Pickford, G. E. (1964) *Bull. Bingham Oceanogr. Collect. 19*, 5-67.

Salvato, B., & Zatta, P. (1983) *Life Chem. Rep., Suppl. Ser. 1*, 139-140.

Takagi, T. (1986) in *Invertebrate Oxygen Carriers* (Linzen, B., Ed.) pp 259-262, Springer-Verlag, Berlin and Heidelberg.

van Bruggen, E. F. J., Schutter, W. G., van Breemen, J. F. L., Bijholt, M. M. C., & Wichertjes, T. (1981a) in *Electron Microscopy of Proteins* (Harris, J., Ed.) pp 1-38, Academic, London.

van Bruggen, E. F. J., Schutter, W. G., Wichertjes, T., & Keegstra, W. (1981b) in *Invertebrate Oxygen Binding Proteins* (Lamy, J., & Lamy, J., Eds.) pp 405-414, Dekker, New York.

van Holde, K. E., & Weischet, W. (1978) *Biopolymers 17*, 1387-1403.

van Holde, K. E., & Miller, K. I. (1982) *Q. Rev. Biophys. 15*, 1-129.

van Holde, K. E., & Miller, K. I. (1985) *Biochemistry 24*, 4577-4582.

Weeke, B. (1973) *Scand. J. Immunol., Suppl. 1*, 1-35.

Wichertjes, T., Gielens, C., Schutter, W. G., Préaux, G., Lontie, R., & van Bruggen, E. F. J. (1986a) in *Invertebrate Oxygen Carriers* (Linzen, B., Ed.) pp 227-230, Springer-Verlag, Berlin and Heidelberg.

Wichertjes, T., Gielens, C., Schutter, W. G., Préaux, G., Lontie, R., & van Bruggen, E. F. J. (1986b) *Biochim. Biophys. Acta 872*, 183-194.

Use of Consensus Oligonucleotides for Detecting and Isolating Nucleic Acids Encoding Calcium Binding Domains of the Troponin C Superfamily†

Susan H. Hardin,‡ Matthew J. Keast,‡ Paul E. Hardin,§ and William H. Klein*

Department of Biochemistry and Molecular Biology, M. D. Anderson Hospital, University of Texas System Cancer Center, Houston, Texas 77030

Received November 17, 1986; Revised Manuscript Received January 23, 1987

ABSTRACT: Proteins belonging to the troponin C superfamily (troponin C, calmodulin, myosin light chains, and parvalbumin) are involved in a wide variety of cellular activities mediated by calcium ions. Most of these proteins bind ionic calcium, and all have calcium binding domains that are conserved to some extent at the nucleic acid level. We made use of the conservation in the third calcium binding domain to synthesize two consensus sequence oligonucleotide probes, one 43 bases and the other 25 bases long. By using cDNA and genomic clones encoding calmodulin, troponin C, parvalbumin, and the sea urchin Spec proteins, we show that these probes hybridize with nucleic acid sequences representing calcium binding domains. In an RNA gel blot analysis of embryonic RNA from the sea urchin *Strongylocentrotus purpuratus*, we show that transcripts which have previously been shown to encode troponin C like proteins hybridize with the consensus sequence probes. Screening sea urchin cDNA and genomic libraries with the 43-base consensus oligonucleotide shows that the probe can be used to isolate cloned nucleic acids. Two such genomic clones from a *Lytechinus pictus* library were isolated and characterized. One clone encodes part of an *L. pictus* calmodulin gene, and the other encodes a member of the superfamily that has not been characterized previously. The consensus oligonucleotides should be valuable probes in the diagnosis and isolation of nucleic acids encoding proteins of the troponin C superfamily.

Troponin C, calmodulin, alkali and regulatory myosin light chains, and parvalbumin belong to a group of proteins termed the troponin C superfamily by M. O. Dayhoff (1978). The diagnostic structural feature of the superfamily is the presence, in each protein, of conserved calcium binding domains that form a distinct helical structure called an EF hand (Kretsinger, 1980). Generally, these proteins function as regulators of cellular processes mediated by calcium ion, and because of this, much attention has been focused on their structure and mechanism of action (Means & Conn, 1987). X-ray crystallography of rat testes calmodulin (Babu et al., 1985) and turkey skeletal muscle troponin C (Herzberg & James, 1985) has recently been performed and has provided detailed structural information. Both proteins consist of two globular lobes containing the calcium binding sites connected by an exposed α helix.

The genes encoding several members of the troponin C superfamily have been isolated, and their nucleotide sequences have been determined (Falkenthal et al., 1985; Putkey et al., 1983; Nabeshima et al., 1984; Robert et al., 1984; Hardin et al., 1985; Salvato et al., 1986). These sequences and those obtained from protein sequencing (Kretsinger, 1980) show that there is weak conservation of amino acid residues among the various proteins. However, the nucleotide sequences encoding amino acids important for calcium ion interaction and for EF-hand helix structure are more strongly conserved.

