Molecular Cloning and Characterization of Human Fetal Liver Tropomodulin

A TROPOMYOSIN-BINDING PROTEIN*

(Received for publication, June 24, 1991)

Lanping Amy Sung‡‡, Velia M. Fowler††, Karel Lambert‡, Mark A. Sussman⁎⁎, David Karr† ††, and Shu Chien‡

From the Departments of Applied Mechanics and Engineering Sciences-Bioengineering and Medicine, and Center for Molecular Genetics, University of California at San Diego, La Jolla, California 92037-0643 and the Department of Molecular Biology, Research Institute of Scripps Clinic, La Jolla, California 92037

Human erythrocyte tropomodulin is a novel tropomyosin regulatory protein that binds to the end of erythrocyte tropomyosin and blocks head-to-tail association of tropomyosin along actin filaments. It has been proposed to play a role in modulating the association of tropomyosin with the spectrin-actin complex in the erythrocyte membrane skeleton. Immunoscreening of a human fetal liver cDNA expression library in λgt11, followed by 5‘-end extension by polymerase chain reaction from the same library, yielded a composite cDNA sequence of 2665 base pairs (bp). It contains a 34-bp 5‘-untranslated region, a 1.6-kilobase (kb) 3‘-untranslated region, and a complete open reading frame of 1077 bp that encodes a protein of 359 amino acids with a calculated molecular mass of 40.6 kDa and a pI of 4.8. Authenticity of the tropomodulin cDNA was confirmed by a complete sequence match of 49 predicted amino acids with the sequences of three tryptic peptides of the erythrocyte tropomodulin. The sequence has no internal repeats and no significant homology with any known proteins. Secondary structure predictions indicate that tropomodulin may consist of a series of seven or eight short α-helical segments and fold into a somewhat compact shape. The tropomyosin binding activity has been mapped to an N-terminal region containing residues 39–138. Nine independent PCR clones, five from a human reticulocyte cDNA library and four from the fetal liver cDNA library, revealed identical N-terminal 103 amino acids, suggesting that the sequence reported here may also be of erythrocyte tropomodulin. Northern analysis of human reticulocyte RNA showed two hybridizing bands of 2.7 and 1.6 kb, indicating that the 2665-bp cDNA sequence reported here was that of the longer transcript.

The membrane skeletal network of human erythrocytes is composed of several interacting proteins that provide mechanical stability and flexibility for the cell membrane. The major proteins in the erythrocyte membrane skeletal network are α- and β-spectrin, ankyrin, band 3, protein 4.1, protein 4.2, actin, and bands 6 and 7, most of which have been cloned in recent years (for review, see Refs. 1 and 2). Counterparts of some of the key proteins in muscle cells (myosin, tropomyosin) have also been identified in the erythrocyte membrane skeleton (3–5). These contractile proteins are relatively minor components, and their functions remain largely unknown and pose interesting questions. In skeletal muscle, a complex of three troponin polypeptides (troponins I, T, and C) binds to tropomyosin and regulates actomyosin interaction during muscle contraction (for review, see Ref. 6). To date, however, analogs of the muscle troponins I and T have not been identified in non-muscle cells.

A tropomyosin-binding protein with an apparent molecular weight (M,) of 43,000 has been identified and purified recently from human erythrocyte membranes (7, 8). This protein has been named tropomodulin based on its ability to inhibit tropomyosin binding to actin (8). Tropomodulin is associated with the membrane skeleton after extraction with the nondenaturing detergent, Triton X-100, and is present at a ratio of one M, 43,000 molecule per two tropomyosin molecules (7). Unlike previously described proteins that inhibit tropomyosin binding to actin, tropomodulin itself does not bind directly to actin. Electron microscopy of rotary-shadowed tropomodulin-tropomyosin complexes reveals that tropomodulin binds to (one of) the ends of the rodlike erythrocyte tropomyosin molecules (8). This, together with tropomodulin ability to abolish cooperative binding of tropomyosin to actin, suggests that tropomodulin weakens tropomyosin-actin interactions by blocking head-to-tail association of tropomyosin molecules along the actin filament. It is conceivable that by modulating the association of tropomyosin with spectrin-actin complexes, tropomodulin may play a role in determining the viscoelastic properties of the erythrocyte membrane.

Erythrocyte tropomodulin may be the first member of a new family of tropomyosin-binding proteins, since antibody to tropomodulin cross-reacts with M, 43,000 polypeptides in non-erythroid cells and tissues such as muscle, brain, lens,
neutrophils, and endothelial cells (8). Antibodies to tropomodulin also cross-react strongly with rat and rabbit muscle troponin I, but not with troponin T (8), raising the possibility that tropomodulin may be a non-muscle counterpart of muscle troponin I.

