Allfrey, V. G. (1983) *Cell (Cambridge, Mass.) 34*, 1033.
Prunell, A. (1983) *Biochemistry 22*, 4887.
Puigdomenech, P., Daban, J. R., Palau, J., Podo, F., Guidoni, L., & Temussi, P. A. (1977) *Biochim. Biophys. Acta 492*, 12.
Rill, R. L., Oosterhof, D. K., Hozier, J. C., & Nelson, D. A. (1975) *Nucleic Acids Res. 2*, 1525.
Rubin, R. L., & Moudrianakis, E. N. (1975) *Biochemistry 14*, 1718.
Thomas, G. J., Prescott, B., & Olins, D. E. (1977) *Science (Washington, D.C.) 197*, 385.
Thomas, J. O., & Kornberg, R. D. (1974) *Science (Washington, D.C.) 184*, 865.
Wingender, E., Maass, K., & Bode, J. (1981) *Int. J. Biol. Macromol. 3*, 114.
Wong, N. T. N., & Candido, E. P. M. (1978) *J. Biol. Chem. 253*, 8263.
Zama, M., Bryan, P. N., Harrington, R. E., Olins, A. L., & Olins, D. E. (1977) *Cold Spring Harbor Symp. Quant. Biol. 42*, 31.

Partial Amino Acid Sequence of Human Thrombospondin As Determined by Analysis of cDNA Clones: Homology to Malarial Circumsporozoite Proteins†

Sentaro Kobayashi,‡ Francine Eden-McCutchan,§ Paul Framson,‡ and Paul Bornstein*,‡

Department of Biochemistry, University of Washington School of Medicine, Seattle, Washington 98195, and Division of Molecular Biology Applications, Meloy Laboratories, Inc., Springfield, Virginia 22151

Received September 9, 1986; Revised Manuscript Received October 2, 1986

ABSTRACT: A λgt 11 library prepared from human umbilical vein endothelial cell RNA was screened for cDNAs encoding thrombospondin. Reagents included a monospecific antibody to human thrombospondin and a mixture of four synthetic oligodeoxyribonucleotides derived from an amino acid sequence near the NH_2 terminus of mature human thrombospondin. Two series of cDNA clones coding for sequences at the 5′ and 3′ ends of thrombospondin mRNA, respectively, were isolated. The nucleotide sequence of a 1.3-kilobase (kb) 5′ clone (λTS-33) coded for 99 bases of 5′ untranslated RNA, a signal peptide of 18 amino acids, and the first 379 amino acids of thrombospondin. Northern blot analysis with λTS-33 detected a single mRNA species of ~6.0 kb in rat aortic smooth muscle cell RNA. Thrombospondin mRNA levels increased rapidly, but transiently, in quiescent smooth muscle cells treated with platelet-derived growth factor. The kinetics of this response were very similar to those of the thrombospondin protein to this growth factor. There was significant homology in amino acid sequence between thrombospondin and a conserved region in the circumsporozoite protein of two malarial sporozoites. This region of thrombospondin may therefore represent a potential recognition site for a cell surface thrombospondin receptor.

Thrombospondin (TS)[1] is a high molecular weight glycoprotein (M_r 450 000) that is both stored in α granules of platelets and secreted by a wide variety of mesenchymal cells [see Silverstein et al. (1986) and Majack and Bornstein (1986) for recent reviews]. There is evidence that TS functions in the secondary, irreversible phase of platelet aggregation (Leung, 1984; Dixit et al., 1985; Silverstein et al., 1986). TS also forms complexes with histidine-rich glycoprotein and plasminogen (Silverstein et al., 1985a) and may regulate the release of plasmin from complexes containing plasminogen and tissue plasminogen activator (Silverstein et al., 1985b).

The role of TS in the extracellular matrix is less well understood. TS binds to cell surfaces (McKeown-Longo et al., 1984; Roberts et al., 1985) and to a variety of matrix macromolecules including collagen, fibronectin, and heparin (heparan sulfate) Dixit et al., 1984; Majack & Bornstein, 1986), but the functional consequences of these interactions are not known. Recently, this laboratory has provided evidence that PDGF and heparin-like glycosaminoglycans act antagonistically in regulating the production of a TS-rich extracellular matrix by rat aortic SMC (Majack et al., 1985). Asch et al. (1986) have also shown that PDGF stimulates the production of TS by cultured human glial cells. Furthermore, TS can act synergistically in conjunction with epidermal growth factor to facilitate SMC growth (Majack et al., 1986). Thus, evidence exists for an autocrine, growth-promoting function for matrix TS.

The regions of the TS protein chain responsible for binding to cell surfaces and to extracellular macromolecules have been identified in a preliminary fashion by a combination of limited protease fragmentation and the use of monoclonal antibodies (Galvin et al., 1985). Limited amino acid sequence information is also available (Coligan & Slayter, 1984; Dixit et al., 1984, 1985; Raugi et al., 1984; Galvin et al., 1985). These findings suggest that TS, like fibronectin, is a member of a class of proteins termed "modular glycoproteins" (Hynes & Yamada, 1982). In all likelihood, the TS molecule consists of three identical chains of M_r 140 000–150 000, but variations in structure resulting from alternative splicing of the TS mRNA, as has been shown for fibronectin (Tamkun et al., 1984), may exist.

† This work was supported by National Institutes of Health Grants AM 11248, HL 18645, and DE 02600 and by a grant from R. J. Reynolds, Inc.
‡ University of Washington School of Medicine.
§ Meloy Laboratories, Inc.

[1] Abbreviations: TS, thrombospondin; PDGF, platelet-derived growth factor; SMC, smooth muscle cells; Tris, tris(hydroxymethyl)aminomethane; SDS, sodium dodecyl sulfate; SSC, standard saline citrate.

0006-2960/86/0425-8418$01.50/0

To expedite the structural analysis of this large protein and to gain an understanding of the structure and regulation of the TS gene, we have screened a λgt 11 cDNA expression library prepared from human umbilical vein RNA both with a monospecific polyclonal antibody to TS and with a mixture of synthetic oligomers. We have identified and partially characterized a number of cDNA clones for this protein and have discovered a significant degree of homology between a sequence of 12 amino acids in TS and a highly conserved region in malarial circumsporozoite proteins that may mediate binding of sporozoites to hepatocyes (Dame et al., 1984). We have also used a TS cDNA clone to document that treatment of rat aortic SMC with PDGF leads to a rapid, but transient, increase in TS mRNA levels.

Materials and Methods

Anti-Thrombospondin. Rabbit anti-human TS was affinity-purified on a column of human platelet TS covalently linked to Sepharose CL-4B and characterized as described by Wight et al. (1985). At a dilution of 1/250, the affinity-purified antibody preparation readily detected 50 pg of TS in a solid-phase assay.

