

# The muscle specific domain of mouse N-CAM: structure and alternative splicing patterns

Marion Hamshere, George Dickson<sup>1</sup> and Ian Eperon\*

Department of Biochemistry, University of Leicester, Leicester LE1 7RH and <sup>1</sup>Department of Experimental Pathology, UMDS, Guys Hospital, London Bridge, London SE1 9RT, UK

Received May 2, 1991; Revised and Accepted August 5, 1991

EMBL accession no. X59306

## ABSTRACT

The neural cell adhesion molecule (N-CAM) is an important mediator of calcium independent cell-cell interactions. Variations in the primary structure of the protein are due to alternative splicing of pre-mRNA in the region encoding the extracellular, trans-membrane and cytoplasmic domains. In order to identify the patterns of exon usage during development of skeletal muscle and brain of the mouse, a coupled reverse-transcriptase/polymerase chain reaction was used to identify the murine homologues of the muscle-specific domain (MSD), located between exons 12 and 13 in human N-CAM mRNA. The cDNAs produced have been cloned and sequenced, or analysed directly. The amplification reactions were shown to maintain the concentration ratios of the initial cDNAs. The results indicate that the mouse homologue to exon MSD1a is under tissue and developmental regulation that is independent of exons MSD1b and MSD1c. The inclusion of the triplet exon AAG is also regulated in a cell- and stage-specific manner, which is independent of the other alternatively spliced exons of this domain.

## INTRODUCTION

Neural cell adhesion molecules (N-CAMs) are a group of cell-surface sialoglycoproteins which belong to the immunoglobulin gene superfamily. Proteins of this superfamily are characterised by extracellular regions of  $\beta$ -pleated sheet stabilised by disulphide bridges ('Ig folds')<sup>1</sup>. N-CAM mediates calcium-independent interactions between cells; homophilic binding of N-CAMs on apposing cells results in enhanced or extended interactions between cells. These interactions are modulated by post-translational modifications involving the developmentally-regulated attachment of an unusual polysialic acid (PSA)-containing carbohydrate structure<sup>2</sup>. Sites involved in homophilic binding have been localised to two regions within the first three Ig-folds<sup>1</sup>.

A number of isoforms of N-CAM have been reported which differ in their mode of membrane attachment and in cytoplasmic domain structure, although only a single N-CAM gene has been

demonstrated in a variety of species. Based on analysis of cDNA clones from a variety of organisms<sup>3,4,5,6</sup>, these isoforms appear to arise from patterns of alternative splicing that are determined in a developmental and tissue-specific manner<sup>7,8</sup>. Alternative splicing gives rise to two additional regions of diversity, both in the extracellular domain. Variants include novel sequence segments between exons 7 and 8<sup>9,10</sup>, involving the fourth Ig domain, and also between exons 12 and 13<sup>8,11,12</sup>, in the membrane proximal 'stem' structure of N-CAM polypeptides. In the latter instance, an alternatively-spliced sequence block of 4 exons (MSD1a, MSD1b, MSD1c, and triplet AAG) has been described in chick and human N-CAM mRNA in striated muscle tissues; this has been termed the muscle specific domain (MSD). The unusual 3 base pair insert (triplet AAG) is not due to allelic variation<sup>13</sup>. Cell transfection studies using full-length cDNA clones have further indicated that expression of isoforms containing the MSD region in cultured muscle cells results in enhanced fusion of mononucleate precursor myoblasts into differentiated multinucleate myotubes. Part of the muscle-specific domain (MSD) introduces a string of proline residues which may induce a bend or hinge in the molecule; this may affect either *trans*-acting homophilic binding or cell-cell interaction via a *cis*-acting effect on interaction with other cell-surface adhesion molecules<sup>14</sup>. The MSD is a site of extensive post-translational modification with O-linked carbohydrate, which is characteristic of hinge regions in other Ig superfamily members<sup>15</sup>. The genomic sequences of two exons of the MSD (1b and 1c) have been identified in the human<sup>16</sup> although only the homologue to exon MSD1a and a putative site for the triplet exon AAG have been identified in a genomic clone from the mouse<sup>9</sup>. The exons of this domain are unusually short: 15, 48, 42 and 3 base pairs long for MSD 1a, MSD 1b, MSD 1c and triplet AAG respectively.

Although the human MSD has been found as a single unit in skeletal muscle, evidence has accumulated for the individual regulation of at least 2 of its constituent exons. This is clearly the case for exon MSD1a in the mouse, which has been found without exons MSD1b and MSD 1c in neural tissue. In order to establish whether the component exons of the MSD are regulated collectively, by splicing during muscle development, we have characterized fully the mouse homologues of the MSD

\* To whom correspondence should be addressed

exons described previously in the chick<sup>17</sup>, human<sup>11</sup> and rat<sup>8</sup> and we have examined in detail the patterns during brain and in particular during skeletal muscle development.

## MATERIALS AND METHODS

### Cell and tissue samples

Skeletal muscle (lower hind limb) and brain tissues were dissected from embryonic (E16), newborn (post natal 3 day), and adult (10 week old) Balb c mice. Samples of denervated adult mouse muscle were obtained 7 days after surgical section of the sciatic nerve as previously described<sup>18</sup>. Mouse C-2 myoblasts were grown as described previously<sup>7</sup>, and mid-fusion and myotube samples were obtained 2 and 4 days following initiation of myogenic differentiation by switching cultures into medium containing horse serum. N2A neuroblastoma, C6 glioma and SWA schwannoma cells<sup>15</sup> were grown to confluence before harvesting for RNA preparation.

