ALTERNATIVE SPLICING OF THE NEURAL CELL ADHESION MOLECULE AND THE DIFFERENTIAL DIAGNOSIS OF NEUROBLASTOMA

Elizabeth Giorsal Phimister

A Thesis Submitted for the Degree of Doctor of Philosophy

The Imperial Cancer Research Fund
Paediatric and Neuro-Oncology Group
Frenchay Hospital
Bristol

August, 1993
ALTERNATIVE SPLICING OF THE NEURAL CELL ADHESION MOLECULE AND THE DIFFERENTIAL DIAGNOSIS OF NEUROBLASTOMA

ABSTRACT

While the survival rate of children diagnosed with cancer has more than doubled in the last 25 years, neuroblastoma, accounting for 7 - 9% of all childhood malignancies, remains a disease with a poor prognosis. In some cases, confusion between neuroblastoma and other "small round cell" tumours, such as rhabdomyosarcoma, contributes to erroneous diagnosis and inappropriate treatment. This thesis addresses the possibility of distinguishing such tumours by differential expression of the neural cell adhesion molecule (NCAM).

Fluorescence-activated flow cytometry revealed expression of NCAM by 6/7 neuroblastoma cell lines tested. Western blotting showed that underlying a uniform sialylation, a heterogeneity of isoform expression exists although neuroblastoma and rhabdomyosarcoma cell lines could not be distinguished by isoform complement.

Analyses of two alternative splicing regions in the extracellular domain suggest that tissue-specific variations of NCAM may exist. Splicing patterns at both of these regions were examined in neuroblastoma and rhabdomyosarcoma cell lines by amplifying cDNA complementary to cellular RNA, using the polymerase chain reaction. Products were analysed by Southern blotting and sequence analysis.

The Variable Alternatively-spliced Exon (VASE), previously thought to be restricted to tissues of the nervous system, was found in all rhabdomyosarcoma cell lines tested. However, its presence in all neuroblastoma cell lines examined, renders it a poor candidate by which to
differentiate the two cell types. The sequence of human VASE was determined.

Evidence suggests that three of the four alternatively-spliced exons constituting the "muscle-specific" domain (MSD1) are spliced into RNA synthesised by some neuroblastoma cell lines. The expression of components of the MSD1 domain may provide a basis on which neuroblastoma and rhabdomyosarcoma may be distinguished; MSD1c was found in NCAM of rhabdomyosarcoma cell lines but not that of neuroblastoma cell lines.

These findings indicate that NCAM expression is more complex than first assumed, and that a detailed study of the expression of NCAM on the "small round cell" tumours can lead to a highly selective approach in designing reagents for distinguishing the two types of tumours analysed here.
ACKNOWLEDGEMENTS

I would like to thank John Kemshead for his unfailing enthusiasm, advice and support throughout the course of the project. Thanks are also due to Peter Beverley and Kalpana Patel - their advice and support were greatly appreciated. I thank Ros Rossell and Anne Hancock, who assisted with the molecular characterisation of rhabdomyosarcoma cell lines, and Ray Hicks, who assisted with the fluorescence-activated flow cytometry. While preparing the manuscript, Sharon Standen stood by at all hours to advise on word-processing and Geoff Freke applied his photographic expertise. I thank them both. The Imperial Cancer Research Fund provided financial support over the course of the project, for which I am grateful. Finally, I thank my family and friends, for providing a rich substratum on which to extend my pre-doctoral neurites.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title Page</td>
<td>1</td>
</tr>
<tr>
<td>Abstract</td>
<td>2</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>4</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>5</td>
</tr>
<tr>
<td>List of Figures</td>
<td>10</td>
</tr>
<tr>
<td>List of Tables</td>
<td>16</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>18</td>
</tr>
<tr>
<td><strong>Chapter 1: THE DIAGNOSIS AND PROGNOSIS OF NEUROBLASTOMA AND THE ROLE OF ADHESION MOLECULES IN TUMOUR BIOLOGY</strong></td>
<td>20</td>
</tr>
<tr>
<td>1.1 Introduction</td>
<td>20</td>
</tr>
<tr>
<td>1.1a Tumour Markers</td>
<td>20a</td>
</tr>
<tr>
<td>1.2 The Differential Diagnosis and Prognosis of Neuroblastoma</td>
<td>21</td>
</tr>
<tr>
<td>(a) Staging</td>
<td>22</td>
</tr>
<tr>
<td>(b) Age at diagnosis</td>
<td>23</td>
</tr>
<tr>
<td>(c) Site of tumour</td>
<td>23</td>
</tr>
<tr>
<td>(d) Pathology</td>
<td>25</td>
</tr>
<tr>
<td>(e) Serum concentrations of neuron-specific enolase and ferritin</td>
<td>26</td>
</tr>
<tr>
<td>(f) Catecholamine expression and secretion</td>
<td>27</td>
</tr>
<tr>
<td>(g) Genetic factors</td>
<td>28</td>
</tr>
<tr>
<td>(h) Immunohistochemistry and immunocytology</td>
<td>28</td>
</tr>
<tr>
<td>1.3 Adhesion Molecules and Tumour Biology</td>
<td>32</td>
</tr>
<tr>
<td>(a) The cadherins</td>
<td>33</td>
</tr>
<tr>
<td>(b) The integrins</td>
<td>36</td>
</tr>
<tr>
<td>(c) Fibronectin</td>
<td>37</td>
</tr>
<tr>
<td>(d) The homing receptors</td>
<td>41</td>
</tr>
<tr>
<td>(e) The immunoglobulin superfamily</td>
<td>42</td>
</tr>
<tr>
<td>(f) The L1 neural cell adhesion molecule</td>
<td>44</td>
</tr>
<tr>
<td>1.4 The Neural Cell Adhesion Molecule (NCAM)</td>
<td>45</td>
</tr>
<tr>
<td>(a) NCAM expression</td>
<td>45</td>
</tr>
<tr>
<td>(b) Genomic organisation and transcriptional regulation of NCAM</td>
<td>47</td>
</tr>
<tr>
<td>(c) NCAM isoforms: composition and characteristics</td>
<td>48</td>
</tr>
<tr>
<td>(d) NCAM and the extracellular matrix</td>
<td>56</td>
</tr>
<tr>
<td>(e) Post-translational modifications of NCAM</td>
<td>58</td>
</tr>
<tr>
<td>(f) Alternative splicing of NCAM in the extracellular domain</td>
<td>61</td>
</tr>
</tbody>
</table>
Chapter 2: MATERIALS AND GENERAL METHODS

2.1 Materials

2.2 Cell Culture

2.3 Fluorescence-activated Flow Cytometric Analysis

2.4 Protein Biochemistry
   (a) Neuraminidase enzyme digestion
   (b) Digestion with phosphatidylinositol phospholipase C
   (c) Cell and tissue lysate preparation
   (d) Determination of protein concentration
   (e) SDS-polyacrylamide gel electrophoresis
   (f) Electrophoretic transfer of proteins from polyacrylamide gel to nitrocellulose membrane
   (g) Analysis of transferred proteins

2.5 Generation of Polymerase Chain Reaction (PCR) Products
   (a) RNA extraction
   (b) Assay of RNA concentration and purity
   (c) Synthesis of cDNA
   (d) Symmetric PCR amplification

2.6 Analysis of PCR Products
   (a) Agarose gel electrophoresis and transfer to membrane
   (b) Acrylamide gel electrophoresis and transfer to membrane
   (c) End-labelling of oligonucleotides
   (d) Random labelling of oligonucleotides
   (e) Determination of radioactivity
   (f) Hybridisation of blot
   (g) Preparation of blot for rehybridisation

2.7 Direct Sequencing of PCR Products
   (a) Isolation of PCR products
   (b) Asymmetric amplification of PCR products
   (c) Purification and quantitative assay of asymmetric PCR products
   (d) Sequencing of PCR products
   (e) Gel analysis of sequencing reaction products

2.8 Cloning and Sequencing of PCR Products
   (a) Preparation of PCR products
   (b) Cloning of PCR products
   (c) Small scale preparation of plasmid DNA
   (d) Screening of clones for relevant inserts
   (e) Sequence analysis of cloned plasmid DNA preparations
Chapter 3: EXPRESSION OF NCAM ON TUMOUR CELL LINES AND TISSUE

3.1 Introduction

3.2 Methods
(a) NCAM expression on selected neuroblastoma cell lines as determined by fluorescence-activated flow cytometry
(b) Establishing experimental parameters: positive and negative controls and a titre of monoclonal antibody ERIC-1 for use in Western blotting
(c) Western blot analysis of extracts of neuroblastoma, rhabdomyosarcoma, neuroepithelioma and myeloid leukaemia cell lines
(d) Western blot analysis of neuroblastoma and rhabdomyosarcoma tissue extracts
(e) Glycosyl-phosphatidylinositol-anchored NCAM isoforms on the JR-1 rhabdomyosarcoma cell line as revealed by Western blotting
(f) NCAM expression and cell density

3.3 Results
(a) NCAM expression on selected neuroblastoma cell lines and a neuroepithelioma cell line as determined by fluorescence-activated flow cytometry
(b) Establishing experimental parameters: positive and negative controls and a titre of monoclonal antibody ERIC-1 for use in Western blotting
(c) Western blot analysis of extracts of neuroblastoma and rhabdomyosarcoma and myeloid leukaemia cell lines
(d) Western blot analysis of neuroblastoma and rhabdomyosarcoma tissue extracts
(e) Glycosyl-phosphatidylinositol-anchored NCAM isoforms of the JR-1 rhabdomyosarcoma cell line as revealed by Western blotting
(f) NCAM expression and cell density

3.4 Discussion
(a) NCAM expression on selected neuroblastoma cell lines and a neuroepithelioma cell line as determined by fluorescence-activated flow cytometry
(b) Western blot analysis of extracts of neuroblastoma rhabdomyosarcoma, neuroepithelioma, and myeloid leukaemia cell lines
(c) Western blot analysis of neuroblastoma and rhabdomyosarcoma tissue extracts
(d) Glycosyl-phosphatidylinositol-anchored NCAM isoforms on the JR-1 rhabdomyosarcoma cell line as revealed by Western blotting
(e) NCAM expression and cell density

3.5 Conclusion

7.
**Chapter 4: EXPRESSION OF THE VARIABLE ALTERNATIVELY SPLICED EXON (VASE) IN NEUROBLASTOMA AND Rhabdomyosarcoma Cell Lines**

### 4.1 Introduction

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>129</td>
</tr>
</tbody>
</table>

### 4.2 Southern Blotting of PCR-amplified cDNA Spanning the Exon 7/8 Splice Junction

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Methods</td>
<td>131</td>
</tr>
<tr>
<td>(b) Results</td>
<td>135</td>
</tr>
<tr>
<td>(i) Agarose gel electrophoresis</td>
<td>135</td>
</tr>
<tr>
<td>(ii) Polyacrylamide gel electrophoresis</td>
<td>137</td>
</tr>
<tr>
<td>(c) Discussion</td>
<td>141</td>
</tr>
</tbody>
</table>

### 4.3 Direct Sequence Analysis of PCR Products

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Methods</td>
<td>144</td>
</tr>
<tr>
<td>(i) Establishing a method</td>
<td>144</td>
</tr>
<tr>
<td>(ii) Sequence analysis of PCR products</td>
<td>145</td>
</tr>
<tr>
<td>(b) Results</td>
<td>146</td>
</tr>
<tr>
<td>(i) Establishing a method</td>
<td>146</td>
</tr>
<tr>
<td>(ii) Sequence analysis of PCR products</td>
<td>149</td>
</tr>
<tr>
<td>(c) Discussion</td>
<td>154</td>
</tr>
</tbody>
</table>

### 4.4 Cloning of PCR Products from the JR-1 Rhabdomyosarcoma Cell Line: Generation of Clone sets 1 and 2

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Methods</td>
<td>159</td>
</tr>
<tr>
<td>(b) Results</td>
<td>160</td>
</tr>
<tr>
<td>(c) Discussion</td>
<td>170</td>
</tr>
</tbody>
</table>

### 4.5 Sequence Analysis of Selected Samples from Clone sets 1 and 2

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Methods</td>
<td>173</td>
</tr>
<tr>
<td>(b) Results</td>
<td>173</td>
</tr>
<tr>
<td>(c) Discussion</td>
<td>183</td>
</tr>
</tbody>
</table>

### 4.6 Conclusion

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>190</td>
</tr>
</tbody>
</table>
Chapter 5: EXPRESSION OF THE MUSCLE SPECIFIC
DOMAIN MSD1 IN NEUROBLASTOMA AND
RHABDOMYOSARCOMA CELL LINES

5.1 Introduction

5.2 Generation of PCR Products which Span the Exon 12/13 Splice Junction and their Analysis by Southern Blotting

(a) Methods

(b) Results

(i) Southern blotting using the MSD1a oligonucleotide probe

(ii) Southern blotting using the MSD1b oligonucleotide probe

(iii) Southern blotting using the MSD1c oligonucleotide probe

(iv) Southern blotting using an oligonucleotide probe complementary to the MSD1 domain

(v) Southern blotting using cDNAs generated by random priming of λ9.5, as a collective probe

(c) Discussion

5.3 Cloning of PCR Products from the NB1 and Kelly Neuroblastoma Cell Lines: Generation of Clone sets a and b

(a) Methods

(b) Results

(c) Discussion

5.4 Sequence Analysis of Selected Samples from Clone sets a and b

(a) Methods

(b) Results

(c) Discussion

5.5 Conclusion

Chapter 6: SUMMARY AND SUGGESTIONS FOR FURTHER STUDIES

6.1 Summary

6.2 Suggestions for Further Studies

6.3 Conclusion

References

Addendum

Publications arising from the thesis
| Figure 1.1 | Expression of NCAM, N-cadherin and E-cadherin during morphogenesis | 35 |
| Figure 1.2 | Structure and alternative splicing patterns of fibronectin RNA | 38 |
| Figure 1.3 | Structure and alternative splicing patterns of (a) NCAM RNA and (b) corresponding protein | 49 |
| Figure 1.4 | Approximate topology of NCAM isoforms generated by alternative splicing | 50 |
| Figure 1.5 | Hypothetical structure of the NCAM N-linked glycosylation unit | 59a |
| Figure 3.1 | FACs profiles generated by anti-NCAM antibodies and human neuroblastoma cell lines | 99 |
| Figure 3.2 | Establishing a positive control - Western blot analysis of adult brain | 101 |
| Figure 3.3 | Western blot analysis of SK-N-BE (2C) and GOTO neuroblastoma cell lines | 102 |
| Figure 3.4 | Western blot analysis of the Kelly neuroblastoma and NB100 neuroepithelioma cell lines | 104 |
| Figure 3.5 | Western blot analysis of the IMR-32, SK-N-SH and NB1 neuroblastoma cell lines | 106 |
| Figure 3.6 | Western blot analysis of the JR-1, Rhab-1 and RD618 rhabdomyosarcoma cell lines | 108 |
| Figure 3.7 | Western blot analysis of the Kg1a myeloid leukaemia cell line | 110 |
Figure 3.8  Western blot analysis of neuroblastoma tissues  114

Figure 3.9  Western blot analysis of rhabdomyosarcoma tissues  116

Figure 3.10  Western blot analysis of the JR-1 rhabdomyosarcoma cell line and NCAM isoforms released from the cell surface by PIPLC  118

Figure 4.1  a) Rodent sequence of VASE with flanking sequences and corresponding amino acids  130

b) Schematic representation of a 140 kDa polypeptide encoded by RNA containing VASE  130

Figure 4.2.1  Exons 7 and 8, and the positions of primers and the non-VASE probe  133

Figure 4.2.2  Southern blot of PCR products spanning the exon 7/8 splice site from a selection of cell lines; probed with the VASE oligonucleotide and prepared by agarose gel electrophoresis  136

Figure 4.2.3  Southern blot of PCR products spanning the exon 7/8 splice site from a selection of cell lines; probed with the non-VASE oligonucleotide and prepared by agarose gel electrophoresis  138

Figure 4.2.4  Southern blot of PCR products spanning the exon 7/8 splice site from a selection of cell lines; probed with the VASE oligonucleotide and prepared by polyacrylamide gel electrophoresis  139

Figure 4.2.5  Southern blot of PCR products spanning the exon 7/8 splice site from a selection of cell lines; probed with the non-VASE oligonucleotide and prepared by polyacrylamide gel electrophoresis  140
Figure 4.3.1 Amplified extracts of the 194 bp PCR product; extraction procedures and primer ratios compared

Figure 4.3.2 Sequence of symmetrically- and asymmetrically-amplified PCR products compared

Figure 4.3.3 Southern blot of isolated PCR products

Figure 4.3.4 Sequence of 194 bp PCR product of the Kelly neuroblastoma cell line

Figure 4.3.5 Sequence of 224 bp PCR product of Kelly and NB1 neuroblastoma cell lines

Figure 4.3.6 Human and rodent VASE sequences compared

Figure 4.3.7 Sequence of 224 bp PCR product of the JR-1 rhabdomyosarcoma cell line

Figure 4.3.8 Sequence of 240 bp PCR product of the Kelly NB1 neuroblastoma cell lines and the JR-1 rhabdomyosarcoma cell line

Figure 4.3.9 Sequence of 660 and 900 bp PCR products of the Kelly neuroblastoma cell line

Figure 4.3.10 Potential structure of a VASE-related PCR-generated artifact

Figure 4.4.1 Structure of the pCR1000™ plasmid

Figure 4.4.2 Southern blot of selected samples from clone set 1; hybridised with the VASE oligonucleotide

Figure 4.4.3 Southern blot of selected samples from clone set 1; hybridised with the non-VASE oligonucleotide
Figure 4.4.4  Southern blot of selected samples from clone set 2; hybridised with the VASE oligonucleotide

Figure 4.4.5  Southern blot of selected samples from clone set 2; hybridised with the non-VASE oligonucleotide

Figure 4.5.1  Sequence film of a cloned PCR product

Figure 4.5.2  (a) sequence obtained from sample 45, clone set 1
(b) sequence obtained from sample 15, clone set 1
(c) sequence obtained from sample 30, clone set 1
(d) sequence obtained from sample 20, clone set 1
(e) sequence obtained from sample 46, clone set 1

Figure 4.5.3  (a) sequence obtained from sample 28, clone set 2
(b) sequence obtained from sample 38, clone set 2
(c) sequence obtained from sample 9, clone set 2
(d) sequence obtained from sample 32, clone set 2

Figure 4.5.4  Comparison of VASE sequences obtained from the JR-1 rhabdomyosarcoma cell line

Figure 5.1  a) Human sequence of MSD1 with flanking sequences and corresponding amino acids
(b) Schematic representation of a 125 kDa polypeptide encoded by RNA containing MSD1

Figure 5.2.1  Exons 12 and 13 and the positions of primers used in PCR amplification and sequencing
Figure 5.2.2  Southern blot of PCR products spanning the exon 12/13 splice site from a selection of cell lines; hybridised with the MSD1a oligonucleotide

Figure 5.2.3  Southern blot of PCR products spanning the exon 12/13 splice site from a selection of cell lines; hybridised with the MSD1b oligonucleotide

Figure 5.2.4  Southern blot of PCR products spanning the exon 12/13 splice site from a selection of cell lines; hybridised with the MSD1c oligonucleotide

Figure 5.2.5  Southern blot of PCR products spanning the exon 12/13 splice site from a selection of cell lines; hybridised with the MSD1 oligonucleotide

Figure 5.2.6  Southern blot of PCR products spanning the exon 12/13 splice site from a selection of cell lines; hybridised with the "λ9.5" oligonucleotides

Figure 5.3.1  Southern blot of selected samples from clone set a; hybridised with MSD1 oligonucleotide

Figure 5.3.2  Southern blot of selected samples from clone set b₁; hybridised with MSD1 oligonucleotide

Figure 5.3.3  Southern blot of selected samples from clone set b₂; hybridised with MSD1 oligonucleotide

Figure 5.3.4  Southern blot of selected samples from clone sets b₁ and b₂; hybridised with MSD1a, MSD1b, MSD1c and "λ9.5" oligonucleotides
Figure 5.4.1

(a) sequence obtained from sample 8, clone set a
(b) sequence obtained from sample 19, clone set a
(c) sequence obtained from sample 27, clone set a
(d) sequence obtained from sample 34, clone set a
(e) sequence obtained from sample 35, clone set a

Figure 5.4.2

(a) sequence obtained from sample 8, clone set b₁
(b) sequence obtained from sample 13, clone set b₁
(c) sequence obtained from sample 24, clone set b₂
(d) sequence obtained from sample 27, clone set b₂
(e) sequence obtained from sample 28, clone set b₂

Figure 6.1

Mouse and Human MSD1 sequences compared: cDNA and corresponding amino acids
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1.1</td>
<td>Incidence of neuroblastoma and survival rate with respect to stage of disease</td>
<td>24</td>
</tr>
<tr>
<td>Table 1.2</td>
<td>Antibodies which have been proposed as agents by which to differentially diagnose neuroblastoma and rhabdomyosarcoma</td>
<td>30</td>
</tr>
<tr>
<td>Table 1.3</td>
<td>Capacities of the domains and tandem domains comprising the neural cell adhesion molecule</td>
<td>52</td>
</tr>
<tr>
<td>Table 3.1</td>
<td>FACs analysis of neuroblastoma, neuroepithelioma, and leukaemia cell lines</td>
<td>98</td>
</tr>
<tr>
<td>Table 3.2</td>
<td>NCAM isoforms identified in human neuroblastoma cell lines</td>
<td>112</td>
</tr>
<tr>
<td>Table 3.3</td>
<td>NCAM mRNA identified in human neuroblastoma cell lines</td>
<td>122</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>VASE expression in human cell lines and distinct rat brain regions</td>
<td>132</td>
</tr>
<tr>
<td>Table 4.2</td>
<td>PCR products putatively spanning the exon 7/8 splice junction: their sizes and blotting profiles</td>
<td>142</td>
</tr>
<tr>
<td>Table 4.4.1</td>
<td>Colonies arising from the cloning of random PCR products; clone set 1</td>
<td>162</td>
</tr>
<tr>
<td>Table 4.4.2</td>
<td>Clone set 1; products from the JR-1 rhabdomyosarcoma cell line, cloned and digested with (Not-1) restriction enzyme</td>
<td>166</td>
</tr>
<tr>
<td>Table 4.4.3</td>
<td>Colonies arising from the cloning of selected PCR products; clone set 2</td>
<td>167</td>
</tr>
<tr>
<td>Table 4.4.4</td>
<td>Clone set 2; products from the JR-1 rhabdomyosarcoma cell line, cloned and digested with (Not-1) restriction enzyme</td>
<td>169</td>
</tr>
<tr>
<td>Table 4.5</td>
<td>Cloned PCR products derived from the JR-1 rhabdomyosarcoma cell line: their size, blotting profiles, and sequences</td>
<td>190</td>
</tr>
<tr>
<td>Table 5.1</td>
<td>Alternative splicing patterns of the MSD1 domain</td>
<td>196</td>
</tr>
<tr>
<td>Table 5.2</td>
<td>Theoretical sizes of potential PCR products generated by amplification across the exon 12/13 splice junction</td>
<td>211</td>
</tr>
<tr>
<td>Table 5.3.1</td>
<td>Colonies arising from the cloning of random PCR products; clone set a</td>
<td>218</td>
</tr>
<tr>
<td>Table 5.3.2</td>
<td>Colonies arising from the cloning of selected PCR products; clone set b₁</td>
<td>221</td>
</tr>
<tr>
<td>Table 5.3.3</td>
<td>Colonies arising from the cloning of selected PCR products; clone set b₂</td>
<td>222</td>
</tr>
</tbody>
</table>
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALL</td>
<td>acute lymphoblastic leukaemia</td>
</tr>
<tr>
<td>approx.</td>
<td>approximately</td>
</tr>
<tr>
<td>°C</td>
<td>degrees centigrade</td>
</tr>
<tr>
<td>CAM</td>
<td>cell adhesion molecule</td>
</tr>
<tr>
<td>CNS</td>
<td>central nervous system</td>
</tr>
<tr>
<td>dATP</td>
<td>2' deoxyadenosine 5' triphosphate</td>
</tr>
<tr>
<td>dCTP</td>
<td>2' deoxycytidine 5' triphosphate</td>
</tr>
<tr>
<td>dGTP</td>
<td>2' deoxyguanosine 5' triphosphate</td>
</tr>
<tr>
<td>dTTP</td>
<td>2' deoxythymidine 5' triphosphate</td>
</tr>
<tr>
<td>ddATP</td>
<td>2' 3'- dideoxyadenosine 5' triphosphate</td>
</tr>
<tr>
<td>ddCTP</td>
<td>2' 3'-dideoxycytidine 5' triphosphate</td>
</tr>
<tr>
<td>ddGTP</td>
<td>2' 3'-dideoxyguanosine 5' triphosphate</td>
</tr>
<tr>
<td>ddTTP</td>
<td>2' 3'-dideoxythymidine 5' triphosphate</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>ECM</td>
<td>extracellular matrix</td>
</tr>
<tr>
<td>FCS</td>
<td>foetal calf serum</td>
</tr>
<tr>
<td>g</td>
<td>gravity</td>
</tr>
<tr>
<td>hr</td>
<td>hr(s)</td>
</tr>
<tr>
<td>Ig</td>
<td>immunoglobulin</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger ribonucleic acid</td>
</tr>
<tr>
<td>m. wt.</td>
<td>molecular weight</td>
</tr>
<tr>
<td>NBL</td>
<td>neuroblastoma</td>
</tr>
<tr>
<td>O/N</td>
<td>overnight</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PNS</td>
<td>peripheral nervous system</td>
</tr>
</tbody>
</table>
CHAPTER 1

THE DIAGNOSIS AND PROGNOSIS OF NEUROBLASTOMA
AND THE ROLE OF ADHESION MOLECULES IN TUMOUR
BIOLOGY

1.1 Introduction

Cancer research over the last three decades has contributed towards a dramatic improvement in the diagnosis and management of childhood malignancies. The cure rate for such malignancies has increased from <20% to over 60% (Cassady, 1991) and has occurred as a consequence of techniques leading to improved diagnosis, staging, and treatment.