Construction of a cDNA Library in λgt 11. Total RNA was extracted from human umbilical vein endothelial cells with a solution containing 4 M guanidinium thiocyanate, 25 mM sodium citrate, pH 7.0, 0.5% *N*-laurylsarcosine, 0.1 M 2-mercaptoethanol, and 0.2% Antifoam A; the RNA was purified as described by Chirgwin et al. (1979). RNA (100 μg) was copied into ds-cDNA with reverse transcriptase and DNA polymerase I (Buell et al., 1978; Wickens et al., 1978). The ds-cDA was purified on an Elutip-D column (Schleicher & Schuell, Keene, NH), made blunt-ended by incubation with S1 nuclease (Ricca et al., 1981), and treated with *Eco*RI methylase and DNA polymerase I (Klenow fragment) as described by Maniatis et al. (1982). After further purification on Elutip-D columns, the DNA was ligated to 0.5 μg of phosphorylated *Eco*RI linkers with T4 DNA ligase (Maniatis et al., 1982). The mixture was then cleaved with *Eco*RI and fractionated on an 8% acrylamide gel in Tris–borate buffer. DNA with a size greater than 1 kilobase (kb) was eluted from the gel, purified twice on Elutip-D columns, and inserted into λgt 11 phage that had been cleaved with *Eco*RI and treated with phosphatase. A library of approximately 1.5×10^6 phage was produced, 63% of which contained inserts. The library was amplified by producing plate lysates at 42 °C on *Escherichia coli* Y 1088 (Young & Davis, 1983a). The titer of the amplified library was determined to be 8.1×10^{10} pfu/mL.

Screening of cDNA Library. The λgt 11 cDNA library was screened with anti-TS antibody as described by Young and Davis (1983a,b). Phage plaques were transferred to IPTG-saturated nitrocellulose. Filters were then blocked with gelatin and incubated with anti-TS serum (1/250 dilution) overnight at 20 °C. Filters were washed and incubated with HRP-conjugated second antibody followed by application of color development reagents (Bio-Rad).

Synthesis of a Mixture of Four 42-Base Oligonucleotides. The basis for the synthesis of a mixture of oligonucleotides complementary to the noncoding strand of TS cDNA is shown in Table I. The sequence of amino acids 11–25 in the heparin-binding fragment of human platelet TS, as determined by Coligan and Slayter (1984), was chosen as the most favorable available sequence for synthesis of a specific oligonucleotide. The heparin-binding fragment of TS has been shown to be derived from the NH_2 terminus of the molecule (Coligan & Slayter, 1984; Dixit et al., 1984; Raugi et al., 1984). The selection of codons was based on the tendency of human coding sequences to avoid CpG base pairs (Lathe, 1985; Bird, 1981). Preferential codon usage (Lathe, 1985) was then used as an additional guide to base selection. Finally, deoxyinosine was used in positions of greatest ambiguity since this base can pair with three bases (A, C, and T) sufficiently well to avoid destabilization of the duplex DNA (Ohtsuka et al., 1985). The next-to-last line in Table I shows the sequences of the four 42-mers that were synthesized as a mixture. Each 42-mer contains four deoxyinosine bases. Oligonucleotide synthesis was performed with an Applied Biosystems Model 380A DNA synthesizer using phosphoramidite precursors. The deblocked oligonucleotide was further purified by electrophoresis through a 10% denaturing polyacrylamide gel followed by elution and Sephadex G-25 chromatography. For preparation of probes, oligonucleotides were labeled with T4 polynucleotide kinase (Pharmacia P-L) and [γ-^{32}P]ATP (Amersham, 7000 Ci/mmol).

DNA Sequence Analysis. DNA was sequenced by the dideoxy procedure (Sanger et al., 1977) with single-strand templates (Sanger et al., 1980). Oligonucleotide primers, 18 nucleotides in length, were used without additional purification (Sanchez-Pescador & Urdea, 1984). Data were analyzed with software from Intelligenetics, Inc. (Palo Alto, CA).

SMC Cell Culture. Rat aortic SMC were grown from explants and cultured as previously described (Majack & Clowes, 1984; Majack et al., 1985). PDGF was obtained from E. Raines and R. Ross (University of Washington) and was purified on phenyl-Sepharose (Raines & Ross, 1982). Purified PDGF showed half-maximal stimulation of [^{3}H]TdR incorporation into Swiss 3T3 cells at 1.0 ng/mL and stimulated DNA synthesis maximally in SMC at 10 ng/mL. SMC, previously made quiescent by culture in plasma-derived serum for 3 days, were exposed to PDGF at 5 ng/mL for timed intervals as described by Majack et al. (1985).

Preparation of RNA from SMC and Northern Blot Analysis. Total cellular RNA was extracted from trypsinized SMC with SDS–proteinase K. Purified RNA was denatured in 50% formamide at 65 °C and separated on 1.2% agarose containing 6.7% formaldehyde. RNA was transferred to nitrocellulose and baked at 80 °C for 3 h. Prehybridization was in 50% formamide, 6× SSC, 2× Denhardt's solution, 50 mM Na_3PO_4, pH 7.0, and 0.1 mg/mL yeast tRNA at 55 °C for 3 h. Nick-translated λTS-33, cloned into pGEM-2, was hybridized at a concentration of 1×10^6 cpm/mL at 55 °C in the same buffer overnight. Nitrocellulose blots were washed twice with 2× SSC and 0.1% SDS at 20 °C and twice with 0.1× SSC and 0.1% SDS at 55 °C.

Southern Blots. Phage DNA was digested with restriction enzymes separated on 1% agarose gels and blotted to Gene Screen-Plus (New England Nuclear) or nitrocellulose (Schleicher & Schuell). Nick-translated probes were hybridized to the blot in 6× SSC, 1× Denhardt's solution, and 0.1% SDS at 68 °C overnight. Generally, 1×10^7 cpm of probe was used together with 1 mg of sheared, denatured carrier DNA. The blot was washed repeatedly in 0.1× SSC at 52 °C and air-dried. Hybridization with ^{32}P end-labeled poly[d(T)$_{13}$GT] was performed in the same way, but washing was at 42 °C.