### RNA preparation

RNA was extracted from cell lines and mouse tissues by the addition of 4ml 3M LiCl /6M urea<sup>19</sup>. Tubes were incubated on ice for at least 2 hours then centrifuged at 10,000rpm for 30 minutes. The pellet was washed with cold LiCl/urea and resuspended in 300 $\mu$ l TE.1/0.5% SDS. Three chloroform/isoamyl alcohol extractions were performed and the RNA finally precipitated with ethanol/sodium acetate (20/1) and stored at  $-70^{\circ}\text{C}$ .

Poly(A)<sup>+</sup> RNA was isolated from total RNA using either the spun columns available from Pharmacia, or with mAP<sup>TM</sup> paper from Amersham, following the protocol for small sample volumes. Both methods gave comparable results for isoform variation in adult mouse brain and adult muscle when analysed by RT-PCR (data not shown).

### RT-PCR

25pmol of oligonucleotide MH2b (GCCGGCGCGGAGCTTTCTGCCCTTCCAGCTTGGGT), complementary to the constitutive downstream exon 13, was added to poly(A)<sup>+</sup> RNA purified from 5 $\mu$ g total RNA in a buffer containing 200 $\mu$ M dNTPs, 50mM KCl, 10mM Tris-HCl pH 8.0, 1.5mM MgCl<sub>2</sub>. The RNA was denatured at 90 $^{\circ}\text{C}$  for 5 minutes then cooled on ice. 20 units of RNAGuard<sup>TM</sup> was added and the primer allowed to anneal at 65 $^{\circ}\text{C}$  for 5 minutes. 10 units of MMuLV Reverse Transcriptase (RT) was added and the reaction incubated at 37 $^{\circ}\text{C}$  for one hour. The first strand cDNA was amplified by the addition of 20pmol MH1b2(CCCCCGCCCGAATCCCACTGAGTTC AAGACACAG) oligonucleotide and a further 10pmol MH2b, and 2 units of Taq polymerase to give a final volume of 100 $\mu$ l in the same buffer. Samples were overlaid with paraffin and amplified by 30 cycles of: 92 $^{\circ}\text{C}$  denaturation for 1.5 minutes, 65 $^{\circ}\text{C}$  annealing for 1.5 minutes, and 72 $^{\circ}\text{C}$  extension for 2 minutes. Final steps comprised annealing at 65 $^{\circ}\text{C}$  annealing for 1.5 minute and extension at 72 $^{\circ}\text{C}$  for 5 minutes to ensure that full length double stranded PCR products were formed.

### Cloning and sequencing of cDNA

The cDNA products were cut with HindIII and EcoRI in standard reaction buffers and cloned into cut mICE 18<sup>20</sup>. Ligations were transfected into competent JM109 cells. 96 plaques from each cDNA cloning were selected and grown in 250 $\mu$ l TY broth in microtitre plates with constant shaking for 9 hours. Single

stranded template was prepared as described<sup>21</sup>, except that the cells were removed by centrifugation rather than filtration. 150 $\mu$ l of supernatant was transferred into a second microtitre plate. ssDNA was precipitated by the addition of 30 $\mu$ l PEG/NaCl (20%/2.5M). After 10 minutes the plates were centrifuged at 3,000 $\times$ g, and the supernatant discarded. Pellets were resuspended in 50 $\mu$ l TE.1/ 1% SDS and incubated at 70 $^{\circ}\text{C}$  for 10 minutes. DNA was finally precipitated by the addition of ethanol/sodium acetate (20/1), washed with 70% ethanol and resuspended in TE.1. Sequencing reactions were as described<sup>22</sup>.

### Direct analysis of amplified products

1/100th of a first round RT-PCR was transferred to a second PCR. The amplification of the substrate was undertaken using 10pmol of a nested oligonucleotide MH ex13 (AACCCAGTGC-ACCCAAG) internal to the cDNA product from the first round amplification, with 10pmol MH2b. 2.5 $\mu$ Ci [ $\alpha$ -<sup>32</sup>P] dATP was added to each reaction containing 200 $\mu$ M dNTPs, 50mM KCl, 10mM Tris-HCl pH 8.0, 1.5mM MgCl<sub>2</sub>. Samples were overlaid with paraffin and amplified for 15 cycles of denaturation (92 $^{\circ}\text{C}$ , 1.5 minutes), annealing(65 $^{\circ}\text{C}$ , 1.5 minutes), extension (72 $^{\circ}\text{C}$ , 2 minutes), followed by 59 $^{\circ}\text{C}$  for 1.5 minutes and 72 $^{\circ}\text{C}$  for 5 minutes. After precipitation with ethanol, samples were loaded onto 10% native polyacrylamide gels.

### Identification of PCR products

The identity of bands separated on native polyacrylamide gels was confirmed by the direct sequencing of re-amplified excised bands. End-labelled oligonucleotide MH ex13 and dideoxynucleoside triphosphate were used in a PCR.