The prognosis for children over one year of age with neuroblastoma remains poor, and the differential diagnosis of the "small round cell" tumours of childhood, including neuroblastoma, continues to be problematic. In testimony to this is the huge effort expended by the scientific community in devising methods to aid in the identification of "small round cell" tumours. The advanced stage at which the neuroblastoma patient typically presents provides an additional challenge to clinical management.

Following leukaemia, non-Hodgkins lymphoma, and tumours of the central nervous system, neuroblastoma is the fourth most common tumour of childhood in the United Kingdom. It is the second most common solid tumour of infancy and childhood. It occurs in approximately 1 in 10,500 children between the ages of 0 and 14 (Innocent et al., 1992) and represents 7 - 9% of all childhood malignancies.

Neuroblastoma, rhabdomyosarcoma, Ewing's sarcoma, Wilms' tumour and lymphoma are embryonic tumours - so-called because they
may present with an undifferentiated embryonic phenotype, making diagnosis difficult. It is important that a speedy and accurate diagnosis be made; primitive tumours are often aggressive, and the optimum management of one tumour, or a certain stage of a given tumour, may be quite different from that of another.

The development of rapid and accurate techniques to aid diagnosis is highly desirable. Some procedures - e.g. electron microscopy - may provide useful diagnostic information, but are not easily implemented in the standard laboratory. In contrast, other procedures such as immunohistochemistry and immunocytology are more easily implemented and provide information which is less susceptible to vagaries of interpretation. Illustrating the need for additional diagnostic tests is a small group of tumours which defy differential diagnosis.

1.1a Tumour Markers

The pursuit of the ideal tumour marker has resulted in many tests for use in the diagnosis and management of cancer - the holy grail being a molecular index with absolute tumour specificity (thereby ensuring accurate differential diagnosis), 100% sensitivity (i.e. detectable on all tumours of a specific type) and easy accessibility. While no marker has met these criteria, several good markers have been established for specific tumours.

Most tumour markers are glycoproteins, although monoclonal antibodies have identified cancer-associated carbohydrates (For examples, see Magnani et al., 1983; Persson et al., 1988; Osako et al., 1993; for a review, see Sell, 1990). Tumour markers are either products of the normal cell from which the malignancy derives (e.g. prostate-specific antigen in prostatic cancer) or onco-developmental proteins (e.g. NCAM in Wilms' tumour), and fail when the tumour ceases to produce the "normal" product, or to express the onco-developmental antigen. Proposed functions of conventional tumor markers include roles in intercellular recognition, intracellular processing of glycoproteins, cell activation and metastatic spread (Sell, 1990).

Markers have several uses in patient care: they can be used for screening, diagnosis, monitoring treatment, detecting relapse, establishing prognosis and as "targets" in experimental therapies. The value of a marker for a particular tumour depends on three things: the frequency with which it is expressed/secreted by the tumour in detectable amounts (sensitivity), the extent to which it is specific to tumour tissue, and the effectiveness of
therapy. By definition, sensitivity and specificity are not mutually compatible. Highly sensitive markers tend to give rise to false positive results, and highly specific markers - to false negatives. Sometimes markers are combined, to improve the sensitivity of tumour detection (for example see Pasanen et al., 1992) although a decline in specificity often results. This is especially the case in distinguishing metastatic from benign tumours.

Other factors limit the utility of markers in specific contexts. For example, low prevalence of cancer incidence in asymptomatic people, taken together with the rate of false negatives and false positives, has rendered all general screening programmes tested to date unfeasible. However, defining a "high risk" selected population increases the predictive value of a given marker. For instance, the National Institutes of Health has recommended regular screening of patients who are positive for hepatitis B surface antigen (a marker in itself) and have active hepatitis or cirrhosis (Di Bisceglie et al., 1988). A combinatorial approach of imaging by ultrasound and measuring α-fetoprotein (AFP) levels still yields high rates of false positive and negative results - 93% and 30% respectively in one study (Cottone et al., 1988) but two factors make such testing worthwhile: small (and resectable) hepatocellular carcinomas secrete AFP in detectable amounts, and larger hepatocellular carcinomas have an extremely poor prognosis. The former feature is unusual for tumour markers, and illustrates a general limitation: markers tend to be proportional to tumour bulk - and hence early-stage tumours often go undetected.

Without invasive procedure, there are several sources of information by which one can diagnose cancer: site of presentation, physical symptoms, imaging, and the presence of soluble tumour markers in the blood or urine. Generally speaking, soluble tumour markers have been disappointing in aiding the differential diagnosis of malignant tumours - especially from their benign counterparts. In this context, onco-developmental markers tend to be of greater use than cell-type markers. Currently, there are only a few soluble markers used in the diagnosis of tumours: elevated levels of human chorionic gonadotropin (hCG) indicate gestational trophoblastic neoplasia (Bagshawe, 1992), α-fetoprotein correlates with hepatocellular carcinoma (Tang et al., 1989) and CA 125 indicates ovarian cancer (Zurawski et al., 1988). The diagnostic value of each of these markers may be augmented by tumour imaging using ultrasound.
If a biopsy is taken, diagnostic and prognostic markers can include cytogenetic and pathological features, and cell markers. The characterisation of cytogenetic markers suggests that while aberrations may prove useful in exploring etiology and molecular mechanisms of disease (e.g. in retinoblastoma), the generally poor correlation between defined features and specific cancers is such as to render them useless for diagnosis. Pathological features and cell markers are more reliable and have been used extensively in establishing diagnoses from biopsies. Examples of these are given in sections 1.2: (d) and (h). Most well-characterised cell markers are specialised structural components typical of the tissue of origin. More recently, nuclear transcription factors have also come under scrutiny with promising results. For example, the myf transcription factors apparently demonstrate high sensitivity and absolute specificity for embryonal and primitive rhabdomyosarcoma (Clark et al., 1991). Mutations in p53, a tumour suppressor gene encoding a transcription factor which blocks the cell cycle, have been used to differentiate metastatic from multifocal hepatocellular carcinoma (Oda et al., 1992). p53 protein accumulation has also been proposed as a prognostic indicator for breast cancer (Thor et al., 1992). It seems likely that p53 will feature more prominently as a prognostic indicator in the future, considering the multiple phenotypic alterations caused by different mutations, and the high frequency of these mutations in human cancers.

Historically, tumour markers have proven most successful in the arena of monitoring tumour progression/regression. Two of the "oldest" and most successful tumour markers are hCG and carcinoembryonic antigen (CEA), as expressed by gestational trophoblastic tumours and colorectal cancer respectively, although not exclusively. While these are used primarily in monitoring the course of disease, a brief review of these markers illustrates how detailed characterisation of antigen expression and function may refine their clinical application, and shed light on tumour behaviour.

hCG was the first major tumour marker to be identified. It is used as a marker for a group of tumours: gestational choriocarcinoma, hydatiform mole, and testicular tumours, but is most effective in the diagnosis and monitoring of choriocarcinoma. It is secreted in urine and is a heterodimer comprised of α and β subunits, both of which also appear as single chains; antibodies distinguish these subunits from each other, as well as from intact hCG. As the α subunit is highly homologous with that of luteinizing
hormone, care must be taken in selecting an appropriate antibody for its
detection. Variations in hCG expression also necessitate caution in the
interpretation of antibody assays, and invite further investigation of their
clinical significance. For instance "clipped" hCG with altered conformation
and reduced biological activity occurs when central peptide linkages
between the two chains are missing. During pregnancy, the extent of
sialylation of hCG changes, retarding its rate of clearance and augmenting
its biological activity - it may be that changes are also associated with cell
transformation. Evidence also suggests that the ratio of free β subunit to
intact hCG may be a prognostic indicator in patients with choriocarcinoma.
Thus, while hCG is a fine tumour marker in its own right, detailed
investigation of antigen expression has led to improvements in assay
design, and may lead to more refined diagnostic and prognostic
procedures.

Like hCG, CEA is most effectively used in monitoring the course of
disease, and has been used in the management of several tumours - notably
breast, lung and colorectal carcinoma. However, it is most useful as a
serum marker for colorectal carcinoma; its up-regulation concurs with
tumour growth and metastasis (Boucheret al., 1989) and this concurrence
accords with its known function - i.e. it is a homophilic cell surface
adhesion molecule, a member of the immunoglobulin superfamily. It is not
known exactly how CEA enhances metastatic potential, but experiments
suggest that it may do so by three possible means. a) Its expression may
result in clumps of cells instead of single cells breaking away from the
primary tumour - clumps are more likely to metastasise successfully than
single cells. This has been established in the case of melanoma (Updyke
and Nicholson, 1986), where the expression of ICAM-1 correlates with
metastatic potential (Johnson et al., 1989). b) It may disrupt other adhesion
processes required for integrity of tissue mass (Benchimol et al., 1989), as
has been suggested for sialylated NCAM (Yang et al., 1992). c) It may
mediate adhesion of metastases to their target organs (Hostetter et al.,
1990).

No discussion of tumour markers would be complete without
mention of the mucins - a group of heavily O-glycosylated glycoproteins
which have only recently received attention as potential tumour markers.
Several experiments suggest they make good candidates. Immuno-staining
sections of normal, lactating, benign and malignant sections of breast tissue
with a monoclonal antibody to the intact MUC-1 mucin stained almost all
sections. In contrast, a monoclonal antibody which reacts with the core protein of MUC-1 yielded significant signal with the malignant sections only, suggesting that this epitope is obscured in non-neoplastic tissue (Gendler et al., 1991). A similar trend has been noted in comparing pancreatic tumours vs. normal pancreatic tissue (Osako et al., 1993) and colorectal cancer vs. normal colon (Itzkowitz et al., 1989), although in these cases, antibodies to di- and core oligosaccharides (instead of the core protein) showed preferential staining with transformed tissue, indicating aberrant and incomplete glycosylation. Other work has demonstrated tissue-specific differences in the glycosylation of MUC-1 expressed by adenomas derived from different tissues (Dahiya et al., 1993).

The contribution of tumour markers to clinical management and knowledge of tumour biology argues the case for a more detailed characterisation of tumour markers, as does the fact that good tumour markers make good candidates for targeted therapy. While the markers themselves are still fairly "crude" in that it is usually the entire antigen (via a common epitope) which serves as an index, further researches into modulation of markers by neoplastic vs. "normal" and benign tissue may reveal highly reliable epitopes on what are considered to be not overly reliable antigens.

1.2 The Differential Diagnosis and Prognosis of Neuroblastoma

A solid tumour derived from progenitor cells of the sympathetic nervous system, neuroblastoma has been described as the prototype of the "small round cell tumours of childhood", a group of tumours which also includes rhabdomyosarcoma, Ewing's sarcoma, Wilms' tumour, and lymphoma. The differential diagnosis of these tumours can be extremely difficult, and much effort has been made to facilitate the task. It largely rests upon an approach investigating the pathology, biochemistry, cytogenetics, and immunohistochemistry/cytology of the tumour and derived cell lines. Some parameters used in the differential diagnosis of neuroblastoma are also used as prognostic indicators; thus it is convenient to discuss both prognosis and differential diagnosis where these interests coincide.

Over the years, various parameters and systems have been explored as aids in determining the diagnosis and prognosis of neuroblastoma.
These are discussed below.

(a) **Staging**

A variety of staging systems are used to describe the clinical status of different disease states. One of the first widely used staging systems is that of Evans (Evans *et al.*, 1971). A modification of this is the International Neuroblastoma Staging System (Brodeur *et al.*, 1988) and is regarded as one of the most reliable means for determining the prognosis of neuroblastoma. Outlined below, it is based on the extent of tumour development.

- **Stage 1.** Localised tumour confined to area of origin; identifiable lymph nodes negative microscopically.

- **Stage 2a.** Local spread through the capsule of the structure or organ of origin, but spread limited to same side as that of origin; identifiable lymph nodes negative microscopically.

- **Stage 2b.** As in 2a, but ipsilateral involvement of lymph nodes; contralateral lymph nodes negative microscopically.

- **Stage 3.** Spread through capsule of origin, with some spread to contralateral side, but no generalized spread. Lymph nodes may be involved on both sides. Alternatively, a unilateral tumour with contralateral lymph node involvement, or a midline tumour with bilateral lymph node involvement.

- **Stage 4.** Remote spread, involving skeleton, other organs, soft tissues or distal lymph node groups.

- **Stage 4s.** In patients under one year of age; a small primary tumour with remote spread to liver, skin and/or bone marrow, but no evidence of skeletal spread on complete skeletal survey.

Predictably, patients with Stages 1 and 2 disease have a comparatively good prognosis, while those with Stages 3 and 4 have a poor prognosis. Quixotically, Stage 4s has a favourable prognosis, despite metastases and in some instances, widespread subcutaneous nodules. In the majority of cases which present with 4s, dispersed metastases
disappear while any remaining tissue differentiates from immature neuroblasts into mature ganglion cells. Stage 1 and 2 cases of neuroblastoma, where the patient is under one year of age, also have a tendency to differentiate into ganglioneuromas. However, most patients present at Stages 3 or 4, contributing to the overall poor prognosis of the disease (Table 1.1).

(b) Age at diagnosis

The influence of age at diagnosis on prognosis is striking, and has been documented in many studies (Evans, 1980; Evans et al., 1987; Davis et al., 1987; Innocent, 1992; Oppendal et al., 1988; Shimada et al., 1984). Children who present with neuroblastoma and are under the age of one year, are much more likely to survive than those who present at a later stage of development.

This trend might be partly explained by the fact that Stages 1, 2 and 4s are better represented in children under one year, than in those older than one year, while Stages 3 and 4 are more commonly observed in the latter age group. However, age is an independent variable in the prognostic equation. For example, a patient with Stage 2a disease at six months is more likely to survive than a patient with Stage 2a disease at five years. A higher rate of spontaneous regression in younger patients may contribute to this trend. Supporting this theory is the finding that autopsies of infants who have died from unrelated causes and analyses of foetuses have revealed much higher incidences of neuroblastoma than those encountered in vivo (Guin et al., 1969; Shanklin and Sotelo-Avila 1969).

(c) Site of tumour

The site at which the primary tumour establishes, influences prognosis. It has emerged over time that tumours arising at sites above the diaphragm have a better prognosis than those which arise below. Adrenal
Table 1.1

Incidence of neuroblastoma and survival rate with respect to stage of disease
(printed with permission from J. Kemshead)
Table 1.1

<table>
<thead>
<tr>
<th>Good prognosis</th>
<th>All Neuroblastomas (%)</th>
<th>3-year survival (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1</td>
<td>5</td>
<td>100</td>
</tr>
<tr>
<td>Stage 2a</td>
<td>5</td>
<td>90-100</td>
</tr>
<tr>
<td>Stage 4s</td>
<td>5-10</td>
<td>70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intermediate prognosis</th>
<th>All Neuroblastomas (%)</th>
<th>3-year survival (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 2b</td>
<td>5</td>
<td>50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Poor prognosis</th>
<th>All Neuroblastomas (%)</th>
<th>3-year survival (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 3</td>
<td>15-20</td>
<td>30-40</td>
</tr>
<tr>
<td>Stage 4</td>
<td>60</td>
<td>15-20</td>
</tr>
</tbody>
</table>
primary tumours arising in the abdomen, constituting a large proportion of neuroblastomas, have a particularly poor prognosis.

(d) Pathology

In the differential diagnosis of the "small round cell" tumours of childhood, pathology is useful in the majority of cases, where distinct features which correspond with cell type can be observed either by light or electron microscopy. Features such as catecholamine granules, neural processes, and neurofilaments are regarded as indicative of neuroblastoma. However, primitive neuroblastomas often lack these components. In addition, catecholamine granules can be confused with lysosomes. Some neuroblastoma cell lines can be induced to undergo differentiation under the influence of agents such as retinoic acid (Sidell et al., 1983), phorbol esters (Pahlman et al., 1984), nerve growth factor (Pavelic and Spevanti, 1987) or vasoactive intestinal peptide (Pence and Shorter, 1990) whereupon characteristic neurites are easily observed. Other neuroblastoma cell lines however, are resistant to differentiating agents and some of the other small cell tumours (e.g. Ewing's sarcoma, which is now believed to be of neural crest origin) will demonstrate neural differentiation in response to such agents.

Histological factors have been found to be of great use in determining the prognosis of specific cases of neuroblastoma, as noted by Shimada et al. in 1984. These investigators classified tumours from 295 patients on the basis of histology and correlated histological grade with stage (as described above), age at diagnosis, primary site and survival status. It was discovered that histological grade proved to be an additional useful prognostic indicator, complementing Evans' staging system.

Tumour cells observed to be rich in stroma had a favourable prognosis, unless they presented a nodular phenotype, in which nodules of undifferentiated neuroblasts are found amidst mature matrix. "Stroma-
poor tumour cells had a comparatively unfavourable prognosis. However, the authors discovered that the application of other prognostic parameters - the extent of cellular differentiation, age at diagnosis, and nuclear pathology - resulted in two sub-groups of "stroma-poor" cells; one with a favourable outlook (84% survival), and the other with an unfavourable prognosis (4.5% survival). The Shimada system illustrates how diagnosis may be refined by including biological features in the clinical assessment. As the system is labour-intensive and a certain degree of subjectivity is inherent in the assessment of tumour pathology, it is rarely used these days. However, its parameters are of interest in that they demonstrate how prognosis is affected by biological indices.

(e) Serum concentrations of neuron-specific enolase and ferritin

The serum level of (γγ) neuron-specific enolase (NSE) has been cited as a potential prognostic parameter (Zeltzer et al., 1983). Its level of expression has been noted in some cases to correlate with the degree of tumour differentiation. Other studies have failed to observe such a correlation (Tsokos et al., 1984), and have in fact suggested the term "neuron-specific" to be a misnomer as the γ enolase chain has been demonstrated to be present in other tissues, albeit at lower levels. However, expression of NSE may be of use in the differential diagnosis of neuroblastoma from rhabdomyosarcoma when used in conjunction with other parameters [see section (h)].

An increased level of serum ferritin (100 - 150 ng ml⁻¹) is correlated with a poor prognosis (Hann et al, 1985; Evans et al., 1987) and like NSE concentration, is easy to detect by immuno-assay. Ferritin also appears to be secreted by the tumour itself; nude mice inoculated with human neuroblastoma cells were found to contain human ferritin in their serum (Hann et al., 1985). Physiologically, it differs from liver ferritin, in that it
inhibits rosette formation of lymphocytes. Increased serum ferritin is not always detected; its absence is more likely in the older patient, and in those with Stage 4s neuroblastoma. It is of little use as a differentiating agent, as other tumours have been found to secrete it.

(f) Catecholamine expression and secretion

By definition, cells which engage in amine precursor uptake and decarboxylase activity (otherwise known as APUD cells), metabolise amines giving rise to catecholamines. Neuroblastoma is an "APUDoma"; consequently the presence of amines in tumour tissue, serum, and urine can be used as a prognostic index. Less differentiated tumours tend to secrete higher levels of the comparatively "primitive" homovanillic acid, while tumours exhibiting a greater degree of differentiation predominantly secrete vanillylmandelic acid. It follows that a high ratio (>1.0) of the concentration of vanillylmandelic acid to that of homovanillic acid correlates with a favourable prognosis, whereas a low ratio predicts a poor one. Not all neuroblastoma patients secrete catecholamines into the serum or urine. Older patients especially, in whom diagnosis is often difficult, are unlikely to have detectable levels.

Mass screening programmes have been implemented in Japan and Canada, in which the catecholamine metabolites in urine of infants under 1 year are assayed, resulting in an increased detection rate of children with low grade neuroblastoma. Such programmes have a possible disadvantage - at earlier ages of screening, a higher proportion of children with naturally-differentiating neuroblastoma are likely to be identified and unnecessarily treated. However, results suggest that the programme implemented in Japan may result in a decrease in the incidence of metastatic neuroblastoma although results are preliminary (Craft and Parker, 1992).
(g) Genetic factors

The observation of double minutes and/or homogeneously staining regions on karyotype analysis, implies gene amplification. However, amplification of the N-myc oncogene can only be confirmed by molecular methods. N-myc amplification correlates well with rapid progression of neuroblastoma and a poor prognosis (Seeger et al., 1985). Recent work on the distribution of the N-myc protein in quail neural crest (Wakamatsu et al., 1993) implies that its intracellular location might correlate with state of differentiation, and by implication, with the prognosis of neuroblastoma. N-myc amplification in other "small round cell" tumours e.g. rhabdomyosarcoma, renders it ineffective as an index by which to differentiate them from one another (Garson et al., 1986; Hayashi et al., 1990). It is possible that an oncogene analogue, smg p25A may prove of use in the differential diagnosis of neuroblastoma (Sano et al., 1990), but its characterisation is, as yet, preliminary. Similarly, expression of the neuronal form of pp60c-src may prove to be a diagnostic and/or prognostic indicator (Bjelfman et al., 1990).

Translocations involving chromosome 1 are observed in a significant proportion of neuroblastomas, although it is a poor diagnostic parameter, as many other tumours also have translocations and other gross aberrations of genetic integrity in this chromosome. Ploidy is also a useful prognostic indicator as patients with either aneuploid or polyploid tumours tend to fare much better than those with a diploid complement.

(h) Immunohistochemistry and immunocytology

Since their discovery in 1975 by Kohler and Milstein, monoclonal antibodies have revolutionised cell biology. They have proven invaluable in experiments exploring functional aspects of proteins and facilitated the study of immunohistochemistry. They have proven useful in the identification of tumour types and the detection of metastatic spread to
tissues such as the blood and bone marrow. *In vivo*, they are used to image tumours by immunoscintigraphy, and also as agents with which to target toxic agents to tumours.