Given the conservation of the calcium binding sites, it seemed likely that a consensus probe could be developed to detect nucleic acid sequences encoding members of the troponin C superfamily. Such a probe could then be used for isolating novel genes encoding troponin C related proteins that have not yet been characterized. The troponin C superfamily appears to be a much larger group of proteins than previously

† This work was supported by NIH Grant HD 22619. W.H.K. is the recipient of NIH Research Career Development Award HD735.

* Author whom correspondence should be addressed.

‡ Present address: Program in Molecular, Cellular and Developmental Biology, Indiana University, Bloomington, IN 47405.

§ Present address: Program in Genetics, Indiana University, Bloomington, IN 47405.

0006-2960/87/0426-3518$01.50/0

thought, with specialized members present in distinct organisms and tissues. Recent examples of specialized members are the vitamin D induced calbindin from chickens (Wilson et al., 1985) and the sea urchin Spec proteins (Carpenter et al., 1984). We report here the characterization of two synthetic oligonucleotides that can be used as probes for the cloning and analysis of genes encoding calcium binding proteins.

Our work was prompted by our ongoing studies on the Spec gene family in the sea urchin *Strongylocentrotus purpuratus*. These genes encode 10–12 calcium binding proteins that are members of the troponin C superfamily and are expressed exclusively in aboral ectoderm cells of the sea urchin embryo (Bruskin et al., 1982; Carpenter et al., 1984; Lynn et al., 1983). Using sea urchin Spec and calmodulin DNA sequences as well as sequences encoding chicken slow muscle troponin C and rat parvalbumin, we demonstrate that the oligonucleotide probes readily hybridize to these DNAs, albeit with different affinities. The oligonucleotides also hybridize to predicted RNAs by using an RNA gel blot analysis.

We show that the probes can be used to isolate cDNA clones and genes encoding calcium binding domains. Exons from two such genes are described, both from the genome of the sea urchin *Lytechinus pictus*. One corresponds to the third calcium binding domain of calmodulin and the other to a calcium binding domain from an uncharacterized gene. This latter gene does not encode any known troponin C related protein, and its isolation demonstrates the validity of using the consensus oligonucleotides as a means for isolating novel nucleic acids encoding calcium binding domains.

Experimental Procedures

Oligonucleotide Synthesis. The solid-phase phosphoramidite method was used to synthesize both the 43- and 25-base oligonucleotides (M. D. Anderson Tumor Biology Program).

Genomic and cDNA Clones. The Spec1 genomic clone was isolated from an *S. purpuratus* λEMBL3 genomic library (Dr. Eric Davidson, California Institute of Technology) and has been discussed in detail (Hardin et al., 1985). The Spec1, Spec2a, and Spec2c cDNA clones (Carpenter et al., 1984) and the sea urchin calmodulin cDNA clone (Floyd et al., 1986) were isolated from a λgt10 library (Dr. Terry Thomas, Texas A&M University). The rat parvalbumin and the chicken slow muscle troponin C cDNA clones were gifts of Dr. Anthony Means, Baylor College of Medicine.

Radiolabeling of Probes. One microgram of the 43- or 25-base oligonucleotide was 5′ end labeled with [γ-^{32}P]ATP (3000 Ci/mmol) by using T4 polynucleotide kinase according to Maniatis et al. (1982). Reactions were monitored by trichloroacetic acid (TCA) precipitation, and reactions with specific activities $\geq 5 \times 10^7$ cpm/μg of DNA were used. Labeled oligonucleotides were ethanol precipitated, washed twice with 70% ethanol, resuspended in 50 μL of 5 × SET [0.75 M NaCl, 0.15 M tris(hydroxymethyl)aminomethane hydrochloride (pH 8.0), 10 mM ethylenediaminetetraacetic acid], and incubated at 65 °C for 5 min before being added to hybridization buffer (Bruskin et al., 1981).

An M-13 clone containing most of the reading frame of Spec1 was labeled as follows. Plus-strand M-13 DNA was isolated, annealed with the hybridization primer (New England Biolabs), and elongated with the large fragment of *Escherichia coli* DNA polymerase I in the presence of [α-^{32}P]dCTP, -dATP, -dTTP, and -dGTP for 30 min at 30 °C. The labeled probe had a specific activity of 10^8 cpm/μg.

A pUC8 clone containing a 1.4-kilobase (kb) *Eco*RI *S. purpuratus* calmodulin cDNA insert (pCal 8) was digested with *Eco*RI, and the fragments were separated on a 1% agarose gel. The calmodulin cDNA fragment was then electroeluted in 0.1 × TBE (Maniatis et al., 1982) and nick translated according to Bruskin et al. (1981) to a specific activity of 3×10^7 cpm/μg.

Filter Hybridizations. DNA or RNA was blotted onto nitrocellulose filters from 1% agarose or 1% agarose–formaldehyde gels, respectively. Hybridizations with the oligonucleotides were carried out at 37 °C as described by Bruskin et al. (1981) with the following exceptions: the hybridization buffer did not contain formamide; the probes were preabsorbed with a piece of nitrocellulose for 30–60 min to remove nonspecific nitrocellulose binding contaminants before addition to the blotted filter. Briefly, the filters were placed in sealable plastic bags and prewashed in 10 × Denhardts [1.0% bovine serum albumin, 1.0% Ficoll, 1.0% poly(vinylpyrrolidone) 40], 0.1% sodium dodecyl sulfate, 0.02 M sodium phosphate (pH 7.0), and 5 × SET at 68 °C for 1 h. The filters were hybridized at 37 °C in 10 mL of 10% dextran sulfate, 5 × SET, 0.02 M sodium phosphate (pH 7.0), 1 × Denhardts, and 100 μg/mL sheared calf thymus DNA.

