At Least 27 Alternatively Spliced Forms of the Neural Cell Adhesion Molecule mRNA Are Expressed during Rat Heart Development

ANTONIO A. REYES, STEPHEN J. SMALL,[†] AND RICHARD AKESON*

Division of Basic Research, Children's Hospital Research Foundation, Cincinnati, Ohio 45229-2899

Received 5 September 1990/Accepted 11 December 1990

The major membrane-associated or transmembrane isoforms of the neural cell adhesion molecule (NCAM) are generated by alternative splicing at the 3' end of the mRNA. Further diversity in NCAM structure is observed in the extracellular region of the polypeptide, where the insertion of additional amino acid residues can result from alternative splicing events occurring at the exon 7-exon 8 and exon 12-exon 13 junctions. Here we report the characterization of tissue-specific patterns of alternative splicing at the exon 12-exon 13 junction by using the polymerase chain reaction. Nine alternatively spliced sequences in rat heart between exon 12 and exon 13 were identified. Each sequence consisted of different combinations of the three small exons (15, 48, and 42 bp in length) and the AAG triplet that make up MSD1, the 108-bp muscle-specific sequence found in human skeletal muscle NCAM (G. Dickson, H. J. Gower, C. H. Barton, H. M. Prentice, V. L. Elsom, S. E. Moore, R. D. Cox, C. Quinn, W. Putt, and F. S. Walsh, *Cell* 50:1119-1130, 1987). Although the rat equivalent of MSD1 (designated 15⁺48⁺42⁺3⁺) was detected in all ages of heart examined, it was only one of four or five major splice combinations at any given age. The only alternatively spliced sequence found in the exon 7-exon 8 junction of heart NCAM mRNA was the 30-bp variable alternatively spliced exon previously identified in rat brain. Twenty-seven NCAM forms with distinct sequences were found by analysis of individual NCAM transcripts from postnatal day 1 heart tissue for alternative splicing at the exon 7-exon 8 junction, the exon 12-exon 13 junction and the 3' end. Several combinations of splicing patterns in these three different regions of the gene appeared to be preferentially expressed. The observation that the expression of alternatively spliced forms of NCAM is developmentally regulated suggests a role for NCAM diversity in cardiac development.

Regulation of gene expression can occur at the transcriptional, posttranscriptional, translational, and post-translational levels. One control mechanism that occurs posttranscriptionally in eukaryotes is alternative splicing. Alternative splicing of a pre-mRNA transcribed from a single gene can result in protein isoforms that differ in function or localization (recently reviewed in reference 33). In some cases, alternative splicing appears to be stochastic, while in others it is regulated in a tissue- and/or developmental stage-specific manner (33). In this case, alternative splicing can allow the fine modulation of gene expression such that protein isoforms with functional differences are expressed in the proper spatiotemporal fashion.

The neural cell adhesion molecule (NCAM) is a cell surface molecule that mediates cell-cell and cell-matrix adhesion via a homophilic binding mechanism (8, 25). A heparin-binding domain has also been identified in NCAM, and several lines of evidence suggest that NCAM-heparin binding is necessary for NCAM function (6, 23). NCAM-mediated interactions are believed to play an integral role in a number of developmental events, including neurite outgrowth, retinal histogenesis, and nerve-muscle interaction. The exact mechanisms by which NCAM modulates these events are not known, although differences in abundance, cell surface distribution, and extent of glycosylation of NCAM have all been suggested as possible mechanisms that regulate cell adhesiveness (8, 14).

Recent evidence indicates that the NCAM gene is alter-

natively spliced. The NCAM gene consists of 20 exons in both mammals and birds (3, 18). The 5'-most exon, which encodes the 5'-untranslated and leader sequences, is designated exon 0 (5, 12). The remaining exons, which are spread out over ~75 kb of genome, are numbered 1 through 19 as proposed for the chicken NCAM gene (18). Examination of cDNA and genomic sequences showed that transcripts of the single-copy NCAM gene differ in their pattern of exon splicing at the 3' end of the gene (4, 18). Thus, depending on the choice of 3' end exon combinations, three forms of NCAM polypeptide are possible: membrane-linked via phosphatidylinositol, transmembrane with a small cytoplasmic domain, and transmembrane with a large cytoplasmic domain (see Fig. 1). The corresponding molecular masses of these NCAM isoforms in mammals are approximately 120, 140, and 180 kDa, respectively (3, 35). NCAM mRNAs encoding the 120-kDa isoform may also differ in poly(A) addition site usage, resulting in short or long 3'-untranslated regions (3).

Alternative splicing has been observed in two other regions of the NCAM gene. The presence of a 30-bp alternatively spliced exon termed VASE (30) between exons 7 and 8 was detected in rat and mouse brain cDNAs (22, 32). A 108-bp sequence termed MSD1 was found between exons 12 and 13 in human skeletal muscle cDNA (7). Genomic cloning of the MSD1 sequence revealed that it consists of three smaller exons of 15, 48, and 42 bp arranged in that order in the intron between exons 12 and 13 (27, 34). The 3'-terminal AAG triplet of MSD1 could be encoded by a 3-bp exon (27), although the apparent location of this exon in the 12-13 intron does not agree with the prediction from the cDNA sequences. In embryonic chicken heart, a 93-bp sequence

* Corresponding author.