As with some of the methods described above, there have been many false starts with antibodies and the identification of "small round cell" tumours; almost as soon as an antibody is believed to be specific for a given tumour type, it is shown to either cross-react with other types of tumour, or to react with only a subset of specimens of a given tumour type. This is partly due to heterogeneity in the expression of cell surface antigens, as has been documented in neuroblastoma (Malpas *et al.*, 1982; Schönmann *et al.*, 1986). However, the use of panels of antibodies helps to compensate for the diagnostic pitfalls resulting from antigen heterogeneity, and maximises the chances of tumour detection in the clinical context.

Antigens and monoclonal antibodies proposed for use in the differential diagnosis of rhabdomyosarcoma and neuroblastoma tissues are summarised in Table 1.2. While they are indispensable in the typing of a large majority of these tumours, differential diagnosis of some primitive tumours remains problematic. For instance, of the cytoplasmic markers, the intermediate filaments are some of the most diagnostically useful tumour markers. Expression of vimentin, desmin, keratin, neurofilament and glial fibrillary acidic protein accord with tissues of mesenchymal, muscle, epithelial, neuronal and glial tissue respectively. Their expression is restricted to mature cells; accordingly, their use is limited in the characterisation of very primitive tumours, although desmin is found in most embryonal tumours arising from muscle. In contrast with normal tissue, primitive tumours not only do not express antigens typical of their tissue-type, they also sometimes express antigens atypical of their tissue of origin. For instance rhabdomyosarcoma tissue has been found to express neurofilament (Meittinen and Rapola, 1989), while desmin and α-actin have
Table 1.2

Antibodies which have been proposed as agents by which to differentially diagnose neuroblastoma and rhabdomyosarcoma
Table 1.2

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Antigen</th>
<th>General Tissue-specificity</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>anti-NSE</td>
<td>neuron specific enolase</td>
<td>neural</td>
<td>Tsokas et al., 1984</td>
</tr>
<tr>
<td>anti-neurofilament (all isoforms)</td>
<td>neurofilament</td>
<td>neural</td>
<td>Molenaar, et al., 1990</td>
</tr>
<tr>
<td>SY 38</td>
<td>synaptophysin</td>
<td>neuro-endocrine</td>
<td>Wiedenmann et al., 1986</td>
</tr>
<tr>
<td>LK2H10</td>
<td>chromogranin A</td>
<td>neuro-endocrine</td>
<td>Lloyd and Wilson, 1983</td>
</tr>
<tr>
<td>HSAN 1.2</td>
<td>unidentified</td>
<td>neural</td>
<td>Reynolds and Smith, 1980</td>
</tr>
<tr>
<td>KP-NAC 8</td>
<td>unidentified</td>
<td>neural</td>
<td>Matsumura et al., 1987</td>
</tr>
<tr>
<td>PI153/3</td>
<td>unidentified</td>
<td>neural</td>
<td>Kemshead et al., 1983</td>
</tr>
<tr>
<td>anti-desmin</td>
<td>desmin</td>
<td>muscle</td>
<td>Osborn and Weber, 1983</td>
</tr>
<tr>
<td>anti-MyoD1</td>
<td>myoD1 protein</td>
<td>muscle</td>
<td>Dias et al., 1990</td>
</tr>
<tr>
<td>(polyclonal)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>anti-myoglobin</td>
<td>myoglobin</td>
<td>muscle</td>
<td>Osborn and Weber, 1983</td>
</tr>
<tr>
<td>anti-α-sarcomeric</td>
<td>α-sarcomeric actin</td>
<td>muscle</td>
<td>Schüch et al., 1987</td>
</tr>
<tr>
<td>actin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHF35</td>
<td>α and γ actin</td>
<td>muscle</td>
<td>Tsukada et al., 1987</td>
</tr>
</tbody>
</table>
been detected in neuroblastoma cell lines (Sugimoto et al., 1991).

Various leukocyte differentiation antigens have been proposed as markers by which to differentially diagnose leukaemia and lymphoma from non-lymphoid neoplasias - most notably members of the "small round cell" tumours. These include CD30 and CD45 in the case of lymphoma (Mason and Gatter, 1987; Pilkington and Pallesen, 1989; Smith et al., 1993) and CD53, found on both leukaemias and lymphomas (Mechtersheimer et al., 1993). CD10 - the common acute lymphoblastic leukaemia-associated antigen - has been described as a marker for leukaemia (Feickart et al., 1989; Ebener et al., 1990) although research suggests that it may be expressed by some neuroblastoma cell lines (Sugimoto et al., 1985; Pilkington and Pallesen, 1989).

Antigen profile can also be used prognostically. For instance, neurofilaments in neuroblasts are of comparatively small size (68 kDa) with respect to those expressed in mature neurons (160 and 200 kDa). Neurofilament expression in neuroblastomas of varying degrees of differentiation accords with this profile (Molenaar et al., 1990) and antibodies specific to the three different neurofilament proteins are available (Sternberger and Sternberger, 1983).

As with other tumours, the presence of P-glycoprotein in neuroblastoma has been found to correlate with multi-drug resistance, poor response to chemotherapy, and hence a poor prognosis (Chan et al., 1991; Gazitt et al., 1992). The way in which this occurs, however, remains to be elucidated: resistance to cyclophosphamide and cis-platin, commonly used as chemotherapeutic agents in the treatment of neuroblastoma, is thought to be dissociated from P-glycoprotein expression (Dalton et al., 1989). Furthermore, the expression of P-glycoprotein by neuroblastoma is itself controversial. Some workers have failed to detect its presence on scanning panels of tumour samples (O'Meara et al., 1992); others suggest that its expression on neuroblastoma directly correlates with a limited pattern of differentiation (Bates et al., 1991).

It is thus seen that a wide range of techniques has been developed for the purpose of improving the differential diagnosis and prognosis of neuroblastoma. While most neuroblastoma tumours are readily identified using some of the techniques described, a residue defies diagnosis. It is for this reason that research continues in an attempt to establish new, reliable procedures which allow their identification. In this regard, characterisation of cell surface molecules is a promising avenue of research; these molecules are accessible to external probes, rendering them good candidates for
targeting as well as diagnostic indicators. Their potential for morphoregulatory function suggests that they may be indicative of cell type.

1.3 Adhesion Molecules and Tumour Biology

Cell surface molecules play a fundamental role in many biological processes. Embryogenesis, organogenesis, homeostasis, and wound repair all depend on interactions between cell and local environment (See Edelman, 1984 and Springer, 1990 for reviews). As surface molecules lie at the interface between cell and environment, it is inevitable that they play a communicative role between the two. As such, they can be regarded as transmembrane linkers between intra- and extra-cellular environs.

Tumour cells are aberrant in that they contravene the laws of homeostasis, and assume a behavioural phenotype more typical of an embryonic cell. The homing tendencies normally associated with immune surveillance are also observed in the capacity of the cancer cell to arrest in the vessels of "foreign" tissue and extravasate. On considering the role of adhesion molecules in tumour biology, it may thus be informative to consider their parallel roles in related processes. Conversely, advances in the characterisation of adhesion molecules in the context of oncology may impinge upon other fields of study.

It is important to recognise that interactions between cell surface molecules and those of the extracellular environment cross a spectrum which ranges from tight adhesion to molecular "indifference", and possibly repulsion. The cellular response to cell surface events is similarly varied - cell cycle, rate of mitosis, metastasis and apoptosis can all be influenced by environmental conditions which are communicated via the cell surface. Such responses are likely to result from the integration of a variety of incoming signals. Therefore, a comprehensive understanding of
how extracellular influences give rise to changes in cell behaviour will only be obtained when it is understood how a mixture of "messages" from the cell surface are integrated.

Augmenting the challenge of relating form to function is the fact that new surface recognition molecules are discovered almost weekly - in this respect we are still learning the "alphabet" of recognition machinery. To date, three major classes of molecule have been found to mediate cell adhesion: proteoglycans, glycolipids, and glycoproteins. Cellular glycoproteins involved in adhesion can be subdivided into five main groups: cadherins, selectins, integrins, the immunoglobulin superfamily and the cartilage link family. Integrins and members of the immunoglobulin superfamily are known to interact with components of the extracellular environment - such as extracellular matrix proteins. Recent evidence has uncovered associations both within and between members of these groups. Advances in cloning and antibody technology have given rise to a wealth of publications over the last decade, relating to these classes of molecules and their behaviour. It would thus be impractical to attempt a general review; instead selected examples which pertain to tumour biology and especially neuroblastoma are discussed.

(a) The cadherins

Cadherins are molecules which mediate cell adhesion and are resistant to protease activity via (a) calcium ion-dependent mechanism(s) (Takeichi, 1988). They are generally trans-membrane molecules, and interact by homophilic binding. Three cadherins have been identified for some time. Their structure, function and expression have been extensively characterised. They are: placental cadherin (p-cadherin; Nose and Takeichi, 1986), epithelial cadherin (e-cadherin; Nagafuchi et al., 1987), and neural cadherin (n-cadherin; Hatta et al., 1987). It should be noted however, that p-cadherin derives its name from its expression in the

33.
mouse placenta. It has yet to be detected in the human placenta.

The p-, e- and n-cadherins exhibit roughly 50% homology, indicating an evolutionary relationship. They bind to other cadherins of the same type via a homophilic binding mechanism, although weaker heterophilic interactions between different types of cadherin have been observed \textit{in vitro} (Friedlander, \textit{et al.}, 1989). Recently, several novel cadherins have been discovered, one of which is unique in that it is attached to the cell membrane via a phosphatidylinositol anchor (Vestal and Ranscht, 1992).

N-cadherin is a primary adhesion molecule; it is expressed on all three germ layers in the developing embryo. It is expressed in neural and muscle tissues and expression is modulated during early events in embryogenesis, including formation of the neural tube and sensory ganglia. The dynamic shift in expression of E-cadherin, N-cadherin, and the neural cell adhesion molecule during embryonic cell migration suggests morphoregulatory roles for these molecules (See Figure 1.1).

Comparative studies using monoclonal antibodies to investigate cadherin expression in tumour and normal tissues generally indicate an inverse correlation between cadherin expression and tumourigenicity (For examples see Navarro \textit{et al.}, 1991; Shiozaki \textit{et al.}, 1991). This would appear to be consistent with the perceived role of cadherins in mediating adhesion between homologous cell types, and suggests that the initial event in metastasis - i.e. the breaking away of a tumour cell from the primary tumour - may be mediated by cadherin down-regulation. An embryonic analogy to this sequence of events is the down-regulation of N-cadherin on migrating neuroblasts during formation of the sympathetic nervous system.
Figure 1.1

Expression of NCAM, N-cadherin and E-cadherin during morphogenesis

The co-ordinated expression of cell adhesion molecules with morphogenetic events suggests a functional correlation between the two. 

a) formation of the primitive streak, in which cells lining the streak do not express NCAM or E-cadherin, in contrast to those found in the rest of the ectoderm. Note that migrating "epiblasts" do not express NCAM or the E- and N-cadherins. 
b) Condensation of the "epiblasts" results in the formation of the notochord and somites; N-cadherin and NCAM are expressed in these tissues. They are also expressed by cells comprising the nascent neuroderm. 
c) The neuroderm invaginates, forming the neural tube. d) NCAM and N-cadherin are initially expressed in all regions of the neural tube (not shown). Prior to migration of neural crest cells, the cells in the dorsal part of the neural tube lose N-cadherin. NCAM is found on migrating neural crest cells during the initial stages of migration. As migration proceeds, the cells resemble migrating "epiblasts" in that they do not express NCAM, N-cadherin or E-cadherin. Condensation of the migrating neural crest cells results in the formation of sensory ganglia, in which NCAM and N-cadherin are re-expressed. N.B. While N-cadherin and NCAM are expressed by the endoderm during the early stages of chick embryogenesis, modulation of these molecules by the endoderm is not depicted in this figure. (Adapted from Linnemann and Bock, 1989).

Key:

- NCAM and E-cadherin
- E-cadherin
- NCAM and N-cadherin
- NCAM
- NCAM, N-cadherin and E-cadherin
Figure 1.1

a)

b)

c)

d)

ectoderm
primitive streak
migrating epiblast
endoderm

neural tube
neuroderm
somatic ectoderm
notochord
mesoderm
somite

sensory ganglia
migrating neural crest cell
neural tube
(b) The integrins

Integrins comprise a large family of transmembrane heterodimers which have been found to interact primarily with extracellular matrix constituents including fibronectin, vitronectin, laminin, fibrinogen, von Willebrand factor, and several collagens, sometimes via an arginine-glycine-glutamine (RGD) sequence on the ligand (Pierschbacher and Ruoslahti, 1984; for review, see Hynes, 1992). Integrins have also been found to bind members of the immunoglobulin superfamily; for instance the lymphocyte functional antigen - 1 (LFA-1) binds intercellular adhesion molecules 1 and 2 (ICAM-1 and ICAM-2) (Staunton et al., 1989), the very late antigen-4 (VLA-4) binds the vascular cell adhesion molecule (VCAM-1) (Elices et al., 1990), and Mac-1 binds the third domain of ICAM-1 (Diamond et al., 1991).

The integrin heterodimer consists of one α-subunit and one β-subunit which are bound non-covalently to form a functional receptor, the specificity of which depends on the particular combination of α- and β subunits. To date, three integrin families are defined by the type of β-subunit its members possess.

Integrins containing β₁ are found in a wide variety of cell types and constitute the very late antigen (VLA) family. These predominantly mediate cell attachment to the extracellular matrix, although integrins VLA-1 and VLA-4 have been implicated in mediating T-cell cytotoxicity in the context of melanoma (Pandolfi et al., 1992) and VLA-4 in the formation of myotubes (Rosen et al., 1992). Integrins containing β₂ are involved in intercellular interactions within the immune system and are consequently described as Leu-CAMs (Kishimoto et al., 1989).

Like the cadherins, the expression of some integrins appears to disappear with the advent of the malignant phenotype [for example, see section (c)]. Upregulation of some integrins however, correlates with the
progression of specific tumours. This stands to reason, as many integrins - e.g. those of the β1 family - mediate interactions with the extracellular matrix, and on contemplating the metastatic cascade, one would expect an increase in the expression of molecules which may mediate extravasation, intravasation and arrest.

(c) Fibronectin

Of the extracellular matrix molecules with which integrins react, fibronectin is one of the most extensively characterised, and has been shown to mediate cell adhesion and differentiation. It is found along pathways of neural cell migration and is associated with the basement membranes of peripheral nerves (Paetau et al., 1980). A brief review of its structure and expression illustrates how its molecular structure (and presumably function) can vary from tissue to tissue.

Fibronectins are large, adhesive glycoproteins comprising two disulphide-linked subunits of approximately 250 kilodaltons (kDa). They play an important role in embryonic cell migration, adhesion, and differentiation (for review, see Humphries and Yamada, 1989). Amino acid sequence analysis has shown that these subunits are made up of repeating sequences which show homology to one another; on this basis repeating sequences can be divided into three homology groups (I, II, and III). The molecules can also be divided into domains which bind fibrin, heparin, collagens, some bacteria, and DNA, as well as cell surfaces (see Figure 1.2). The tenth fibronectin type III repeat lies within the cell-binding domain and contains a GRGDS sequence believed to interact with integrins.

Cellular fibronectin, produced by fibroblasts and other cell types, differs from plasma fibronectin in that it is locally produced and insoluble, whereas the latter is soluble and produced mainly by the liver (Tamkun and Hynes, 1983). Certain regions expressed in cellular fibronectins are not
Figure 1.2

Structure and alternative splicing patterns of fibronectin RNA

Three regions of the fibronectin protein are alternatively-spliced: ED-A, ED-B and the type III connecting segment (IIICS). IIICS is encoded by a single exon with internal acceptor sites; as a consequence, different parts of the exon (CS1, V64, and CS5) can be included or excluded from the final mRNA transcript. Variations in splicing pattern of the IIICS region are indicated directly below its position in the transcript. Fibronectin binds various components of the extracellular matrix and VLA-4. These are indicated above the exons encoding the portion of polypeptide to which they bind.

Key:

- \(=\) fibronectin type I repeat
- • = fibronectin type II repeat
- ◯ = fibronectin type III repeat
- = CS1
- = V64
- = CS5
expressed in plasma fibronectin. This is achieved by developmentally-regulated alternative splicing of fibronectin messenger RNA (mRNA). It would seem that sufficient information is provided by fibronectin pre-mRNA to give rise to alternatively-spliced products (Mardon and Sebastio, 1992). However, cell type-specific trans-acting factors, including growth factors, influence their relative proportions (Sekiguchi and Titani, 1989; Barone et al., 1989; Magnuson et al., 1991).

To date, three regions of the fibronectin molecule have been shown to be alternatively spliced: ED-A, ED-B and IIICS (see Figure 1.2). ED-A and ED-B are fibronectin type III repeats - each encoded by an exon which can be spliced out of mRNA (Zardi et al., 1987; Paolella et al., 1988). Currently, it is not possible to assign putative functions to ED-A and ED-B on the basis of amino acid sequence, although it is possible to speculate on the basis of their tissue distribution. As such, it has been proposed that ED-B plays a role in cell migration and/or proliferation prior to terminal differentiation. Consistent with this is its expression in the developing chicken embryo (Ffrench-Constant and Hynes, 1989a), its oncodevelopmental expression in fetal and cancerous lung tissues (Oyama et al., 1990), and its re-expression during wound-healing in the adult rat (Ffrench-Constant et al., 1989b). Similarly, it has been suggested that ED-A may be involved with matrix assembly and/or as a negative regulator of cellular differentiation; its excision appears to be down-regulated in malignant liver tumours (Oyama et al., 1989).

The IIICS region in its entirety distinguishes cellular fibronectin from plasma fibronectin. In the former both constituent polypeptides contain IIICS, while the latter is a heterodimer with respect to this alternatively-spliced domain. The IIICS domain connects two type III fibronectin repeats and comprises three mini-exons - CS1, V64 and CS5 - defined initially in the rat fibronectin gene by internal donor and acceptor
sites (Tamkun et al., 1984).

Recent investigations exploring the function of these exons (or parts thereof) demonstrate how alternative splicing of an extracellular matrix molecule may affect the migration and adhesion of certain cell types. For example, sympathetic neurons from the chick embryo extend neurites either on peptides comprising the cell binding domain or CS1; the effects of these peptides on neurite outgrowth are additive (Humphries et al., 1988). CS1 can also mediate the migration of neural crest cells (Dufour et al., 1988) and bind the VLA-4 integrin on lymphoid cells (Guan and Hynes, 1990). Like CS5, it has been shown to mediate the attachment of melanoma cells (Kocher et al., 1990). The active site of CS5 is believed to be the tetrapeptide REDV, as synthetic REDV inhibits melanoma cell spreading on fibronectin (Humphries et al., 1986a). With respect to melanoma metastasis, the cell binding domain containing GRGDS would seem more influential than CS1 and CS5. Synthetic peptides comprising CS1, GRGDS and REDV were co-injected with melanoma cells into the tail veins of syngeneic mice; the GRGDS pentapeptide inhibited lung colonisation while the other mice injected with the other two polypeptides had no survival advantage over the control (Humphries et al., 1986b). However, it must be remembered that GRGDS may mimic active sites in other ECM molecules as well as that of fibronectin.

Studies have shown that poorly differentiated tumours often have few or undetectable fibronectin receptors (Plantefaber and Hynes, 1989) in keeping with the concept that cancer cells have reduced potential for contact with certain components of the extracellular matrix. Predictably, over-expression of fibronectin receptors results in a decrease in anchorage independent growth (Giancotti and Ruoslahti, 1990). Furthermore, an antisense construct of the promoter and 5' region effects a decrease in cell surface fibronectin of hybrid melanoma/fibroblasts, with a concomitant
increase in tumorigenicity (Steel and Harris, 1989). Neuroblastoma cell lines have been found to variably express fibronectin receptors VLA-4, -5 and -6; expression of VLA-5 correlates with a phenotype suggesting reduced tumourigenicity (Yoshihara et al., 1992).

Other extracellular matrix molecules also influence neural behaviour or are associated with the migration of neural crest cells. These include laminin, tenascin, various collagens, and a variety of glycosaminoglycans such as heparan sulphate, chondroitin sulphate and hyaluronate - the latter three are often found in the context of a proteoglycan. However, it is not within the scope of this review to attempt description of all known molecules which may play a role in the biology of neuroblastoma; rather to illustrate via selected examples, how such molecules may influence the biology of the neuroblastoma cell and hence the course of disease.

(d) The homing receptors

The selectins comprise a subset of the homing receptors. They are transmembrane glycoproteins, the extracellular part of which comprises regions with homology to the complement binding proteins and epidermal growth factor. They terminate in a lectin-like domain. (For a review, see Berg et al., 1989). They mediate binding between unstimulated leukocytes and the luminal surface of endothelial cells prior to extravasation and tissue surveillance. Some are found on lymphocytes (e.g. LECAM-1) and others, on epithelial cells (e.g. GMP-140 and ELAM-1). Their known ligands are carbohydrates, to which the N-terminal lectin is believed to bind.

CD44, a member of the cartilage link protein family, also functions as a homing receptor and is of especial interest in the context of tumour biology. Extensive and varied post-translational modifications combined with alternative splicing confers potential for heterogeneity in isoform expression, perhaps altering ligand specificity. It is known to bind
hyaluronate, an almost omnipresent extracellular matrix molecule which is also involved in cell-cell interactions and aggregation. CD44 is involved in homing lymphocytes and monocytes to lymph nodes and sites of inflammation and has also been implicated in triggering lytic activity in cytotoxic T lymphocytes (Haynes et al., 1989; Seth et al., 1991; Pandolfi et al., 1992). Variants have been detected on keratinocytes, melanoma, squamous lung carcinoma, fibrosarcoma, pancreatic carcinoma, and mammary adenocarcinoma.

CD44 is of great interest to molecular oncologists: a variant which correlates with and appears to induce metastasis has been identified. The overexpression of a variant protein in a non-metastasising rat pancreatic carcinoma cell line establishes metastatic behaviour (Günthert et al., 1991). Moreover, co-injection of the metastasising cell line with monoclonal antibodies specific to the variant protein reduces metastatic spread (Reber et al., 1990). Sequence analysis of cDNA encoding the human homologue has since been performed, and metastatic potential found to correlate with the alternative splicing of two exons proximal to the plasma membrane (Hofmann et al., 1991). The functional characterisation of these exons should prove to be most interesting.

(e) The immunoglobulin superfamily

The third class of cognitive cell-surface glycoproteins, the immunoglobulin superfamily, is believed to have evolved from a single ancestral gene by duplication, fusion, and exon shuffling (Williams, 1987). It includes the immunoglobulins themselves, CD2, CD3, CD4, CD8, P0, myelin-associated glycoprotein (MAG), Thy-1, carcinoembryonic antigen (CEA; Benchimol et al., 1989), intercellular adhesion molecules (ICAM-1, ICAM-2 and ICAMR; Vazeux et al., 1992), MUC18 (Lehmann et al., 1989), the vascular cell adhesion molecule (VCAM-1; Osborn et al., 1989), F3 (Gennarini et al., 1989), the neuron-glia CAM (NgCAM; Burgoon et al.,
1991), the neuron-glia related CAM (NrCAM; Mauro et al., 1992), L1 (Rathjen and Schachner, 1984) and the neural cell adhesion molecule (NCAM; Cunningham et al., 1987).