To determine the optimal wash conditions for either oligonucleotide, a DNA blot containing various troponin C family clones was washed in 1 × SET, 1 × Denhardts, 0.1% sodium pyrophosphate, and 0.025 M phosphate buffer (Bruskin et al., 1981) for 1 h at 15 °C below the calculated T_m for the least stable duplex that would be formed with the DNA present on the blot. The calculation was made by using the known sequences and the formula described by Maniatis et al. (1982). The blot was exposed to Kodak XRP film to observe the hybridization patterns and then rewashed at increasing 5 °C temperature increments to optimize for hybridization signal with fragments containing troponin C related sequences and minimize for nonspecific signals. It was found that for both consensus oligonucleotides the optimal wash temperature was 42 °C for hybridization to the known troponin C related sequences shown in Figure 2.

Genomic Library. A genomic library was constructed from the sea urchin *L. pictus*. Sperm DNA was partially digested with *Sau*IIIa and fractionated on a 10–40% sucrose gradient, and fragments with a mean size of 17 kb were then ligated to λEMBL3 arms (Vector Cloning Systems, San Diego, CA) and packaged with commercial extracts (Gigapack; Vector Cloning Systems).

Screening cDNA and Genomic Libraries. Thirty thousand phage from an *S. purpuratus* gastrula stage λgt10 cDNA library (Dr. Terry Thomas, Texas A&M University) were screened with the 43-base oligonucleotide probe and washed at 42 °C, 0.15 M sodium ion.

Two hundred eighty thousand phage from the *L. pictus* genomic library (1.5 genome equiv) were screened with the 43-base oligonucleotide probe under the same conditions used for the filter hybridization as described above. Eleven hybridizing plaques were rescreened, and of those, five hybridized strongly. Two strongly hybridizing plaques were analyzed in detail and are discussed below.

Sequencing of Genomic Clones. Clones that hybridized strongly to the 43-base oligonucleotide were mapped, blotted to nitrocellulose, and rehybridized to the 43-base oligonucleotide. The smallest hybridizing fragments were subcloned into M-13mp18 and M-13mp19 (Pharmacia, Piscataway, NJ) and sequenced according to the chain termination method of Sanger.

Results and Discussion

Consensus Oligonucleotides Representing the Third Calcium Binding Domain of Several Troponin C Related Pro-

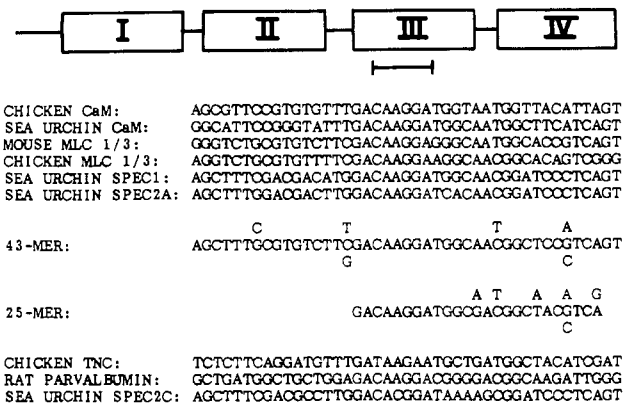

FIGURE 1: Derivation of 43- and 25-base consensus oligonucleotides. The top line shows a schematic representation of a protein of the troponin C superfamily. The boxes with Roman numerals indicate the four calcium binding domains. The lines in between represent amino acid residues not in the domains. The bar underneath is $19^1/_3$ amino acids (43 bases) in length and represents the region of the protein sequence from which the 43- and 25-base oligonucleotides derive. The position of the 43-base oligonucleotide with respect to the calcium binding loop (i.e., the residues that interact with calcium ion) is depicted in Figure 6. The following six lines are sequences of this region that were used to derive the 43-base consensus sequence: chicken and sea urchin calmodulin (CaM), mouse and chicken myosin light chains 1 and 3 (MLC), and sea urchin Spec1 and Spec2a. The next two lines are the consensus 43- and 25-base oligonucleotides and the following two lines are the sequences that were used with the above sequences to derive the 25-base sequence: chicken slow muscle troponin C (TnC) and rat parvalbumin (Parv). In the case of parvalbumin, the fourth calcium binding domain is compared. The last line is the sequence of sea urchin Spec2c, whose hybridization with the consensus probes is described in the text.

teins. There are four calcium binding domains present in most (but not all) proteins belonging to the troponin C superfamily. Because of the internal homology within a given protein, it is thought that the four domains arose from two duplications of a single ancestral calcium binding domain (Dayhoff, 1978). In the case of the alkali myosin light chains, the ability of the protein to bind calcium ion has been lost but the calcium binding domains have been reasonably well conserved (Robert et al., 1984).