## RESULTS

### Cloning of mouse homologue to human MSD region

The region of murine N-CAM equivalent to that between exons 12 and 13 (containing the possible MSD homologue) was amplified from cDNA by PCR (Figure 1). The products were cloned and the sequences determined for 78 clones, 34 of which were derived from myoblast RNA and 44 from myotube RNA. Ten of the clones incorporated all of the exons homologous to the human MSD1a, MSD1b, MSD1c and triplet AAG exons. Figure 2 shows such a cDNA sequence, flanked by regions of the constitutive exons 12 and 13, aligned with the MSD exons identified in the human and chicken. The other cDNA clones contained various combinations of these exons, and are illustrated in Figure 3.

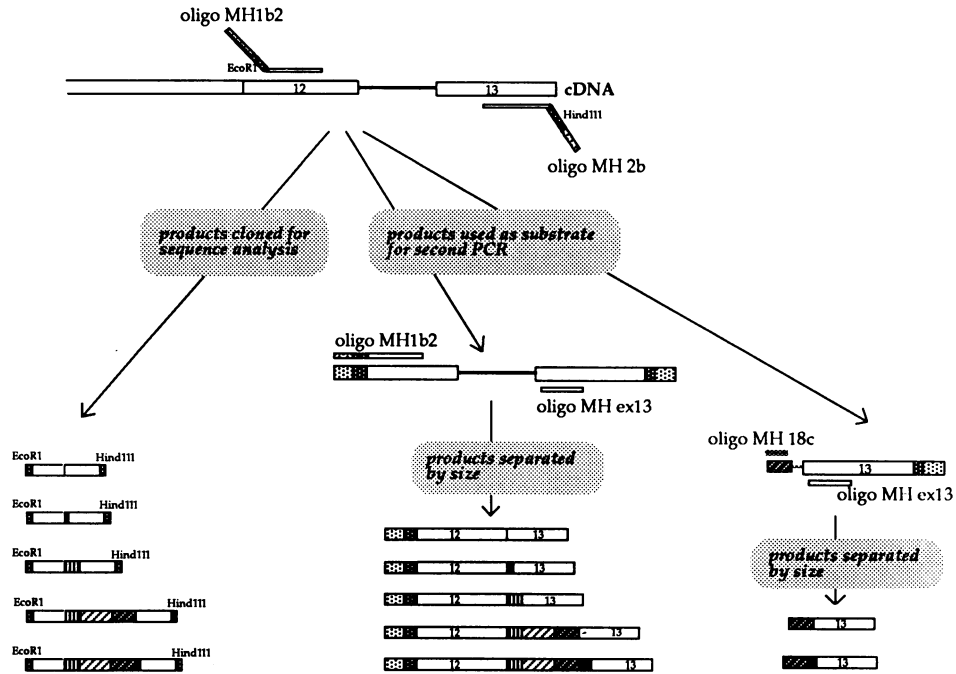
The sequence analysis produced statistically significant figures for the proportions of the major mRNA forms in myoblasts and myotubes. Figure 3 indicates the putative alternative splicing patterns required to produce the alternative forms, and the distribution of variant forms within each cell type. Significant differences ( $P < 0.005$ ) are noted for the exon usage in undifferentiated myoblasts and differentiated myotubes. These clones include several unreported combinations of exons between the constitutive exons 12 and 13, in the region homologous to human MSD (12-a-b-c-13, 12-a-b-c-AAG-13, 12-a-c-13, and 12-c-AAG-13).

### Quantification of PCR products

The analysis described above involved reverse transcription, amplification and cloning reactions. Although reverse transcription is unlikely to preserve perfectly the original

proportions of isoforms, especially when they differ in length, it is a standard method for determining the ratios of mRNA and difficult to circumvent. However, it was particularly important to validate the amplification reaction, both because of the potential for gross distortions of the proportions of isoforms and because we wanted to use a second amplification reaction, with the products of the first amplification reaction as substrates, in order

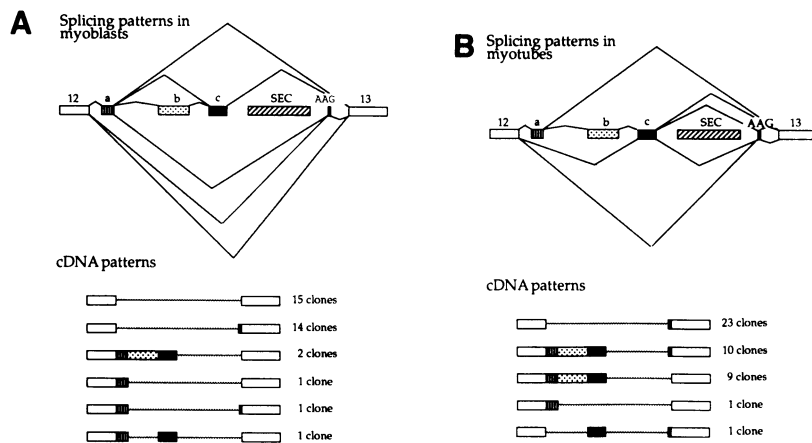
to visualise the ratios of the products directly rather than via cloning and sequence analysis. If valid, this approach would allow more significant values to be determined for rare events. In order to test whether the proportions of substrate cDNA were maintained during amplification, independent of the starting concentrations of the substrate, two cDNA sequences were amplified for 15 cycles in a series of reactions where their relative