Results

Detection of cDNA Clones by Screening with Anti-TS Antibody and with Oligonucleotide Probes. One library equivalent, consisting of 7.5×10^5 plaques, was screened with anti-TS antibody, and several recombinants were identified. The positive phages were plaque-purified, and their DNA was isolated. The inserted DNA in these phages shows sequence

Table I: Synthetic Oligodeoxyribonucleotides Used as Hybridization Probes

	11													25
Amino acid[a]	SER	VAL	PHE	ASP	ILE	PHE	GLU	LEU	THR	GLY	ALA	ALA	TRP	LYS	GLY
Codons	N	GUN	UU$^{U}_{C}$	GA$^{U}_{C}$	AU$^{U}_{C}$A	UU$^{U}_{C}$	GA$^{A}_{G}$	CUN UU$^{A}_{G}$	ACN	GGN	GCN	GCN	UGG	AA$^{A}_{G}$	GGN
No CpG sequence	A U G	GUN	UUU	GA$^{U}_{C}$	AU$^{U}_{C}$A	UUU	GA$^{A}_{G}$	CUN UU$^{A}_{G}$	AC$^{A}_{U}$	GGA G U	GC$^{A}_{U}$	GCN	UGG	AA$^{A}_{G}$	GGN
Preferential codon usage[b]	U (0.72)			GAC (0.62)	AUC (0.64)		GAG (0.60)	CU$^{C}_{G}$ (0.86)			GCU (0.65)				
Substitution of deoxyinosine		I								I		I		I	
Best oligonucleotide[c]	T	GTI	TTT	GAC	ATC	TTT	GAG	CT$^{C}_{G}$	AC$^{A}_{T}$	GGI	GGT	GCI	TGG	AAI	GG
Sequencing primer[d]		3'	AAA	CTG	TAG	AAA	CTC	GA	5'						

[a] Sequence from Coligan and Slayter (1984). [b] The number in parentheses indicates the fractional codon usage. [c] A mixture of four 42-mers was synthesized. [d] Primer used to establish the nucleotide sequence 5′ to the cDNA sequence recognized by synthetic oligonucleotide.

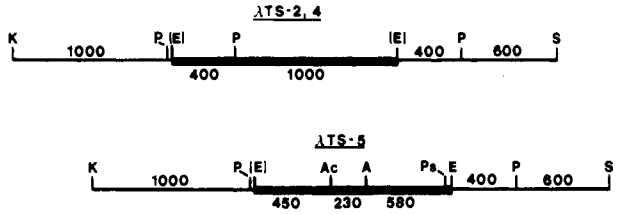

FIGURE 1: Thrombospondin cDNA clones identified by antibody screening. The heavy line indicates the extent of the cDNA inserted in the *Eco*RI site of λgt 11. Sites in parentheses indicate restriction sites lost during the process of addition of linkers and cloning in the vector. The insert in λTS-5 was released by digestion with *Pvu*II and *Eco*RI, purified, and nick-translated. The nick-translated probe was hybridized to Southern blots of the other three clones, digested with either *Eco*RI and *Pvu*II or *Kpn*I and *Sac*I, to demonstrate the extent of overlap of the cDNA inserts. λTS-2 and -4 appear to be siblings. A, *Ava*I; Ac, *Acc*I; E, *Eco*RI; K, *Kpn*I; P, *Pvu*II; *Ps*, PstI; S, *Sac*I.

homology, indicating that the anti-TS antibody detected a specific series of clones. Restriction analysis confirmed the relationship of these clones and indicated that the insert in λTS-5 could be recloned as a *Pvu*II–*Eco*RI fragment, 1250 base pairs (bp) in length. This fragment was subcloned into pGEM-2 for further analysis, and a partial restriction map of the insert was prepared (Figure 1).

The 1250-bp *Pvu*II–*Eco*RI-cleaved fragment of λTS-5, labeled by nick translation, was used to probe a Northern blot of rat SMC cell RNA. A single mRNA with an estimated size of 6.0 kb was identified (data not shown). An mRNA of this size could adequately provide for translation of a protein of about 1400 amino acids; thus, it is a good candidate for TS mRNA. However, since the identity of the antibody-identified cDNA clones depended almost entirely on the specificity of the antibody preparation, these results needed independent verification.

We elected to rescreen the λgt 11 library using an oligonucleotide probe. This probe was designed with an amino acid sequence near the NH_2 terminus of the TS chain (Coligan & Slayter, 1984) (see Table I for the rationale of oligonucleotide development). Approximately 1×10^6 plaques of an endothelial cell cDNA library in λgt 11 were screened with the mixture of oligonucleotide probes shown in Table I. The initial screen, performed by hybridization in 6× SSC at 42 °C, with washing in 2× SSC at 37 °C, revealed about 500 positive plaques. Surprisingly, most of these phages were still positive when the filters were washed in 2× SSC at 55 °C. Ten phages were selected and subjected to two additional rounds of plaque purification. DNA was isolated from these phages and digested with *EcoRI*. Inserts ranging in size from 1.3 to 1.7 kb were released. To determine relationships among these inserts, one of them, λTS-33 with an insert size of 1.3 kb, was subcloned into pGEM-2, and the isolated insert was nick-translated. Southern blot analysis indicated that the λTS-33 probe hybridized to five other clones of identical length. These six clones are probably siblings. In addition, λTS-33 hybridized to clone λTS-1211 with an insert length of 1.6 kb and to λTS-4119, the insert of which was not released by *Eco*RI digestion. Thus the oligonucleotide probe also detected a specific group of clones.

The insert in λTS-33, subcloned into pGEM-2, was partially sequenced in the form of double-stranded DNA with the primer indicated on the last line of Table I. The sequence upstream of the primer translated to the amino acid sequence Asn-Arg-Ile-Pro-Glu-Ser-Gly-Gly-Asp-Asn-Ser, matching exactly the amino-terminal sequence of human TS (Coligan & Slayter, 1984; Dixit et al., 1984; Raugi et al., 1984). This sequence unequivocally identifies λTS-33, and the other clones with which this clone hybridizes, as authentic TS cDNA clones.

The nick-translated 1.3-kb insert of λTS-33, cloned into pGEM-2, was then used in Northern blot analysis of rat aortic SMC RNA. An mRNA identical in size with that found with the antibody-identified λTS-5 probe was identified (see Figure 3). This finding strongly supports the conclusion that the antibody-identified clones are also TS clones. Since there was no cross-hybridization between the synthetic oligomer- and antibody-identified clones, the latter are most probably derived from a more 3′ region of the TS mRNA. To determine whether the antibody-identified clones included the poly(A) sequence in TS mRNA, the pGEM-2 plasmid that contained the cDNA insert from λTS-5 was digested with the combinations of enzymes *Hin*dIII–*Ava*I, *Eco*RI–*Ava*I, and *Hin*dIII–*Acc*I. The *Hin*dIII site exists in the polylinker region of pGEM-2 just upstream of the altered *Eco*RI site in λTS-5. The digests were resolved on a 1.5% agarose gel and trans-

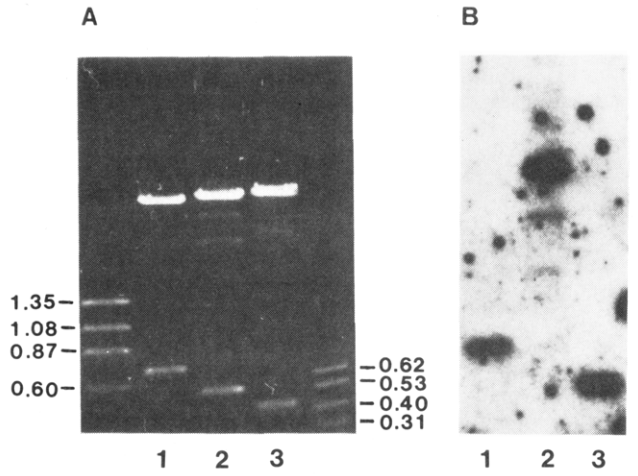

FIGURE 2: (A) Agarose (1.5%) gel electrophoretogram of restriction enzyme digests of the 1260-bp insert of λTS-5, excised with *Eco*RI and *Pvu*II (see Figure 1) and cloned into the *Eco*RI and *Hin*cII sites of pGEM-2. (Lane 1) *Hin*dIII–*Ava*I; (lane 2) *Eco*RI–*Ava*I; (lane 3) *Hin*dIII–*Acc*I. Molecular weight markers (in kb) were obtained by digestion of ϕX 174 with *Hae*III and pBR with *Msp*I. The gel was stained with ethidium bromide. (B) Southern blot. DNA from a gel identical with that shown in panel A was transferred to nitrocellulose paper and probed with ^{32}P end-labeled poly[d(T)$_{13}$GT].