Although there is some sequence conservation among the troponin C related proteins, a nucleic acid probe encoding one protein of the superfamily cross-reacts weakly or not at all with other superfamily members (see below). Since a consensus sequence would minimize differences, our goal was to find a probe that, under the appropriate conditions, would show significant cross-reaction with any of the known troponin C related nucleic acids. Such a probe could then be used in attempts to isolate unidentified superfamily members. To find conserved regions among the various troponin C related proteins, we searched for maximum similarities using five genomic or cDNA clones for which the nucleotide sequences of the complete reading frame were known. These were the chicken calmodulin cDNA (Putkey et al., 1983), the mouse and chicken mysoin light chain genes 1 and 3 (Robert et al., 1984; Nabeshima et al., 1984), and the sea urchin Spec1 and Spec2a genes (Carpenter et al., 1984). Several regions of similarity could be found, but one of the strongest regions represents a portion of the third calcium binding domain, including most of the residues that interact with calcium ion (see below). Figure 1 displays a 43-nucleotide consensus sequence derived from this region. The consensus sequence is between 81 and 88% similar to the five input sequences, and it seemed likely that an oligonucleotide complementary to this sequence would readily hybridize with them. However, comparisons with

Table I: Comparison of 43- and 25-Base Consensus Sequences with Corresponding Regions of Various Troponin C Related Sequences

	matching bases (% match)	
sequence[a]	43-base consensus	25-base consensus
CaM (chicken)	37 (86)	22 (88)
MLC (mouse)	38 (88)	23 (92)
MLC (chicken)	35 (81)	22 (88)
Spec1 (sea urchin)	37 (86)	23 (92)
Spec2a (sea urchin)	36 (84)	21 (84)
Spec2c (sea urchin)	33 (77)	18 (72)
TnC (chicken)	27 (63)	21 (84)
Parv (rat)[b]	22 (51)	21 (84)

[a] CaM is chicken calmodulin (Putkey et al., 1983). MLC is myosin light chains 1 and 3 [data for mouse taken from Robert et al. (1984) and data for chicken MLC taken from Nabeshima et al. (1984)]. Spec1, Spec2a, and Spec2c are all from *S. purpuratus* as described by Carpenter et al. (1984). TnC and Parv are chicken slow muscle troponin C (Putkey et al., 1987) and rat parvalbumin (Epstein et al., 1986). [b] The parvalbumin sequence shows a stronger similarity to the fourth calcium binding domain than to the third, when compared with either the 43- or 25-base oligonucleotides. The comparison with the fourth binding domain is given in the table.

cDNA for chicken slow muscle troponin C and rat parvalbumin show far less similarity, 63 and 51%, respectively (Table I). (These sequences were unavailable for use at the time the other sequences were compared.) It would be difficult to devise hybridization conditions to detect specific sequences with this level of similarity, but closer inspection of the 43-base region shows that a 25-base subregion is even more conserved among the various genes. The 25-base consensus sequence is shown in Figure 1, and the match to the various nucleic acids ranges from 72 to 92% (Table I). The complementary strands of the 43- and 25-base oligonucleotides shown in Figure 1 were synthesized and used in a variety of assays to ascertain their specificity.

Southern Hybridization of 43- and 25-Base Consensus Oligonucleotides with Calmodulin, Spec1, Spec2c, Troponin C, and Parvalbumin Nucleic Acid Sequences. To determine whether the consensus oligonucleotides would hybridize with nucleic acid sequences corresponding to the third calcium binding domain, we carried out Southern blot analysis using several cDNA and genomic clones. The oligonucleotides were radiolabeled with ^{32}P at their 5′ ends by using T4 polynucleotide kinase. Hybridization conditions were chosen to be a few degrees below the calculated T_m of the duplexes, and the nitrocellulose filters were washed under conditions close to the T_m for the least stable duplex (see Experimental Procedures).

A cDNA clone corresponding to the *S. purpuratus* calmodulin mRNA has been previously described (Floyd et al., 1986). This clone contains sequences from amino acid 68 to amino acid 148 and includes the region that shows similarity to the 43- and 25-base consensus oligonucleotides. As shown in lane 1 of Figure 2 (panels B and C), strong hybridization is seen with a 1.4-kb *Eco*RI fragment containing the calmodulin sequences, but no detectable signal is observed with the 2.7-kb pUC8 vector.