[†] Present address: Department of Biology, Fairchild Center, Columbia University, New York, NY 10027.

was found at the exon 12-exon 13 junction in a single cDNA clone (20). Like MSD1, the chicken heart insert is encoded by 3 small exons (15, 33, and 42 bp in size) and an AAG triplet is found at the 3' end of the insert. The 42-bp exons of MSD1 and the chicken heart insert have a high degree of sequence similarity. A fifth alternatively spliced exon has been found in the exon 12-exon 13 junction of human skeletal muscle NCAM. The use of this exon, called SEC, would result in the synthesis of an NCAM polypeptide lacking a membrane-linked or transmembrane domain, thus generating a secreted form (10).

The observation that MSD1 is encoded by at least three smaller exons has raised the possibility that these small exons can be spliced independently of each other. In support of this, the 15-bp exon and the AAG triplet of MSD1 were found either together or singly between exons 12 and 13 in mouse brain NCAM cDNAs (27). Indirect evidence from S1 nuclease protection assays also suggested that the 15-bp exon was spliced independently in human brain NCAM mRNAs (34). Preliminary evidence for developmental regulation of exon 12-exon 13 splicing came from S1 analysis of human muscle cell lines (34). Fusion of myoblasts to form myotubes, at least in vitro, was accompanied by the appearance of the MSD1 sequence in the majority of NCAM transcripts.

To date, direct evidence for NCAM alternative splicing has come from the analysis of a small number of cDNA clones. In order to more systematically identify multiple NCAM forms generated by alternative splicing events, we have used the polymerase chain reaction (PCR) (26) to analyze tissue- and developmental age-specific expression of multiple NCAM transcripts in rat heart tissue. Novel combinations of small alternatively spliced exons in the exon 12-exon 13 junction were identified. Overall splicing patterns at the exon 7-exon 8 junction, the exon 12-exon 13 junction, and the 3' end in postnatal day 1 (P1) heart were also determined. The identification of multiple NCAM forms and determination of precisely which forms are expressed by cells during development serve as first steps toward understanding the role of NCAM diversity in the regulation of cell adhesive interactions.

MATERIALS AND METHODS

RNA and cDNA preparations. Tissues were dissected from Sprague-Dawley rats. The rat muscle cell line L6 and the rat pheochromocytoma cell line PC12 were provided by Dave Schubert and Kenneth Neet, respectively. The liver cell lines H4TG and H4IIE were obtained from American Type Culture Collection. Total RNA was prepared by the guanidine-CsCl method (15). First-strand cDNA was synthesized from 20 µg of total RNA as described previously (30), with the following modifications. cDNA to be used for the amplification of the exon 7-exon 8 or exon 12-exon 13 junction was prepared with a mixture of oligo(dT) and oligonucleotide (oligo) 89.3 (5'-CATCTGCCCTCTGCTTGGGTGCGCTGGG3') as primers. Oligo 89.3 is complementary to NCAM mRNA (except for one mismatch to create a *Hind*III site) just 3' of the exon 12-exon 13 junction (see Fig. 1). For amplification of the entire region extending from exon 7 to the 3' end, cDNA synthesis was primed with oligo(dT) alone. Between 1/10 and 1/20 of the cDNA reaction mixture was used as a template for each PCR.

PCR. To determine alternative splicing at the exon 7-exon 8 or exon 12-exon 13 junction, asymmetric PCRs were performed as described previously (30), except that 10 ng of

unlabeled primer and 0.5 ng of labeled primer were used in a reaction volume of 10 µl. The primers used for the exon 7-exon 8 junction amplification were ³²P-17.1 and unlabeled 17.2 (30). For the exon 12-exon 13 junction, the primers used were ³²P-89.2 (5'-GGGCGAGATCAGTCAGCCACTGAGTTC3') and unlabeled 89.3 (see Fig. 1). PCR products were analyzed on 6% sequencing gels. Individual bands were excised from the gel, electroeluted, purified on NEN sorb columns (New England Nuclear), and sequenced by the Maxam and Gilbert procedure (16). For some of the minor bands, a second round of PCR amplification was performed on isolated fragments to increase the yield of DNA for sequencing.

Fragments corresponding to the region extending from exon 7 to the 3' end were generated by PCR using oligo 89.1 as the upstream primer and oligo 89.8, 89.7, or 89.11 as the downstream primer (see Fig. 1). Oligo 89.1 (5'-GGAGAACGTCCACCCGAAACATCAGCAGTGA3') has one mismatch with NCAM message (to generate a *Sal*I site) and is located immediately 5' of the exon 7-exon 8 junction. The sequence of oligo 89.8 (5'-GTGTGCGCGCGCGTGTGTGTTTCAA GTTCC3') was obtained from pR13.2, a cDNA clone encoding rat 120-kDa NCAM (29); oligo 89.7 (5'-CTACAGGACCTTCTCGGGCTCTG3') was obtained from pR18, which encodes the rat 140-kDa isoform (32); and oligo 89.11 (5'-AG AATTTGACTCTGGCAGTTGTGG3') was obtained from p5B4, which contains a partial cDNA sequence of the rat 180-kDa NCAM (21). Oligo 89.8 is complementary to a portion of the 3'-untranslated region of 120-kDa NCAM immediately downstream of the stop codon; oligo 89.7 straddles the exon 17-exon 19 junction in the 140-kDa NCAM; and oligo 89.11 is complementary to a region of exon 18 in the 180-kDa NCAM. PCR conditions were the same as those in the asymmetric PCR experiments, except that 75 ng of each primer was used in a reaction volume of 20 µl, the annealing step was at 65 or 69°C for 1 min, and the extension step was at 72°C for 4 min.