Here we report the molecular cloning and sequencing of human tropomodulin from a human fetal liver cDNA library. Human fetal liver is a blood producing organ in early life, and our evidence suggests that the sequence reported here may also be that of erythrocyte tropomodulin. This work has been reported previously in abstract form (32). The homology search indicates that tropomodulin does not belong to any existing family of tropomyosin-binding proteins, including muscle troponins I and T. Tropomodulin, therefore, appears to be a new protein that functions to regulate tropomyosin-actin associations in erythrocytes and, perhaps, in other cells as well. We have mapped the tropomyosin-binding domain to an N-terminal region of tropomodulin (amino acids 39–138) using fusion protein constructs expressed in bacteria. The availability of the cDNA for tropomodulin will allow a more detailed characterization of the tropomyosin-binding site, as well as investigations of the tissue distribution, expression and function(s) of tropomodulin, and its possible isoforms.

**EXPERIMENTAL PROCEDURES**

Screening of λgt11 cDNA Library—Affinity-purified rabbit-antihuman erythrocyte tropomodulin IgG (8) was used to screen a human fetal liver cDNA expression library in λgt11 provided by Dr. B. G. Forget (9, 10). Immunoscreening of the λgt11 library was performed according to Huynh et al. (11), except that positive clones were identified with goat-anti-rabbit IgG conjugated with horseradish peroxidase (Bio-Rad). For each 150-mm Petri dish, 5 × 10⁶ plating units were used. Immunoreactive clones were further tested with two anti-peptide antibodies described below.

Anti-peptide Antibodies—Purified erythrocyte tropomodulin (7) was digested with trypsin. The tryptic peptides were separated on high pressure liquid chromatography, and three peptides (1, 2, and 3) were sequenced by Bill Lane (Harvard Biological Laboratories). Two synthetic 15-mer peptides (1B and 2B), designed according to the C-terminal amino acid sequences of peptides 1 and 2 (see Figs. 1 and 2), were conjugated to hemocyanin and used as immunogens to raise antibodies in rabbits (12). Antibodies were affinity-purified using purified tropomodulin-Sepharose as described (13) and were shown to be monospecific for tropomodulin on blots of whole erythrocyte and DELDPDNALLPAGLR, respectively.

Verification of cDNA Clone by Primer Extension Sequencing—The C-terminal seven amino acids of peptide 1 (25 amino acids long) were used to design a mixed oligonucleotide primer (see Fig. 2, single dashed line with arrowhead). The cDNA insert from λ phase clone 10 (see Fig. 1) was subcloned into pBlueScript II SK (–) plasmid (Stratagene, La Jolla, CA) and used as the template. The sequencing primer was annealed to clone 10 DNA at 45°C for 30 min. DNA sequencing reactions (14) were carried out under conditions optimized to reveal the amino acid sequence adjacent to the primer.

Subcloning and Sequence Analysis—cDNA inserts from all four positive phage clones were subcloned into pBlueScript II SK (–). Unidirectional deletion clones of clone 10 were generated using exonuclease III (15). All cDNA fragments were sequenced using T3 and T7 primers by the dideoxy chain termination method (14). Sequence analysis and GenBank data base searches were performed using IBI MacVector sequence analysis software (International Biotechnologies, Inc., New Haven, CT). An on-line search of the National Biomedical Research Foundation Protein Identification Resource protein sequence data base was also performed.

5'-End Extension of cDNA by PCR—The additional 5′-end sequence was obtained from cDNA libraries by polymerase chain reaction (PCR) using a pair of primers: a tropomodulin primer and a λgt11 primer. The two libraries used were the human fetal liver cDNA library (9, 10), from which clone 10 was obtained, and a sickle cell patient reticulocyte cDNA library (16). The tropomodulin primer corresponded to nucleotides 174–194 of clone 10 (see Fig. 2, double dashed line with arrowhead, antisense orientation) with the addition of an EcoRI site at the 5′-end plus two additional nucleotides (TG). The λgt11 primer (a 15-mer upstream primer, obtained from Clontech, Palo Alto, CA) has a sequence of 5′-GACTCTTGGACCCG-3′.