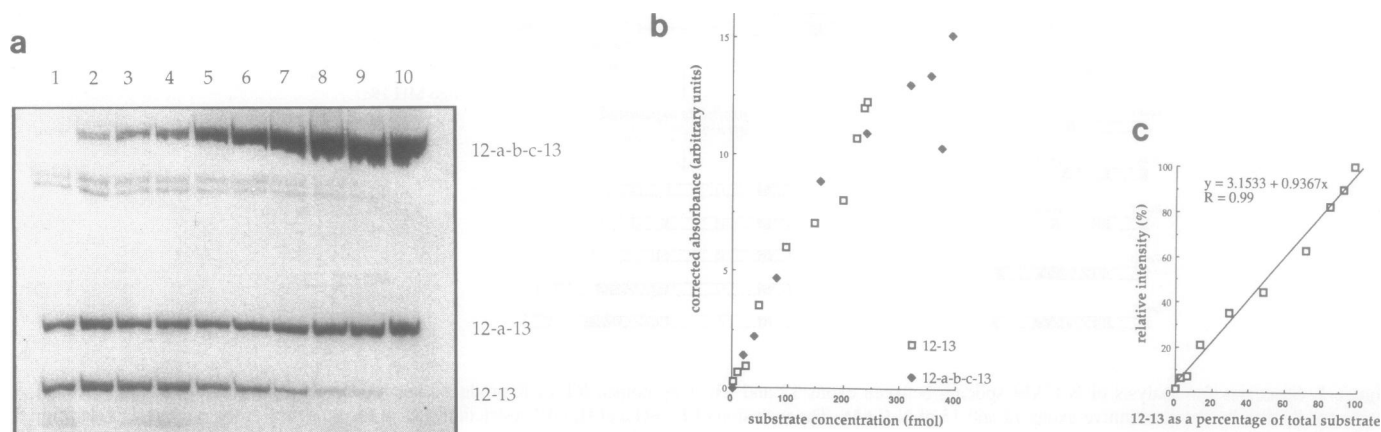
**Figure 1.** Strategies for analysis of N-CAM splicing between exons 12 and 13. A combined RT-PCR produced heterogeneous fragments by the amplification of sequences flanked by the constitutive exons 12 and 13 of N-CAM. The inclusion of EcoRI and HindIII restriction sites in the sequence of the oligonucleotide primers allowed for the subsequent ligation of products into a vector for sequence analysis. An aliquot of the heterogeneous PCR product was also used as a substrate for two second round PCRs using primers directed against part of exon 13 in conjunction with the primer against exon 12, or with a primer directed against exon MSD1c. The incorporation of  $\alpha^{32}\text{P}$  dATP into these reactions enabled the products to be visualised by their separation on native polyacrylamide gels. Oligonucleotide sequences are represented as clear for regions of exon complementarity; solid boxes for mismatched restriction sites; and shaded boxes for flanking GC-rich 'clamps'. Exon sequences are shown as clear boxes for the constitutive exons 12 and 13, vertical hatching for MSD1a, pale hatching for MSD1b, dark hatching for MSD1c, and solid boxes for the triplet exon AAG.

<b>EXON 12</b>																	
mouse	..G	TTC	AAG	ACA	CAG	CCA	GTC	C									
human	..G	TTC	AAG	ACG	CAG	CCA	GTC	C									
chicken	...	...	...	ACA	CAG	CCA	GTT	C									
<b>EXON MSD1a</b>																	
mouse	AT	AGC	CCT	CCT	CCA	C											
human	AT	AGC	CCT	CCT	CCA	C											
chicken	GT	ATC	CCT	CAC	TCA	C											
<b>EXON MSD1b</b>																	
mouse	AG	GCT	CCT	GCT	AAC	TCT	TCC	ACC	CAT	GTT	CCG	TTG	TCT	CCT	CGA	GCT	A
human	CG	GCA	TCT	GCT	AGC	TCG	TCT	ACC	CCT	GTT	CCA	TTG	TCT	CCA	CCA	GAT	A
chicken						CA	AGT	GCT	GCT	GCC	CCC	ACC	TGT	TTA	GCA	GAG	A
<b>EXON MSD1c</b>																	
mouse	CA	ACC	TGG	CCT	CTT	CCT	GTC	CTT	CCA	ACC	ACA	GAC	CTG	TCT	A		
human	CA	ACT	TGG	CCT	CTT	CCT	GCC	CTT	GCA	ACC	ACA	GAA	CCA	GCT	A		
chicken	CA	ACT	CAG	CCT	CTC	CCT	GCA	TCC	GAA	AGC	ACA	GAG	CCA	CCA	A		
<b>EXON AAG</b>																	
mouse	AA	G															
human	AA	G															
chicken	AA	G															
<b>EXON 13</b>																	
mouse	GG	GAA	CCC	AGT	GCA	CCC	AAG	C..									
human	GG	GAA	CCC	AGT	GCA	CCT	AAG	C..									
chicken	GG	GAA	CCC	AGT	...	...	...	...									

**Figure 2.** cDNA sequence of the mouse exons homologous to human MSD1a, MSD1b, MSD1c and the triplet AAG, flanked by regions of the constitutive exons 12 and 13 of N-CAM. The sequence has been aligned with the human<sup>11</sup> and chicken<sup>17</sup> sequences already elucidated.



**Figure 3.** Patterns of exon usage in C-2 myoblasts and C-2 myotubes, assigned by the numbers of clones isolated and sequenced for each isoform. Exons are patterned as in Figure 1. The putative splicing patterns required to produce these combinations are also shown.