Cloning of PCR products. PCR fragments resulting from the amplification of the region from exon 7 to the 3' end were "polished" by adding 1 to 2 U of *Escherichia coli* DNA polymerase Klenow fragment (New England BioLabs), increasing the [Mg²⁺] to ~7 mM, and incubating the mixture at room temperature for 15 min (38). The DNA polymerase was then inactivated by heating the reaction mixture at 70°C for 15 min and then performing phenol-chloroform extraction. The DNA fragments were then digested with *Sal*I (New England BioLabs) and electrophoresed on a 1% low-melting-point agarose gel. Bands corresponding to the predicted lengths of the PCR products were cut from the gel. DNA fragments were purified by phenol-chloroform extraction. Resulting fragments which had one *Sal*I end and one blunt end were ligated with *Sal*I- and *Sma*I-digested M13mp19, and the ligation mixture was used to transform competent JM109 cells. Individual plaques were picked, and phage polyethylene glycol precipitates were prepared (17). The phage pellets were resuspended in 10 mM Tris-HCl-1 mM EDTA and spotted directly onto GeneScreen Plus (New England Nuclear) with a dot blot manifold (Schleicher & Schuell). The membrane was hybridized sequentially with the following alternative exon-specific oligonucleotide probes: oligo 2 (complementary to the 30-bp VASE exon; exon 7-exon 8 junction) (32); oligo 90.1 (5'-GTGGAGGAGG GCTAT3'; complementary to the complete 15-bp exon; exon 12-exon 13 junction); oligo 90.2 (5'-GTTGAAAGGACAGG AAGAGG3'; complementary to a portion of the 42-bp exon; exon 12-exon 13 junction); and oligo 90.3 (5'-GTTGAAAG

FIG. 1. NCAM alternative splicing. The NCAM polypeptide contains five immunoglobulinlike loops in the extracellular portion of the molecule. The correspondence between exons and polypeptide regions is shown at the top. L designates the leader sequence. The thick vertical bar represents the cell membrane. The region of the polypeptide proximal to the membrane is shown as a dashed line to indicate alternative splicing. Differential splicing between exons 7 and 8, between exons 12 and 13, and at the 3' end is shown in detail at the bottom. The 180- and 140-kDa isoforms are transmembrane polypeptides (exon 16 codes for the transmembrane domain), while the 120-kDa isoform is linked to the membrane via a phosphatidylinositol bridge. The length of the 3'-untranslated region (UT) in the 120-kDa isoform depends on which of two poly(A) addition sites is used. The 140- and 180-kDa isoforms share the same 3'-untranslated region, which is different from that found in the 120-kDa isoform. The positions of oligonucleotides used as PCR primers are indicated by the arrows. Arrows pointing from 5' to 3' represent sense strand primers, while those pointing from 3' to 5' are antisense. Exon sizes are not drawn to scale.

GACAGGAAGAGG3'; complementary to a portion of the 48-bp exon; exon 12-exon 13 junction). Hybridization probes were labeled at the 5' end with T4 polynucleotide kinase (Pharmacia) (32). Hybridizations were performed as described elsewhere (22) at the following temperatures: oligo 2, 60°C; oligo 90.1, 40°C; oligo 90.2, 50°C; and oligo 90.3, 55°C. Posthybridization washes were done at the hybridization temperature in three changes of 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.05% sodium dodecyl sulfate for a total of 15 min. The number of independent clones analyzed for each 3' end class was as follows: 120 kDa, 92; 140 kDa, 67; and 180 kDa, 49.

RESULTS

Alternative splicing at exon 12-exon 13 junction. Alternative splicing at the exon 12-exon 13 junction of NCAM has previously been described primarily for muscle tissue (7, 10, 34). To determine whether tissue- and cell-type-specific patterns of exon 12-exon 13 alternative splicing exist, various rat tissues and cultured cell lines were analyzed by PCR. The primers used for the PCR amplification, oligos 89.2 and 89.3, flank the exon 12-exon 13 junction (Fig. 1). Analysis of PCR products showed that P4 brain tissue yielded bands corresponding to estimated insert sizes of 0, 3, 15, and 18 nucleotides (nt), while the pheochromocytoma line PC12

yielded only the 0- and 3-nt bands (Fig. 2). In contrast, P4 heart tissue and the muscle cell line L6 (postfusion) had more complex patterns, indicating tissue-specific expression of numerous NCAM forms. In addition to the four bands already mentioned, higher-molecular-weight bands were seen in the heart and L6 lanes (Fig. 2), with the major ones at approximately 60, 66, and 108 nt. Liver tissue and the liver cell line H4IIE, both of which express little or no NCAM, yielded only trace amounts of PCR products.

To identify the different alternatively spliced exons in heart and examine in more detail their expression during development, total RNA was prepared from embryonic day 15 (E15), E19, P1, P4, and adult rat hearts and analyzed for exon 12-exon 13 splicing (Fig. 3). Individual PCR product bands were isolated and sequenced. The sequence of the largest fragment (108 nt) is shown in Fig. 4. It is homologous to MSD1, which is alternatively spliced into the exon 12-exon 13 junction in NCAM transcripts from human skeletal muscle (7). The human MSD1 sequence consists of three small exons 15, 48, and 42 bp in size and an AAG triplet (34). Overall, the human and rat sequences are 85 and 73% homologous at the DNA and amino acid levels, respectively. However, the sequence of the 15-bp exon, which codes for three consecutive prolines, is perfectly conserved not only in human and rat cDNA but also in mouse cDNA (27). The respective 48- and 42-bp exons of rat and human cDNA are

FIG. 2. Tissue-specific alternative splicing at the exon 12-exon 13 junction. cDNAs prepared from the indicated tissues and cell lines were subjected to PCR as described in Materials and Methods. Reaction products were separated on a 6% sequencing gel. The 0-nt insert size fragment corresponds to a PCR product of 79 nt.

less conserved, but in both species they code for sequences that are markedly rich in proline, serine, and threonine. Finally, the AAG triplet is found in human, rat, mouse, and chicken cDNA (20, 27, 34). Further studies will be necessary to determine the roles of the highly conserved and less conserved portions of the MSD1 domain.