Genetic loci of these molecules are spread widely through the chromosomes, but there are a number of linkage groups (see Williams and Barclay, 1988 for review). In the human genome, NCAM is linked with Thy-1 amongst other members of the superfamily. It has been suggested that this linkage group is of primitive origin; the immunoglobulin domains of NCAM are each encoded by two exons, the latter of which are believed to have fused later in evolution as the Ig domains of most other members are encoded by a single exon. Furthermore, the Ig-like C2 domains in NCAM (and L1) are intermediate between the variable and constant domains of the classic immunoglobulin - suggesting that NCAM and its "C2" relatives probably started evolving prior to the duplicative event giving rise to the separate V and C domains of the Igs. A high degree of conservation throughout evolution is also demonstrated by the expression of all of the major NCAM isoforms (of 120, 140 and 180 kDa) in the CNS of species as far removed as the frog, mouse, chicken and human. Consequently, a large range of animals has been used to investigate NCAM expression, especially in the areas of embryogenesis and organogenesis, with a certain degree of confidence that the findings of such investigations are either representative of or at least relevant to NCAM expression in general.

A classic theory of tumourigenesis is that development of stem cells is blocked, and instead of differentiating into specialised tissue types, they maintain an embryonic phenotype. As L1 and NCAM are expressed during embryogenesis and tissue differentiation, they are of interest in the context of oncology. In addition, both L1 and NCAM are adhesion molecules and as such, may play a critical role in tumour metastasis.
The L1 neural cell adhesion molecule

L1 was identified more recently than NCAM, and hence its characterisation is comparatively limited. It has however, been cloned and sequenced (Kobayashi et al., 1991). A transmembrane protein, it contains five immunoglobulin-like domains and 6 fibronectin type III repeats. Its protein expression has been documented in the mouse (Rathjen and Schachner, 1984), rat (Salton et al, 1983), and human (Wolff et al., 1988). The neuroglian molecule found in Drosophila (Bieber et al., 1989) and NgCAM in chicken (Grum et al. and Edelman 1984) exhibit a similar distribution, and extensive homology to L1. Adhesion assays using purified L1 suggest that it supports neurite outgrowth in the mouse and chick via heterophilic and homophilic binding mechanisms (Kuhn et al., 1991; Lemmon et al., 1989). NrCAM, closely related to L1, also exhibits both heterophilic and homophilic binding (Mauro et al., 1992). Other functions of L1 include neuron-neuron adhesion, fasciculation (Linnemann et al., 1987) and mediating neurite extension along other neurites and ensheathing cells (Wood et al., 1990).

Like NCAM, the expression of L1 was initially believed to be limited to the nervous system. Also like NCAM, it is more widely expressed than initially supposed; its presence has been demonstrated in melanoma, rhabdomyosarcoma and mouse leukocytes (Linnemann et al., 1989a; Kowitz et al., 1992; Reid and Hemperly, 1992). More predictably, it is expressed by neuroblastoma cells. While NCAM appears at early stages in the development of the CNS, PNS and other tissues, L1 expression occurs later. In the nervous system, it is restricted to post-mitotic neurons in early histogenesis and at later stages, non-myelinated axons (Persohn and Schachner, 1987). In mouse brain, the protein has been shown to strongly associate with a protein kinase activity (Sadoul et al., 1989). cDNA analysis reveals that L1 undergoes alternative splicing of a 12 bp exon in the
cytoplasmic domain (Miura et al., 1991; Reid and Hemperly, 1992) although the functional significance of this event remains to be elucidated.

There is preliminary evidence that L1 may participate in the extracellular environment; it is found in supernatants of cultured peripheral neurons (Sweadner, 1983) and is detectable in basement membranes and collagen fibres in proximity to murine sciatic nerve (Martini and Schachner, 1988). It bears two RGD sequences in the Ig-like domain 6 (Moos et al., 1988) and thus has the potential to interact with certain integrins; however whether the orientation of these sequences within the larger context of the molecule permits such interaction remains to be seen.

1.4 The Neural Cell Adhesion Molecule (NCAM)

NCAM exhibits remarkable homology with L1 in that they are both transmembrane glycoproteins which comprise Ig-like C2 domains and fibronectin type III-like repeats. They are both homophilic, intercellular adhesion molecules (Rutishauser et al., 1982; Lemmon et al., 1989). Their similarity in structure may allow for cis- NCAM/L1 interactions which result in a more potent receptor for L1 on another cell (Kadmon et al., 1990). Consistent with this suggested liaison is the similarity in reduction of the lateral mobility of both L1 and NCAM-180 with long-term morphological differentiation, and the finding that they co-localise at sites of cell contact (Pollerberg et al., 1990b). It has been proposed that cis-interactions between NCAM isoforms may facilitate multivalent binding; if this is proved to be correct, then NCAM/NCAM cis-binding may occur via a similar mechanism.

(a) NCAM expression

NCAM is one of the most thoroughly characterised members of the immunoglobulin superfamily. It has been investigated in reptiles, birds and mammals, and also shows remarkable homology with fasciclin II, an
NCAM has been implicated in a wide range of events. These include axonal bundling (Thanos et al., 1984), promotion of gap-junction formation (Keane et al., 1988), neural migration along glial pathways (Silver and Rutishauser, 1984), myoblast fusion (Knudsen et al., 1989, 1990), innervation (Covault and Sanes, 1985), and stabilising synapses (Persohn et al., 1989). Its role in mediating immune responses to leukaemic targets (Lanier et al., 1991; Palucka et al., 1992) and myelination (Bhat and Silberberg, 1990; Seil and Herndon, 1991) remains controversial. Both NCAM and N-cadherin have been implicated in modulating early events during morphogenesis; especially in the migration of neural crest cells and their condensation into sympathetic ganglia. While appropriate expression of NCAM is essential for later stages in neural development (Fraser et al., 1984; Silver and Rutishauser, 1984), it seems that N-cadherin may play a more influential role in the initiation and cessation of neural crest cell migration; changes in its expression correlate more closely with these events than do those of NCAM (Akitaya and Bronner-Fraser, 1992). However, aberrant NCAM expression has been documented in the neural crest of "splotch neural tube defect" mouse embryos, in which deficiencies of neural crest cell derivatives are observed (Moase and Trasler, 1991), implying that NCAM does play a critical role in neurulation and neural crest cell migration.

Initially, expression of NCAM was believed to be restricted to the nervous system (Brackenbury et al., 1977). However, it was later discovered to be expressed by certain glial cells, myoblasts, mature skeletal muscle tissue on denervation or paralysis (Daniloff et al., 1986; Schubert et al., 1989) and at neuromuscular synapses (Covault and Sanes, 1985). While NCAM expression is largely expressed in the central and peripheral nervous systems in the adult vertebrate, recent findings have established its expression further afield. Its presence has been noted in cardiac muscle
(Gordon et al., 1990), foetal kidney (Roth et al., 1988), chicken osteoblasts (Lee and Chuong, 1992) certain thymocytes in the thymus of embryonic and newborn mice (Brunet et al., 1989), murine endocrine cells (Langley et al., 1989; Mayerhofer et al., 1991, 1992a, 1992b), natural killer cells (Husman et al., 1989; Lanier et al., 1989), a subset of bone marrow stromal cells (F. Gibson, personal communication) and infiltrating T-lymphocytes in premalignant cervical lesions (M. Stanley, personal communication). The functional significance of NCAM in most of these cases has yet to be clearly defined.

Such widespread expression is consistent with the fact that NCAM is a primary adhesion molecule. Differential expression of NCAM is achieved by three means: rearrangement of sequences in mRNA (by alternative splicing and differential selection of poly-A sites), post-translational modification, and transcriptional control of the NCAM gene.

(b) Genomic organisation and transcriptional regulation of NCAM

In the human, NCAM is found on chromosome 11, q23. It spans over 70 kb, and comprises at least 25 exons, some of which are alternatively-spliced. Analysis of the 5' promoter region of NCAM in mice (Hirsch et al., 1991) and rats (Chen et al., 1990) suggests that alternative splicing is regulated independently of promoter choice. Transfection assays using hybrid promoter-reporter gene constructs and DNase footprinting suggested that this region allows for intricate control of transcription regulation (Hirsch et al., 1990; Mann et al., 1990). Within 840 base pairs upstream of the main transcription initiation site lie regions which either promote or repress initiation (depending on their context) and eight domains which interact with nuclear proteins. Sequences within two of these domains conform to regulatory elements in other promoters. One of these is the ATTA motif, typical of elements which bind homeobox
proteins; its presence is hardly surprising considering the major role NCAM would appear to play in the early events of embryogenesis.

Similarly, the fact that nerve growth factor (Prentice et al., 1990), osteogenic protein 1 (Perides et al., 1992), transforming growth factor β1 (TGFβ1) (Roubin et al., 1990) and thyroxine (Levi et al., 1990) have been demonstrated to up-regulate NCAM expression is consistent with their roles in morphogenesis or metamorphosis. Such regulation is likely to be effected by trans-interactions with the promoter region. In fact transfection studies similar to those used in the characterisation of the NCAM 5' promoter region suggest that this is the case where TGFβ1 is concerned (Roubin et al., 1990).

(c) NCAM isoforms: composition and characteristics

Initial investigations into the genetic basis of NCAM expression demonstrated that the three major isoforms expressed in chicken neural tissue, of 120, 140 and 180 kDa, (NCAM-120, NCAM-140 and NCAM-180 respectively) are generated by alternative splicing (Owens et al., 1987). The way in which this is achieved is illustrated in Figures 1.3 and 1.4. They have different carboxy-terminal regions; the transmembrane NCAM-180 and -140 polypeptides having longer and shorter cytoplasmic domains respectively. NCAM-120 is attached to the lipid bilayer via a glycosylphosphatidylinositol (GPI) anchor (Hemperly et al., 1986a), presumably recruited by a lipid attachment site encoded by exon 15. The polypeptides generated in this manner share a similar amino-terminal region, comprising five immunoglobulin domains which probably present as disulphide-bonded loops, and two fibronectin type III repeats.

Frei et al. (1992) engineered proteins expressing individual domains and tandem sets of domains to investigate the possibility that these might subserve different functional roles. On assaying these proteins for their capacity in promoting adhesion, spreading, and neurite outgrowth of
Figure 1.3

Structure and alternative splicing patterns of (a) NCAM RNA and (b) corresponding protein

(a) NCAM RNA
The exons of NCAM which are alternatively spliced can be classified as belonging to one of two groups:
1) those which determine whether the molecule is attached to the plasma membrane (SEC) or determines the length of the cytoplasmic tail (exons 15 and 18)
2) those found within the "extracellular" part of the molecule (MSD1 domain and VASE).
Alternative splicing patterns are indicated by the inverted "V"s, with the exception of the MSD1 domain; the different combinations of its constitutive exons are summarised in Table 5.1.

Key:
- ☐ = sequence encoding hydrophobic amino acids
- Δ = alternatively-spliced exon
- AAA = polyadenylation site
- ⬤ = stop codon

(b) NCAM protein
NCAM-140 is illustrated. This isoform lacks the alternatively-spliced exons 15 and 18.

Key:
- ☐ = protein encoded by constitutively-spliced exon
- ☐☐ = fibronectin type III-like repeat
- ■ = protein encoded by alternatively-spliced exon and affecting means of membrane attachment or length of cytoplasmic tail
- ☐☐ = protein encoded by MSD1 exons
- □ = protein encoded by VASE
- ☐☐ = protein encoded by MSD1 exons
- ☐☐ = plasma membrane
- ★ = potential site of N-glycosylation
Figure 1.3

a) 

b)
**Figure 1.4**

**Approximate topology of NCAM isoforms generated by alternative splicing**

NCAM-180, NCAM-140, NCAM-120 and secreted NCAM are illustrated. Electron microscopy studies of NCAMs expressed by chicken brain reveals that these molecules comprise two linear segments joined at an angle (Hall and Rutishauser, 1987).

**Key:**

- [ ] = protein encoded by constitutively-spliced exon
- [■] = protein encoded by alternatively-spliced exon as in Figure 1.3 b
- [■■] = protein encoded by MSD1 exons
- [■■] = plasma membrane
- [○] = glycosyl phosphatidylinositol anchor
Figure 1.4

NCAM - 180
NCAM-140
NCAM-120
secreted
NCAM
cerebellar neurons they found their hypothesis to be correct (see Table 1.3). Epitope mapping, and binding studies using polypeptides and deletion mutants have recently demonstrated that the homophilic binding site is situated within the third immunoglobulin domain (Rao et al., 1992), in apparent contradiction with results obtained by Frei et al. It is difficult to reconcile these contradictory results. It is possible that adhesion of murine cerebellar neurons to NCAM and its derivatives differs mechanistically from that between NCAM transfectants.

As a consequence of its glycolipid anchor, NCAM-120 is thought to have a heightened lateral mobility within the plasma membrane. For molecules such as NCAM, which appear to be involved in guiding direction of growth, a mobile population of isoforms may assist rapid adhesion to pre-existing neurons. The GPI-anchor has also been shown to permit release of NCAM-120 on cultured rat glial cells into the extracellular environment (He et al., 1987). Such behaviour, however, cannot be said to be a necessarily constitutive property of GPI-anchors. Release may be triggered by endogenous phosphatidylinositol phospholipase C (PIPLC) (Low, 1987; Fouchier et al., 1990), an enzyme which cleaves the inositol moiety from hydrophobic diacylglycerol. The latter two appear to be important second messengers for stimulation of cell growth (Berridge and Irvine, 1984; Nishizuka, 1986).

It has been proposed that NCAM-120 may be used by the cell primarily to communicate information abroad, i.e. to other cells, while the larger transmembrane isoforms also provide an informative service to the cell on which they are found (Doherty and Walsh, 1991). Evidence from experiments exploring the dynamic relationship between cleavage of GPI-anchored proteins and NCAM synthesis indirectly supports this proposal (Theveniau et al., 1991). The GPI-anchor may elevate the molecule higher above the membrane glycocalyx (P. Doherty, personal communication)
Table 1.3

Capacities of the domains and tandem domains comprising the neural cell adhesion molecule
(an adaptation of a table presented by Frei et al., 1992)

**Key:**

- Ig = immunoglobulin-like domain
- Fn = fibronectin type III-like domain
Table 1.3

<table>
<thead>
<tr>
<th>Domain</th>
<th>Cell Adhesion</th>
<th>Neurite Outgrowth</th>
<th>Cell Spreading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ig I</td>
<td>+++</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>Ig II</td>
<td>++</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Ig III</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ig IV</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ig V</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ig I-V</td>
<td>-</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>Fn I-II</td>
<td>+</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>Brain NCAM</td>
<td>-</td>
<td>+++</td>
<td>++</td>
</tr>
</tbody>
</table>
thus exposing or re-orienting critical functional regions (e.g. the fibronectin type III repeats) to the extracellular milieu.

Transfection experiments have demonstrated that the GPI-linked NCAM-120 isoform is directed to the apical surface of polarised epithelial cells, while transmembrane isoforms are expressed on the basolateral surface (Powell et al., 1991). Thus, it appears that alternative splicing of the GPI anchor may influence the intracellular trafficking of NCAM, although isoform-specific glycosylation has not been excluded as a critical signal. The alternatively-spliced variant of fibronectin containing ED-A is also found to localise preferentially to the apical surface of polarised epithelial cells; that it continues to do so (albeit in smaller quantities) in the presence of tunicamycin (Wang et al., 1991) suggests that glycosylation does not inform trafficking machinery in this context.

It is interesting to note that serum starvation has been found to increase both the number and intensity of a subset of GPI-linked proteins in insulin-sensitive cell types. Reapplication of serum or insulin induces acute loss of these proteins (Lisanti et al., 1989), illustrating how the environment can affect the status of GPI-linked proteins, and how the GPI-anchor facilitates rapid release of proteins. Similarly, elevation of GPI-turnover by either growth factor stimulation or transformation have been reported - the latter in many cases (see Kaplan et al., 1987). This may be related to the increase of autoantibodies which bind phosphatidylinositol in the sera of patients with epithelial tumours (Faiderbe et al., 1993).

Little is known about how the NCAM isoforms might differ in function and how the cytoplasmic domains interact with the cytoplasm. Recent work suggests that intracellular signalling by L1, N-cadherin and the transmembrane forms of NCAM during neurite outgrowth is mediated by G protein-dependent activation of Ca\(^{2+}\) channels. Monolayers of fibroblasts transfected with either transmembrane NCAM, N-cadherin, or
L1 induce neural differentiation in PC12 cells, but not in the presence of pertussis toxin or calcium channel blocking agents (Doherty et al., 1991; Williams et al., 1992). Corroborating this is data obtained by Frei et al. (1992) - intracellular calcium, pH, and inositol phosphate 1 are all increased on addition of soluble NCAM to cultures of rat cerebellar neurons.

It has recently been demonstrated that both NCAM-180 and N-cadherin are substrates for (an) endogenous calcium-activated protease(s), the action of which probably disrupts cytoskeletal binding (Covault et al., 1991). Thus, a self-regulating mechanism may exist whereby the activation of Ca\(^{2+}\) channels by NCAM and N-cadherin may influence the extent to which the latter "communicate" with cytoskeleton. The two molecules may differ in their activation of pp60\(^{c-src}\)-dependent phosphorylation of tubulin: antibodies to NCAM trigger this reaction, while those to N-cadherin do not (Atashi et al., 1992) although negative results from antibody perturbation studies can be inconclusive.

An intriguing link exists between G protein activation and the phosphatidylinositol (PI) cycle. One of the major signal transduction agents stimulated by activated G proteins is phospholipase C, which hydrolyses phosphatidylinositol anchors giving rise to diacylglycerol and inositol-1,4,5-trisphosphate (IP\(_3\)). It remains to be established whether the GPI-linked NCAM isoform is affected in this manner. Transmembrane isoforms in both neural and muscular tissue have been shown to be phosphorylated (Sorkin et al., 1984; Moore et al., 1987; Ramos and Ellis, 1989), consistent with the finding that the kinase inhibitor K-252b blocks differentiation induced by NCAM transfectants (Doherty et al., 1991).

NCAM-180 can associate with the cytoskeleton-membrane linker protein spectrin, and has a reduced lateral mobility compared with NCAM-140 (Pollerberg et al., 1986). A possible consequence is the finding that neurite outgrowth on cells transfected with NCAM-180 is less than that on
NCAM-140 transfectants (Doherty et al., 1992c), although it remains a possibility that the extended cytoplasmic domain of NCAM-180 may influence the conformation of the external portion of the molecule in such a way as to modulate binding characteristics and hence its ability to support neurite outgrowth.

That cytoplasmic domains influence extracellular binding is illustrated by a set of experiments which investigate the binding of E-cadherin. Transfectants bearing E-cadherin were observed to bind each other, in preference to those bearing chimeric molecules comprising the cytoplasmic domain of NCAM-140 and the extracellular domain of E-cadherin (Jaffe et al., 1990). In addition, the propensity for homophilic binding of E-cadherin is eliminated by deleting of a small region in the cytoplasmic domain (Nagafuchi and Takeichi, 1988). Such cytoplasmic integrity is obviously not requisite amongst all the cadherins, as is demonstrated by the recent discovery of GPI-linked T-cadherin which exhibits homophilic binding despite its lack of transmembrane and cytoplasmic domains.

Isoform expression in the chicken is temporally regulated: NCAM-140 is expressed during very early stages of embryogenesis, while NCAM-180 appears during neurogenesis (Murray et al., 1986) and NCAM-120 after hatching. This is similar to isoform expression in humans, where expression of NCAM-140 is followed by the appearance of NCAM-180, which is described as being restricted to the CNS (Edelman and Chuong, 1982; Keilhauer et al., 1985; Prieto et al., 1989; Persohn and Schachner, 1990) and down-regulated with development. Conversely, the expression of NCAM-120 increases, while that of NCAM-140 remains level. Immunolocalisation of NCAM in differentiated mouse neuroblastoma cells demonstrates a concentration of NCAM-180 at points of contact between adjacent cells whereas NCAM-140 maintains a more diffuse distribution in
the plasma membrane (Pollerberg et al., 1986, 1987). Consistent with this is the finding that NCAM-180 is restricted to post-synaptic membranes in the mouse cerebella, while NCAM-140 is found on both pre- and post-synaptic membranes (Persohn et al., 1989).

Myoblasts in culture predominantly express a transmembrane isoform of 145 kDa, which is highly homologous to the 140 kDa isoform found in neural tissue (Dickson et al., 1987). They also express a minor 120 kDa isoform which is assumed to be GPI-linked as it does not label with phosphorous-32. It has recently been discovered that fusion-competent myoblasts, if treated with PIPLC, do not fuse to form myotubes (Knudsen et al., 1989) suggesting that GPI-linked isoforms might play a critical role in myoblast adhesion or another process involved in myotube formation. It cannot be ruled out that (an)other GPI-linked molecule(s) is/are responsible for this observation. That NCAM plays a role in this event is supported by the fact that anti-NCAM antibodies inhibit the rate of myotube formation (Knudsen et al., 1990) while transfection with a gene construct encoding NCAM enhances it (Dickson et al., 1990). N-cadherin however, may play a more critical role - its expression correlates more closely with myoblast fusion, and perturbation of N-cadherin function affects the final extent of myotube formation (Mége et al., 1992).

Upon fusion (or just prior to fusion) of myoblasts, new GPI-linked NCAM isoforms of 125 and 155 kDa are expressed, with down-regulation of the transmembrane isoform. The expression of NCAM isoforms on cardiac muscle follows a similar trend (Wharton et al., 1989). And so crudely speaking, muscle assumes a molecular identity separate from that of neural tissue with respect to NCAM expressssion.

(d) NCAM and the extracellular matrix

Experiments suggest that NCAM participates in heterophilic interactions with the extracellular matrix. For instance, culture supernatant
of a teratocarcinoma cell line promotes neurite outgrowth of cerebellar neurons, but not in the presence of antibodies to NCAM (and L1) (Werz and Schachner, 1988). The same supernatant, coated on latex beads, was later demonstrated to induce accumulation of NCAM-180 on the cell bodies, neurites, and growth cones of a neuroblastoma cell line. In contrast, NCAM- and L1-coated beads evoked no accumulation, suggesting it unlikely that extracellular NCAM (and L1) plays an active role in inducing accumulation (Pollerberg et al., 1990a).

The best-characterised interaction between NCAM and the extracellular matrix is that between a region comprising 17 basic amino acids in immunoglobulin-like domain 2, and heparin or heparan sulphate (Cole and Akeson, 1989). Synthetic peptides comprising this domain inhibit retinal cell adhesion to an NCAM substrate, but those in which some amino acid residues are altered do not (Reyes et al., 1990). Other experiments also suggest that a heparin derivative stabilises NCAM-mediated cell-cell interaction (Cole et al., 1986). This finding suggests a possible indirect relationship between NCAM and fibronectin; the heparin-binding domain of fibronectin promotes neurite outgrowth (Perris et al., 1989). That NCAM binds heparan sulphate is consistent with its role in development; antibody studies suggest that heparan sulfate is prevalent at sites of active morphogenesis, and its expression is developmentally regulated (David et al., 1992).

Experiments using a 110-kDa amino-terminal fragment of NCAM (NCAM-110) arising from protease digestion suggest that NCAM binds collagen types I-VI and IX, and that such binding may be inhibited by heparin and chondroitin sulphate (Probstmeier et al., 1989). This finding contrasts with earlier work in which a teratocarcinoma cell line expressing NCAM failed to bind collagen IV (Werz and Schachner, 1988). Post-translational modifications may account for this difference. Additional
work suggests that NCAM-110 may influence fibrillogenesis by interacting with collagen I (Probstmeier et al., 1992).

That NCAM has been shown to bind components of the extracellular matrix implies that NCAM has potential not only in mediating cell-extracellular matrix interaction, but also as an active component of the extracellular matrix. Supporting this theory is the finding that NCAM is found in the ECM of peripheral nerves (Martini and Schachner, 1988; Rieger et al., 1988), surrounding blood vessels (Gulbenkian et al., 1989) and in interstitial spaces near synaptic sites after denervation of rat skeletal muscle (Sanes et al., 1986). Evidence also suggests that substratum-bound NCAM may bind a ligand on mouse neuroblastoma cells via a heterophilic mechanism (Murray and Jensen, 1992).