**Figure 4.** Demonstration that PCR maintains the original ratio between N-CAM cDNA isoforms. a). cDNAs were co-amplified in a series of PCRs to assess the linearity of PCRs. The initial concentrations of the cDNAs incorporated as standards are shown in the table below. The absorbance reading obtained by laser densitometry was corrected for the number of A + T residues and for loading differences as assessed by the absorbance reading of the internal standard, and is also shown below.

	Track number									
	1	2	3	4	5	6	7	8	9	10
12-13 (fmol)	242	233	220	196	147	98.0	49.0	24.5	12.25	2.45
corrected absorbance	12.3	12.0	10.7	8.00	7.05	5.95	3.5	0.96	0.70	0.23
12-a-b-c-13 (fmol)	3.95	19.8	39.5	79.0	158	237	316	355	375	391
corrected absorbance	0	1.36	2.25	4.70	8.77	11.0	13.0	13.4	10.3	15.2

b). The initial concentration of cDNA standard was plotted against the absorbance reading, after correction for A+T residues and internal standard. Open boxes represent results obtained from amplification of cDNA encoding the isoform 12-13, circles represent results obtained from amplification of cDNA encoding the isoform 12-a-b-c-13. The concentration of each cDNA was determined by the absorbance of an original stock solution at 260nm. These were then diluted to produce the initial standard solutions. c). The percentage of one substrate as a proportion of the total substrate plotted against the relative intensity of the band produced on a gel (after correction for A+T content). The straight line has the equation  $y = 3.1533 + 0.9367x$ .

starting concentrations were varied systematically. A third substrate (cDNA 12-1a-13) was included at a constant concentration as a control. The results are shown in Figure 4a. Incorporation of <sup>32</sup>P during the reaction allowed quantification by laser densitometry of film.

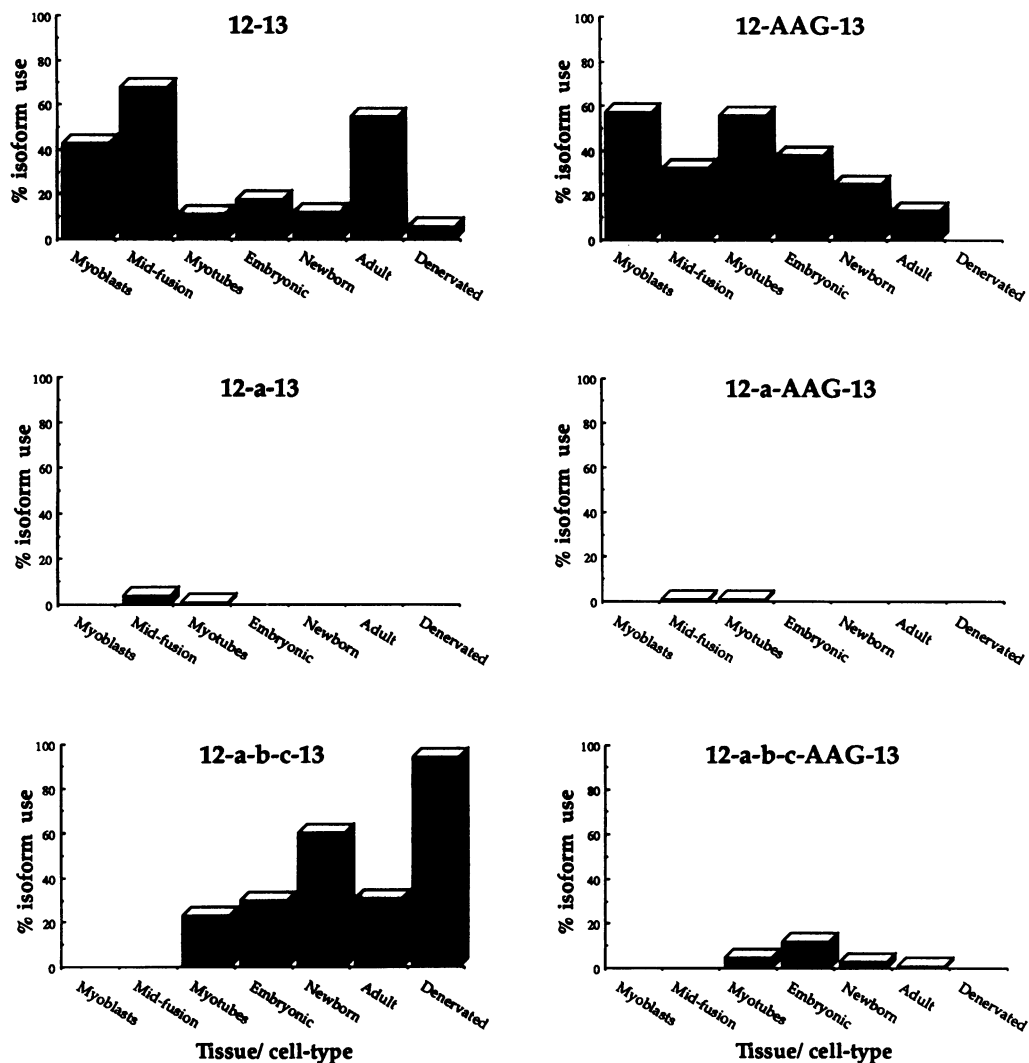
Figure 4b shows the graphs of product yield (absorbance)

versus substrate concentration obtained after analysis of Figure 4a. For both substrates the relationship was linear and the substrates were amplified with equivalent efficiencies. This point is emphasized in Figure 4c, which shows that the starting ratio of cDNA substrates was maintained during amplification. From other work (data not shown), we have found that the