All of the other rat heart PCR fragments analyzed gave sequences that were various combinations of the three small exons and the AAG triplet. In addition to the 108-bp sequence (designated $15^+48^+42^+3^+$), the major exon 12-exon 13 splicing patterns in rat heart were $15^-48^-42^-3^-$, $15^-48^-42^-3^+$, $15^+48^-42^-3^+$, and $15^+48^+42^-3^+$ (Fig. 3). The following minor forms were also detected: $15^+48^-42^-3^-$, $15^-48^+42^-3^+$, $15^+48^+42^-3^+$, and $15^-48^+42^+3^+$. The relative band intensities within each lane in Fig. 3 demonstrate developmental regulation of expression of the different splice patterns. Thus, while the relative abundance of some alternatively spliced forms appeared approximately constant at all ages examined (for example, $15^+48^-42^-3^+$), that of others increased with age ($15^+48^-42^+3^+$ and $15^+48^+42^+3^+$), but with different times of onset. The relative abundance of other forms decreased during development ($15^+48^+42^-3^+$). Transcripts with $15^-48^+42^+3^+$ and $15^-48^+42^-3^+$ combinations had low abundance in embryonic heart, peaked around P1, and decreased thereafter. The expression of $15^-48^-42^-3^-$ also varied with age, but changes in its relative abundance did not exhibit a simple pattern. It should be noted that although the MSD1 sequence ($15^+48^+42^+3^+$) was present in a significant fraction of NCAM transcripts at all ages examined, it was only one of four or five major spliced forms at any one age. The amino acid residues encoded by the different combinations of alternatively spliced exons are shown in Table 1.

FIG. 3. Alternative splicing at the exon 12-exon 13 junction during heart development. cDNAs prepared from heart tissues of rats of different ages were subjected to PCR as described in Materials and Methods. Sequencing reactions were run in parallel lanes to provide molecular weight markers (not shown). The 0-nt insert size fragment corresponds to a PCR product of 79 nt. R5.6 and R8.2 are L-cell fibroblast lines stably transfected with $30^+/15^-48^-42^-3^-$ and $30^-/15^-48^-42^-3^-$ forms of 140-kDa rat NCAM cDNAs, respectively.

Alternative exon 12-exon 13 PCR products were also detected in rat brain. However, only products with sizes of 0, 3, 15, and 18 nt were observed (Fig. 2). These were sequenced and found to correspond to the spliced forms $15^-48^-42^-3^-$, $15^-48^-42^-3^+$, $15^+48^-42^-3^-$, and $15^+48^-42^-3^+$, respectively. Similarly, only the 15-bp exon and the AAG triplet have been observed in human and mouse brain NCAM (27, 34). Alternative splicing choices at the exon 12-exon 13 junction were even less extensive in the cell line PC12, where only the $15^-48^-42^-3^-$ and $15^-48^-42^-3^+$ patterns were observed. Restricted differential splicing at the exon 12-exon 13 junction may therefore be characteristic of neural NCAMs, whereas more heterogeneous patterns are found in muscle NCAMs.

Alternative splicing at the exon 7-exon 8 junction. Alternative splicing of the 30-bp VASE exon at the exon 7-exon 8 junction of NCAM transcripts was observed first in rat brain (32) and later in mouse brain (27). To determine whether a similar situation exists in rat heart, splicing at the exon 7-exon 8 junction was analyzed with rats of different ages by PCR (Fig. 5). Only two bands, corresponding to insert sizes of 0 and 30 bp, were seen. The 30-bp insert in heart NCAM was shown to be identical to VASE (30). The proportion of heart NCAM transcripts with VASE increased with age, from ~10% in E15 heart to ~50% in adult heart. This result is analogous to the situation in rat brain, where the expression of VASE increased dramatically from 3% of NCAM transcripts in E15 heart to 43% in adult brain (31).

NCAM diversity resulting from alternative splicing permutations. In addition to the exon 7-exon 8 and exon 12-exon 13 junctions, alternative splicing occurs at the 3' end of the NCAM message (4, 18). To systematically determine the frequency of mRNAs containing specific splicing combinations at each of these three regions, PCR experiments amplifying the entire region between exon 7 and the 3' end of

FIG. 4. Alternatively spliced sequences at the exon 12-exon 13 junction of NCAM mRNAs. The rat and human (MSD1) inserts at the exon 12-exon 13 junction are compared. Identical positions are indicated by dots. The boundaries of the three smaller exons (MSD1a through MSD1c in human mRNA) and the AAG triplet are also indicated.

the NCAM transcript were performed (Fig. 1). Three downstream PCR primers were chosen, with one specific for each of the 120-kDa (oligo 89.8), 140-kDa (oligo 89.7), and 180-kDa (oligo 89.11) 3' end sequences (Fig. 1). The same upstream primer (oligo 89.1) located immediately 5' of the exon 7-exon 8 junction was used for each reaction. P1 heart was used as source of NCAM message because all nine exon 12-exon 13 alternative splice forms were present at this age. The fragments resulting from PCR amplification, which had the predicted lengths of 1.2 kb (oligo 89.8 primed), 1.4 kb (oligo 89.7 primed), and 1.6 kb (oligo 89.11 primed), were isolated and cloned into M13 (data not shown; see Materials and Methods). Thus, three populations of PCR clones were obtained, with each population representing a specific 3' end class. Individual plaques were isolated and characterized by sequential hybridization to probes complementary to the 30-bp VASE exon at the exon 7-exon 8 junction or to the 15-, 48-, or 42-bp exon at the exon 12-exon 13 junction. The AAG triplet was too small to be detected by this method.