**RESULTS AND DISCUSSION**

Isolation of Tropomodulin cDNA Clones—Immunoscreening of about 5 × 10⁵ recombinant phages of a human fetal liver cDNA library with affinity-purified anti-human erythrocyte tropomodulin antibody yielded four immunoreactive clones. The sizes of these inserts were 2.5 kb (clone 10), 1.9 kb (clone 5), 1.2 kb (clone 7), and 0.9 kb (clone 1) (Fig. 1). These four clones cross-hybridized with one another. Restriction and partial nucleotide sequence analyses of these clones further demonstrated that they were overlapping clones of the same cDNA (Fig. 1).

These four immunoreactive clones were further tested for reactivity with two additional antibodies that were made against synthetic peptides corresponding to sequences of two different erythrocyte tropomodulin tryptic peptides (peptides 1B and 2B, see "Experimental Procedures"). Clone 10, the longest λgt11 expression cDNA clone was the only one that reacted strongly with both anti-peptide antibodies. Therefore, this clone was chosen for initial nucleotide sequencing to verify that its cDNA corresponded to the authentic tropomodulin cDNA.

Verification of Tropomodulin Clone—Primer extension sequencing was used to verify the authenticity of clone 10 before the entire 2.5-kb sequence was determined. A mixed oligonu-

---

1 The abbreviations used are: PCR, polymerase chain reaction; kb, kilobase(s); bp, base pair(s).
The complete amino acid sequence of human tropomodulin is shown in Fig. 2. The entire 2.5 kb of clone 10 was then sequenced using Tc and Tp primers as described (14).

**5'-End Extension of cDNA**—The most N-terminal amino acid sequence predicted from the cDNA sequence of clone 10 matches the C-terminal 11 amino acids of the 21-amino acid erythrocyte tropomodulin peptide 2 (Fig. 2), clearly indicating the absence of a 5'-terminal sequence in clone 10. The additional 5'-sequence beyond clone 10 was obtained by PCR amplification from cDNA libraries using a tropomodulin-specific primer (see Fig. 2) and a λgt11 primer. Five PCR clones from a human reticulocyte cDNA library had identical 342 nucleotide sequences (encoding N-terminal 103 amino acids) and had an expected 194-bp overlap with the 5'-end of clone 10. The sequences of four independent PCR clones amplified from the human fetal liver cDNA library were also identical in sequence to the reticulocyte PCR clones, except that the fetal liver clones were shorter at the 5'-end by 32, 32, 41, and 57 bp. The two fetal liver clones starting at nucleotide -2 (two nucleotides upstream from the putative initiator ATG codon, nucleotides 1-3, see below) would encode peptides of the same size, whereas the two cloned starting at nucleotides 7 or 23 could potentially encode a smaller protein of 25 kDa that begins at Met-137 (Fig. 2). However, we have not detected an immunoreactive 25-kDa protein in Western blots of mature human erythrocytes (7) or developing erythroblasts isolated from the spleens of anemic mice. These results suggest that the difference in lengths of the PCR clones isolated from the fetal liver library may reflect the length heterogeneity of existing clones in the library and that the tropomodulin proteins expressed in adult reticulocytes and erythropoietic fetal liver are the same.

**Sequence Analysis of Tropomodulin cDNA**—The complete nucleotide sequence of the tropomodulin cDNA and the derived amino acid sequence are shown in Fig. 2. A comparison of this amino acid sequence with the sequences of three erythrocyte tropomodulin trypic peptides demonstrated complete identity (Fig. 2, boxed regions), further confirming the authenticity of this cDNA. The sequence contains 2665 bp of nucleotide sequence of the tropomodulin cDNA and the deduced amino acid sequence are shown in Fig. 2. The position where the primer sequence was derived for PCR extension is double-underlined with an arrowhead. Position where the mixed oligonucleotide primer was derived (from peptide 1) for initial verification sequencing is single-underlined with an arrowhead. 2B and 1B are labeled with dotted underlines. The position where the primer sequence was derived for PCR extension is double-underlined with an arrowhead. Position where the mixed oligonucleotide primer was derived (from peptide 1) for initial verification sequencing is single-underlined with an arrowhead.

**Fig. 2.** The nucleotide sequence of human tropomodulin cDNA and its deduced amino acid sequence. The first nucleotide of the putative ATG initiation codon is designated as nucleotide position 1. The deduced amino acid sequence is shown using the single letter code and is numbered starting at the initiation codon. Boxed residues are those matched with the amino acid sequences obtained from three trypsin-cleaved peptides of purified erythrocyte tropomodulin (1, 2, 3) from N to C terminus. Peptides 1B and 2B are labeled with dotted underlines. The position where the primer sequence was derived for PCR extension is double-underlined with an arrowhead. Position where the mixed oligonucleotide primer was derived (from peptide 1) for initial verification sequencing is single-underlined with an arrowhead. 2B and 1B are labeled with dotted underlines. The position where the primer sequence was derived for PCR extension is double-underlined with an arrowhead. Position where the mixed oligonucleotide primer was derived (from peptide 1) for initial verification sequencing is single-underlined with an arrowhead.
of the tropomodulin cDNA sequence could potentially give rise to an mRNA of 1.6 kb, the size for the detected shorter transcript. It has been known that sometimes more than one adenylation signal is present in the primary RNA transcript, leading to the formation of additional species of mRNA with 3'portions differing in length. This potential 1.6-kb cDNA would have the identical coding sequence and, therefore, encode the identical tropomodulin protein. Alternatively, the two different messages may represent isoforms of the protein (8) or differentially processed forms of a primary transcript.