Spec1 and Spec2c are two different but related sea urchin genes whose structures and sequences are known (Hardin et al., 1985). The genes are split into six exons, and exon 4 contains the region of similarity with the consensus sequences. In the Spec1 gene, exon 4 is contained within a 2-kb *Hin*dIII–*Sst*I fragment. This fragment, as well as a 0.38-kb *Eco*RI fragment from a Spec1 cDNA clone, hybridizes with the consensus oligonucleotide probes (Figure 2B,C, lanes 2 and 3). Other fragments, including those that contain other calcium binding domains, do not hybridize. This is not always

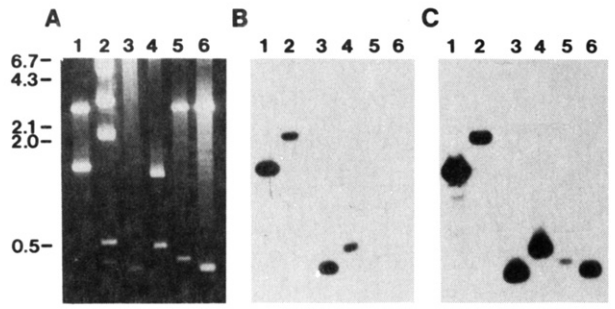

FIGURE 2: Southern blot hybridization of 43- and 25-base consensus oligonucleotides with nucleic acids encoding calcium binding proteins: lane 1, *Eco*RI-digested sea urchin calmodulin cDNA clone (in pUC8); lane 2, *Hin*dIII–*Sst*I-digested ϕ4 DNA (an *S. purpuratus* Spec1 genomic clone in λEMBL3); lane 3, *Eco*RI-digested Spec1 cDNA clone (in λgt10); lane 4, *Eco*RI-digested Spec2c cDNA clone (in λgt10); lane 5, *Pst*I-digested chicken slow muscle troponin C cDNA clone (in pUC8); lane 6, *Eco*RI-digested rat parvalbumin cDNA clone (in pUC8). (A) Ethidium bromide stained gel. (B) Hybridization with ^{32}P-labeled 43-base consensus oligonucleotide. (C) Hybridization with ^{32}P-labeled 25-base consensus oligonucleotide.

true. Since the four calcium binding domains arose from a common ancestral domain, in some cases more than one domain within a given gene is well matched to the consensus oligonucleotides. This is true for the first and third calcium binding domains of the Spec2a gene (data not shown).

A Spec2c 0.5-kb *Eco*RI cDNA clone containing all four calcium domains hybridizes with both the 43-base probe and with the 25-base probe (Figure 2B,C, lane 4).

Two other cDNA clones, one encoding chicken slow muscle troponin C and the other encoding rat parvalbumin, were given to us by Dr. Anthony Means of Baylor College of Medicine. The troponin C clone is cleaved into a 0.43- and a 0.23-kb fragment when digested with *Pst*I, and the parvalbumin clone contains an insert fragment of 0.35 kb when digested with *Eco*RI. When hybridized with the 43-base consensus oligonucleotide, very weak or no detectable hybridization is observed in either case (Figure 2B, lanes 5 and 6). Longer exposures show a weak signal for troponin C, presumably due to its 63% sequence similarity (Table I). When the cDNA clones are hybridized with the 25-base consensus probe, positive signals are readily observed (Figure 2C, lanes 5 and 6). On the basis of the matches with the consensus sequences, it is not clear why the intensity of the troponin C signal is not equal to that of the parvalbumin signal. Other hybridization reactions also fail to yield the expected quantitative pattern. For example, the Spec1 genomic sequence hybridizes to either consensus oligonucleotide with higher intensity than the corresponding Spec1 cDNA clone (compare Figure 2B,C, lanes 2 and 3). Thus, while the 43- and 25-base censensus sequence probes hybridize specifically to a variety of troponin C related sequences, they cannot be used quantitiatively. Nevertheless, it is reasonable to assume that in combination these probes could be used to detect an analogous calcium binding domain represented in any nucleic acid sequence.

It was of interest to determine whether a particular cDNA clone encoding a troponin C related protein could hybridize to other members of the superfamily if hybridization conditions were sufficiently mild. The filter used for hybridization with the 25-base oligonucleotide shown in Figure 2C was washed to remove radioactivity and rehybridized with ^{32}P-labeled Spec1 cDNA clone. The conditions used were identical with the hybridization with the 43- and 25-base oligonucleotide probes (see Experimental Procedures). The Spec1 probe hybridized strongly to the corresponding Spec1 genomic and cDNA clones and to the closely related Spec2c cDNA clone.

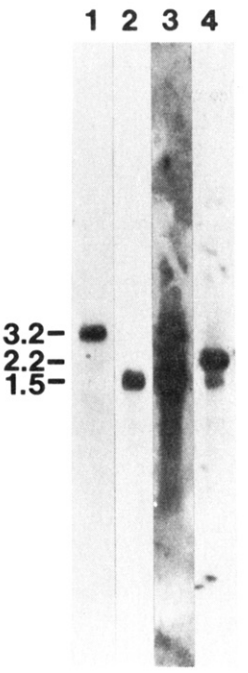

FIGURE 3: RNA gel blot analysis of 43- and 25-base consensus oligonucleotides with *S. purpuratus* gastrula RNA. Ten micrograms of total RNA from gastrula-stage embryos were electrophoresed on agarose–formaldehyde gels, transferred to nitrocellulose, and hybridized with lane 1, pCAL8 cDNA clone; lane 2, Spec1 cDNA clone; lane 3, 43-base consensus oligonucleotide; lane 4, 25-base consensus oligonucleotide. Probes were labeled with ^{32}P as described under Experimental Procedures.

However, no hybridization was observed with calmodulin, troponin C, or parvalbumin sequences (data not shown). This result emphasizes the fact that the sequence similarities between the various family members are usually not high enough to obtain cross-hybridization and that the consensus oligonucleotides minimize the differences and allow for specific hybridization.