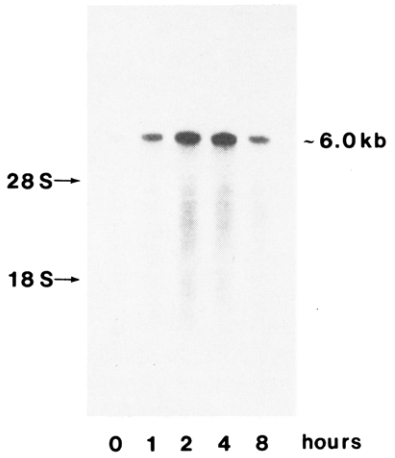

FIGURE 3: Northern blot analysis of RNA from PDGF-treated SMC. A total of 20 μg of RNA from control SMC, and SMC treated with PDGF for 1–8 h, was electrophoresed on 1.2% agarose and transferred to nitrocellulose paper. Blots were probed with the nick-translated 1.3-kb insert of λTS-33 as described under Materials and Methods. The positions of 28S and 18S ribosomal RNA markers were determined by ethidium bromide staining.

ferred to nitrocellulose, and the blot was probed with ^{32}P end-labeled poly[d(T)$_{13}$GT]. As shown in Figure 2, the poly(A) sequence in TS mRNA appeared to be located in the 450 bases separating the *Hin*dIII and *Acc*I sites, since only the *Hin*dIII–*Acc*I and *Hin*dIII–*Ava*I fragments hybridized to the probe.

Induction of TS mRNA by PDGF. The expression of TS is influenced by PDGF. It has been shown that TS protein increases in SMC treated with PDGF (Majack et al., 1985). We have explored whether this increase is a result of increased TS mRNA levels. SMC, made quiescent by culture in plasma-derived serum, were treated with PDGF at 5 ng/mL for 0–8 h, and total RNA was then prepared from cell layers. RNA was resolved on formaldehyde–agarose gels and transferred to nitrocellulose, and the blot was probed with the nick-translated 1.3-kb λTS-33 insert. As shown in Figure 3, there was a substantial induction of TS mRNA by PDGF. TS mRNA levels peaked 2–4 h after PDGF treatment and were substantially reduced 8 h after addition of the growth factor. This response agrees well with the kinetics of the response of metabolically labeled TS to PDGF (Majack et al., 1985). Thus it is likely that the increase in TS protein levels seen after PDGF treatment results principally, if not entirely, from increased levels of TS mRNA. Nuclear runon experiments are in progress to determine whether the observed increase in mRNA results from an increase in the rate of TS gene transcription or from increased mRNA stability.

Nucleotide Sequence of the 1.3-kb λTS-33 cDNA Insert. The strategy used to determine the DNA sequence of λTS-33 is shown in Figure 4; the nucleotide sequence of the 1290 bases comprising the insert, together with the translated amino acid sequence, is shown in Figure 5. The previously determined NH_2-terminal amino acid of TS (Asn) is found at amino acid 1. The subsequent 24 translated amino acids agree with the sequence determined on the heparin-binding fragment of TS by Coligan and Slayter (1984), with the exception of amino acid 23 for which the translated amino acid is Arg instead of Trp. Preceding the sequence of the mature protein is a typical highly hydrophobic signal sequence of 18 amino acids. The putative initiating ATG codon (bases 100–102) is preceded by the sequence ACC and followed by the sequence GG, providing a favorable context for initiation of translation as suggested by the consensus sequences compiled by Kozak (1984). Codon usage does not differ significantly from the most frequently used codons in translated human DNA (Lathe, 1985).

The sequence in Figure 5 includes 99 bases of 5′ untranslated sequence. Although translation is possible in all three reading frames, there are no ATG codons in this sequence; we therefore presume that translation starts with the synthesis of the putative signal sequence (amio acids −18 to −1). The open reading frame described in Figure 5 correctly predicts the amino-terminal sequence of mature thrombospondin (amino acids 1–25) as well as the sequence of 11 amino acids (241–251) previously determined by Galvin et al. (1985) at