(c) Post-translational modifications of NCAM

Post-translational modifications of NCAM include sulphation, phosphorylation, glycosylation and sialylation. Six sites which are potentially N-glycosylated have been identified (See Figure 1.3). Knowledge of NCAM carbohydrate composition is scant, but evidence for oligosaccharide microheterogeneity exists, implying influence on function.

The sulphation site best characterised is a glycosyl moiety expressed on a subset of neural NCAM molecules excluding NCAM-120, and recognised by monoclonal antibody HNK-1 (Kruse et al., 1984; Sanes et al., 1986). HNK-1 partially inhibits NCAM-mediated adhesion; appropriately, its binding site has been localised to the third Ig-like domain (Cole and Schachner, 1987). Its epitope, a sulphated carbohydrate, was first characterised on human natural killer cells and is also expressed by other adhesion molecules, including L1. Molecular modelling (Santoni et al., 1989a) and adhesion assays (Riopelle et al., 1986; Kunemund et al., 1988) suggest that it plays a functional role in NCAM binding. Other antibodies have also been found to recognise immunologically distinct subsets of
NCAM molecules in addition to those of other antigens (Patel et al., 1989a; Naegele and Barnstaple, 1991; Williams et al., 1985).

Another carbohydrate which is narrowly restricted to a subset of NCAMs was discovered on characterisation of a set of antibodies raised against olfactory epithelium of rats and frogs (Key and Akeson, 1990). Sensory olfactory neurons are unusual in that they undergo continual growth, death, and regeneration in the adult vertebrate and it is thus of interest that they should bear a unique carbohydrate epitope. Evidence suggests that they may retain the highly polysialated "embryonic" NCAM (E-NCAM) throughout adulthood, unlike most other neurons which undergo desialylation with age.

Where biochemistry is concerned, the most unusual feature of NCAM, and that which has a profound influence on cell adhesion and probably tumour biology, is the attachment of long α2-8 linked chains of sialic acid to glycosyl moieties. Indeed, polysialic acid chains of at least 55 residues have been detected on what is believed to be NCAM on neuroblastoma (Livingstone et al., 1988; Finne et al., 1983). These are thought to be N-linked to one of three residues on Ig-like domain 5. The occurrence of N-linked homopolymers of α2-8 linked sialic acid is very rare in the mammalian cell; there is only one other known example (Zuber et al., 1992). Most N-linked carbohydrates as a rule, only contain sialic acid residues as terminal, non-reducing sugars attached by an α2-3 or α2-6 linkage to the galactosyl residues of bi-, tri- or tetra-antennary sugar chains (see Figure 1.5). These linkages also occur between sialic acid and glycosphingolipids (rendering them gangliosides).

Due to heterogeneity of polysialic acid chain length, E-NCAM presents a broad smear from approximately 140 to over 250 kDa on analysis by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). As α2-8 linked PSA is also found at lower levels on the α subunit of sodium channels in the rat brain (Zuber et al., 1992), the presence of PSA might not necessarily imply the presence of NCAM, although it would appear that NCAM is its major carrier.

Sialic acid content of NCAM is subject to marked developmental regulation; it constitutes 30% by weight of embryonic NCAM decreasing to 10% in adult NCAM. Furthermore, this conversion has been shown to
Figure 1.5

**Hypothetical structure of the NCAM N-linked glycosylation unit**

N-acetyl neuraminic (sialic acid) residues are terminally linked to galactosamine residues via $\alpha2$-$3$ or $\alpha2$-$6$ linkages, or are linked in long chains via $\alpha2$-$8$ linkages. The exoneuraminidase used in the experiments described in Chapter 3, sequentially cleaves residues linked by all of these bonds. In contrast, Endo N, which has also been used in the characterisation of NCAM by others (e.g. Vimr *et al.*, 1989) cleaves $\alpha2$-$8$ linkage specifically, with a requisite of residues proximal to the cleavage site - indicated by an asterisk. This diagram is an adaptation of one presented by Regan *et al.*, 1991.

**Key:**

- Asn = asparagine residue
- GlcNAc = N-acetyl glucosamine residue
- Man = mannose residue
- Gal = galactose residue
- Fuc = fucose residue
- Sia = N-acetyl neuraminic (sialic) acid residue
effect an increase in strength of homophilic binding (Hoffman and Edelman, 1983; Sadoul et al., 1983) and is necessary for normal neural development (Edelman and Chuong, 1982). Corroborating the proposed inverse relationship between degree of adhesiveness and extent of polysialylation is the finding that mutants of group B streptococcal \( N. \) meningitidis which lack \( \alpha 2-8 \) linked polysialic acid, adhere up to three times more strongly to target epithelial cells than the PSA-bearing parent strain (Swartley and Spellman, 1992). In this context, it is interesting to note that IgM antibodies directed against E-NCAM, have been found in patients suffering from group B streptococcal meningitis - perhaps a causative factor of the neurologic sequelae which occur in over 50% of people who survive the infection (Nedelec et al., 1990).

Contrasting with these findings, however, is the negligible increase in homophilic binding effected by removal of PSA from purified NCAM (Hall and Rutishauser, 1987; Hall et al., 1990). Resolving these apparently contradictory conclusions are two recent findings which suggest that the steric effects of PSA on NCAM are such as to have a potent effect on the apposition of cell surfaces and intercellular space, thereby influencing other cell-cell interactions (Acheson et al., 1991; Yang et al., 1992). Thus, more PSA results in a greater distance between cell membranes, and fewer opportunities for intermolecular contact.

An immunohistochemical study of polysialylated NCAM during postnatal development and in the adult rat brain supports the theory that polysialic residues may contribute to neuronal plasticity (Aaron and Chesselet, 1989; Theodosis et al., 1991). Also consistent with this proposal is the fact that neurite outgrowth is dependent on the presence of PSA on neuronal NCAM (Doherty et al., 1992a).

It is seen that biochemical modifications alone provide great potential for antigenic modulation. Alternative splicing and association
with other molecules (cytoplasmic, extracellular, and within the plane of the plasma membrane) provide additional means of heterogeneity. NCAM is also a primary adhesion molecule and subject to developmental and spatial regulation. These features are consistent with the fact that tissues which express and interact with NCAM during dynamic stages of development - e.g. rat hippocampal neurons - exhibit developmental changes in their responsiveness to it. That these neurons possess a concomitant and converse trend in their sensitivity to N-cadherin illustrates how dynamic events correlate with multiple trends in molecular sensitivity (Doherty, et al., 1992d).

(f) **Alternative splicing of NCAM in the extracellular domain**

Alternative splicing of exons 15 and 18 result in modifications of NCAM which affect the gross structure of the C-terminal domain. Examination of NCAM mRNA sequences in muscle and neural tissues have revealed two splicing patterns in the extracellular part of the molecule, one of which appears to be tissue-specific. Sequencing of human muscle cDNA clones demonstrated that a novel sequence of 108 base pairs (bp) is expressed between exons 12 and 13 of mRNAs encoding GPI-linked isoforms (Dickson et al., 1987). The sequence is found in GPI-linked isoforms of muscle tissue only, and hence called the muscle-specific domain 1 (MSD1; see chapter 5). Functional significance has yet to be proven. However the position of this splice site - between the two fibronectin type III repeats - implies that it may modulate cell interactions. A transfectant bearing the two repeats in tandem is a more effective substrate for neuronal adhesion than those bearing single fibronectin domains (Frei et al., 1992); the imposition of MSD1 between them may alter their conformation and biological activity. It must be noted however, that transfectants bearing entire NCAM molecules both with and without the MSD1 domain show no significant difference in inter-transfectant
adhesion (Rowett et al., 1990).

A subpopulation of mRNAs in rat brain encoding NCAM-120, -140, and -180 was found to include an alternative splicing site between exons 7 and 8, in which a 30 bp exon (VASE - an acronym for Variable Alternatively-Spliced Exon) had been inserted (Small et al., 1988). The inclusion of this exon was found to be more prevalent with development in the central nervous system and also spatially controlled (Small and Akeson, 1990). The hydrophilic nature of the encoded amino acid sequence and its position in Ig-like domain 4, suggest that it may be involved in intermolecular interaction. Indeed, transfection studies have demonstrated that NCAM-140 incorporating VASE has a reduced capacity for promoting neurite outgrowth (Doherty et al., 1992b).

Yet another exon, SEC, found immediately 3' of MSD1 in muscle, was found to be alternatively spliced in cDNA derived from a skeletal muscle mRNA. Compelling evidence suggests that its protein product - a truncated protein generated by an in-frame stop codon and no potential for insertion into the plasma membrane - is secreted into the extracellular environment (Gower et al., 1988). It is thus an excellent candidate as a source of extracellular NCAM; indeed, it has been hypothesised that the SEC-encoded polypeptide may anchor the molecule to the extracellular matrix (G. Dickson; personal communication).

Soluble NCAM has been detected in cerebro-spinal fluid (CSF) and plasma in humans (Jørgensen, 1988; Plioplys et al., 1989) and rats (Dalseg et al., 1989; Krog et al., 1992) and in the supernatant of cultured brain cells (Nybroe et al., 1989), suggesting that SEC may also be expressed in some neural cells. However, the origin of soluble isoforms is difficult to ascertain until it is possible to differentiate soluble isoforms generated by either proteolysis, cleavage of the GPI-anchor or inclusion of the SEC sequence.
At the beginning of this project, analysis of NCAM expression and splicing by tumour cells was limited. The last couple of years have witnessed a marked acceleration of papers published in this field. Existing data suggests that some tumours of neuroectodermal origin, e.g. Wilms' tumour (Roth et al., 1988; Zuber and Roth, 1990), pheochromocytoma (Prentice et al., 1990), melanoma (Linnemann et al., 1989a; Pandolfi et al., 1992), small cell lung carcinoma (Aletsee-Ufrecht, 1990a; Moolenaar et al., 1990) and neuroblastoma (Figarella-Branger et al., 1990) express heavily sialylated NCAM isoforms. NCAM is also found on tumours of mesodermal origin - e.g. rhabdomyosarcoma (Mechtersheimer et al., 1991) and myeloid leukaemia (Lanier et al., 1991; Barker et al., 1992). (For further discussion, see Chapter 3).

In light of the fact that NCAM appears to be a good marker for both neuroblastoma and rhabdomyosarcoma (i.e. it is present in a large majority of tumours analysed) it follows that tissue-specific forms of NCAM may be of value in the differential diagnosis of these malignancies. The aim of this thesis is to test the prediction that tissue-specific expression of NCAM is maintained in tumours, as a preliminary approach to their identification. Initial characterisation of NCAM focused on its protein expression in neuroblastoma and rhabdomyosarcoma. Further investigations elucidated patterns of alternative splicing patterns which indeed appear to exhibit a consistent difference between the two tumours.
# Chapter 2

## Materials and General Methods

### 2.1 Materials

<table>
<thead>
<tr>
<th>Chemical Reagent</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetic Acid</td>
<td>British Drugs Houses (BDH)</td>
</tr>
<tr>
<td>Acrylamide</td>
<td>BioRad Laboratories</td>
</tr>
<tr>
<td>Agarose</td>
<td>BioRad</td>
</tr>
<tr>
<td>Ammonium acetate</td>
<td>Sigma Chemical Company, Ltd.</td>
</tr>
<tr>
<td>Ammonium persulphate (APS)</td>
<td>BioRad</td>
</tr>
<tr>
<td>β-mercaptoethanol (β-ME)</td>
<td>Sigma</td>
</tr>
<tr>
<td>Bisacrylamide</td>
<td>BioRad</td>
</tr>
<tr>
<td>(N'N-methylene bisacrylamide)</td>
<td></td>
</tr>
<tr>
<td>Boric acid</td>
<td>Fisons, PLC</td>
</tr>
<tr>
<td>Bovine serum albumin (BSA)</td>
<td>Sigma</td>
</tr>
<tr>
<td>Bromophenol Blue (bpb)</td>
<td>Sigma</td>
</tr>
<tr>
<td>Butan-1-ol</td>
<td>Fisons, PLC</td>
</tr>
<tr>
<td>Calcium chloride</td>
<td>BDH</td>
</tr>
<tr>
<td>Chloroform</td>
<td>Fisons, PLC</td>
</tr>
<tr>
<td>Coomassie Brilliant Blue</td>
<td>Pharmacia LKB Biotechnology (LKB)</td>
</tr>
<tr>
<td>Cyanol blue</td>
<td>Sigma</td>
</tr>
<tr>
<td>Denatured salmon sperm DNA</td>
<td>Sigma</td>
</tr>
<tr>
<td>Dextran sulphate</td>
<td>Pharmacia</td>
</tr>
<tr>
<td>Dimethyl formamide</td>
<td>BDH</td>
</tr>
</tbody>
</table>
Dimethyl sulphoxide (DMSO) Fisons, PLC
Dithiothreitol (DTT) Sigma
Ethanol Hayman, Ltd
Ethylene diamine tetra-acetate (EDTA) BDH
Ethidium bromide Sigma
Fast Red TR Salt (5-chloro-2-toluene diazonium chloride hemi[zinc chloride]) Sigma
Foetal Calf Serum Gibco Laboratories
Formamide BDH
Glacial Acetic Acid Fisons, PLC
Glucose BDH
L-Glutamine Flow Laboratories
Glycerol Fisons, PLC
Glycine BDH
Guanidium Isothiocyanate Sigma
Hepes Sigma
Human serum albumin (HSA) Sigma
Isoamyl alcohol Fisons, PLC
Kanamycin Boehringer Mannheim
Leupeptin Sigma
Levamisole (2,3,5,6-tetrahydro-6-phenyl-imidaso (2,1-6)triazole) Sigma
Lithium chloride Sigma
Magnesium acetate Sigma
Magnesium choride Sigma
Marvel milk powder Premier Brands
Mercuric choride Sigma
Methanol BDH
Naphthol AS-MX phosphate Sigma
Penicillin Flow Laboratories
Phenol Rathburn Chemicals
Phenyl-methyl-sulphonyl-fluoride (PMSF) Sigma
Potassium acetate Sigma
Potassium chloride Sigma
Potassium phosphate BDH
Sodium acetate BDH
Sodium chloride BDH
Sodium citrate Fisons, PLC
Sodium hydroxide BDH
Sodium dodecyl sulphate (SDS) BDH
Sodium phosphate BDH
Spermidine Sigma
Streptomycin Flow Laboratories
Sucrose Fisons, PLC
TEMED (NNN'N'-tetra methyl ethylene diamine) BioRad
Tris (Trishydroxy-methylamine) Sigma
Urea BDH

**Immunological Reagents**

Alkaline phosphatase-conjugated F(ab') fragment of rabbit Ig to mouse Ig Dakopatts (a/s)

Fluorescein-conjugated F(ab') fragment of rabbit Ig to mouse Ig Dakopatts (a/s)
Monoclonal Antibodies
ERIC-1
M340
UJ13A
anti-PSA

Genetic cDNA Clones
λ9.5
N1

Molecular Biology Reagents
Avian myeloid leukemia virus reverse transcriptase (AMV-RT)
DNA dipstick™ assay kit
DNase-free RNase
Deoxynucleotides (dNTPs)
Deoxyxynucleotides (ddNTPs)
γ-32P dATP
α-32P dCTP
Isogene kit
Klenow fragment
1 kb molecular weight markers
RNase-free DNase
RNAguard
Random hexadeoxynucleotides [d(N)₆]
Restriction enzyme buffer (x10)

Bourne et al., 1991
Bourne et al., 1989
Patel et al., 1989b
Rougon et al., 1982
Dickson et al., 1987
Barthels et al., 1988

Northumbria Biologicals Ltd. (NBL)
Invitrogen Corporation
Boehringer Mannheim
Pharmacia
Amersham International
Amersham
Invitrogen
Pharmacia
Bethesda Research Laboratories
Boehringer Mannheim
Pharmacia
NBL
\(^{35}\text{S} \text{DNA markers} \quad \text{Amersham} \\
\text{Sequenase-2 & Sequencing kit} \quad \text{United States Biochemicals Corporation (USB)} \\
\text{TA cloning}^{\text{TM}} \text{ kit} \quad \text{Invitrogen} \\
\text{T. aquaticus (Taq) polymerase} \quad \text{NBL} \\
\alpha^{35}\text{S} \text{ dATP} \quad \text{Amersham} \\
\text{T4 polynucleotide kinase} \quad \text{NBL} \\
\text{X-galactosidase (5-bromo-4-chloro-3-indolyl-} \\
\beta\text{-D-galactopyranoside)} \quad \text{Sigma}

\textbf{Culture Media}

\text{Complete RPMI 1640} \quad \text{Gibco;} \\
\text{(Roswell Park Memorial Institute)} \quad \text{supplemented with} \\
10\% \text{ heat-inactivated FCS} \\
2\text{mM L-glutamine} \\
100 \text{ IU/ml penicillin} \\
100 \text{ mg/ml streptomycin} \\
\text{LB (Luria-Bertani) Medium} \quad 1\% \text{ bacto-tryptone} \\
\text{(supplied by ICRF)} \quad 0.5\% \text{ bacto-yeast extract} \\
\text{} \quad 1\% \text{ NaCl} \\
\text{SOC medium} \quad \text{Invitrogen;} \\
\text{} \quad 2\% \text{ bacto-tryptone} \\
\text{} \quad 0.5\% \text{ bacto-yeast extract} \\
\text{} \quad 10\text{mM NaCl} \\
\text{} \quad 2.5\text{mM KCl} \\
\text{} \quad 10\text{mM MgCl}_2 \\
\text{} \quad 10\text{mM MgSO}_4 \\
\text{} \quad 20\text{mM glucose}

\textbf{Tumour Cell Lines}

<table>
<thead>
<tr>
<th>Cell Line</th>
<th>Tumour Type</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>GH1</td>
<td>T-ALL</td>
<td>Minowada \textit{et al.}, 1972</td>
</tr>
<tr>
<td>GOTO</td>
<td>NBL</td>
<td>Sekiguchi, \textit{et al.}, 1979</td>
</tr>
<tr>
<td>IMR-32</td>
<td>NBL</td>
<td>Tumilowicz \textit{et al.}, 1970</td>
</tr>
<tr>
<td>JR-1</td>
<td>RMS</td>
<td>Clayton \textit{et al.}, 1986</td>
</tr>
<tr>
<td>Kg1a</td>
<td>myeloid leukaemia</td>
<td>Lanier \textit{et al.}, 1991</td>
</tr>
</tbody>
</table>
Kelly NBL Schwab et al., 1983
NB1 NBL Imashuku et al., 1973
NB100 neuroepithelioma Schlesinger et al., 1976
PCF NBL Sugimoto, unpublished
RD618 RMS McAllister et al., 1969
Rhab-1 RMS Wheldon, unpublished
SK-N-BE (2C) NBL Ciccarone et al., 1989
SK-N-SH NBL Montgomery et al., 1983
SK-N-DZ NBL Beidler et al., 1973

**Miscellaneous**

Catering Clingfilm Amersham
Heavy white mineral oil Sigma
Hybridisation mesh DuPont
Hypodermic needles Sabre
Inoculation loops Sterilin (U.K.) Ltd.
Micropore filters Gelman Sciences
Molecular weight Rainbow Markers Amersham
NAP-5 column Pharmacia
Neuraminidase, type X Sigma
Nick TE column Pharmacia
Nitrocellulose sheets, 0.2mM Schleicher & Schuell
Nylon membrane “GeneScreen Plus” DuPont
Plastic disposable tissue culture materials Falcon
Syringe barrels Falcon
X-ray film Fuji
## Buffers and Solutions

<table>
<thead>
<tr>
<th>Buffer Type</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acrylamide separating gel (DNA)</td>
<td>26.6 ml 30% acrylamide</td>
</tr>
<tr>
<td></td>
<td>10 ml x10 TBE</td>
</tr>
<tr>
<td></td>
<td>62.7 ml ddH₂O</td>
</tr>
<tr>
<td></td>
<td>0.7 ml 10% APS</td>
</tr>
<tr>
<td></td>
<td>35 µl TEMED</td>
</tr>
<tr>
<td>Acrylamide separating gel (protein)</td>
<td>7 ml 30% acrylamide</td>
</tr>
<tr>
<td></td>
<td>11.2 ml 1M Tris, pH8.8</td>
</tr>
<tr>
<td></td>
<td>11.5 ml ddH₂O</td>
</tr>
<tr>
<td></td>
<td>0.3 ml 10% SDS</td>
</tr>
<tr>
<td></td>
<td>0.1 ml 10% APS</td>
</tr>
<tr>
<td></td>
<td>20 µl TEMED</td>
</tr>
<tr>
<td>Acrylamide stacking gel</td>
<td>1.7 ml 30% acrylamide</td>
</tr>
<tr>
<td></td>
<td>1.3 ml 1M Tris, pH6.8</td>
</tr>
<tr>
<td></td>
<td>7 ml ddH₂O</td>
</tr>
<tr>
<td></td>
<td>0.1 ml 10% SDS</td>
</tr>
<tr>
<td></td>
<td>50 µl APS</td>
</tr>
<tr>
<td></td>
<td>10 µl TEMED</td>
</tr>
<tr>
<td>Annealling mix</td>
<td>80mM Tris, pH7.5</td>
</tr>
<tr>
<td></td>
<td>0.5M MgCl₂</td>
</tr>
<tr>
<td></td>
<td>0.1M NaCl</td>
</tr>
<tr>
<td></td>
<td>20% DMSO</td>
</tr>
<tr>
<td>Cell lysate sample buffer</td>
<td>63mM Tris, pH6.8</td>
</tr>
<tr>
<td></td>
<td>12.5% glycerol</td>
</tr>
<tr>
<td></td>
<td>2.5mM PMSF</td>
</tr>
<tr>
<td></td>
<td>1.25 mg/ml leupeptin</td>
</tr>
<tr>
<td></td>
<td>1.25 mM EDTA</td>
</tr>
<tr>
<td></td>
<td>2% SDS</td>
</tr>
<tr>
<td>Chase mix</td>
<td>0.25mM dATP</td>
</tr>
<tr>
<td></td>
<td>0.25mM dCTP</td>
</tr>
<tr>
<td></td>
<td>0.25mM dGTP</td>
</tr>
<tr>
<td></td>
<td>0.25mM dTTP</td>
</tr>
<tr>
<td></td>
<td>50 mM NaCl</td>
</tr>
<tr>
<td></td>
<td>10% DMSO</td>
</tr>
<tr>
<td>Coomassie Brilliant Blue R250 staining solution</td>
<td>0.5g Coomassie</td>
</tr>
<tr>
<td></td>
<td>Brilliant Blue R250</td>
</tr>
<tr>
<td></td>
<td>40% methanol</td>
</tr>
<tr>
<td></td>
<td>7% acetic acid</td>
</tr>
<tr>
<td></td>
<td>53% ddH₂O</td>
</tr>
<tr>
<td>Destain solution: as for above, minus Coomassie Blue</td>
<td></td>
</tr>
</tbody>
</table>
Elution buffer

0.5M NH₄ acetate
10mM Mg acetate
1mM EDTA
0.1% SDS

Fast-Red substrate
(filtered through a 0.2 mM millipore filter prior to use)

48mM naphthol
AS-Mx phosphate
0.004% dimethylformamide
3.9mM Fast Red TR salt
98mM Tris, pH8.2

Gel fixing buffer

10% methanol
10% glacial acetic acid

Gel loading buffer (glb); type IV
(filtered through 0.2mM millipore filter prior to use)

0.25% bpb
40% sucrose

Hybridisation buffer

10% dextran sulphate
1M NaCl
1% SDS

x10 labelling buffer

0.5M Tris, pH6
1M MgCl₂
50mM dithiothreitol
1mM spermidine
1mM EDTA

Labelling Mix

2.5 mCi/ml ^35^SdATP
50mM DTT
0.65U/ml Sequenase 2

Oligo-labelling buffer

0.25M Tris, pH8
25mM MgCl₂
1M Hepes, pH6.6
0.1mM dATP
0.1mM dGTP
0.1mM dTTP
1.8% β-ME
27 U/ml d(N)₆

PBS (Phosphate buffered saline)

0.17M NaCl
3mM KCl
1mM Na₂HPO₄
1.8mM KH₂PO₄
pH7.4

RNA lysis buffer

5M guanidine isothiocyanate
50mM Tris
10mM EDTA
<table>
<thead>
<tr>
<th>Buffer/Solution</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNA solubilisation buffer</td>
<td>10mM Tris, pH7.5, 0.1% SDS, 1mM EDTA</td>
</tr>
<tr>
<td>Reverse transcriptase x10 buffer</td>
<td>0.5M Tris, pH8.3, 0.1M MgCl₂, 0.4M KCl</td>
</tr>
<tr>
<td>Running buffer for SDS-PAGE</td>
<td>0.25M Tris, pH8.3, 0.192M glycine, 0.1% SDS</td>
</tr>
<tr>
<td>SSC (x1)</td>
<td>0.15M NaCl, 0.15M sodium citrate</td>
</tr>
<tr>
<td>Sequencing gel solution A</td>
<td>6% acrylamide, 2.5 x TBE, 7.3M urea, 0.3M sucrose, 0.005% bpb</td>
</tr>
<tr>
<td>Sequencing gel solution B</td>
<td>6.3% acrylamide, TBE buffer, 7M urea</td>
</tr>
<tr>
<td>Solution I (autoclaved and filtered, prior to use)</td>
<td>50mM glucose, 25mM Tris, pH8, 10mM EDTA</td>
</tr>
<tr>
<td>Solution II</td>
<td>0.2M NaOH, 1% SDS</td>
</tr>
<tr>
<td>Solution III (autoclaved and filtered, prior to use)</td>
<td>3M potassium acetate, 5M acetic acid</td>
</tr>
<tr>
<td>TBE buffer</td>
<td>0.89M Tris, 0.89M boric acid, 0.25M EDTA</td>
</tr>
<tr>
<td>TBS (Tris buffered saline)</td>
<td>0.14M NaCl, 5mM KCl, 24mM Tris</td>
</tr>
<tr>
<td>TE buffer (Tris-EDTA buffer)</td>
<td>25mM Tris, 10 mM EDTA</td>
</tr>
</tbody>
</table>
Termination mixes
50mM NaCl
10% DMSO
80mM dCTP
80mM dGTP
80mM dTTP
8mM ddCTP
or ddGTP
or ddTTP
or 0.16mM ddATP

Tissue lysate buffer
63mM Tris, pH6.8
12.5% glycerol
2.5mM PMSF
1.25mg/ml leupeptin
1.25mM EDTA

Transfer buffer
38mM glycine
47mM Tris
0.04% SDS
20% methanol
2.2 Cell Culture

Cells were grown at 37°C in a 6% CO₂ incubator using complete RPMI 1640 medium (cRPMI). When approaching confluence, adherent cell lines were subcultured using trypsin/versene (1:3) to remove the adherent cells from the tissue culture flasks, while loosely-adherent cell lines were removed by agitation. Those removed by trypsin/versene were diluted with 1 volume of cRPMI, and all were collected by centrifuging at 200 x g (IEC Centra-8R general purpose centrifuge; International Equipment Company) for five min. Cells were resuspended in 10 ml of cRPMI, their number estimated using a haemocytometer (Weber Scientific International Ltd.), and sub-cultured 1:2 - 1:4 as appropriate.