Hybridization of 43- and 25-Base Consensus Sequences with S. purpuratus Embryonic RNA. To determine whether the consensus probes would hybridize to RNAs by using RNA gel blot hybridization procedures, we made use of total *S. purpuratus* gastrula RNA. On the basis of our previous studies, we expected that at least three transcript size classes would hybridize with the probes: calmodulin mRNA, which is 3.2 kb in length and represents about 0.1% of the mRNA of the gastrula (Figure 3, lane 1) (Floyd et al., 1986); Spec1 mRNA, which is 1.5 kb in length and represents 0.6% of the mRNA (Figure 3, lane 2) (Bruskin et al., 1982; Lynn et al., 1983); and Spec2 mRNAs, of which there are about 10 related 2.2-kb mRNAs representing about 0.06% of the mRNA (Bruskin et al., 1982). Other troponin C related mRNAs have not been characterized in this system, but some may be expected to hybridize to the consensus probes. Troponin C itself, a striated and cardiac muscle specific protein, is unlikely to be present in detectable amounts since sea urchin gastrula stage embryos have little or no differentiated muscle tissue.

Hybridization of the 43-base consensus probe with gastrula RNA shows two discrete bands at 3.2 and 1.5 kb comigrating with the calmodulin and Spec1 messages, respectively, and nondiscrete hybridization above, between, and below these bands (Figure 3, lane 3). The hybridization between the 3.2- and 1.5-kb bands is attributed to the 2.2-kb Spec2 mRNA, which characteristically hybridize as a broad nondiscrete band (Bruskin et al., 1982; Carpenter et al., 1984). These data support the notion that the above-mentioned three transcript

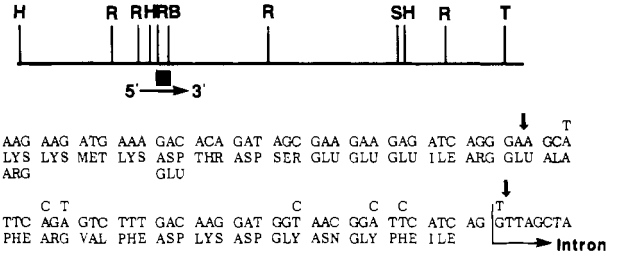

FIGURE 4: Analysis of λLpCaM1. The top line shows a partial restriction map of the λLpCaM1 insert. H, *Hin*dIII; R, *Eco*RI; B, *Bam*HI; S, *Sal*I; T, *Sst*I. (Not all the sites are mapped.) The box below the map indicates the 0.25-kb *Eco*RI–*Bam*HI fragment that hybridizes with the 43-base consensus oligonucleotide. The sequence beginning just to the right of the *Eco*RI site and continuing to the end of calmodulin sequence match is shown in the second part of the figure. Line A indicates the 43-base consensus sequence. Bases are indicated only where they differ from the λLpCaM1 sequence. Arrows pointing downward depict the boundaries of the 43-base oligonucleotide. Line B is the λLpCaM1 sequence. Line C is the translation of the nucleic acid sequence. Line D is the residues that are different in the previously identified *S. purpuratus* calmodulin sequence. The putative exon/intron junction interrupts a codon beginning AG that is probably a serine residue.

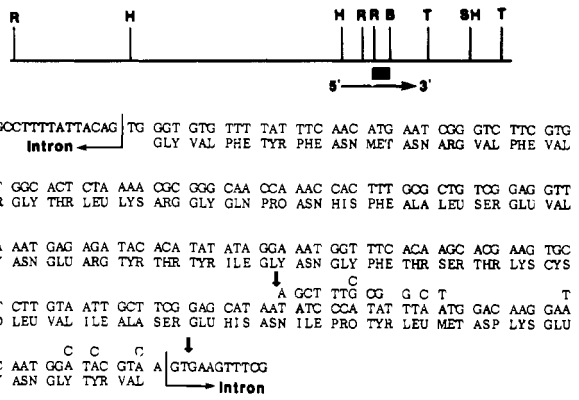

FIGURE 5: Analysis of λLp14-7. The top line shows a partial restriction map of the λLp14-7 insert. The symbols are the same as described in Figure 4. Not all the sites are mapped. The box below the map is the 0.45-kb *Eco*RI–*Bam*HI fragment that hybridizes with the 43-base consensus probe. The partial sequence of this fragment is shown in the second half of the figure. Line A indicates the 43-base consensus sequence. Only bases that are different from λLp14-7 are indicated. The downward arrows depict the boundaries of the 43-base oligonucleotide. Line B is the λLp14-7 sequence; the putative introns are indicated. Line C is the translation of the nucleic acid sequence.

size classes are the major mRNAs encoding troponin C related proteins in sea urchin gastrula staged embryos. This conclusion is complicated by the background hybridization and by the results shown in Figure 3, lane 4, where the 25-base consensus probe is hybridized to total gastrula RNA. With this probe, a strong band is seen at approximately 2.2 kb and a weaker one at 1.5 kb. Very weak but detectable hybridization is seen in the region where calmodulin mRNA migrates. The 2.2- and 1.5-kb RNAs are enriched in preparations of ectoderm RNA relative to endoderm/mesoderm RNA (data not shown), and this suggests that they represent the ectoderm-specific Spec1 and Spec2 transcripts. The reason for the dramatic difference in the hybridization of the 43- and 25-base probes is not clear and, as with the Southern blot hybridization discussed above, indicates that the probes cannot be utilized for quantitative comparisons. However, both probes show hybridization to transcripts that are of the same length as mRNAs encoding calcium binding proteins, and these results suggest that the oligonucleotides can be used to detect mRNAs at levels of less than 0.1% of a heterogeneous mRNA population.