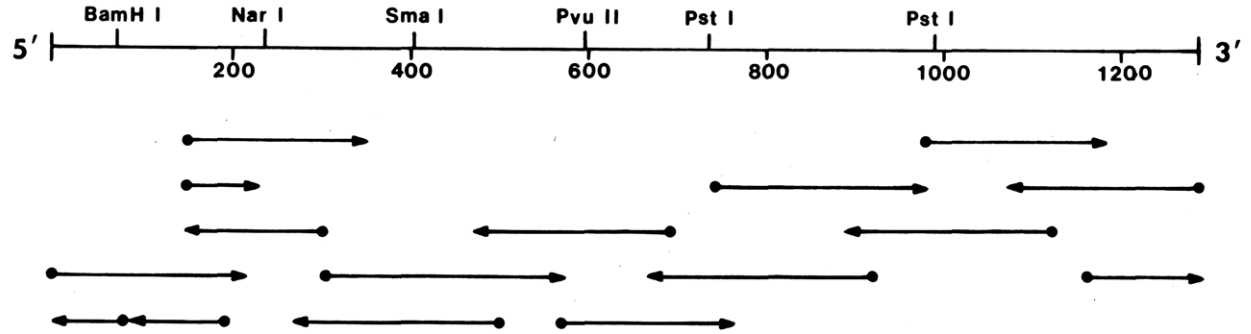

FIGURE 4: Sequencing strategy. The human TS clone λTS-33, 1.29 kb in length, was cloned into M13, and the sequence was obtained on both strands with single-stranded templates. The beginning of each arrow indicates the location of hybridization of a synthetic oligonucleotide used to prime the sequence reaction. Major restriction enzyme sites are shown.

```
                                                30                                  60                                  90
          GAATTCCGCG CCCCTCCAGC CCTCGCCGCC CTCGCCACCG CTCCCGGCCG CCGCGCTCCG GTACACACAG GATCCCTGCT GGGCACCAAC AGCTCCACC

                                  120                                 150                                 180
ATG GGG CTG GCC TGG GGA CTA GGC GTC CTG TTC CTG ATG CAT GTG TGT GGC ACC AAC CGC ATT CCA GAG TCT GGC GGA GAC AAC AGC GTG
Met Gly Leu Ala Trp Gly Leu Gly Val Leu Phe Leu Met His Val Cys Gly Thr Asn Arg Ile Pro Glu Ser Gly Gly Asp Asn Ser Val
-18                             -10                                  1                                  10

                                  210                                 240                                 270
TTT GAC ATC TTT GAA CTC ACC GGG GCC GCC CGC AAG GGG TCT GGG CGC CGA CTG GTG AAG GGC CCC GAC CCT TCC AGC CCA GCT TTC CGC
Phe Asp Ile Phe Glu Leu Thr Gly Ala Ala Arg Lys Gly Ser Gly Arg Arg Leu Val Lys Gly Pro Asp Pro Ser Ser Pro Ala Phe Arg
                                20                                  30                                  40

                                  300                                 330                                 360
ATC GAG GAT GCC AAC CTG ATC CCC CCT GTG CCT GAT GAC AAG TTC CAA GAC CTG GTG GAT GCT GTG CGG GCA GAA AAG GGT TTC CTC CTT
Ile Glu Asp Ala Asn Leu Ile Pro Pro Val Pro Asp Asp Lys Phe Gln Asp Leu Val Asp Ala Val Arg Ala Glu Lys Gly Phe Leu Leu
                                50                                  60                                  70

                                  390                                 420                                 450
CTG GCA TCC CTG AGG CAG ATG AAG AAG ACC CGG GGC ACG CTG CTG GCC CTG GAG CGG AAA GAC CAC TCT GGC CAG GTC TTC AGC GTG GTG
Leu Ala Ser Leu Arg Gln Met Lys Lys Thr Arg Gly Thr Leu Leu Ala Leu Glu Arg Lys Asp His Ser Gly Gln Val Phe Ser Val Val
                                80                                  90                                 110

                                  480                                 510                                 540
TCC AAT GGC AAG GCG GGC ACC CTG GAC CTC AGC CTG ACC GTC CAA GGA AAG CAG CAC GTG GTG TCT GTG GAA GAA GCT CTC CTG GCA ACC
Ser Asn Gly Lys Ala Gly Thr Leu Asp Leu Ser Leu Thr Val Gln Gly Lys Gln His Val Val Ser Val Glu Glu Ala Leu Leu Ala Thr
                               110                                 120                                 130

                                  570                                 600                                 630
GGC CAG TGG AAG AGC ATC ACC CTG TTT GTG CAG GAA GAC AGG GCC CAG CTG TAC ATC GAC TGT GAA AAG ATG GAG AAT GCT GAG TTG GAC
Gly Gln Trp Lys Ser Ile Thr Leu Phe Val Gln Glu Asp Arg Ala Gln Leu Tyr Ile Asp Cys Glu Lys Met Glu Asn Ala Glu Leu Asp
                               140                                 150                                 160

                                  660                                 690                                1020
GTC CCC ATC CAA AGC GTC TTC ACC AGA GAC CTG GCC AGC ATC GCC AGA CTC CGC ATC GCA AAG GGG GGC GTC AAT GAC AAT TTC CAG GGG
Val Pro Ile Gln Ser Val Phe Thr Arg Asp Leu Ala Ser Ile Ala Arg Leu Arg Ile Ala Lys Gly Gly Val Asn Asp Asn Phe Gln Gly
                               170                                 180                                 190

                                  750                                 780                                 810
GTG CTG CAG AAT GTG AGG TTT GTC TTT GGA ACC ACA CCA GAA GAC ATC CTC AGG AAC AAA GGC TGC TCC AGC TCT ACC AGT GTC CTC CTC
Val Leu Gln Asn Val Arg Phe Val Phe Gly Thr Thr Pro Glu Asp Ile Leu Arg Asn Lys Gly Cys Ser Ser Ser Thr Ser Val Leu Leu
                               200                                 210                                 220

                                  840                                 870                                 900
ACC CTT GAC AAC AAC GTG GTG AAT GGT TCC AGC CCT GCC ATC CGC ACT AAC TAC ATT GGC CAC AAG ACA AAG GAC TTG CAA GCC ATC TGC
Thr Leu Asp Asn Asn Val Val Asn Gly Ser Ser Pro Ala Ile Arg Thr Asn Tyr Ile Gly His Lys Thr Lys Asp Leu Gln Ala Ile Cys
                               230                                 240                                 250

                                  930                                 960                                 990
GGC ATC TCC TGT GAT GAG CTG TCC AGC ATG GTC CTG GAA CTC AGG GGC CTG CGC ACC ATT GTG ACC ACG CTG CAG GAC AGC ATC CGC AAA
Gly Ile Ser Cys Asp Glu Leu Ser Ser Met Val Leu Glu Leu Arg Gly Leu Arg Thr Ile Val Thr Thr Leu Gln Asp Ser Ile Arg Lys
                               260                                 270                                 280

                                 1020                                1050                                1080
GTG ACT GAA GAG AAC AAA GAG TTG GCC AAT GAG CTG AGG CGG CCT CCC CTA TGC TAT CAC AAC GGA GTT CAG TAC AGA AAT AAC GAG GAA
Val Thr Glu Glu Asn Lys Glu Leu Ala Asn Glu Leu Arg Arg Pro Pro Leu Cys Tyr His Asn Gly Val Gln Tyr Arg Asn Asn Glu Glu
                               290                                 300                                 310

                                 1110                                1140                                1170
TGG ACT GTT GAT AGC TGC ACT GAG TGT CAC TGT CAG AAC TCA GTT ACC ATC TGC AAA AAG GTG TCC TGC CCC ATC ATG CCC TGC TCC AAT
Trp Thr Val Asp Ser Cys Thr Glu Cys His Cys Gln Asn Ser Val Thr Ile Cys Lys Lys Val Ser Cys Pro Ile Met Pro Cys Ser Asn
                               320                                 330                                 340

                                 1200                                1230                                1260
GCC ACA GTT CCT GAT GGA GAA TGC TGT CCT CGC TGT TGG CCC AGC GAC TCT GCG GAC GAT GGC TGG TCT CCA TGG TCC GAG TGG ACC TCC
Ala Thr Val Pro Asp Gly Glu Cys Cys Pro Arg Cys Trp Pro Ser Asp Ser Ala Asp Asp Gly Trp Ser Pro Trp Ser Glu Trp Thr Ser
                               350                                 360                                 370

                                 1290
TGT TCT ACG AGC TGT GGC AAT
Cys Ser Thr Ser Cys Gly Asn
```

FIGURE 5: Nucleotide sequence of cDNA clone λTS-33 and its predicted amino acid sequence. The amino acid sequence corresponding to the mature protein is numbered from 1 to 379, and the amino acid sequence that corresponds to the signal peptide is represented by minus numbers. The overlined sequence corresponds to the sequence recognized by a mixture of four 42-mers synthesized on the basis of the amino acid sequence published by Coligan and Slayter (1984). The amino acid sequence within the box shows a high degree of homology with malarial circumsporozoite proteins.