2.3 Fluorescence-activated Flow Cytometric Analysis

Cells were harvested; adherent cell lines were scraped from culture flasks using a "rubber policeman" while loosely-adherent cell lines were removed by agitation. To facilitate disaggregation, cells were washed twice in PBS/1% HSA, incubated in versene at 37°C for 5 min, washed, and incubated in 2mM EDTA/PBS/1% HSA under the same conditions. Following a further two washes, aliquots of 10⁶ cells were incubated with MoAb ERIC-1 (52 µg/ml), MoAb UJ13A (52 µg/ml) or PBS, in a 0.5 ml volume for 1 hr at 4°C. After two washes, the cells were incubated with fluorescein-conjugated rabbit anti-mouse Ig under similar conditions. After a further two washes, the cells were resuspended and fixed in 0.5 ml PBS containing 1% paraformaldehyde. Cells were analysed using a Becton Dickinson Facscan™ reading fluorescence intensity from 5 x 10³ recorded events and a laser frequency of 488nm.
2.4 Protein Biochemistry

(a) Neuraminidase enzyme digestion

3 x 10^7 cells were harvested, washed twice in PBS, and resuspended in 0.3 ml of 150mM NaCl, containing 50mM NaOAc, pH 4.5, neuraminidase type X (acylneuraminyldolase; 5U/ml) and the following inhibitors: PMSF (2mM), leupeptin (10 mg/ml) MgCl_2 (1mM) and CaCl_2 (1mM). PMSF inhibits serine proteases, leupeptin inhibits cathepsin B, and magnesium and calcium ions inhibit phosphatidylinositol-specific phospholipase C. The latter is an unlikely exogenous contaminant, but activation of endogenous enzyme upon disruption of cell culture is a possibility (see Theveniau et al., 1991). Supernatants resulting from PIPLC digestion and tissue homogenates were digested under identical conditions. Digestion was carried out at 37°C for 5 hr, with occasional mixing by inversion. Cells were washed twice in PBS/1% HSA, prior to lysate preparation.

(b) Digestion with phosphatidylinositol phospholipase C

2.5 x 10^7 cells were harvested, washed twice in PBS, and resuspended in 0.4 ml of PBS containing 1 unit of commercial PIPLC (Peninsula) or 10% (v/v) Low's PIPLC. An incubate containing 0.05M HgCl_2 constituted a negative control. Incubation was carried out at 37°C for one hour; crude supernatants were obtained after centrifuging cells at 200 x g for five min. Cellular debris was further removed by centrifuging at 10,000 x g for 1.5 hr, followed by aspiration of the supernatant.

(c) Cell and tissue lysate preparation

3 x 10^7 cells were harvested, washed twice in PBS and solubilised in cell lysate buffer kept at 4°C. DNA was sheared by passing the lysate through a series of needles with lumen diameters of 1.1, 0.8 and 0.6 mm; the resultant lysate was boiled for three minutes. In instances where lysates remained viscous after boiling, RNase-free DNase (23U) were added, and mixed briefly by inversion at RT. Lysate preparations were stored at -20°C until needed.

Tissue extracts were obtained by homogenising tissue in a round-bottomed tube with 2 volumes of tissue lysate buffer, using an electrical homogeniser (UltraTurax). This procedure was carried out on ice to minimise degradation, and homogenising limited to a duration of 5 seconds.
Tissue debris was eliminated by centrifugation at 11,000 x g and 4°C for 5 min, and the supernatant removed and retained.

(d) **Determination of protein concentration**

An appropriate volume of protein assay dye (BioRad) was diluted 1:4, and filtered through Whatman no. 1 filter paper. 5 ml of diluent were added to a 5% solution of lysate, made up in 100 µl with PBS. Each lysate was assayed in duplicate. After vortexing and allowing to stand for 5 minutes, protein content was assayed by measuring absorbence at 595 nm with a spectrophotometer (Cecil Instruments; EC 292 Digital U.V. Spectrophotometer). Concentration was calculated by plotting the absorbence value against a standard curve obtained from assaying varying quantities of bovine serum albumin (20, 40, 60, 80 and 100 mg) in an identical fashion.

(e) **SDS-polyacrylamide gel electrophoresis**

Proteins were analysed by SDS-PAGE (Laemmli, 1970) followed by Western blotting, using MoAb ERIC-1. An 120 x 140 x 1.5 mm, 7% separating gel was poured and polymerisation allowed to proceed for 1 hr, with an overlay of butanol. The butanol and unpolymerised gel layer were removed and the surface of the gel rinsed with 0.5M Tris-HCl, pH6.8. Directly above this, a stacking gel of 5% acrylamide was prepared containing an 1.5 mm well-forming comb. Polymerisation was complete in approximately 45 minutes. On removal of the comb, wells were rinsed twice with 0.5M Tris-HCl, pH6.8; buffer was removed after each wash with a canicula. Wells were filled with running buffer.

Samples (300 µg extract unless specified otherwise) were augmented by 10% volume with 1M DTT/ 0.1% bromophenol blue (bpb), boiled for 3 min, and applied to the wells with feather-tipped pipette tips. An aliquot (20µl) of Rainbow molecular weight markers (14.3-200 kDa) was electrophoresed alongside the protein extracts. Samples were
electrophoresed at 50V through the stacking gel and at 280V through the separating gel, until approximately 1 hr after the bpb marker dye had run off the bottom of the gel.

(f) **Electrophoretic transfer of proteins from polyacrylamide gel to nitrocellulose membrane**

Separated proteins were transferred to nitrocellulose by horizontal electrophoresis (Towbin *et al.*, 1979). After SDS-PAGE of cell or tissue lysates, the stacking gel was removed, and the separating gel briefly rinsed in transfer buffer. 9 sheets of filter paper (LKB) were individually wetted with transfer buffer and stacked on the wet horizontal anode of a 2117-250 Novablot electrophoretic transfer kit (LKB). A nitrocellulose filter was wetted with transfer buffer and layered on the stacked filter paper, followed by the separating gel, and 9 additional sheets of wetted filter paper. Care was taken to ensure that filter paper, nitrocellulose, and gel were layered in such a way so as to exclude air bubbles. The wetted cathode was placed on top of this sandwich, and a constant current of 134mA was applied for 1 hr. On completion of transfer, the nitrocellulose sheet was washed in PBS/5% Marvel O/N on a platform shaker. The gel was stained in Coomassie Brilliant Blue protein stain O/N, and subsequently destained in destain solution for 1-2 hr.

(g) **Analysis of transferred proteins**

Nitrocellulose strips were incubated with either MoAb anti-PSA (1:400), anti-NCAM MoAb ERIC-1 monoclonal antibody or the "irrelevant" MoAb M340 for 30 min at RT. MoAbs ERIC-1 and M340 were used at a concentration of 26 μg/ml unless indicated otherwise. Incubations were carried out on Parafilm in petri dishes. The strips were individually washed 2 times for 5 min in PBS. After subsequent incubations with alkaline phosphatase-conjugated rabbit anti-mouse Ig (1:250) for 30 min at RT, the strips were again washed twice in PBS, and once in TBS/1mM levamisole.
The strips were developed by incubating for 15 min at RT in the dark with Fast Red substrate and washed in TBS/1mM levamisole. Additional incubations with substrate were carried out if necessary.

The molecular weights of proteins detected by Western blot analysis were calculated on construction of a standard curve from the log_{10} molecular weights versus the relative mobilities of standard proteins (Weber and Osborn, 1969).

### 2.5 Generation of Polymerase Chain Reaction (PCR) products

#### (a) RNA extraction

2 x 10^7 cells were harvested and washed twice in PBS. Cells were lysed in ice-cold RNA lysis buffer (2 x 0.5 ml), and the DNA sheared by passing the lysate through a succession of needles with narrowing diameters of 1.1, 0.8 and 0.6 mm. 7 mls of ice-cold 4M LiCl was added to the lysate, and the mixture stored at 4°C in the dark, O/N.

The suspension was centrifuged for 90 min at 14,000 x g, and 4°C. On decanting the supernatant, the sedimented precipitate was resuspended in 15 mls of ice-cold 3M LiCl. The resultant suspension was centrifuged for 60 min at 14,000 x g and 4°C. On decanting the supernatant, the sedimented precipitate was dissolved in 2 mls of solubilisation buffer. Dissolution was encouraged by freezing, thawing, and vortexing the mixture; this sequence of events was repeated 2 or 3 times as necessary. The mixture was divided into three aliquots.

Phenol, previously equilibrated with TE buffer, and chloroform/4% isoamylalcohol were added to each aliquot in a 1:1 v/v ratio, so that the combined volume equalled approximately that of the crude RNA extract. Mixtures were vortexed and centrifuged, the upper aqueous layer extracted, and the lower organic layer discarded. This procedure was repeated, and a third extraction performed using chloroform/4% isoamylalcohol. RNA was
precipitated by adding 50 µl of saturated sodium acetate to each aliquot, followed by two volumes of cold ethanol. Mixing was effected by several inversions, and precipitation allowed to proceed at -20°C overnight.

Precipitate was collected by centrifuging at 10,000 x g and 4°C for 10 min. On discarding the supernatant ethanol, the pellet was washed three times in 90% ethanol, dried in a vacuum dryer (SpeedVac Concentrator SVC100H, Stratech Scientific), and resuspended in an appropriate volume of sterile TE buffer.

(b) Assay of RNA concentration and purity

RNA was diluted 1:200 with water; the resulting dilution was checked to ensure neutral pH. Concentration was assayed by spectrophotometer at 260 nm using a quartz cuvette with a path length of 1 cm. RNA concentration was calculated using the specific extinction coefficient (absorbancy index) of 0.025 ml mg⁻¹. The extinction coefficient relates absorbance to concentration under specified conditions (e.g. using a 1 cm path length) and is based on Beer's law (i.e. the amount of light absorbed is proportional to the absorbing species; see equation (1)). If the molecular weight of the absorbing species is known, it is sometimes expressed as a molar absorption coefficient.

\[
(1) \quad K = \frac{d}{l} \quad \text{K = extinction co-efficient}
\]
\[
\text{d = absorbance}
\]
\[
\text{l = concentration (mg/ml)}
\]

Protein concentration was measured at 280 nm. As the wavelengths resulting in absorption maxima used in these experiments (260 and 280 nm) are relatively close, some "cross-contamination" will occur; RNA will partly contribute to absorbances obtained at 280 nm, and protein, to those obtained at 260 nm. Expressing the absorbances as a ratio negates this effect and provides a direct indication of the purity of RNA insofar as protein and phenol contamination is concerned. (Phenol also has an absorbance peak at approx. 280 nm).
(c) **Synthesis of cDNA**

Only RNA which yielded an RNA : protein absorbence ratio of 1.8 or greater was used as a template for the synthesis of cDNA. cDNA was generated in a 20 µl volume containing 2 µg RNA, 1mM dNTP mix (Pharmacia), 18U avian myeloid leukaemia virus reverse transcriptase (AMV-RT), 1 µM 3' primer, 37U RNAguard, and 2 µl reverse transcriptase x10 buffer. Synthesis was carried out at 42°C for 1 hr and cDNA stored on ice prior to amplification by the polymerase chain reaction (PCR).

(d) **Symmetric PCR amplification**

18 µl of cDNA, generated as described above, was amplified in a 100 µl volume containing 20pmole 3' primer, 20pmole 5' primer, 200µM dNTP, 1x reaction buffer and 2.5U of Taq polymerase. A negative control, comprising the above reagents with no cDNA, was constituted alongside every set of samples amplified by PCR. Samples were layered with 100 µl heavy white mineral oil, and amplified in a thermal cycler (Techne; model PHC-2). Samples were heated at 94°C for 3 minutes to ensure effective denaturation, followed by 30 cycles consisting of three segments of 94°C for one minute, 60°C for one minute and 72°C for 90 seconds. This was followed by a final single cycle of 94°C for one minute, 60°C for 1 minute and 72°C for 7 minutes. Samples were extracted with 4% isoamylalcohol in chloroform and visualised by electrophoresing a 10µl aliquot through a 2.5% agarose mini-gel permeated with 0.5 µg/ml ethidium bromide.
2.6 Analysis of PCR Products

(a) Agarose gel electrophoresis and transfer to membrane

10 µl aliquots of amplified cDNA plus 2 µl of type IV gel loading buffer (glb) were electrophoresed through a 2.5% agarose gel containing 1x TBE and 0.5 µg/ml ethidium bromide, at 150mA for approximately 4 hr. Samples were electrophoresed alongside appropriate positive controls which served as a means by which to gauge efficiency of hybridisation.

1 kb molecular weight markers (0.4 mg) were used as reference markers for estimating the molecular weight of the separated DNA species. Electrophoresis was carried out in TBE buffer containing 0.5 µg/ml ethidium bromide. The gel was viewed under ultraviolet light, photographed and the DNA therein denatured by gently shaking in 0.4M NaOH containing 0.6M NaCl for 30 min at RT. This was followed by neutralisation in 1.5M NaCl containing 0.5M Tris, pH7.5 for 30 min at RT. DNA was transferred from the gel to a GeneScreen nylon filter by placing the gel, face down, on a horizontal filter paper wick, the ends of which were immersed in x10 SSC. The nylon filter was layered on top of the gel, intervening bubbles were eliminated, and two sheets of filter paper (Whatman; 3M) were positioned, followed by stacks of paper towel, a glass plate and a weight. Passive transfer was allowed to proceed O/N. Transferred DNA was denatured in 0.4M NaOH, and neutralised with 2 x
SSC containing 0.2M Tris, pH7.5. Both washes were carried out at RT for 30 seconds. The blot was allowed to air dry.

(b) **Acrylamide gel electrophoresis and transfer to membrane**

An 8% acrylamide gel containing a 1.5 mm well-forming comb, was poured and allowed to polymerise. On removal of the comb, the wells were rinsed twice and filled with TBE buffer. 40 µl aliquots of PCR product were combined with 8 µl of type IV gel and electrophoresed in TBE buffer until approximately 1 hr after the marker dye had run off the bottom of the gel.

The gel was stained in 0.5 µg/ml ethidium bromide in TBE buffer for 30 min at RT, destained in TBE buffer under the same conditions, viewed under U.V. light, and photographed.

DNA was transferred to Genescreen using a method analogous to that used for the transfer of proteins to nitrocellulose in the preparation of a Western blot. 9 sheets of filter paper (LKB) were wetted with 50mM TBE and stacked carefully on the anode of a 2117-250 Novablot electrophoretic transfer kit (LKB), followed by a sheet of GeneScreen cut to an appropriate size and pre-soaked in 50mM TBE. The gel was placed on top, followed by 9 sheets of filter paper wetted with 50mM TBE, and the cathode. A constant current of 134mA was applied for 90 min, whereupon the filter was denatured with 0.4M NaOH and neutralised with 2 x SSC containing 0.2M Tris; each wash was applied for 30 sec at RT. The blot was allowed to air-dry.

(c) **End-labelling of oligonucleotides**

0.5 µg of oligonucleotide probe was end-labelled at the 5' terminus by incubating with 7U T4 polynucleotide kinase, 50 µCi γ³²P dATP (Amersham) and 1 µl x10 labelling buffer in a 10 µl volume. The reaction was allowed to proceed at 37°C for 45 minutes, and terminated by placing on ice for 5 min. Labelled oligonucleotide was purified by diluting the
reaction mixture to 100 μl, removing 1 μl for later analysis, and applying the remainder to the top of a NAP-5 column which had previously been rinsed with TE buffer. Four 0.5 ml fractions were collected on further additions of TE buffer to the top of the column.

(d) Random labelling of oligonucleotides

25 ng of oligonucleotide was random-labelled by boiling the oligonucleotide for 3 min, followed by quenching on ice and incubating with 10 μl of oligo-labelling buffer, 2 mg BSA, 50 μCi α^32P dCTP, and 5.5 U Klenow fragment, in a 50 μl volume. Labelling was allowed to proceed at RT, overnight, whereupon the radio-labelled probe was purified by eluting from a Nick TE column; otherwise, purification and radio-assay were carried out in an identical manner to that in which end-labelled probes were purified and assayed.

(e) Determination of radioactivity

Quantitative measurement of beta radioactivity was performed using a 1209 Rackbeta automated beta counter (LKB Wallac). Results are given in counts per minute (cpm). The activity in MBq is given by:

\[ \text{Mbq} = \frac{\text{cpm} \times F}{60 \times 10^6} \]

where

\[ F \]

is the conversion factor for counts to disintegrations.

For ^32P, \( F = 2.86 \), given a beta counter efficiency of 35%

\[ 60 \times 10^6 \]

is the number of disintegrations per minute in 1 MBq.

The gamma counter was regularly calibrated for ^32P by repeated measurement of a standardised source of ^32P-labeled dATP. Calibration using a standard ^32P capsule was impracticable, owing to the rapid half-life
of the isotope. Counting efficiencies of approximately 35% were obtained periodically, and remained unchanged throughout the duration of this project.

(f) Hybridisation of blot

A hybridisation mesh was cut so that its size matched that of the dried blot, and the two were immersed in 2 x SSC for 15 min. With the mesh juxtaposed with the blotted side of the Genescreen membrane, the two were rolled and deposited in a hybridisation bottle. These were rinsed with 2 x SSC so that the blot/mesh unfurled and rested flush against the sides of the bottle, and bubbles arising between the membrane and mesh were eliminated. Using a dual hybridisation oven (Hybaid), the membrane was prehybridised for a minimum of 1.5 hr in 10 ml hybridisation buffer and denatured salmon sperm (1 mg), after which denatured oligonucleotide probe with an activity of approximately $10^7$ cpm was added.

An appropriate temperature of hybridisation was calculated for each probe, using equations which take into account the length and base composition of the probe in question (Wahl et al., 1987). All blots were hybridised overnight.

Blots were washed in 2 x SSC containing 1% SDS for 5 min at RT, 2 x SSC containing 1% SDS for 15 min at the temperature of hybridisation, and 2 x SSC containing 0.1% SDS for 15 min, again at the temperature of hybridisation. Each wash was performed twice. The washed blot was wrapped in clingfilm, placed in a cassette with an intensifying screen, and exposed to X-ray film.

(g) Preparation of blot for rehybridisation

In instances where it was necessary to analyse the same blot with different oligonucleotides, old oligonucleotide was stripped from the blot by incubating the latter in 0.4M NaOH at 42°C on a platform shaker for 30 min, followed by neutralising with 0.1 x SSC containing 0.1% SDS and 0.2M Tris, pH7.5 under the same conditions. These incubations were repeated,
and the blot wrapped in clingfilm and exposed to film for an appropriate period of time.

2.7 Direct Sequencing of PCR Products

(a) Isolation of PCR products

Approximately 400 ng of PCR product were electrophoresed either through a TBE/8% polyacrylamide gel at 200V, for approximately 4 hr, after which the gel was stained in TBE/ethidium bromide (0.5 μg/ml) for half an hour. On destaining in TBE for 30 min, individual bands were located under U.V. light, and excised. The DNA therein was eluted by the "crush and soak" method (Sambrook et al., 1989) as follows.

Two volumes of elution buffer were added to the gel slice and incubated at 37°C on a rotating wheel O/N. The sample was centrifuged at 11,000 x g for 1 min at 4°C, and the supernatant transferred to a fresh eppendorf. An additional volume of elution buffer was added to the gel slice, vortexed, and the supernatants combined. 2 volumes of cold ethanol were added, and the solution stored on ice for 30 min. Precipitated DNA was recovered by centrifugation at 11,000 x g for 10 min at 4°C, and redissolved in 20 μl of TE buffer, pH7.6. 25 μl of 3M NaOAc, pH5.2 was added and the DNA reprecipitated with 450 μl of cold ethanol and recovered by centrifugation as previously. The pellet was rinsed with 70% ethanol, and redissolved in 25 μl of TE buffer (pH7.6).

(b) Asymmetric amplification of PCR products

Approximately 20 ng of eluted DNA was subjected to amplification by asymmetric PCR, using 50 pmole 3' primer and 1 pmole 5' primer. Otherwise, conditions of amplification were identical to those used in the generation of symmetrically amplified DNA. Amplified products were extracted with an equal volume of chloroform/4% isoamyl alcohol. A 20 μl aliquot of these products was subjected to a second round of asymmetric amplification, using the same primer ratio and reaction conditions as stated
previously. On extraction, a 10 μl aliquot was augmented with 2 μl type IV glb, and electrophoresed though a 2.5% agarose gel, impregnated with 0.5 μg/ml ethidium bromide, to ensure that PCR amplification had been successful in generating ample quantities of DNA product and that the negative water control appeared negative.