Screening cDNA and Genomic Libraries with 43-Base Consensus Oligonucleotide. The 43-base consensus oligonucleotide was used to screen two DNA libraries to determine whether sequences encoding calcium binding domains could be isolated. We fist screened a λgt10 *S. purpuratus* gastrula cDNA library. On the basis of the above results (Figure 3), we postulated that most of the reacting clones would be Spec sequences, with a smaller number being calmodulin and uncharacterized sequences. In a screen of 30 000 recombinants using either the 43-base consensus oligonucleotide or a Spec sequence as a probe, we found 141 double positives, 26 positives unique to the 43-base oligonucleotide, and 46 positives for the Spec probe (data not shown). The hybridization conditions we employed, using the Spec probe, allowed both Spec1 and Spec2 sequences to cross-react but not other troponin C related sequences. Thus, the expected results were obtained with the consensus probe; most but not all of the positives were Spec sequences.

Another test of the 43-base consensus oligonucleotide was to screen an *L. pictus* genomic λEMBL3 library and isolate and characterize the positive recombinants. Several positive clones were isolated from the screen, and two λ phage were analyzed in detail. One of the isolated clones, λLpCaM1, is part of an *L. pictus* calmodulin gene. Figure 4 shows a partial restriction map of λLpCaM1, the region of hybridization to the 43-base consensus probe, and the partial sequence of a 0.25-kb *Eco*RI–*Bam*HI fragment containing the sequence that hybridizes to the probe. The sequence near the *Eco*RI site represents amino acid residues 74–100 of calmodulin. Only residues 74 and 78 are different from the previously sequenced cDNA clone from *S. purpuratus* (Floyd et al., 1986), and these are conservative changes: a Lys → Arg and an Asp → Glu. The Arg encoded at position 74 of the *L. pictus* calmodulin is also found in all vertebrate calmodulins except eel, which has a Lys at this position (Means & Conn, 1987). As expected, the sequence contains the third calcium binding domain of calmodulin and there is strong similarity (37 of 43 bases) with the 43-base consensus oligonucleotide. A putative exon/intron junction is indicated in Figure 4. Interestingly, this is the precise location of the exon/intron junction in the *S. purpuratus* Spec1 gene (Hardin et al., 1985), although in the chicken calmodulin gene, the boundary is shifted seven amino acids upstream (Means & Conn, 1987).

The other genomic recombinant, λLp14-7, contains a 0.45-kb *Eco*RI–*Bam*HI fragment that has a weak but significant match to the oligonucleotide, 27 of 43 bases (Figure 5). These sequences are contained in what appears to be a 207-base exon bounded by canonical splice sequences at its 5′ and 3′ ends. The putative exon does not correspond to any known gene encoding a calcium binding protein, but comparison of the conserved sequences with Spec1 or calmodulin demonstrates the presence of a putative calcium binding domain (Figure 6). In Figure 6, the λLp14-7 sequence is compared to the third calcium binding domain of Spec1 and chicken calmodulin. Conserved amino acid residues that either interact with calcium ion or are important for EF-hand helix formation are indicated above the sequence (X, Y, Z, G, I, −X, −Z). The λLp14-7 sequence shows strong conservation at these residues. The two differences are conservative replacements; Asp → Glu and Leu/Ile → Val. The putative 5′ exon/intron junction in λLp14-7 precedes the calcium binding domain and is not located at any junction previously reported. However, the putative 3′ exon/intron junction is at the same codon location as the exon/intron junction of the sea urchin Spec1 gene (Figure 6, bottom arrow).

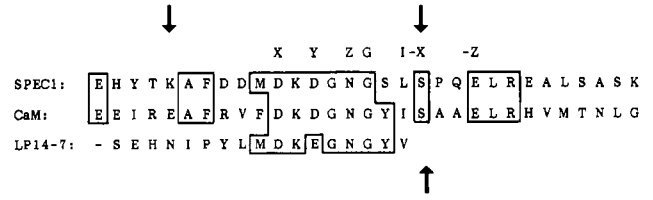

FIGURE 6: Comparison of λLp14-7 with the third calcium binding domains of *S. purpuratus* Spec1 and chicken calmodulin. The protein sequence shown for Spec1 and calmodulin represents the complete third calcium binding domain. The downward-pointing arrows indicate the boundaries of the 43-base consensus oligonucleotide. The upward-pointing arrow indicates the exon/intron junction present in the sea urchin Spec1 and calmodulin genes. Matched residues are boxed. Residues important for calcium ion interaction (X, Y, Z, -X, -Z) or for EF-hand helix structure (G, I) are indicated above the sequences.