The frequency of specific exon 7-exon 8-exon 12-exon 13 splice combinations detected in each class of PCR clones is shown in Table 2. The clones are further divided into two groups according to the presence or absence of the 30-bp VASE exon at the exon 7-exon 8 junction. The predominant exon 12-exon 13 splice patterns in the cloned sequences were the same as those present as major bands in the more localized PCR amplification of the exon 12-exon 13 junction (Fig. 3). Overall, it appears that any exon 12-exon 13 splice pattern can combine with either 30⁺ or 30⁻ at the exon 7-exon 8 junction or with any one of three 3' end classes.

(Although a few clones of the 180-kDa class had splice combinations not found in the 120- and 140-kDa classes, these clones were present at a low frequency. Hence, the inability to detect these in the other 3' end classes could be simply due to statistical variations.) However, the assortment of splice patterns found at the exon 7-exon 8 junction, the exon 12-exon 13 junction, and the 3' end did not appear entirely random. For example, the 15⁻48⁻42⁻ pattern was preferentially spliced to 30⁻ in all three 3' end classes. On the other hand, the 15⁺48⁻42⁻ pattern was more frequently combined with 30⁻ in the 120- and 140-kDa classes but not in the 180-kDa class. The 30⁻/15⁺48⁻42⁻ pattern made up a larger fraction of the 120- and 140-kDa molecules (33% in both) than of the 180-kDa molecules (18%). Finally, 30⁺ clones (group 1 in Table 2) were present in a significantly higher percentage of 180-kDa molecules (47%) than 120-kDa or 140-kDa molecules (32% in both). This quantitative determination of combinations of NCAM splicing thus demonstrates slight preferences for some splicing patterns to occur together, but overall most splicing events appear to occur independently.

DISCUSSION

NCAM expression in developing rat hearts is regulated in a complex and dynamic fashion. Wharton et al. (37) used immunostaining techniques to study NCAM abundance in different rat cardiac cell types and regions. Expression of NCAM in myocardial cells was detected beginning at E12, peaked around the time of birth, and then declined rapidly

TABLE 1. Alternatively spliced sequences at the exon 12-exon 13 junction of rat heart NCAM^a

Exon combination	Amino acid sequence
15 ⁻ 48 ⁻ 42 ⁻ 3 ⁻	Arg
15 ⁻ 48 ⁻ 42 ⁻ 3 ⁺	GlnGly
15 ⁺ 48 ⁻ 42 ⁻ 3 ⁻	HisSerProProProArg
15 ⁺ 48 ⁻ 42 ⁻ 3 ⁺	HisSerProProProGlnGly
15 ⁻ 48 ⁺ 42 ⁻ 3 ⁺	LeuAlaProAlaAsnSerSerThrLeuValProLeuSerProArgAlaLysGly
15 ⁺ 48 ⁻ 42 ⁺ 3 ⁺	HisSerProProProProThrTrpProLeuProValLeuSerThrThrGluLeuSerLysGly
15 ⁺ 48 ⁺ 42 ⁻ 3 ⁺	HisSerProProProLeuAlaProAlaAsnSerSerThrLeuValProLeuSerProArgAlaLysGly
15 ⁻ 48 ⁺ 42 ⁺ 3 ⁺	LeuAlaProAlaAsnSerSerThrLeuValProLeuSerProArgAlaThrThrTrpProLeuProValLeuSerThrThrGluLeuSerLysGly
15 ⁺ 48 ⁺ 42 ⁺ 3 ⁺	See Fig. 4

^a The different alternatively spliced exon combinations at the exon 12-exon 13 junction and the amino acid sequences they encode are shown. The residues in boldface are encoded by codons found at splice junctions. In the absence of any alternative splicing (15⁻48⁻42⁻3⁻), an arginine residue is present at the exon 12-exon 13 junction within the sequence GlnProValArgGluProSer (amino acid residues 588 to 594 in reference 32).

FIG. 5. Alternative splicing at the exon 7-exon 8 junction during heart development. cDNAs prepared from heart tissues of rats of different ages were subjected to PCR as described in Materials and Methods. The 0-nt insert size band corresponds to a PCR product of 88 nt.

postnatally. In contrast, cardiac nerves exhibited a high level of NCAM expression beginning at E18, when cardiac innervation was first observed, and continuing into adulthood. NCAM was also detected in the smooth muscle regions of the aorta and proximal pulmonary arteries but not in coronary branches (37). In the aorta, NCAM staining was detected at E17 and then decreased substantially postnatally (1).

Cardiac NCAM expression is also regulated by alternative splicing. Tissue- and developmental stage-specific alternative splicing in other systems is extensively documented (33). One example of tissue-specific alternative splicing that results in differences in cell adhesion is the splicing of

fibronectin. Fibronectin mRNAs are alternatively spliced in different tissues and also among distinct tumors derived from tissues to generate forms of the polypeptide which vary in the number of adhesive domains expressed (9, 11). An example of developmental stage-specific alternative splicing that results in a change in protein localization is that of the immunoglobulin μ heavy chain gene. Alternative splicing of 3' end exons results in the production of a membrane-bound form in pre-B lymphocytes and a secreted form in mature B cells (2, 24).