the NH_2 terminus of a chymotrypsin-produced fragment of TS (M_r 120000 per chain). The cDNA cloned in λTS-33 can therefore be identified unequivocally as coding for human TS. Mumby et al. (1984) have identified a chymotrypsin-produced fragment of TS (M_r 70000 per chain) that retains an affinity for type V collagen and is interchain disulfide bonded. This fragment has been placed on the COOH-terminal side of the NH_2-terminal heparin-binding domain in the studies of Galvin et al. (1985), and a similar fragment was identified by Lawler et al. (1985). The amino-terminal heparin-binding domain contains approximately 275 amino acids and very little cysteine (Dixit et al., 1984). The amino acid sequence in Figure 5 supports this composition since only four cysteines are present in the first 299 amino acids, whereas the last 80 amino acids

Thrombospondin	368	SER	GLU	TRP	THR	SER	CYS	SER	THR	SER	CYS	GLY	ASN	379
P. falciparum	344	THR	GLU	TRP	SER	PRO	CYS	SER	VAL	THR	CYS	GLY	ASN	355
P. knowlesi	296	THR	GLU	TRP	THR	PRO	CYS	SER	VAL	THR	CYS	GLY	ASN	307

FIGURE 6: Comparison of homologous amino acid sequences from thrombospondin (this work; amino acids 368–379, Figure 5), the circumsporozoite protein of *P. falciparum* (Dame et al., 1984; amino acids 344-355), and the circumsporozoite protein of *P. knowlesi* (Ozaki et al., 1983; amino acids 296–307). Boxed amino acids are identical in all three proteins, and underlined amino acids are identical in two of the three proteins with a conservative substitution in the third protein.

in the sequence contain 12 cysteines. The amino acid sequence 300–379 (Figure 5) may therefore represent the region of a proposed interchain disulfide bond knot in TS (Galvin et al., 1985; Lawler et al., 1985).

Discussion

The antibody-identified and synthetic oligomer identified cDNA clones described in this study are presumed to represent the 5′ and 3′ regions of the human TS mRNA, respectively. Together they encompass approximately two-thirds of the coding sequence in this mRNA. The 5′ clone, λTS-33, for which the sequence is given in Figure 5, can be identified unambiguously as a TS cDNA clone since the DNA sequence predicts previously determined TS amino acid sequences. The identity of the 3′ clones is based on their isolation from a cDNA expression library with highly specific anti-TS antibodies and their hybridization to an mRNA with the same apparent molecular weight (~6.0 kb) as that identified by documented 5′ TS clones in Northern blot analyses. Since only one size species of mRNA can be identified by probes representing both the 5′ and 3′ ends of the mRNA, we can conclude that alternate splicing of mRNA precursors does not occur or, if it does, that the final products are of similar size.

It is of interest that almost 500 positive clones were identified in an initial screen of approximately 1×10^6 plaques with the mixture of 42-mers indicated in Table I. Thus, TS appears to be a relatively high abundance mRNA in human endothelial cells. The mixture of synthetic 42-mers, derived according to the simplifying assumptions outlined in Table I, is identical with the DNA sequence of bases 186–227 (overlined in Figure 5) in 31 of 42 positions. Bases 187–209 provide an exact fit with only one mismatch plus the substitution of one deoxyinosine for deoxyguanine. This accounts for hybridization to TS mRNA even under conditions of high stringency.

Human umbilical vein endothelial cells grown in the presence of serum produce substantial amounts of TS (Sage & Bornstein, 1982). Nevertheless, the relatively high levels of TS mRNA present in the RNA used to construct the human umbilical vein endothelial cell library probably did not result from the action of PDGF since these cells lack PDGF receptors (Bowen-Pope et al., 1985). Umbilical vein and other endothelial cells do not require exogenous PDGF for growth, although these cells release growth-promoting activity into their culture medium, part of which is neutralized by monospecific antibodies to PDGF [reviewed by Ross et al. (1986)]. TS gene expression may therefore be constitutively elevated in umbilical vein endothelial cells.

In contrast, as shown in Figure 3 (lane 1), TS mRNA is nearly undetectable in quiescent rat aortic SMC, whereas PDGF treatment resulted in a rapid but transient increase (lanes 2–5). Peak levels of mRNA were achieved between 2 and 4 h after addition of PDGF. This time course of induction parallels the increase in TS protein production observed in these cells after PDGF treatment (Majack et al., 1985). The kinetics of TS mRNA induction (or stabilization) by PDGF differ from those of a number of growth-related mRNAs that are maximally induced within 20–30 min after exposure of cells to serum or growth factor. These mRNAs, which have been termed "immediate early" RNAs, include c-*myc* and c-*fos* and a large number of uncharacterized genes (Lau & Nathans, 1985). Immediate early RNAs share the property of being superinducible when cycloheximide is added together with the growth factor; i.e., inhibition of protein synthesis appears to enhance transcription or stabilize this mRNA [see Lau and Nathans (1985) for a discussion]. In contrast, preliminary experiments (S. Kobayashi and P. Bornstein unpublished results) suggest that the presence of cycloheximide prevents the appearance of the elevated TS mRNA levels observed with PDGF. PDGF may therefore stimulate the synthesis or activity of a protein whose presence is required for elevated TS mRNA levels. We have recently isolated TS genomic clones and plan to investigate this question at the level of the TS gene.

Although only a third or less of the coding sequence for TS has been established in this study, we have attempted to determine whether significant homologies exist with other known DNA and protein sequences. No significant homologies were found with other matrix and heparin-binding proteins such as fibronectin and vitronectin. In particular, the internal homologies observed in the amino acid sequence of the fibronectin chain (Petersen et al., 1983) do not occur in this region of TS. A surprising homology was discovered between the last 12 amino acids coded by the λTS-33 cDNA clone and amino acid sequences found in the immunodominant circumsporozoite surface proteins of the malarial parasites *Plasmodium falciparum* and *Plasmodium knowlesi*. These sequences are shown in Figure 6. Of the 12 amino acids, 7 (including 2 cysteines and 1 tryptophan) are identical with amino acids in the sporozoite proteins, and of the remaining 5, two represent serine for threonine substitutions. Although limited regions of the proteins are involved, these homologies may be significant since this sequence is one of only two that are highly conserved when the *P. falciparum* and *P. knowlesi* proteins are compared. It has been suggested that this conservation may reflect an important biological function such as adherence to liver cells prior to invasion by the malarial parasite (Dame et al., 1984). Conceivably, the corresponding region in TS may be involved in binding of this protein to cell surfaces.