**Table 1.**  
Proportionate abundance of N-CAM isoforms in muscle cells and tissues.

splicing pattern	C-2 myoblasts	C-2 mid-fusion	C-2 myotubes	Embryonic muscle	Newborn muscle	Adult muscle	Denervated muscle
% 12-13	43	65	11	18	12	55	6
% 12-AAG-13	57	31	57	39	25	13	<0.5
% 12-a-13	<0.5	3	2	<0.5	<0.5	<0.5	<0.5
% 12-a-AAG-13	<0.5	1	3	<0.5	<0.5	<0.5	<0.5
% 12-a-b-c-13	<0.5	<0.5	22	31	60	31	94
% 12-a-b-c-AAG-13	<0.5	<0.5	5	12	3	1	<0.5



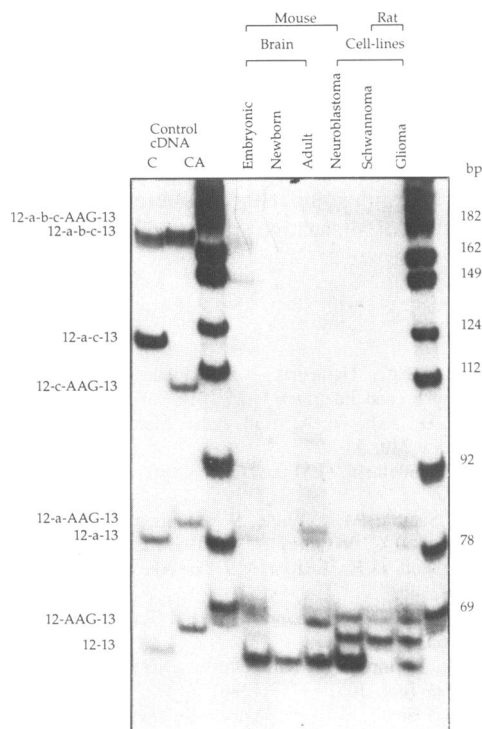
**Figure 6.** Diagrams showing the proportions of each N-CAM isoform in mRNA from muscle cells and tissues. The proportion of each isoform is shown as a percentage of total N-CAM mRNA isolated from C-2 myoblasts, C-2 mid-fusion (2 day post induction of myogenic differentiation), C-2 myotubes (4 day post induction) from cell culture, embryonic, newborn, adult and post surgically denervated skeletal muscle, and analysed by a combined RT-PCR.

shift towards the exon combination 12-a-b-c-13 with reduced levels of inclusion of the triplet AAG exon.

#### Exon usage in neural cells

PCR products produced as described above but from neural cell mRNA were separated on native polyacrylamide gels alongside the cDNA markers. The PCR products derived from the different stages of development showed no significant switch in exon usage although some differences in the prevalence of different isoforms

in normal and transformed cells can be seen (Figure 7). 1) Little or no incorporation of the MSD exons a,b, and c was detected in RNA from the brains of mice at different stages of development. The very low levels of apparent incorporation of these exons in mRNA isolated from embryonic brain may have been derived from muscle due to contamination during dissection. 2) Low levels of the previously identified isoforms 12-a-13 and 12-a-AAG-13<sup>9</sup> were present in embryonic and adult tissues. 3) The isoform 12-AAG-13 was not found in RNA derived from



**Figure 7.** Analysis of N-CAM exons incorporated in mRNA from brain tissues and neural cell lines. PCR products were amplified and separated as previously described. RT-PCR products of RNA from embryonic, newborn and adult mouse brain is compared with RNA isolated from mouse N2A neuroblastoma, rat C6 glioma, and rat SWA schwannoma. Two control cDNAs described previously were run as size markers. cDNA marker C includes the isoforms 12-13, 12-a-13, 12-a-c-13, and 12-a-b-c-13. cDNA marker CA includes 12-AAG-13, 12-a-AAG-13, 12-c-AAG-13 and 12-a-b-c-AAG-13.

normal mouse tissues, but it was detected in RNA isolated from a mouse neuroblastoma cell line. 4) The same isoform was the predominant form in rat Glioma and Schwannoma cell lines.

## DISCUSSION

The exons comprising the MSD region of humans and chick N-CAM have been shown to undergo tissue-specific regulation of alternative splicing, and recently the developmental variation between late embryonic and adult N-CAM isoform use has been elucidated for rat heart muscle. It was not clear whether the exons comprising the MSD were regulated separately or as a unit; any evidence of control of individual exons would imply that each exon makes a separate contribution to the function of N-CAM. We have characterised the patterns of expression of the mouse MSD exons during the development of skeletal muscle and brain, finding clear evidence for major shifts in expression of individual MSD exons. Cloning and sequencing of the mouse exons indicated a high level of DNA sequence similarity between man and mouse: 100%, 77% and 83% for MSD1a, MSD1b and MSD1c respectively.