Alternative 3' end splicing of cardiac NCAM transcripts, which results in transmembrane or membrane-linked isoforms, is developmentally regulated. Thus, immunoblot analysis showed predominant expression of the 140-kDa isoform in embryonic heart and of the 120-kDa isoform postnatally (37). PCR amplification of P1 heart NCAM mRNAs indicates the presence of all three 3' end classes (Table 2). The transcript for the 180-kDa isoform was less abundant than that for either the 120-kDa or 140-kDa isoform (data not shown), consistent with the finding that 180-kDa NCAM is neuron specific (8).

The results presented here show that heart NCAM expression is also regulated by alternative splicing at the exon 12-exon 13 junction. A number of splicing patterns in this region are found in heart tissue but not in other tissues (Fig. 2). At least nine different combinations of the 15-, 48-, and 42-bp exons and the AAG triplet were found to be alternatively spliced. Assuming that each small exon and the triplet can be spliced independently of the others, 16 combinations are theoretically possible. Since all of the detected fragments that were longer than 18 nt contained the AAG triplet, it is possible that the corresponding 3' sequence of each of these fragments was present at a lower abundance but was not resolved from the 3' sequence during electrophoresis. However, some combinations were clearly more abundant than others (Fig. 3), implying that at least some of the combinations not observed were present either at amounts below the detection limit of the PCR conditions used or not at all. Thus, preferred combinations of the three small exons and the AAG triplet exist. Furthermore, the expression of a number of these forms appeared to be developmentally regulated. All these observations indicate that NCAM splicing at the exon 12-exon 13 junction is not stochastic but is regulated, like that at the exon 7-exon 8 junction and the 3' terminus.

Any given exon 12-exon 13 splice pattern can combine with 30⁺ or 30⁻ at the exon 7-exon 8 junction or with any one of three 3' end classes. However, preferred combinations are apparent, at least for age P1. A total of 27 NCAM forms was cloned at this age. If each form can exist as 3⁺ or 3⁻, then 54 forms are possible. (It should also be noted that the minor 15⁻48⁺42⁻3⁺ form detected in Fig. 3 was not represented in any of the PCR clones isolated.) It has been suggested that NCAM have important functions in cardiac morphogenesis and innervation (37). Characterization of the spatiotemporal pattern of expression of the multiple NCAM forms in heart should provide further insight into the roles they play in these processes. Cardiac nerve axons of sympathetic origin presumably do not contribute to the NCAM mRNAs analyzed here, as axons do not contain significant RNA; however, there may be a minor contribution from parasympathetic ganglia. While all muscle cells express NCAM, it cannot be determined from the data reported here whether individual cells or cardiac regions express all the NCAM forms observed or whether some NCAM forms are selectively expressed by some cells or regions.

TABLE 2. Frequency of alternative exon combinations in P1 rat heart NCAM^a

Group and exon combination (exon 7-exon 8 junction/exon 12-exon 13 junction)	% of clones in indicated 3' end class containing exon combination		
	120 kDa	140 kDa	180 kDa
Group 1 (VASE⁺)			
30 ⁺ /15 ⁺ 48 ⁺ 42 ⁺	1	6	6
30 ⁺ /15 ⁺ 48 ⁺ 42 ⁻	10	6	2
30 ⁺ /15 ⁺ 48 ⁻ 42 ⁺	0	0	2
30 ⁺ /15 ⁻ 48 ⁺ 42 ⁺	0	0	2
30 ⁺ /15 ⁺ 48 ⁻ 42 ⁻	16	15	23
30 ⁺ /15 ⁻ 48 ⁻ 42 ⁻	5	5	12
Subtotal	32	32	47
Group 2 (VASE⁻)			
30 ⁻ /15 ⁺ 48 ⁺ 42 ⁺	7	9	4
30 ⁻ /15 ⁺ 48 ⁺ 42 ⁻	13	8	8
30 ⁻ /15 ⁺ 48 ⁻ 42 ⁺	0	0	2
30 ⁻ /15 ⁺ 48 ⁻ 42 ⁻	33	33	18
30 ⁻ /15 ⁻ 48 ⁻ 42 ⁻	15	18	21
Subtotal	68	68	53

^a PCR products resulting from the amplification of the entire NCAM region between exon 7 and the 3' end were cloned into M13, and individual clones were hybridized to alternative exon-specific probes.

It is reasonable to predict that differential splicing can result in NCAM polypeptides with widely different properties. For example, it has been proposed that the insertion of the 30-bp VASE sequence at the exon 7-exon 8 junction changes the structure of the fourth immunoglobulinlike loop of NCAM from constantlike to variablelike (31). Alternative splicing at the exon 12-exon 13 junction results in the insertion of additional amino acids near the end of a partial fibronectin type III repeat (7, 20). Since the other eight alternative splice patterns detected in the exon 12-exon 13 junction of rat heart NCAM consist of different combinations of the small exons found in the 108-bp sequence, their corresponding amino acid sequences are related to that of the full-length 36-residue insert. However, the shorter inserts also contain novel amino acid residues encoded by codons that straddle splice junctions (Table 1). It is interesting that seven of nine alternative splice patterns contain the AAG triplet (Table 1; Fig. 3). In all theoretically possible splice patterns, a 3⁻ pattern encodes an Arg (CGG or AGG) in the C terminus of the insert, whereas a 3⁺ pattern encodes GluGly (CAAGGG) or LysGly (AAAGGG). The functional significance of the preference for a 3⁺ pattern remains to be determined.