A possibly related finding by Roberts et al. (1985b) presents evidence that TS may mediate cytoadherence of erythrocytes that have been parasitized by *P. falciparum*. In infected erythrocytes, malarial parasites are in the form of trophozoites and schizonts, whereas the circumsporozoite proteins are expressed only in the sporozoite stage of the life cycle of the parasite (Roberts et al., 1985b; Ozaki et al., 1983). The ability of TS to bind to parasitized erythrocytes has been correlated with the presence of a high molecular weight malarial protein, which appears on the surface of the infected red blood cells and is responsible for the "knobby" phenotype (Leech et al., 1984; Roberts et al., 1985b). Conceivably, the ability of TS to bind to red cells containing the knobby (K^+) strain of

malarial parasites is unrelated to the homology that we describe with the sporozoite proteins. Further work is clearly required to clarify this point and to ascertain whether the homology to malarial proteins is related to the interaction of TS with fibroblasts and other cells.

Shortly prior to submission of this paper for publication, Dixit et al. (1986) reported the characterization of a partial human TS cDNA clone. The sequence reported by Dixit et al. (1986) is identical in its coding region with that reported by us but lacks the DNA coding for the last 25 amino acids in Figure 5 and therefore lacks the region that demonstrates homology to the sporozoite proteins. A minor discrepancy exists in the 5′ untranslated region in that the sequence for nucleotides 43–48 is reported by Dixit et al. to be CCCGGG. The sequence CCCGGC was shown to be correct by selective restriction enzyme digestion. Finally, Dixit et al. identified an mRNA of ~8 kb in human fibroblast RNA by Northern blot analysis. The discrepancy with a size of ~6 kb reported in this paper for rat aortic SMC RNA may reflect cell and species differences or may have a technical basis.

Added in Proof

Further analysis of the partial amino acid sequence of TS reveals a significant homology to a sequence in the globular NH_2-terminal domain of human type I procollagen (Chu et al., 1984). With the introduction of one gap in the TS sequence, amino acids 300–354 (Figure 5) align with amino acids 40–95 in procollagen. Of 55 amino acids 22 are identical, including all 10 cysteines. This finding strongly suggests a similarity in protein folding in these regions of the two proteins.

Acknowledgments

We thank Susan C. Cook, David Roeder, and Caroline Wilson for expert technical assistance, Dr. John McKay for helpful suggestions, and Donna Stewart for preparation of the manuscript.

References

Asch, A. S., Leung, L. L. K., Shapiro, J., & Nachman, R. L. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 2904–2908.

Bird, A. P. (1981) *Nucleic Acids Res. 8*, 1499–1504.

Bowen-Pope, D. F., Seifert, R. A., & Ross, R. (1985) in *Control of Animal Cell Proliferation* (Boynton, A. L., & Lefferts, H. L., Eds.), Vol. 1, pp 281–312, Academic, New York.

Buell, G. N., Wickens, M. P., Payvar, F., & Schimke, R. T. (1978) *J. Biol. Chem. 253*, 2471–2482.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., & Rutter, W. J. (1979) *Biochemistry 18*, 5294–5299.

Chu et al. (1984) *Nature (London) 310*, 337.

Coligan, J. E., & Slayter, H. S. (1984) *J. Biol. Chem. 259*, 3944–3948.

Dame, J. B., Williams, J. L., McCutchan, T. F., Weber, J. L., Wirtz, R. A., Hockmeyer, W. T., Maloy, W. L., Hayness, J. D., Schneider, I., Roberts, D., Sanders, G. S., Reddy, E. P., Diggs, C., & Miller, L. H. (1984) *Science (Washington, D.C.) 225*, 593–599.

Dixit, V. M., Grant, G. A., Santoro, S. A., & Frazier, W. A. (1984) *J. Biol. Chem. 259*, 10100–10105.

Dixit, V. M., Haverstick, D. M., O'Rourke, K. M., Hennessy, S. W., Grant, G. A., Santoro, S. A., & Frazier, W. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 3472–3476.

Dixit, V. M., Hennessy, S. W., Grant, G. A., Rotwein, P., & Frazier, W. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5449–5453.

Galvin, N. J., Dixit, V. M., O'Rourke, K. M., Santoro, S. A., Grant, G. A., & Frazier, W. A. (1985) *J. Cell Biol. 101*, 1434–1441.

Kozak, M. (1984) *Nucleic Acids Res. 12*, 857–872.

Lathe, R. (1985) *J. Mol. Biol. 183*, 1–12.

Lau, L. F., & Nathans, D. (1985) *EMBO J. 4*, 3145–3151.

Lawler, J., Derick, L. H., Connolly, J. E., Chen, J.-H., & Cho, F. C. (1985) *J. Biol. Chem. 260*, 3762–3772.

Leech, J. H., Barnwell, J. W., Miller, L. H., & Howard, R. J. (1984) *J. Exp. Med. 159*, 1567–1575.

Leung, L. L. K. (1984) *J. Clin. Invest. 74*, 1764–1772.

Majack, R. A., & Clowes, A. W. (1984) *J. Cell. Physiol. 118*, 253–256.

Majack, R. A., & Bornstein, P. (1986) in *Cell Membranes: Methods and Reviews*, Plenum, New York (in press).

Majack, R. A., Cook, S. C., & Bornstein, P. (1985) *J. Cell Biol. 101*, 1059–1070.

Majack, R. A., Cook, S. C., & Bornstein, P. (1986) *Proc. Natl. Acad. Sci. U.S.A.* (in press).

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning, A. Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Mumby, S. M., Raugi, G. J., & Bornstein, P. (1984) *J. Cell Biol. 98*, 646–652.

Ohtsuka, E., Matsuki, S., Ikehara, M., Takahashi, Y., & Matsubara, K. (1985) *J. Biol. Chem. 160*, 2605–2608.

Ozaki, L. S., Svec, P., Nussenzweig, R. S., Nussenzweig, V., & Godson, G. N. (1983) *Cell (Cambridge, Mass.) 34*, 815–822.

Petersen, T. E., Thøgersen, H. C., Skorstengaard, K., Vibe-Petersen, K., Sahl, P., Sottrup-Jensen, L., & Magnusson, S. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 137–141.

Raines, E. W., & Ross, R. (1982) *J. Biol. Chem. 257*, 5154–5160.

Raugi, G. J., Mumby, S. M., Ready, C. A., & Bornstein, P. (1984) *Thromb. Res. 36*, 165–175.

Ricca, G. A., Hamilton, R. W., McLean, J. W., Conn, A., Kalinyak, J. E., & Taylor, J. M. (1981) *J. Biol. Chem. 256*, 10362–10368.

Roberts, D. D., Haverstick, D. M., Dixit, V. M., Frazier, W. A., Santoro, S. A., & Ginsberg, V. (1985a) *J. Biol. Chem. 260*, 1059–1070.

Roberts, D. D., Sherwood, J. A., Spitalnik, S. L., Panton, L. J., Howard, R. J., Dixit, V. M., Frazier, W. A., Miller, L. H., & Ginsburg, V. (1985b) *Nature (London) 318*, 64–66.

Ross, R., Raines, E. W., & Bowen-Pope, D. F. (1986) *Cell (Cambridge, Mass.) 46*, 155–169.