### Correlation of results between cloning and direct analysis

Characterization of the patterns of splicing was achieved by a coupled RT-PCR, the products of which were either cloned and sequenced or analysed directly by re-amplification and subsequent separation on native polyacrylamide gels. The results of the two methods were compared by a  $\chi^2$  analysis, using the values

obtained for the re-amplification method as expected values for the cloning method. The comparison showed no significant difference in the results obtained by the two methods for RNA isolated from C-2 myoblasts, but did show a difference for those obtained from RNA isolated from C-2 myotubes. The difference in the results obtained for the two methods from RNA isolated from C-2 myotubes appears to be due to a relative lack of the isoform 12-a-b-c-AAG-13 and an excess of the isoform 12-13 in the direct re-amplification method. The lack of comparability between the results obtained by the cloning of RT-PCR products, and the direct re-amplification and subsequent separation on polyacrylamide gels for RNA isolated from C-2 myotubes may be due to several factors. A clear possibility is that one or both of the methods may exhibit poor reproducibility. However the direct re-amplification procedure was clearly shown to yield linear and reproducible results, and good correlation with direct cloning strategies has been shown in other studies<sup>23</sup>. Alternatively, as the N-CAM isoform distribution was shown to alter during development, and, although for both methods RNA was isolated from cells 4 days after myogenic induction, other factors such as cell density and quality of sera may affect the exact time course of differentiation.

### Discrete and coordinated regulation of the MSD exons

In mononucleate myoblast C-2 cells the predominant isoforms of N-CAM excluded the MSD exons, although there were substantial levels of the triplet exon AAG. During fusion of the cells the incorporation of AAG declined transiently and MSD1a appeared at a low level; later, in myotubes, the MSD1b and MSD1c were incorporated, almost exclusively together and in conjunction with MSD1a. During muscle development the MSD exons were present only as a coordinate unit, but in neural cells exon MSD1a was seen at low levels in brain tissue, with no mRNA incorporating all the MSD exons. Although other alternative splicing patterns were found within the MSD, they were uncommon (12-c-AAG-13, 12-a-c-13). We conclude that the incorporation of exon MSD1a is regulated independently of exons MSD1b and MSD1c, but that inclusion of the latter two requires MSD1a. Incorporation of the triplet AAG exon does not appear to require or be coordinated with the MSD exons.

The decline in MSD-containing RNA in adult skeletal muscle, and the appearance of low levels of MSD1a alone, might arise from an increased contribution of Schwann cells and muscle satellite cells to the mRNA being analysed. This would be expected from the nerve-dependent reduction in N-CAM gene expression as skeletal muscle cells mature, which results in the loss of N-CAM from all regions except the neuromuscular junction and the muscle satellite stem cells<sup>24</sup>.

The most likely process by which the levels of the N-CAM mRNA isoforms are determined is alternative splicing of the pre-mRNA. However, different stabilities of the isoforms might contribute also. The levels appear to be controlled in a tissue- and stage-specific manner. With the muscle cell lines, such control could reflect either specific regulation of the determining process or a general change in the efficiency of the process. In the case of the tissue samples, changes in the levels of isoforms during development could arise in the same ways or as a consequence of altered proportions of different cell types in the tissue, where each cell type produces a specific set of isoforms. The question of regulation is particularly intriguing for the AAG exon, a separate function for which is implied by its independent variation in incorporation.

### Effect of denervating muscle

In RNA isolated from skeletal muscle seven days after surgical denervation the predominant isoform was 12-a-b-c-13; little inclusion of the triplet exon AAG was detected, and the isoform 12-13 had almost disappeared. Surgical denervation of adult skeletal muscle is known to cause a marked increase in N-CAM expression<sup>25</sup>. The marked shift in exon usage seems to indicate that N-CAM expression is under post-transcriptional control as well. The transcriptional regulation involves a single regulatory unit, which exhibits a GC-rich content and a high frequency of the dinucleotide CpG but lacks the normal TATA or CCAAT motifs<sup>26</sup>. This regulation is poorly understood, although changes in tissue-specific trans-acting factors, modulation of methylation state of the region, or the effect of continued electrical stimulation of cells by nerve impulses may prove to be important. It is very likely that the apparent switch in splicing results in part from the nerve-dependent transcriptional activation of multinucleated myofibres, such that these cells increase their contribution to the N-CAM mRNA isoforms isolated from the tissue. Likewise, the contributions of Schwann and satellite cells might be expected to alter in response to nerve degeneration and muscle atrophy. Further experiments would be required to show whether splicing was regulated directly.

### Possible role for the MSD1a, MSD1b, MSD1c and AAG exons

All combinations of splicing patterns of the MSD exons would maintain the reading frame, but it is impossible to predict whether the minor isoforms produced by alternative splicing within this domain would encode functional N-CAM proteins. The physiological role of the protein domain encoded by the exons of the MSD has not been studied in detail, although incorporation of MSD1a may introduce a 'hinge' into the molecule, and the domain as a whole is known to undergo significant post-translational modification, notably O-linked glycosylation. Inclusion of the MSD as a unit has also been shown to induce enhanced myogenesis in myoblasts which have been transfected with full length cDNA encoding the MSD<sup>27</sup>. Inclusion of the triplet AAG exon results in an insertion of a single amino acid into the primary structure. In place of arginine, which is found in all the isoforms that do not incorporate AAG, glutamine and glycine are encoded by the isoforms 12-AAG-13 and 12-a-AA-G-13, and lysine and glycine by 12-a-b-c-AAG-13. If the inclusion of the AAG exon is demonstrated to be regulated, these subtle changes in the amino acid structure may represent an important point for modulation of cell-cell interaction which is mediated by the regulation of alternative splicing of an unusually short exon in N-CAM.