The 108-bp human MSD1 sequence and its rat homolog code for 36 amino acid residues that are rich in proline, serine, and threonine. The 15-bp exon is perfectly conserved in rat and human cDNA and codes for three consecutive prolines. When spliced together, the 15- and 42-bp exons in rat cDNA encode four prolines in a row. Such stretches of proline residues could result in the introduction of a bend in the region of the molecule between the fifth immunoglobulinlike loop and the membrane-linked or transmembrane domain (20, 34). Furthermore, an O-linked oligosaccharide has been mapped to MSD1 (36). It has been suggested that the presence of this carbohydrate molecule could affect NCAM-NCAM side-by-side interaction, or even protein turnover, thus affecting cell-surface distribution or the abundance of NCAM (35). The presence of this sugar molecule on NCAM could also affect the interaction between NCAM and the adhesion molecule L1 (13). Finally, the differences in membrane attachment modes resulting from alternative 3' end splicing could affect intracellular signaling (28) or interaction with the cytoskeleton (19).

Accurate quantitation of the relative amounts of specific transcripts present in an RNA population by the PCR method requires that the different mRNAs that are analyzed be copied into the corresponding cDNAs with the same efficiency and that the amount of amplified PCR product be proportional to the amount of template cDNA. In the experiments described here, these conditions were not tested directly. However, additional lines of evidence suggest that the results of the PCR amplification experiments do reflect the relative abundance of the different NCAM mRNAs. For example, in the analyses of alternative splicing at a single splice junction, it has been shown that, at least for the exon 7-exon 8 junction, PCR conditions similar to those used here yielded VASE⁺ and VASE⁻ product bands with relative intensities consistent with the amounts of transcripts as determined by Northern (RNA) analysis (30).

In the case of the more extended PCR amplifications (Table 2), it is harder to compare the abundance of the 120-, 140-, and 180-kDa isoform messages because a different 3' end primer was used for each amplification. However, when the same set of primers was used to amplify NCAM messages in P4 and adult rat brain under the same PCR conditions (data not shown), the results were consistent with

Northern hybridization data (32); that is, in the adult brain the predominant species detected was the message for the 120-kDa NCAM, while in P4 brain the most abundant NCAM transcript was that encoding the 180-kDa form. Within each 3' end class, product bands resulting from the extended PCR amplifications can vary in length by a maximum of 138 bp (a transcript with the 30-bp VASE exon at the exon 7-exon 8 junction and 108 bp at the exon 12-exon 13 junction, compared with one without any alternatively spliced sequence). This maximum difference is only about 10% of the total length of the various PCR products (1.2 to 1.6 kb) and probably does not affect the efficiency of the amplification reaction.

Thus it is reasonable to conclude that the PCR results presented here demonstrate not only the diversity of NCAM expressed during cardiac development but also the approximate relative abundance of each of those forms. The availability of NCAM clones containing different combinations of alternatively spliced exons makes possible the functional characterization of each of the different NCAM polypeptides expressed in heart tissue and should ultimately aid in elucidating the exact mechanisms by which multiple NCAM forms function in development.

ACKNOWLEDGMENTS

We thank Sandra Degen for assistance with DNA sequencing, Shirley Arnold for technical assistance, and Chris Disney for assistance with manuscript preparation.

This work was supported by NIH grant HD21065 to R.A.

REFERENCES

1. Akeson, R. A., J. R. Wujek, S. Roe, S. L. Warren, and S. J. Small. 1988. Smooth muscle cells transiently express NCAM. *Mol. Brain Res.* 4:107-120.
2. Alt, F. W., A. L. M. Bothwell, M. Knapp, E. Siden, and E. Mather. 1980. Synthesis of secreted and membrane bound immunoglobulin μ heavy chains is directed by mRNAs that differ at their 3' ends. *Cell* 20:293-302.
3. Barbas, J. A., J.-C. Chaix, M. Steinmetz, and C. Goridis. 1988. Differential splicing and alternative polyadenylation generates distinct NCAM transcripts and proteins in the mouse. *EMBO J.* 7:625-632.
4. Barthels, P., M.-J. Santoni, W. Wille, C. Ruppert, J.-C. Chaix, M.-R. Hirsch, J. C. Fontecilla-Camps, and C. Goridis. 1987. Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79,000 polypeptide without a membrane-spanning region. *EMBO J.* 6:907-914.
5. Chen, A., A. Reyes, and R. Akeson. 1990. Transcription initiation sites and structural organization of the extreme 5' region of the rat neural cell adhesion molecule gene. *Mol. Cell. Biol.* 10:3314-3324.
6. Cole, G. J., and R. A. Akeson. 1989. Identification of a heparin binding domain of the neural cell adhesion molecule N-CAM using synthetic peptides. *Neuron* 2:1157-1165.
7. Dickson, G., H. J. Gower, C. H. Barton, H. M. Prentice, V. L. Elsom, S. E. Moore, R. D. Cox, C. Quinn, W. Putt, and F. S. Walsh. 1987. Human muscle neural cell adhesion molecule (N-CAM): identification of a muscle specific sequence in the extracellular domain. *Cell* 50:1119-1130.
8. Edelman, G. M. 1988. Morphoregulatory molecules. *Biochemistry* 27:3533-3543.
9. Ffrench-Constant, C. and R. O. Hynes. 1989. Alternative splicing of fibronectin is temporally and spatially regulated in the chicken embryo. *Development* 106:375-388.
10. Gower, H. J., C. H. Barton, V. L. Elsom, J. Thompson, S. E. Moore, G. Dickson, and F. S. Walsh. 1988. Alternative splicing generates a secreted form of N-CAM in muscle and brain. *Cell* 55:955-964.
11. Hershberger, R. P., and L. A. Culp. 1990. Cell-type-specific expression of alternatively spliced human fibronectin III/CS

