H., and C. Vaise. 1989. cAMP regulates P450sc mRNA gene expres-
31: 61-8.
24: 1728-2823.
23: 371-393.
22: 567-600.
21: 430-435.
20: 704-709.
19: 4017-4022.
18: 958-976.
17: 521-528.
16: 4683-4690.
15: 47-56.
14: 4683--4690.
13: 697-702.
12: 1056-1061.
11: 355-362.
10: 1-14.
9: 351-355.
8: 4107--4114.
7: 410-415.
6: 958-976.
5: 970-976.
4: 503-509.
3: 410-419.
2: 663-678.
1: 663-678.
19. 206:17326-17332.
18. 205:1728-17332.
17. 204:1728-17332.
16. 203: 1263-1270.
15. 202: 1263-1270.
14. 201: 1263-1270.
13. 200: 1263-1270.
12. 199: 1263-1270.
11. 198: 1263-1270.
10. 197: 1263-1270.
9. 196: 1263-1270.
8. 195: 1263-1270.
7. 194: 1263-1270.
6. 193: 1263-1270.
5. 192: 1263-1270.
4. 191: 1263-1270.
3. 190: 1263-1270.
2. 189: 1263-1270.
1. 188: 1263-1270.

The Journal of Cell Biology, Volume 122, 1993 278

adrenal hyperplasia. Proc. Natl. Acad. Sci. USA. 84:5858-5862.
33. Mellon, S. H., and W. L. Miller. 1989. Extracellular steroid 21-
hydroxylation is not mediated by P450sc. J. Clin. Invest. 84: 1497-
32. Mellon, S. H., and C. Vaise. 1989. cAMP regulates P450sc mRNA gene expres-
31: 61-8.
30: 430-434.
29: 8796-7900.
28: 1214-2134.
27: 616-622.
26: 849-863.
25: 61-8.
24: 430-435.
23: 521-528.
22: 697-702.
21: 430-435.
20: 704-709.
19: 4017-4022.
18: 958-976.
17: 521-528.
16: 4683-4690.
15: 47-56.
14: 4683--4690.
13: 697-702.
12: 1056-1061.
11: 355-362.
10: 1-14.
9: 351-355.
8: 4107--4114.
7: 410-415.
6: 958-976.
5: 970-976.
4: 503-509.
3: 410-415.
2: 663-678.
1: 663-678.