**Primary Structure Analysis of Tropomodulin**—The open reading frame of this 2.7-kb cDNA encodes a protein with 359 amino acids, a calculated molecular mass of 40.6 kDa, and a pI of 4.8; the values are in agreement with the composition previously reported for purified erythrocyte tropomodulin (7,8).

There are several potential phosphorylation sites in tropomodulin. However, phosphorylation of tropomodulin has not been detected in intact erythrocytes (7). Several potential N-glycosylation sites can be identified in the sequence, but they are not expected to be utilized since tropomodulin does not have a signal peptide and is not a secreted protein (7). Residues of serine and threonine are in general, evenly distributed throughout the entire molecule, except at positions 151–154 where 4 serine residues in a row are found (Fig. 2). It is not known whether tropomodulin is O-glycosylated.

**Secondary Structure Analysis of Tropomodulin**—Both the Chou-Fasman method (22) and the Robson-Garnier method (23) predict that tropomodulin is predominantly α-helical with very little β-sheet configuration (Fig. 4).

Combining the results of the two predictive methods suggests that tropomodulin may consist of a series of seven or eight short α-helices (15–30 amino acids long) that are interrupted by sequences with turns or other secondary structures (Fig. 4). Inspection of the regions of sequence predicted to be α-helical reveals that these regions of sequence contain short sets of heptad repeats in which the residues at positions a and d tend to be hydrophobic or non-polar and residues at positions b, c, e, f, and g tend to be charged or polar (Fig. 5) (25–28). All of these sets of heptad repeats are relatively short, comprised of two to four heptads in each, and are often bounded by prolines on either side (Fig. 5, prolines indicated by !). The presence of

---

3 M. Sussman and V. M. Fowler, submitted for publication.
these heptad repeats strongly indicates that predictions of α-helical structure are likely to be correct for these regions (29).

Structure determinations of a variety of globular proteins reveal that many of them contain bundles of two to six α-helices that interact with one another (29). Unlike the long α-helices that interact to form the coiled-coil structures of rodlike extended proteins such as tropomyosin (40 heptad repeats, 284 residues, 43 nm long) (25), the heptad repeats of the α-helices in globular proteins occur in relatively short runs (e.g. from two to seven) corresponding to 3–14 turns of α-helix. Similarly, the short α-helices in tropomodulin may also interact with one another to form bundles, leading to a protein with a compact or only moderately extended shape (29). In addition, the charged to apolar residues ratio of 0.75 and a proline content of 6.13% are also consistent with only a moderately extended shape for tropomodulin (29). This agrees with a relatively globular shape for tropomodulin molecules visualized by rotary shadowing electron microscopy (8) and with hydrodynamic measurements of Stokes radius (3.7 nm), sedimentation coefficient ($s_{20,W} = 2.8$), and frictional ratio ($f/f_c = 1.5$) (7).

Tropomyosin-binding Domain—In order to map the tropomyosin-binding domain in the tropomodulin sequence, fusion proteins containing different regions of tropomodulin were expressed in E. coli Y1089 (Fig. 6, panel A). A phage clones 10 (lane 2), 5 (lane 3), and 7 (lane 4) were induced to express approximately the same amount of fusion proteins, as determined by Western blot analysis (Fig. 6, panel B). The fusion protein derived from clone 10 (residues 39–359, panel C, lane 2) was capable of binding $^{125}$I-erythrocyte tropomyosin, whereas neither the fusion protein derived from clone 5 (residues 139–359, panel C, lane 3) nor that from clone 7 (residues 152–359, panel C, lane 4) was able to bind tropomyosin (Fig. 6). These results suggest that the tropomyosin binding activity is located in the N-terminal portion of tropomodulin including residues 39–138. The doublet bands of clones 5 and 7 fusion proteins are likely to be the results of partial degradation. Partial degradation of foreign proteins in bacteria is not an uncommon phenomenon.