(c) Purification and quantitative assay of asymmetric PCR products

85 μl of asymmetrically amplified PCR product was concentrated using the Isogene™ Kit. 2 volumes of sodium iodide reagent were added to the PCR product, the mixture vortexed, and cooled at 4°C for 5 min. 20 μl of resuspended DNA binder was added, and the mixture rotated at RT for 10 min on an end-over-end rotatory mixer, followed by centrifuging at 11,000 x g for two min. The supernatant was discarded, and the bound DNA washed by vortexing with 200 μl Wash Buffer (diluted 1:100 with ethanol). Again, the mixture was centrifuged and the supernatant discarded. A second wash was carried out and on discarding the supernatant, the bound DNA eluted into 30 μl TE buffer. The volume was reduced to 10 μl using a vacuum dryer (model 583, BioRad).

Approximate DNA concentration was ascertained using the DNA Dipstick™ Kit. Dilutions of 1:9, 1:99 and 1:999 were made of the purified PCR product in TE buffer. 1.0 μl of each dilution was dotted on to the dipstick membrane, alongside a 2 ng/ml standard. The membrane was allowed to air dry, after which it was wetted with ddH₂O, and immersed in coupling solution for 10 min. The coupling solution was rinsed from the membrane, and the latter immersed in wash buffer for 2 min, followed by developing solution for 5-10 min. The approximate concentration of DNA was estimated by comparing the colour intensities of each dilution of sample with the 2 ng standard and with a key supplied by the manufacturers.
(d) **Sequencing of PCR products**

Asymmetrically amplified DNA was sequenced in the presence of DMSO; inclusion of the latter has been shown to enhance the specificity of the reaction (Winship, 1989). 2 µl of purified asymmetric DNA template (0.2pmole) generated in the manner described above, was mixed with 1 µl containing 10 pmole of 5' amplification primer and 3 µl of x2 annealing buffer. After boiling for three minutes, and immediately snap-cooling on dry ice, 4 µl of labelling mix was added. The resulting 10 µl mixture was divided equally into four tubes, and 2 µl of ddATP, ddGTP, ddCTP or ddTTP termination mix added to one of the four aliquots. The tubes were incubated at 37°C for five minutes, 2 µl of chase mix was then added, and the reaction was left for a further five minutes at 37°C. The reaction was terminated by adding 5 µl of 0.04% bromophenol blue, 0.04% cyanol blue in formamide, and boiling for 3 minutes. Samples were stored at -20°C.

(e) **Gel analysis of sequencing reaction products**

Radiolabeled products of the sequencing reaction were analysed by denaturing polyacrylamide gel electrophoresis. A 21cm x 50cm x 0.4mm gradient buffer gel was poured, constituted by a 1:5 ratio of solution A : solution B. Prior to polymerisation, a 0.4mm well-forming comb was inserted into the top of the gel. On polymerisation, the gel was allowed to age for a minimum of 3 hr, after which it was secured within the gel-running apparatus (BioRad), and subjected to a constant voltage of 1800V until such time as the gel obtained a temperature of approximately 50°C.

Products from the sequencing reaction were heated at 80°C for 2 min, during which time the well-forming comb was removed, and the wells cleared of urea. 3 µl aliquots of sample were introduced into each well using feather-tip pipette tips, and in the order A, G, C, T, where "A" represents the reaction in which ddATP served as the terminating nucleotide, "C" represents the reaction in which ddCTP terminated chain-
extension, *et cetera*. The gel was run at 1800V in TBE buffer until the marker dye reached the bottom, at which point the gel apparatus was dismantled, and the gel, supported on one glass plate, was soaked in gel fixing buffer for 20 min. The gel was removed from the buffer, allowed to drain, and transferred to a sheet of 3M Whatman filter paper, whereupon it was wrapped in clingfilm and dried at 85°C for 1 hr using a gel dryer. On removal of the clingfilm, the gel was exposed to X-ray film and developed after an appropriate period of time - usually 1-3 days.

2.8 Cloning and Sequencing of PCR Products

(a) Preparation of PCR products

In instances where PCR products could be adequately separated by PAGE, specific DNA species were isolated and prepared for cloning by the "Crush and Soak" method (see above). Isolates were amplified by symmetric PCR amplification, and the resultant mixture assayed using the DNA dipstick kit.

In other instances, PCR products were electrophoresed through a 2% agarose gel impregnated with 0.5 μg/ml ethidium bromide, in TBE buffer. The gel was viewed under U.V. light, relevant bands excised, and the gel photographed. One volume of TE buffer was added to the gel slice, the mixture incubated at 50°C for 10 min, frozen on dry ice, and thawed with intermittent vortexing. This procedure was repeated, at the end of which the gel fragments were sedimented by centrifuging at 11,000 x g and 4°C for 10 min. The supernatant containing eluted DNA was removed, 20 μl of which was amplified by symmetric PCR amplification. In some instances, the number of intermediate cycles in the amplification reaction was reduced from 30 to 15, in attempt to minimise irrelevant products which typically arise under anaemic conditions (Mullis, 1991).
(b) **Cloning of PCR products**

The TA cloning™ kit was employed to clone the PCR products generated as described above. pCR™ plasmid and PCR product were incubated at a 1:1, 1:3 or 1:10 molar ratio with 4U DNA ligase and 1 µl 10x ligation buffer in a 10 µl volume at 12°C, O/N.

A 50 µl vial of competent INVαF+ E.coli cells was thawed on ice, 1 µl of 0.5M β-mercaptoethanol added, and the suspension mixed by gentle tapping. 1 µl of ligation mixture was added, and the suspension mixed as before and stored on ice for 30 min. Heat shock was delivered by rapidly transferring the suspension to a water bath heated at 42°C for 1 min, whereupon the vial was returned to ice for 2 min. 450 µl of SOC medium was added under sterile conditions, and the vial incubated for exactly 1 hr at 37°C and 225 rotations per minute (rpm) in a G42 Environmental Incubator Shaker (New Brunswick Scientific Co. Inc.). 25 µl aliquots of suspension were spread on 10 cm diameter plates containing LB broth/agarose, impregnated with kanamycin (50mg/ml) with an overlay of X-galactose (25ml x 40 mg/ml). Plates were incubated in an inverted position at 37°C for approximately 24 hr.

(c) **Small scale preparation of plasmid DNA**

5 ml aliquots of LB broth were inoculated with positive (white) colonies, and incubated at 37°C and 225 rpm O/N. A single negative (deep-blue) colony was harvested, cultured and extracted alongside the positive colonies.

Suspensions were sedimented by centrifuging at 600 x g for 10 min, the supernatant discarded, and the pellet resuspended in 100 µl of solution I. Solution II (200 µl) was added, and the mixture rapidly inverted 5 times. Solution III (200 µl) was added, the mixture inverted, and bacterial debris collected by centrifuging at 4°C and 11,000 x g for 5 min. The supernatant was transferred to a fresh tube and extracted with phenol : chloroform.
DNA was precipitated with 2 volumes of ethanol at RT, and collected by centrifuging as before. Ethanol was removed by aspiration. The pellet was washed with 70% ethanol, and dissolved in 50 µl TE buffer containing DNase-free pancreatic RNase (20 mg/ml).

(d) Screening of clones for relevant inserts

Approximately 1 mg of DNA preparation generated as described above, was digested with 4U Not-1 restriction enzyme in a 40 µl volume containing 4 µl x10 restriction enzyme buffer at 37°C for 90 min. After chilling on ice for 5 min, 8 µl type IV glb was added to each sample, and the samples electrophoresed alongside 0.4 mg 1 kb markers, or 0.6 mg 35S DNA markers, through a 1% agarose gel impregnated with 0.5 µg/ml ethidium bromide, in TBE buffer. Appropriate positive controls were electrophoresed alongside the digested minipreparations of cloned DNA. When samples had traversed half the length of the gel, electrophoresis was terminated, the gel photographed under UV light, and DNA transferred to nylon Genescreen membranes as previously described. Membranes were hybridised with appropriate oligonucleotides to reveal samples which might contain PCR products of interest.

(e) Sequence analysis of cloned PCR products

Approximately 100 µg plasmid DNA was diluted to 9 µl, and 1 µl of 2M NaOH/2mM EDTA added. The mixture was incubated at RT for 5 min, neutralised with 2 µl of 3M NaOAC, pH4.5, and diluted to a 20 µl volume. Denatured DNA was precipitated on adding 60 µl ethanol and immersing in a bath of dry ice and ethanol for 15 min. Precipitates were collected by centrifuging at 11,000 x g and 4°C for 5 min, and washed with 70% ethanol. On decanting 70% ethanol, the pellet was dried in a vacuum dryer and resuspended in a 10 µl volume containing 2 µl x5 reaction buffer (USB) and 20 pmole of primer. The mixture was incubated at 68°C for 2 min and gradually cooled in a water bath to 35°C on the bench. 5.5 µl of labelling
mix (USB) containing 5 mCi $\alpha^{35}$SdATP was added to each sample, and incubated at RT for 5 min, whereupon 3.5 µl aliquots of this mixture were added to one of four pre-warmed termination mixes (ddATP, ddGTP, ddCTP or ddTTP termination mix; 2.5 µl each). The termination reaction was carried out at 37°C for 5 min, and the reaction stopped by adding 4 µl stop buffer (USB). Samples were stored at -20°C, and analysed by denaturing polyacrylamide gel electrophoresis.
CHAPTER 3

EXPRESSION OF NCAM ON TUMOUR CELL LINES AND TISSUE

3.1 Introduction

It is of interest and clinically relevant to examine the detailed expression of NCAM by tumours which express it; especially tumours which might retain tissue-specific isoforms, and thereby be differentiated from one another. Obvious candidates for such analysis are neuroblastoma and rhabdomyosarcoma - embryonic tumours which are sometimes difficult to diagnose.

NCAM exhibits developmental and cell-type-specific variations in its expression. With respect to isoform composition, two main patterns have emerged over time: one exhibited by tissues of neural origin, and the other, by muscle tissue. Other patterns of NCAM expression have been observed - for example, that exhibited by the liver of *Xenopus laevis* (Tacchetti *et al.*, 1992) but those of neural cells in amphibia and mammals correlate well between species (Kreig *et al.*, 1989; Saint-Jeannet *et al.*, 1989).

Neuronal tissue has major isoforms of 120, 140, and 180 kDa - their relative proportions are observed to vary with development. Isoforms of 120, 125, 145, and 155 kDa are found in developing muscle. Other tissues which express NCAM usually express an isoform of 140 kDa (for examples, see Langley *et al.*, 1989; Watanabe *et al.*, 1989; Lackie *et al.*, 1990; Rouiller *et al.*, 1990; Mayerhofer *et al.*, 1992b). It is sometimes implied to be identical to that expressed by neuronal tissues, although it is possible that they are subject to tissue-specific splicing patterns within the extracellular domain, or post-translational modifications.
In keeping with the supposition that NCAM-180 is neuronally-restricted is the finding that medulloblastoma expresses NCAM-180, in contrast with astrocytoma, glioma, oligodendrogioma and schwannoma tissues which do not (Frost et al., 1991). Some data, however, suggests that NCAM-180 is not subject to the rigid tissue restrictive limitations initially supposed by investigators. For instance, an isoform of approximately 180 kDa has been observed in myoblasts (Tassin et al., 1991) rat islet cells (Møller et al., 1992) and embryonic cardiac tissue (Burroughs et al., 1991). It is not known whether this isoform is identical to that found in neural tissue or differs in its splicing pattern and/or glycosylation. An isoform of 180 kDa has also been detected on cultured mouse Leydig cells (Mayerhofer et al., 1992a) a subset of murine glial cells (Pollerberg et al., 1985; Miragall et al., 1988) and embryonic and adult murine ilea (Probstmeier et al., 1990).

In addition to the major NCAM isoforms, a 170 kDa isoform has been observed on Western blotting tissue extracts from adult brain (Bhat and Silberberg, 1988; Frost et al., 1991; Phimister et al., 1991); it has also been detected in extracts of medulloblastoma. An isoform of 172 kDa has been detected in hen brain upon removal of N-linked oligosaccharides, suggesting that, in at least this instance, a 172 kDa isoform is precursor to the 180 kDa species (Marsh and Gallin, 1992). An isoform of 95 kDa has also been detected in adult brain and on tumours of the central nervous system (Frost et al., 1991). It is not known whether this protein species arises as a precursor, or through alternative splicing, post-translational modification, or degradation. A secreted isoform of 95 kDa has been observed to arise from in vitro translation of NCAM cDNA containing the SEC exon. However, it would appear that this isoform approximates more closely to a molecular weight of 115 kDa in vivo, as suggested by transfection experiments (Gower et al., 1988).
Immunoblot studies of protein expression are often accompanied by Northern blot profiles as a complementary analysis - correlating trends in RNA and protein expression implies that one gives rise to the other. Additional studies (for examples see Barbas et al., 1988; Andersson et al., 1990) have further corroborated the purported relationship between most mRNA species and isoforms to which they give rise. In rodent brain, four mRNA species of 7.4, 6.7, 5.2, 4.3 and 2.9 kb have been identified. Isoforms of similar or identical size have been detected in human neuronal tissue. The 7.4 and 6.7 kb species give rise to protein isoforms of 180 and 140 kDa respectively (Gennarini et al., 1986; Hemperly et al., 1986b; Barbas et al., 1988; Santoni et al., 1989b; Andersson et al., 1990). The 5.2 and 2.9 kb species appear to encode NCAM-120. Alternative poly(A) addition signals in exon 15 give rise to their different sizes (see Figure 1.3) although the way in which their corresponding proteins might differ in function remains to be elucidated. That their co-ordinated expression is not indiscriminate (Moore et al., 1987) suggests that such a difference exists. Secreted NCAM is also encoded by a mRNA transcript of 5.2 kb (Gower et al., 1988).

At the beginning of this project, Lipinsky et al. (1987) had analysed NCAM expression of Ewing's sarcoma (ES) cell lines and the IMR-32 neuroblastoma cell line. Different levels of protein and polysialylation were observed between the two types of cell, with IMR-32 expressing more of each. On desialylation and immunoblotting, isoforms of 120, 140 and 180 kDa were observed in extracts of ES cell lines and of IMR-32, indicating that these cell lines demonstrate a "neuronal" phenotype with respect to NCAM expression. Corresponding mRNAs, however, could only be found in preparations of IMR-32; the 7.0 kb species believed to encode NCAM-180 was not detected in ES cell lines. The murine neuroblastoma cell line N2a
was observed to contain NCAM RNA species of 2.9, 4.3, 6.7 and 7.4 and protein isoforms of 120, 140 and 180 kDa (Gennarini et al., 1986).

In an attempt to further define the molecular identity of "small round cell" tumours, the diagnosis of which can be difficult, NCAM expression by seven neuroblastoma cell lines and three rhabdomyosarcoma cell lines was investigated. Fluorescence-activated cytometric analysis established that NCAM is expressed by neuroblastoma cells - a more detailed profile was obtained on immunoblotting cell extracts of native and desialylated cell lines.

3.2 Methods

(a) NCAM expression on selected neuroblastoma cell lines as determined by fluorescence-activated flow cytometry

A panel of neuroblastoma cell lines subjected to fluorescence-activated flow cytometric (FACS) analysis after prior incubation with primary monoclonal antibodies ERIC-1 or UJ13A, followed by a secondary fluorescein-conjugated rabbit anti-mouse antibody. Each cell line was also incubated with secondary antibody alone. A T-ALL line GH1, previously demonstrated not to express NCAM (K. Patel, personal communication) was used as a negative control cell line.

(b) Establishing experimental parameters: positive and negative controls and a titre of monoclonal antibody ERIC-1 for use in Western blotting

Prior to investigating NCAM expression by cell lines and tumour tissue, adult brain was Western blotted to establish its suitability as a positive control. Previous work suggests strong expression of NCAM by this tissue (Chuong and Edelman, 1984; Bhat and Silberberg, 1988). MoAb M340, like MoAb ERIC-1, is an antibody of class IgG1, but its antigen remains to be identified. It was used as a negative control antibody.
Aliquots of adult brain homogenate were digested with neuraminidase (acylneuraminyl hydrolase; 0.5 and 1.0 units/ml) or incubated under identical conditions but lacking enzyme (henceforth referred to as a "mock" digestion). The neuraminidase used catalyses the hydrolysis of terminal α2-3, α2-6 and α2-8 ketosidic bonds. Thus products of digestion are likely to comprise desialylated glycolipids and oligosaccharides, as well as high-mannose and complex glycoproteins. While the neuraminidase used cleaves α2-3 linkages more rapidly than α2-8 or α2-6 linkages, complete desialylation should be achieved, bearing in mind the unit definition of the enzyme under the conditions used, the incubation time, and the likely levels and relative proportions of terminal mono-residues and polymers (see Finne et al., 1983).

Samples - undigested, digested, and "mock" digested - were prepared with sample buffer. 150 μg of each sample was analysed by gel electrophoresis followed by Western blotting with either MoAbs ERIC-1 or M340 at a dilution of 1:30 (approx. 43 μg/ml) which was arbitrarily chosen.

To establish an appropriate titre of MoAbs ERIC-1 monoclonal antibody for Western blotting, samples (150 μg) of undigested adult brain were separated by gel electrophoresis and analysed by Western blotting with MoAb ERIC-1, varying in dilution from 1/2 to 1/10,000. One sample was also developed using MoAb M340 at a concentration of 0.65 mg/ml as a negative control - a concentration identical to the highest titre of MoAb ERIC-1 (1/2).

(c) Western blot analysis of extracts of neuroblastoma, rhabdomyosarcoma, neuroepithelioma and myeloid leukaemia cell lines

Six neuroblastoma and three rhabdomyosarcoma cell lines were analysed by Western blotting, to investigate patterns of NCAM expression on these cell lines. Protein expression in the myeloid leukaemia cell line Kglα and the neuroepithelioma cell line NB100 were also investigated. Whole-cell digests and extracts were prepared, electrophoresed and immunoblotted as described in Chapter 2. Each cell line was immunoblotted twice, on different occasions, using different cell stocks. In early experiments, vertical electrophoresis was carried out until the dye "front" reached the bottom of the gel. Later, in order to obtain better separation of NCAM isoforms, electrophoresis was continued until the Rainbow marker™ of 69 kDa reached the bottom of the gel. Adult brain lysate was used as a positive control. Extracts of native cells developed with MoAb M340 served as a negative control. Both MoAbs ERIC-1 and
M340 were used at a concentration of 26 \( \mu \text{g/ml} \). Each cell line was analysed by Western blotting on two separate occasions, using cell lysate prepared from separate cell stocks. The results presented here are representative of the replicates unless indicated otherwise.

(d) Western blot analysis of neuroblastoma and rhabdomyosarcoma tissue extracts

Tissues from neuroblastoma (Stage 4) and rhabdomyosarcoma tumours were snap-frozen shortly after excision and stored at -80°C until required, whereupon they were homogenised as described in Chapter 2. An aliquot of each sample was digested with neuraminidase. Rhabdomyosarcomatous tissue was also "mock" digested. Three different tumour samples were analysed in an identical fashion to that in which cell lysates were analysed, although each sample was Western blotted once. By error, a positive control was omitted in the analysis of neuroblastoma tissue lysates; the negative control in each case comprised undigested tissue extract incubated with MoAb M340 as a primary antibody.

(e) Glycosyl-phosphatidylinositol-anchored NCAM isoforms on the JR-1 rhabdomyosarcoma cell line as revealed by Western blotting

Digests were prepared as described in Chapter 2. Native cell extract, supernatants from PIPLC digests (200 \( \mu \text{l} \) volume) and neuraminidase digests of the latter were electrophoresed and immunoblotted with MoAb ERIC-1. Adult brain was included as a positive control and JR-1 cell lysate blotted with MoAb M340 as a negative control. This experiment was performed in duplicate and the results presented here are representative of the replicate experiment.

(f) NCAM expression and cell density

IMR-32 cells were cultured under conditions outlined in Chapter 2. Cultures were harvested at one of two points - either at subconfluency; 1.0 \( \times 10^7 \) cells/flask, or at high density; 4.7 \( \times 10^7 \) cells/flask. Neuraminidase
digestion and lysate preparation were carried out as described in Chapter 2, followed by electrophoretic separation and immunoblotting. Adult brain served as a positive control, and immunoblotting using MoAb M340 as a primary antibody served as a negative control.

3.3 Results

(a) NCAM expression on selected neuroblastoma cell lines and a neuroepithelioma cell line as determined by fluorescence-activated flow cytometry

Monoclonal antibodies UJ13A and ERIC-1 specifically bind to all of the human neuroblastoma cell lines analysed here except PCF (see Table 3.1). The neuroepithelioma cell line NB100 was also observed to specifically bind the anti-NCAM antibodies. Median fluorescence values obtained after incubating PCF with both MoAbs UJ13A and ERIC-1 did not vary significantly from that obtained on omission of the primary antibody. Heterogeneity in the degree of staining is observed both within and between the different cell lines (see Figure 3.1). For instance, the median FACs values for the neuroblastoma cell lines SK-N-SH and PCF on staining with MoAb UJ13A are 380 and 192 respectively. MoAb UJ13A yields consistently stronger signals than MoAb ERIC-1.

Autofluorescence and/or non-specific binding between cell lines and anti-mouse Ig was observed to vary. For example, in the absence of primary Ig, the Kelly neuroblastoma cell line generated a median fluorescence value of 147, approximately half of the average value obtained with anti-NCAM Igs. In contrast, the G610 neuroblastoma cell line recorded a value of 66 - less than one quarter of the average value obtained with MoAbs UJ13A and ERIC-1. FACs signals generated by cell lines incubated with secondary antibody alone tended to exceed that observed upon incubating the T cell line GH1 with either MoAbs UJ13A or ERIC-1,
Table 3.1

FACs analysis of neuroblastoma, neuroepithelioma, and leukaemia cell lines
Cell lines were incubated with MoAbs UJ13A, ERIC-1, or PBS before a secondary incubation with fluorescein-conjugated rabbit anti-mouse Ig. The median values of FACs profiles are as indicated. The ratio of FACs values compare the median values obtained with MoAbs UJ13A and ERIC-1, and take into account the autofluorescence/non-specific binding of secondary Ig.

Key: ng = negligible
<table>
<thead>
<tr>
<th>Cell Line</th>
<th>Cell Type</th>
<th>Primary Antibody</th>
<th>Median Fluorescence</th>
<th>Ratio of FACs values</th>
</tr>
</thead>
<tbody>
<tr>
<td>GH1</td>
<td>T-ALL</td>
<td>-</td>
<td>50</td>
<td>ng</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UJ13A</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ERIC-1</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>NB100</td>
<td>neuroepithelioma</td>
<td>-</td>
<td>200</td>
<td>1.72</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UJ13A</td>
<td>302</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ERIC-1</td>
<td>259</td>
<td></td>
</tr>
<tr>
<td>IMR-32</td>
<td>neuroblastoma</td>
<td>-</td>
<td>94</td>
<td>1.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UJ13A</td>
<td>245</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ERIC-1</td>
<td>203</td>
<td></td>
</tr>
<tr>
<td>SK-N-SH</td>
<td>neuroblastoma</td>
<td>-</td>
<td>107</td>
<td>1.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UJ13A</td>
<td>380</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ERIC-1</td>
<td>275</td>
<td></td>
</tr>
<tr>
<td>NB1</td>
<td>neuroblastoma</td>
<td>-</td>
<td>89</td>
<td>1.58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UJ13A</td>
<td>278</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ERIC-1</td>
<td>209</td>
<td></td>
</tr>
<tr>
<td>GOTO</td>
<td>neuroblastoma</td>
<td>-</td>
<td>66</td>
<td>1.76</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UJ13A</td>
<td>330</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ERIC-1</td>
<td>216</td>
<td></td>
</tr>
<tr>
<td>PCF</td>
<td>neuroblastoma</td>
<td>-</td>
<td>193</td>
<td>ng</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UJ13A</td>
<td>192</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ERIC-1</td>
<td>195</td>
<td></td>
</tr>
<tr>
<td>Kelly</td>
<td>neuroblastoma</td>
<td>-</td>
<td>147</td>
<td>1.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UJ13A</td>
<td>364</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ERIC-1</td>
<td>262</td>
<td></td>
</tr>
<tr>
<td>SK-N-DZ</td>
<td>neuroblastoma</td>
<td>-</td>
<td>165</td>
<td>1.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UJ13A</td>
<td>250</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ERIC-1</td>
<td>230</td>
<td></td>
</tr>
</tbody>
</table>
Figure 3.1

FACs profiles generated by anti-NCAM antibodies and human neuroblastoma cell lines

Cells were examined by light microscopy, and analysed with a fluorescence activated flow cytometer.