Hybridization of *L. pictus* gastrula poly(A+) RNA, using a probe containing the encoded calcium binding domain of λLp14-7, yields no positive signal (data not shown). It is possible that corresponding transcripts may be present in other developmental stages or in adult tissues, but these have not been tested. A reasonable tissue to investigate would be adult muscle. If transcripts could be found hybridizing with λLp14-7 from muscle or other tissues, cDNA clones could be obtained and the sequence of the complete translational reading frame determined. This, in turn, would allow for the isolation of antibodies which could be used to analyze the properties of the putative protein. Computer searches of data banks for homologous sequences show that λLp14-7 does not correspond to any known troponin C related sequence. While further characterization of this putative gene must be performed, the above analysis indicates that the 43-base oligonucleotide can be used to isolate DNA fragments from genomic libraries that appear to encode genes of the troponin C superfamily.

CONCLUSIONS

The similarities among the various troponin C superfamily members are rarely strong enough to allow for detectable cross-reaction at the nucleic acid level. Attempts to cross-hybridize Spec cDNAs with other troponin C related sequences, even under mild hybridization conditions, yield negative results. However, our results show that the 43- and 25-base consensus oligonucleotides can be used to detect and to isolate nucleic acids that encode members of the superfamily. Because of their general reactivity, the consensus oligonucleotides can be employed either to diagnose a sequence suspected of being a gene for a calcium bnding protein or to isolate new genes with novel features that are related to the known superfamily members.

It is now apparent that the troponin C related calcium binding proteins are a large and diverse group. A large number of metabolic events involving calcium ions may be driven by these proteins. The consensus oligonucleotides defined here should be of value in finding genes for such proteins.

ACKNOWLEDGMENTS

We thank Dr. Anthony Means and his co-workers at Baylor College of Medicine for supplying us with the chicken slow muscle troponin C and rat parvalbumin cDNA clones. We thank Kate Beckingham, John Putkey, and Elizabeth Eldon for their careful reading and criticism of the manuscript.

Registry No. Ca, 7440-70-2.

REFERENCES

Babu, Y. S., Sack, J. S., Greenbough, T. J., Bugg, C. E., Means, A. R., & Cook, W. J. (1985) *Nature (London) 315*, 37–40.

Bruskin, A. M., Tyner, A. L., Wells, D. E., Showman, R. M., & Klein, W. H. (1981) *Dev. Biol. 87*, 308–318.

Bruskin, A. M., Bedard, P. A., Tyner, A. L., Showman, R. M., Brandhorst, B. P., & Klein, W. H. (1982) *Dev. Biol. 91*, 317–324.

Carpenter, C. D., Bruskin, A. M., Hardin, P. E., Keast, M. J., Anstrom, J., Tyner, A. L., Brandhorst, B. P., & Klein, W. H. (1984) *Cell (Cambridge, Mass.) 36*, 663–671.

Dayhoff, M. O. (1978) *Atlas of Protein Sequence and Structure*, Vol. 5, Suppl. 3, pp 273–283.

Epstein, P., Means, A. R., & Berchtold, M. W. (1986) *J. Biol. Chem. 261*, 5886–5891.

Falkenthal, S., Parker, V. P., & Davidson, N. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 449–453.

Floyd, E. E., Gong, Z., Brandhorst, B. P., & Klein, W. H. (1986) *Dev. Biol. 113*, 501–511.

Hardin, S. H., Carpenter, C. D., Hardin, P. E., Bruskin, A. M., & Klein, W. H. (1985) *J. Mol. Biol. 186*, 243–255.

Herzberg, O., & James, M. N. G. (1985) *Nature (London) 313*, 653–659.

Kretsinger, R. H. (1980) *CRC Crit. Rev. Biochem. 8*, 119–174.

Lynn, D. A., Angerer, L. M., Bruskin, A. M., & Klein, W. H. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 2656–2660.

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

Means, A. R., & Conn, P. M., Eds. (1987) *Methods Enzymol. 139*.

Nabeshima, Y., Fujii-Kuriyama, Y., Muramatsu, M., & Ogata, K. (1984) *Nature (London) 308*, 333–337.

Putkey, J. A., Tsui, K. F., Tanaka, T., Lagace, L., Stein, J. P., Lai, E. C., & Means, A. R. (1983) *J. Biol. Chem. 258*, 11864–11870.

Putkey, J. A., Carroll, S. L., & Means, A. R. (1987) *Mol. Cell. Biol. 7*, 1549–1553.

Robert, B., Daubas, P., Akimenko, M.-A., Cohen, A., Garner, J., Guenet, J.-L., & Buckingham, M. (1984) *Cell (Cambridge, Mass.) 39*, 129–140.

Salvato, M., Sulston, J., Albertson, D., & Brenner, S. (1986) *J. Mol. Biol. 190*, 281–290.

Wilson, P. W., Harding, M., & Lawson, D. E. M. (1985) *Nucleic Acids Res. 13*, 8867–8881.