In a separate experiment, we compared the $^{125}$I-tropomyosin binding ability of the clone 10 fusion protein (lacking residues 1–38) with that of purified erythrocyte tropomyosin. After correcting for the amount of each protein transferred to nitrocellulose (measured by antibody binding), we found that clone 10 fusion protein bound about the same amount of $^{125}$I-tropomyosin as did the purified erythrocyte protein (data not shown).

The findings of these tropomyosin binding experiments are summarized in Fig. 7 and show clearly that tropomodulin molecules lacking the N-terminal 38 residues retain binding ability for erythrocyte tropomyosin. Deletion of an additional 100 amino acids from the N terminus, however, completely abolishes the tropomyosin binding ability. Therefore, the region containing residues 39–138 appears to be required for tropomyosin binding. Experiments addressing questions like whether the two heptad repeats located within this region have any functional role for tropomyosin binding activity and whether residues 39–138 are by themselves sufficient for binding to tropomyosin are in progress.

Homology with Other Proteins—A homology search of GenBank (release 68) and National Biomedical Research Foundation-Protein Identification Resource protein sequence data base (release 36.0, updated August, 1990) using the Lipman and Pearson algorithm FASTA (30) revealed that the tropomodulin nucleotide and amino acid sequences have no significant homology with any existing sequences in the data bases. Proteins that are highly α-helical in structure, such as myosin, nuclear lamins, tropomyosin, and keratins, are among those that scored the highest points on the list, reflecting the high α-helical content of tropomodulin (Figs. 4 and 5). However, these scores were not statistically significant. Thus, tropomodulin is a unique tropomyosin-binding protein, distinct from previously identified tropomyosin-binding proteins such as actin, caldesmon, troponin I, or troponin T. The previously reported cross-reactivity of tropomodulin antibodies with muscle troponin I (8) may be due to common features of secondary or tertiary structure recognized by the antibodies, perhaps features that are shared by their respective tropomyosin-binding domains.

Scatchard analysis of solid phase binding assays shows that binding of non-erythrocyte tropomyosins (brain, platelet, striated muscle) to tropomodulin decreases in affinity and/or capacity in comparison with erythrocyte tropomyosin, indicating that the interaction of tropomodulin with tropomyosin is isoform-specific. Together with the existence of immunologically cross-reactive tropomodulin polypeptides in a variety of non-erythroid cells and tissues, these data suggest the existence of a family of tropomodulin-like proteins specific for different tropomyosin isoforms (31).

Although the reported tropomodulin sequence is derived from a fetal liver cDNA, all of the available information indicates that the sequence is likely to be the same as that of erythrocyte tropomodulin. This includes identical N-terminal 103 amino acids predicted from fetal liver and reticuloocyte cDNA clones, a perfect match of predicted amino acid se-
quences to N- and C-terminal erythrocyte tropomodulin peptides (21 and 25 amino acids long, respectively), a close match of amino acid composition (8), matched mRNA/cDNA length, and comparable binding activities toward human erythrocyte tropomyosin. Together with the fact that liver is a blood-producing organ in the human fetus, this evidence indicates the fetal liver sequence reported here is very likely to be also of the human erythrocyte. Although there remains the formal possibility that fetal liver and erythrocyte each have a slightly different tropomodulin isoform, the two isoforms should have only very minor differences, especially in view of the identical N-terminal 103 amino acids which constitute ≥65% of the tropomyosin-binding domain (which may actually be smaller after further refining the boundaries). To address these questions, we will be isolating cDNAs corresponding to tropomodulin-like proteins from various tissues and cells, analyzing their tropomyosin-binding domains, and their possible tropomyosin isoform specificity.

Acknowledgments—We thank Dr. B. G. Forget for providing access to the human fetal liver cDNA library, Dr. J. Conboy for providing the human reticulocyte cDNA library, and Bill Lane at the Harvard Biological Laboratories for isolation and sequencing of tropomodulin. Together with the fact that liver is a blood-producing organ in the human fetus, this evidence indicates the fetal liver sequence reported here is very likely to be also of the human erythrocyte. Although there remains the formal possibility that fetal liver and erythrocyte each have a slightly different tropomodulin isoform, the two isoforms should have only very minor differences, especially in view of the identical N-terminal 103 amino acids which constitute ≥65% of the tropomyosin-binding domain (which may actually be smaller after further refining the boundaries). To address these questions, we will be isolating cDNAs corresponding to tropomodulin-like proteins from various tissues and cells, analyzing their tropomyosin-binding domains, and their possible tropomyosin isoform specificity.

REFERENCES