- mRNAs. *Mol. Cell. Biol.* **10**:662–671.
12. Hirsch, M.-R., L. Gaugler, H. Deagostini-Bazin, L. Bally-Cuif, and C. Goridis. 1990. Identification of positive and negative regulatory elements governing cell-type-specific expression of the neural cell adhesion molecule gene. *Mol. Cell. Biol.* **10**: 1959–1968.
 13. Kadmon, G., A. Kowitz, P. Altevogt, and M. Schachner. 1990. The neural cell adhesion molecule N-CAM enhances L1-dependent cell-cell interactions. *J. Cell Biol.* **110**:193–208.
 14. Key, B., and R. A. Akeson. 1990. Olfactory neurons express a unique glycosylated form of the neural cell adhesion molecule (N-CAM). *J. Cell Biol.* **110**:1729–1743.
 15. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 16. Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**:499–560.
 17. Messing, J. 1983. New M13 vectors for cloning. *Methods Enzymol.* **101**:20–78.
 18. Owens, G. C., G. M. Edelman, and B. A. Cunningham. 1987. Organization of the neural cell adhesion molecule (N-CAM) gene: alternative exon usage as the basis for different membrane-associated domains. *Proc. Natl. Acad. Sci. USA* **84**:294–298.
 19. Pollerberg, G. E., M. Schachner, and J. Davoust. 1986. Differentiation state-dependent surface mobilities of two forms of the neural cell adhesion molecule. *Nature (London)* **324**:462–465.
 20. Prediger, E. A., S. Hoffman, G. M. Edelman, and B. A. Cunningham. 1988. Four exons encode a 93-base-pair insert in three neural cell adhesion molecule mRNAs specific for chicken heart and skeletal muscle. *Proc. Natl. Acad. Sci. USA* **85**:9616–9620.
 21. Ramos, P., R. Safaei, C. Kayalar, and L. Ellis. 1989. Isolation and sequence of lambda-gt10 cDNA encoding the 5B4 antigen expressed on sprouting neurons. *Mol. Brain Res.* **5**:297–303.
 22. Reyes, A. A., and R. Akeson. 1988. Generation of multiple independent substitution mutants by M13 in vitro mutagenesis using a single mutagenic oligonucleotide. *DNA* **7**:579–584.
 23. Reyes, A. A., R. Akeson, L. Brezina, and G. J. Cole. 1990. Structural requirements for neural cell adhesion molecule (NCAM)-heparin interaction. *Cell Regul.* **1**:567–576.
 24. Rogers, J., P. Early, C. Carter, K. Calame, and M. Bond. 1980. Two mRNAs with different 3' ends encode membrane bound and secreted forms of immunoglobulin m chain. *Cell* **20**:303–312.
 25. Rutishauser, U., and T. M. Jessell. 1988. Cell adhesion molecules in vertebrate neural development. *Physiol. Rev.* **68**:819–857.
 26. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**:487–491.
 27. Santoni, M. J., D. Barthels, G. Vopper, A. Boned, C. Goridis, and W. Wille. 1989. Differential exon usage involving an unusual splicing mechanism generates at least eight types of NCAM cDNA in mouse brain. *EMBO J.* **8**:385–391.
 28. Schuch, U., M. J. Lohse, and M. Schachner. 1989. Neural cell adhesion molecules influence second messenger systems. *Neuron* **3**:13–20.
 29. Small, S. J. 1988. Ph.D. thesis. University of Cincinnati, Cincinnati, Ohio.
 30. Small, S. J., and R. Akeson. 1990. Expression of the unique NCAM VASE exon is independently regulated in distinct tissues during development. *J. Cell Biol.* **111**:2089–2096.
 31. Small, S. J., S. L. Haines, and R. Akeson. 1988. Polypeptide variation in an N-CAM extracellular immunoglobulin-like fold is developmentally regulated through alternative splicing. *Neuron* **1**:1007–1017.
 32. Small, S. J., G. E. Shull, M.-J. Santoni, and R. Akeson. 1987. Identification of a cDNA clone that contains the complete coding sequence for a 140-kD rat NCAM polypeptide. *J. Cell Biol.* **105**:2335–2345.
 33. Smith, C. W. J., J. G. Patton, and B. Nadal-Ginard. 1989. Alternative splicing in the control of gene expression. *Annu. Rev. Genet.* **23**:527–577.
 34. Thompson, J., G. Dickson, S. E. Moore, H. J. Gower, W. Putt, J. G. Kenimer, C. H. Barton, and F. S. Walsh. 1989. Alternative splicing of the neural cell adhesion molecule gene generates variant extracellular domain structure in skeletal muscle and brain. *Genes Dev.* **3**:348–357.
 35. Walsh, F. S., and G. Dickson. 1989. Generation of multiple N-CAM polypeptides from a single gene. *Bioessays* **11**:83–88.
 36. Walsh, F. S., R. B. Parekh, S. E. Moore, G. Dickson, C. H. Barton, H. J. Gower, R. A. Dwek, and T. W. Rademacher. 1989. Tissue specific O-linked glycosylation of the neural cell adhesion molecule (N-CAM). *Development* **105**:803–811.
 37. Wharton, J., L. Gordon, F. S. Walsh, T. P. Flanigan, S. E. Moore, and J. M. Polak. 1989. Neural cell adhesion molecule (N-CAM) expression during cardiac development in the rat. *Brain Res.* **483**:170–176.
 38. Williams, J. F. 1989. PCR questions and answers. *Amplifications* **3**:19.