Sage, H., & Bornstein, P. (1982) *Arteriosclerosis 2*, 27–36.

Sanchez-Pescador, R., & Urdea, M. S. (1984) *DNA 3*, 339–343.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H., & Roe, B. A. (1980) *J. Mol. Biol. 143*, 161–178.

Silverstein, R. L., Leung, L. L. K., Harpel, P. C., & Nachman, R. L. (1985a) *J. Clin. Invest. 75*, 2065–2073.

Silverstein, R. L., Nachman, R. L., Leung, L. L. K., & Harpel, P. C. (1985b) *J. Biol. Chem. 260*, 10346–10352.

Silverstein, R. L., Leung, L. L. K., & Nachman, R. L. (1986) *Arteriosclerosis 6*, 245–253.

Tamkun, J. W., Schwarzbauer, J. E., & Hynes, R. O. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 5140–5144.

Wickens, M. P., Buell, G. N., & Schimke, R. T. (1978) J. Biol. Chem. 253, 2483–2495.
Wight, T. N., Raugi, G. J., Mumby, S. M., & Bornstein, P. (1985) J. Histochem. Cytochem. 33, 295–302.
Young, R. A., & Davis, R. W. (1983a) Science (Washington, D.C.) 222, 778–782.
Young, R. A., & Davis, R. W. (1983b) Proc. Natl. Acad. Sci. U.S.A. 80, 1194–1198.

Expression of Apoaequorin Complementary DNA in *Escherichia coli*†

Satoshi Inouye,‡,§ Yoshiyuki Sakaki,‡ Toshio Goto,‖ and Frederick I. Tsuji*,⊥

Research Laboratory for Genetic Information, Kyushu University, Fukuoka 812, Japan, Department of Agricultural Chemistry, Nagoya University, Nagoya 464, Japan, Marine Biology Research Division, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093, and Veterans Administration Medical Center Brentwood, Los Angeles, California 90073

Received March 24, 1986; Revised Manuscript Received July 29, 1986

ABSTRACT: Aequorin is a photoprotein that reacts with Ca^{2+} to yield light, CO_2, and a blue fluorescent protein. The blue fluorescent protein is dissociable into apoaequorin and a small organic molecule, coelenteramide. Cloning and sequence analyses of the cDNA for apoaequorin show that it is made up of 189 amino acid residues and contains three Ca^{2+}-binding sites. Various expression plasmids for apoaequorin cDNA were constructed and expressed in *Escherichia coli*. One plasmid, piQ5, was found to give excellent expression in *E. coli* D1210 and BNN103. Best expression was obtained when the growth temperature was shifted from 37 to 42 °C immediately after the addition of inducer (isopropyl β-D-thiogalactopyranoside) to the culture medium. Aequorin was regenerated from the expressed apoaequorin by incubating with coelenterazine, 2-mercaptoethanol, and ethylenediaminetetraacetic acid. The results show that large amounts of apoaequorin may be prepared by the procedure described.

Aequorin is a chromophore-containing protein present in the outer margin of the umbrella of the jellyfish, *Aequorea victoria* (Shimomura et al., 1962, 1963). When mixed with a trace amount of Ca^{2+}, aequorin undergoes an intramolecular oxidation reaction, yielding as products light (λ_{max} = 470 nm), CO_2, and a blue fluorescent protein [Shimomura et al., 1962, 1963, 1974; for a review, see Johnson and Shimomura (1978)]. Because aequorin has a high specificity for Ca^{2+}, it has been used as a biological indicator of Ca^{2+} (Blinks et al., 1978). The blue fluorescent protein may be dissociated into apoaequorin (apoprotein) and coelenteramide by gel filtration or by treatment with acid or ether (Shimomura & Johnson, 1976). Coelenteramide, the product of the oxidation of the chromophore coelenterazine, serves as the emitter in the reaction (Shimomura & Johnson, 1973). Aequorin may be regenerated from apoaequorin by incubation with coelenterazine, 2-mercaptoethanol (2-ME),[1] and ethylenediaminetetraacetic acid (EDTA) (Shimomura & Johnson, 1975).

Recently, the cDNA for apoaequorin has been cloned and the primary structure of the protein deduced from the nucleotide sequence (Inouye et al., 1985). The primary structure has also been determined by sequencing the protein (Charbonneau et al., 1985). Apoaequorin is composed of 189 amino acid residues, has a molecular weight of 21 400, and has three EF hand structures that are characteristic of Ca^{2+}-binding sites. Electron paramagnetic resonance and proton nuclear magnetic resonance investigations have also been carried out on aequorin in order to understand the molecular mechanism of the light-emitting reaction (Kemple et al., 1984; Ray et al., 1985). Although crystal structure studies have not yet been performed on aequorin, recently several Ca^{2+}-binding proteins have undergone such studies (Herzberg & James, 1985a,b; Babu et al., 1985; Sundralingam et al., 1985). In order to carry out analogous studies on aequorin, and even to develop better methods for using aequorin in Ca^{2+} assay, relatively larger amounts of aequorin are necessary than are presently available. Site-directed mutagenesis studies also require amino acid substitutions to be made and the modified protein produced in sufficient amounts for further study. In order to meet these requirements, a study of the expression of cDNA for apoaequorin in *Escherichia coli* was undertaken. This paper describes procedures for obtaining good expression of the cDNA and for regenerating aequorin from apoaequorin.

Materials and Methods

Bacterial Strains and Plasmids. The *E. coli* strains used were JM83 (Vieira & Messing, 1982), D1210 carrying *lac*I^q and *lac*^{Y} (deBoer et al., 1983), and a protease-deficient strain BNN103 (Young & Davis, 1983). The plasmids employed were pUC9 (Vieira & Messing, 1982) and pUC9-1 and pUC9-2 (Hanna et al., 1984). Plasmid pDR540 (Russell & Bennet, 1982) was used as a source of the *tac* promoter and was obtained from P-L Biochemicals.

Enzymes and Chemicals. All restriction endonucleases, *E. coli* T_4 DNA ligase, Klenow enzyme (from DNA polymerase I), and T_4 polynucleotide kinase were purchased from Takara

†This work was supported in part by a research grant (DMB 85-17578) from the National Science Foundation to F.I.T. and by a grant-in-aid (60223023) from the Ministry of Education, Culture and Sciences of Japan to Y.S.

‡Kyushu University.
§On leave of absence from the Chisso Corporation, Tokyo, Japan.
‖Nagoya University.
⊥University of California, San Diego, and Veterans Administration Medical Center Brentwood.

[1] Abbreviations: EDTA, ethylenediaminetetraacetic acid; Tris, tris(hydroxymethyl)aminomethane; PIPES, piperazine-*N*,*N*′-bis(2-ethanesulfonic acid); 2-ME, 2-mercaptoethanol; IPTG, isopropyl β-D-thiogalactopyranoside; bp, base pair(s).

0006-2960/86/0425-8425$01.50/0