### CONCLUSIONS

Eight patterns of alternative splicing were noted in RNA derived from C-2 mouse muscle cells. These were cloned and sequenced and resulted in the elucidation of the DNA sequence of the mouse homologues to exons MSD1a, MSD1b, MSD1c and triplet AAG in the human. Analysis of PCR amplification of known amounts of N-CAM cDNA indicated that at low substrate concentrations the method was linear. This led to the quantification of isoform use which revealed a significant shift in isoform use toward inclusion of MSD1a, MSD1b and MSD1c, and a decrease in use of triplet AAG during skeletal muscle development. A marked switch in isoform use was also seen between pre- and post-denervated skeletal muscle.

These data indicate that the post-transcriptional processing, probably due to alternative splicing, of four short exons located between the constitutively spliced exons 12 and 13 of N-CAM are under both tissue- and developmental-specific regulation; one of these exons, the triplet AAG, is only 3 base pairs long and is regulated independently of the other exons. These findings lead to interesting questions about the possible mechanisms for the regulated control of alternative splicing of such unusually short exons.

### REFERENCES

- Cunningham, B.A., Hemperly, J.J., Murray, B.A., Prediger, E.A., Brackenbury, R. and Edelman, G.M. (1987) *Science* **236**, 799–806.
- Rutishauser, U., Acheson, A., Hall, A.K., Mann, D.M. and Sunshine, J. (1988) *Science* **240**, 53–57.
- Owens, G.C., Edelman, G.M. and Cunningham, B.A. (1987) *P.N.A.S.* **84**, 294–298.
- Bartels, D., Santoni, M.J., Willie, W., Ruppert, C., Chaix, J.C., Hirsch, M.R., Fontecil, J.C. and Goridis, C. (1987) *E.M.B.O. J.* **6**, 907–914.
- Small, J.S., Shull, G.E., Santoni, M.J. and Akeson, R. (1987) *J. Cell Biol.* **105**, 2335–2345.
- Barbas, J.A., Chaix, J.C., Steinmet, M. and Goridis, C. (1988) *E.M.B.O. J.* **7**, 625–632.
- Moore, S.E., Thompson, J., Kirkness, V., Dickson, J.G. and Walsh, F.S. (1987) *J. Cell Biol.* **105**, 1377–1386.
- Reyes, A.A., Small, S.J. and Akeson, R. (1991) *Mol. Cell. Biol.* **11**, 1654–1661.
- Santoni, M.J., Barthels, D., Vopper, G., Boned, A., Goridis, C. and Willie, W. (1989) *E.M.B.O. J.* **8**, 385–392.
- Small, S.J. and Akeson, R. (1990) *J. Cell Biol.* **111**, 2089–2096.
- Dickson, J.D., Gower, H.J., Barton, C.H., Prentice, H.M., Elsom, V.L., Moore, S.E., Cox, R.D., Quinn, C., Putt, W. and Walsh, F.S. (1987) *Cell* **50**, 1119–1130.
- Gower, H.J., Barton, C.H., Elsom, V.L., Thompson, J., Moore, S.E., Dickson, G. and Walsh, F.S. (1988) *Cell* **55**, 955–964.
- Walsh, F.S. and Dickson, G. (1989) *Bioessays* **11**, 83–88.
- Kadmon, G., Kowitz, A., Altevogt, P. and Schachner, M. (1990) *J. Cell Biol.* **110**, 193–208.
- Walsh, F.S., Prakh, R.B., Moore, S.E., Dickson, G., Barton, C.H., Gower, H.J., Dwek, R.A. and Rademacher, T.W. (1989) *Development* **105**, 803–811.
- Thompson, J., Dickson, G., Moore, S.E., Gower, H.J., Putt, W., Kenimer, J.G., Barton, C.H. and Walsh, F.S. (1989) *Genes Dev.* **3**, 348–357.
- Prediger, E., Hoffman, S., Edelman, G. and Cunningham, B. (1988) *P.N.A.S.* **85**, 9616–9620.
- Dickson, G., Pizzey, J.A., Elsom, V.E., Love, D., Davies, K.E. and Walsh, F.S. (1988) *FEBS Lett.* **242**, 47–52.
- Auffray, C. and Rougeon, F. (1980) *Eur. J. Biochem.* **107**, 303–314.
- Eperon, I.C., (1986) *Nucl. Acids Res.*, **14**, 2830.
- Eperon, I.C., (1986) *Anal. Biochem.* **156**, 406–412.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *P.N.A.S.* **74**, 5463–5467.
- Garde, J., Bell, S.C. and Eperon, I.C. (1991) *P.N.A.S.* **88**, 2456–2460.
- Covault, J. and Sanes, J.R. (1986) *J. Cell Biol.* **102**, 716–730.
- Moore, S.E. and Walsh, F.S. (1986) *Neuroscience* **18**, 499–505.
- Barton, C.H., Mann, D.A. and Walsh, F.S. (1990) *Biochem. J.* **268**, 161–168.
- Dickson J.G., Peck D., Moore S.E., Barton C.H. and Walsh F.S. (1990) *Nature* **344**, 348–351.