(I) cells incubated with fluorescein-conjugated rabbit anti-mouse Ig alone

(II) cells incubated with anti-NCAM antibody followed by fluorescein-conjugated rabbit anti-mouse Ig
Figure 3.1

Relative fluorescence intensity

Number of cells
confirming previous studies which suggest that GH1 does not express NCAM.

(b) Establishing experimental parameters: positive and negative controls and a titre of monoclonal antibody ERIC-1 for use in Western blotting

Four strong bands corresponding to proteins with molecular weights of approximately 120 - 200 kDa were observed in all samples of human adult brain (see Figure 3.2). While it was difficult to assign precise molecular weights as optimum protein separation was not obtained, they are assumed to correspond with isoforms of 180, 170, 140, and 120 kDa previously observed in human brain tissue (Silberman and Bhat, 1988).

The appearance of samples digested with neuraminidase (lanes 2 and 4) did not differ significantly from native samples (lanes 1 and 5) or a sample subjected to "mock" digestion (lane 3), in agreement with previous findings which suggest that NCAM of adult brain is poorly sialylated (Rougon et al., 1982; Sunshine et al., 1987). No signal was observed on developing adult brain using MoAb M340 as the primary antibody when Western blotting (lane 6).

Adult brain immunoblotted with different dilutions of MoAb ERIC-1 yielded a similar pattern with most dilutions of antibody (blot not displayed). A titre of 1/50 (26 μg/ml) was selected as optimum for the detection of NCAM; this was used in the development of subsequent immunoblots. No signal was observed on developing adult brain with MoAb M340, the proposed negative control.

(c) Western blot analysis of extracts of neuroblastoma, rhabdomyosarcoma, neuroepithelioma and myeloid leukaemia cell lines

All extracts of undigested cell lines developed with MoAb ERIC-1 yielded a smear which ranged in size from approximately 140 - >250 kDa (Figures 3.3 - 3.7). No binding was observed on blotting the same extracts with MoAb M340. Adult brain, used as a positive control for each
Figure 3.2

Establishing a positive control - Western blot analysis of adult brain

Aliquots containing 150 μg of adult brain were separated by electrophoresis and immunoblotted with MoAbs ERIC-1 or M340. Samples and their respective primary antibodies are as indicated below.

<table>
<thead>
<tr>
<th>lane</th>
<th>sample</th>
<th>primary antibody</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>adult brain</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>2</td>
<td>adult brain</td>
<td>ERIC-1</td>
</tr>
<tr>
<td></td>
<td>+ 0.5 U/ml neuraminidase</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>3</td>
<td>adult brain; &quot;mock&quot; digest</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>4</td>
<td>adult brain</td>
<td>ERIC-1</td>
</tr>
<tr>
<td></td>
<td>+ 1.0 U/ml neuraminidase</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>5</td>
<td>adult brain</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>6</td>
<td>adult brain</td>
<td>M340</td>
</tr>
</tbody>
</table>

The molecular weight of NCAM isoforms was determined by comparison with Amersham International Rainbow molecular weight markers, comprising myosin (200 kDa), β-galactosidase (116 kDa), phosphorylase-β (92.6 kDa), bovine serum albumin (69 kDa), ovalbumin (46 kDa) and carbonic anhydrase (30 kDa). The positions of constituent markers are indicated in kDa on the left of the figure; calculated molecular weights (to the nearest 5 kDa) are indicated on the right.
Figure 3.3

Western blot analysis of SK-N-BE (2C) and GOTO neuroblastoma cell lines
Both cell lines were analysed in their native state and after digestion with neuraminidase. Adult brain served as a positive control. Native cell extract incubated with MoAb M340 served as a negative control. Samples and their respective primary antibodies are as indicated below. The molecular weights of NCAM isoforms were calculated using Amersham Rainbow markers as a guide (see Figure 3.2). The latter are indicated in kDa on the left of lanes 1 and 5; the former on the right of lanes 4 and 8.

<table>
<thead>
<tr>
<th>lane</th>
<th>sample</th>
<th>primary antibody</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SK-N-BE(2C)</td>
<td>M340</td>
</tr>
<tr>
<td>2</td>
<td>SK-N-BE(2C)</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>3</td>
<td>SK-N-BE(2C)</td>
<td>ERIC-1</td>
</tr>
</tbody>
</table>
| 4    | + neuraminidase | adult brain      | ERIC-1
| 5    | GOTO            | M340             |
| 6    | GOTO            | ERIC-1           |
| 7    | GOTO            | ERIC-1           |
| 8    | + neuraminidase | adult brain      | ERIC-1
**Figure 3.4**

**Western blot analysis of the Kelly neuroblastoma and NB100 neuroepithelioma cell lines**
Both cell lines were analysed in their native state and after digestion with neuraminidase. Adult brain served as a positive control and native cell extract incubated with MoAb M340 as a primary antibody served as a negative control. Samples and their respective primary antibodies are as indicated below. The molecular weights of NCAM isoforms were calculated using Amersham Rainbow markers as a guide (see Figure 3.2). The latter are indicated in kDa on the left of lanes 1 and 5; the former on the right of lanes 4 and 8.

<table>
<thead>
<tr>
<th>lane</th>
<th>sample</th>
<th>primary antibody</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Kelly</td>
<td>M340</td>
</tr>
<tr>
<td>2</td>
<td>Kelly</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>3</td>
<td>Kelly + neuraminidase</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>4</td>
<td>adult brain</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>5</td>
<td>NB100</td>
<td>M340</td>
</tr>
<tr>
<td>6</td>
<td>NB100</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>7</td>
<td>NB100 + neuraminidase</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>8</td>
<td>adult brain</td>
<td>ERIC-1</td>
</tr>
</tbody>
</table>
Figure 3.5

Western blot analysis of the IMR-32, SK-N-SH and NB1 neuroblastoma cell lines

All cell lines were analysed in their native state and after digestion with neuraminidase. Adult brain served as a positive control and native cell extract incubated with MoAb M340 as a primary antibody served as a negative control. Samples and their respective primary antibodies are as indicated below. The molecular weights of NCAM isoforms were calculated using Amersham Rainbow markers as a guide (see Figure 3.2). The latter are indicated in kDa on the left of lanes 1, 5 and 9; the former on the right of lanes 4, 8 and 12.

<table>
<thead>
<tr>
<th>lane</th>
<th>sample</th>
<th>primary antibody</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>IMR-32</td>
<td>M340</td>
</tr>
<tr>
<td>2</td>
<td>IMR-32</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>3</td>
<td>IMR-32</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>4</td>
<td>adult brain</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>5</td>
<td>SK-N-SH</td>
<td>M340</td>
</tr>
<tr>
<td>6</td>
<td>SK-N-SH</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>7</td>
<td>SK-N-SH</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>8</td>
<td>adult brain</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>9</td>
<td>NB1</td>
<td>M340</td>
</tr>
<tr>
<td>10</td>
<td>NB1</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>11</td>
<td>NB1</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>12</td>
<td>adult brain</td>
<td>ERIC-1</td>
</tr>
</tbody>
</table>
**Figure 3.6**

**Western blot analysis of the JR-1, Rab-1 and RD618 rhabdomyosarcoma cell lines**

Cell lines were analysed in their native state and after digestion with neuraminidase. Adult brain served as a positive control and native cell extract incubated with MoAb M340 as a primary antibody served as a negative control. Samples and their respective primary antibodies are as indicated below. The molecular weights of NCAM isoforms were calculated using Amersham Rainbow markers as a guide (see Figure 3.2). The latter are indicated in kDa on the left of lanes 1, 5 and 9; the former on the right of lanes 4, 8 and 12.

<table>
<thead>
<tr>
<th>lane</th>
<th>sample</th>
<th>primary antibody</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>JR-1</td>
<td>M340</td>
</tr>
<tr>
<td>2</td>
<td>JR-1</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>3</td>
<td>JR-1</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>4</td>
<td>JR-1</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>5</td>
<td>Rhab-1</td>
<td>M340</td>
</tr>
<tr>
<td>6</td>
<td>Rhab-1</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>7</td>
<td>Rhab-1</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>8</td>
<td>Rhab-1</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>9</td>
<td>RD618</td>
<td>M340</td>
</tr>
<tr>
<td>10</td>
<td>RD618</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>11</td>
<td>RD618</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>12</td>
<td>RD618</td>
<td>ERIC-1</td>
</tr>
</tbody>
</table>
Western blot analysis of the Kg1a myeloid leukaemia cell line
The Kg1a cell line was analysed in its native state and after digestion with neuraminidase. Adult brain served as a positive control. Native cell lysate developed with MoAb M340 served as a negative control. Samples and their respective primary antibodies are as indicated below. The molecular weights of NCAM isoforms were calculated using Amersham Rainbow markers as a guide (see Figure 3.2). The latter are indicated in kDa on the left of the figure and the former, on the right.

<table>
<thead>
<tr>
<th>lane</th>
<th>sample</th>
<th>primary antibody</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Kg1a</td>
<td>M340</td>
</tr>
<tr>
<td>2</td>
<td>Kg1a</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>3</td>
<td>Kg1a + neuraminidase</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>4</td>
<td>adult brain</td>
<td>ERIC-1</td>
</tr>
</tbody>
</table>
Figure 3.7
experiment, exhibited variation in degree of sialylation between different samples of adult brain. Where lack of sialylation revealed distinct bands, isoforms of approx. 180, 170, 140 and 120 kDa were observed. A hitherto undetected band of approximately 150 kDa was observed in one sample (Figure 3.3, lane 8). It is not possible to assign the molecular "source" of this band - e.g. precursor, degradation product or novel isoform - without further investigation.

Digestion with neuraminidase revealed distinct isoforms. The neuroblastoma cell lines SK-N-BE(2C), GOTO, Kelly, and the neuroepithelioma cell line NB100 express isoforms of 120 and 140 kDa (See Figures 3.3 and 3.4). The isoform of 180 kDa was observed in cell lines IMR-32, SK-N-SH and NB1 in addition to those of 120 and 140 kDa (See Figure 3.5 and Table 3.2). Minor bands of 110 and 135 are also observed in some cell lines: e.g. Kelly (Figure 3.4, lane 3). A minor band of approximately 175 kDa was observed in the lane containing the desialylated NB100 neuroepithelioma cell line (lane 7), although this was not seen on the replicate. In this particular context, use of the word "minor" indicates that the appearance of these bands is sporadic (between different cell lines and different blots of the same cell line). In addition their "molecular source" has not been characterised. In blots of some cell lines, background smearing was observed after neuraminidase treatment (e.g. Kelly; Figure 3.4, lane 3) preventing the detection of small amounts of high molecular weight proteins, should such species be present.

Rhabdomyosarcoma cell lines gave rise to a similar immunoblot profile to that of some neuroblastoma cell lines - on desialylation, bands of approximately 120 and 140 kDa were observed (See Figure 3.6). Minor bands of 110 and 135 were also apparent in immunoblots of desialylated JR-1. Kg1a, a myeloid leukaemia cell line, was seen to express an isoform of 140 kDa (See Figure 3.7).
Table 3.2

NCAM isoforms identified in human neuroblastoma cell lines
Western blotting analyses of desialylated neuroblastoma cell lines revealed a heterogeneity of expression between cell lines (see Figures 3.3 - 3.5).
Table 3.2

<table>
<thead>
<tr>
<th>Cell Lines</th>
<th>Major NCAM isoforms observed after neuraminidase treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NCAM-180</td>
</tr>
<tr>
<td>NB100</td>
<td>✓</td>
</tr>
<tr>
<td>NB1</td>
<td>✓</td>
</tr>
<tr>
<td>Kelly</td>
<td>✓</td>
</tr>
<tr>
<td>IMR-32</td>
<td>✓</td>
</tr>
<tr>
<td>SK-N-SH</td>
<td>✓</td>
</tr>
<tr>
<td>SK-N-BE(2C)</td>
<td>✓</td>
</tr>
<tr>
<td>GOTO</td>
<td>✓</td>
</tr>
</tbody>
</table>
The 170 kDa form was not detected in any cell line analysed here with the possible exception of NB100 - although the high molecular weight species observed in this cell line could not be demonstrated to be reproducible. Signal characteristic of adult brain was observed in all positive controls; in contrast, no signal was observed in any of the negative controls.

(d) Western blot analysis of neuroblastoma and rhabdomyosarcoma tissue extracts

In most respects, immunoblot analysis of neuroblastoma (stage 4) tissues reveals a similar profile to that observed in cell lines. Extensive smearing was observed in two of the three samples of undigested tissue (See Figure 3.8; lanes 2 and 8) with a lesser degree of smearing in the third (lane 5). Despite sialylation, hazy bands corresponding to protein species of approximately 130 and 145 could be discerned in two of these samples (lanes 2 and 5).

Desialylation by neuraminidase disclosed a faint band of 140 kDa in all tissue samples with stronger bands corresponding with isoforms of 110 and 120 kDa (lanes 3, 6 and 9). An additional isoform of 180 kDa was observed in one sample (lane 9). Thus isoform heterogeneity is observed between the tissue samples, as it is between the cell lines.

Heavy sialylation of rhabdomyosarcoma tissue samples is suggested by the smear observed in blots of untreated samples (Figure 3.9). "Mock" digestion revealed partial desialylation of both samples investigated in this fashion (lanes 3 and 8). This may be due to the fact that polysialic acid is labile in acidic conditions. More complete digestion was effected by the addition of neuraminidase, revealing isoforms of 120 and 140 kDa (lanes 4 and 9). The general pattern of NCAM expression by rhabdomyosarcoma tissue appears to mimic that of cell lines, although the minor isoforms of
Figure 3.8

**Western blot analysis of neuroblastoma tissues**

Tissue samples were analysed in their native state and after digestion with neuraminidase. Blots were developed with MoAbs ERIC-1 or M340 - the latter as a negative control. Samples and primary antibodies are as indicated below. The molecular weights of NCAM isoforms were calculated using Amersham Rainbow markers as a guide (see Figure 3.2). The latter are indicated in kDa on the left of lanes 1, 5 and 9; the former on the right of lanes 4, 8 and 12.

<table>
<thead>
<tr>
<th>lane</th>
<th>sample</th>
<th>primary antibody</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>tissue 1</td>
<td>M340</td>
</tr>
<tr>
<td>2</td>
<td>tissue 1</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>3</td>
<td>tissue 1</td>
<td>ERIC-1</td>
</tr>
<tr>
<td></td>
<td>+ neuraminidase</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>tissue 2</td>
<td>M340</td>
</tr>
<tr>
<td>5</td>
<td>tissue 2</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>6</td>
<td>tissue 2</td>
<td>ERIC-1</td>
</tr>
<tr>
<td></td>
<td>+ neuraminidase</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>tissue 3</td>
<td>M340</td>
</tr>
<tr>
<td>8</td>
<td>tissue 3</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>9</td>
<td>tissue 3</td>
<td>ERIC-1</td>
</tr>
<tr>
<td></td>
<td>+ neuraminidase</td>
<td></td>
</tr>
</tbody>
</table>
Western blot analysis of rhabdomyosarcoma tissues

Two tissue samples were analysed in their native state, after mock digestion, and after digestion with neuraminidase. Samples were developed with MoAbs ERIC-1 or M340 - the latter as a negative control. The molecular weights of NCAM isoforms were calculated using Amersham Rainbow markers as a guide (see Figure 3.2). The latter are indicated in kDa on the left of the figure and the former on the right of lane 5. Samples are as indicated below.

<table>
<thead>
<tr>
<th>lane</th>
<th>sample</th>
<th>primary antibody</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>tissue 1</td>
<td>M340</td>
</tr>
<tr>
<td>2</td>
<td>tissue 1</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>3</td>
<td>tissue 1; &quot;mock&quot;-digested</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>4</td>
<td>tissue 1 + neuraminidase</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>5</td>
<td>adult brain</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>6</td>
<td>tissue 2</td>
<td>M340</td>
</tr>
<tr>
<td>7</td>
<td>tissue 2</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>8</td>
<td>tissue 2; &quot;mock&quot; digested</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>9</td>
<td>tissue 2 + neuraminidase</td>
<td>ERIC-1</td>
</tr>
</tbody>
</table>
approximately 110 and 135 kDa are not observed in Western blots of rhabdomyosarcoma tissues.

(e) **Glycosyl-phosphatidylinositol-anchored NCAM isoforms on the JR-1 rhabdomyosarcoma cell line as revealed by Western blotting**

To further investigate NCAM expression by rhabdomyosarcoma, the JR-1 cell line was digested with phosphatidylinositol-specific phospholipase C (PIPLC); the supernatants and cell lysates were analysed by electrophoretic separation followed by immunoblotting.

As observed in section (3) of this chapter, NCAM expression by native JR-1 cells presents as a high molecular weight, polydisperse smear, presumably arising as a consequence of heterogeneous polysialylation (Figure 3.10, lane 1). Digestion with neuraminidase resolved two major isoforms of 120 and 140 kDa, and a minor one of approximately 135 kDa (lane 2). Considerable smearing remained, suggesting incomplete desialylation. Immunoblotting the supernatants of PIPLC digests (lanes 3 and 5) again revealed polydisperse material. A lesser amount of polysialylated material also appears to be spontaneously released by JR-1 cells (lane 7). However, immunoblots of neuraminidase digests suggest that the released isoforms are different - those arising from unhindered PIPLC digestion are of approximately 125 and 135 kDa in size, while those resulting from inhibited digestion conform more closely to the "standard" isoforms in size (lane 8). Positive and negative controls yielded signal and no signal respectively.

(f) **NCAM expression and cell density**

No major difference was observed between immunoblot profiles of IMR-32 grown to different densities although the quality of the blot was not ideal (data not displayed). Lysates of both cell cultures gave rise to characteristic smearing. On desialylation, three major isoforms were observed in each profile - NCAM-180, NCAM-140, and NCAM-120, as in
**Figure 3.10**

**Western blot analysis of the JR-1 rhabdomyosarcoma cell line and NCAM isoforms released from the cell surface by PIPLC**

JR-1 cell lysates were prepared from both native cells and cells digested with neuraminidase. Supernatants aspirated from cells digested with Low's PIPLC, commercial PIPLC, and the latter in the presence of mercuric chloride. These were also analysed. Prior to electrophoresis, samples of each supernatant were digested with neuraminidase. Adult brain was used as a positive control. JR-1 lysate developed with MoAb M340 served as a negative control. The molecular weights of NCAM isoforms were calculated using Amersham Rainbow markers as a guide (see Figure 3.2). The latter are indicated in kDa on the left of lanes 1, 5 and 9; the former on the right of lanes 4, 8 and 12. Samples are as indicated below.

<table>
<thead>
<tr>
<th>lane</th>
<th>sample</th>
<th>1° antibody</th>
<th>lane</th>
<th>sample</th>
<th>1° antibody</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>JR-1</td>
<td>ERIC-1</td>
<td>6</td>
<td>as in lane 5 + neuraminidase</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>2</td>
<td>JR-1 + neuraminidase</td>
<td>ERIC-1</td>
<td>7</td>
<td>supernatant of JR-1 + commercial PIPLC in the presence of HgCl₂</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>3</td>
<td>supernatant of JR-1 digested with Low's PIPLC</td>
<td>ERIC-1</td>
<td>8</td>
<td>as in lane 7 + neuraminidase</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>4</td>
<td>as in lane 3 + neuraminidase</td>
<td>ERIC-1</td>
<td>9</td>
<td>adult brain</td>
<td>ERIC-1</td>
</tr>
<tr>
<td>5</td>
<td>supernatant of JR-1 digested with neuraminidase</td>
<td>ERIC-1</td>
<td>10</td>
<td>JR-1</td>
<td>M340</td>
</tr>
</tbody>
</table>
Figure 3.5, lane 3. A profile typical of adult brain was observed in the positive control; no signal was observed in the negative control.

3.4 Discussion

(a) NCAM expression on selected neuroblastoma cell lines and a neuroepithelioma cell line as determined by fluorescence-activated flow cytometry

The fact that MoAb UJ13A consistently yields a stronger fluorescence signal suggests that it either binds with greater avidity than MoAb ERIC-1, or alternatively, that the epitope of MoAb ERIC-1 is partially masked (see Table 3.1). Extrapolation of these results illustrates why negative results of FACs analysis should be interpreted with care; the absence of epitope and consequent fluorescence does not necessarily equate with the absence of antigen. This is especially relevant to the investigation of alternatively-spliced molecules, which can assume different conformations according to splicing pattern. It is therefore of interest to identify antibody epitopes, and to characterise the way in which they may be modified by post-transcriptional and post-translational modifications. Use of panels of antibodies known to recognise different epitopes of the same antigen, and parallel molecular analyses safeguard against "false" negatives.

That both MoAbs UJ13A and ERIC-1 bind most of the cell lines analysed here suggests that the panel of cell lines is appropriate for further investigation of NCAM expression, with the possible exception of PCF. Clinical testimony to this finding is the fact that MoAb UJ13A has been used in conjunction with other antibodies in the identification of neuroblastoma cells in bone marrow and their subsequent removal (Kemshead et al., 1986).

An immunohistochemical study of a wide variety of tumour and normal tissue samples also confirms the findings presented here: 16/16 neuroblastoma tissue sections examined were found to be strongly positive (Garin-Chesa et al., 1991). Figarella-Branger et al., (1990), Molenaar et al.,
(1991) and Mechtersheimer et al. (1991) also detected NCAM on all neuroblastoma samples tested, in contrast with Feickart et al. (1989) who detected the antigen in only 6/11 neuroblastoma tissues. In this context, it should be noted that variable results are sometimes obtained by different investigators using the same antibodies - a problem recognised by attendants of leucocyte antigen workshops.

Where the molecular topology of NCAM is concerned it is interesting that the ratio of FACs values for MoAbs ERIC-1 and UJ13A varies from cell line to cell line. For instance, the ratio of binding between MoAbs UJ13A and ERIC-1 and the Kelly neuroblastoma cell line is 1.89, compared with that of 1.31 in the case of SK-N-DZ. This suggests that the epitopes of MoAbs ERIC-1 and UJ13A are variably exposed, which in turn infers a difference in "net" topology of NCAM isoforms expressed by the two cell lines.

(b) Western blot analysis of extracts of neuroblastoma, rhabdomyosarcoma, neuroepithelioma and myeloid leukaemia cell lines

Generally, extracts of cell lines were found to immunoblot less effectively than adult brain tissue. Indeed, it was difficult to obtain blots of good quality. Extracts from cell lines were considerably more viscous than tissue extracts and often prone to distortions during vertical electrophoresis. Factors contributing to irregularities in vertical separation may also influence the horizontal transfer of proteins from gel to nitrocellulose. DNase was added to some samples in attempt to reduce viscosity, although this seemed to have little effect on the integrity of the bands observed (data not presented). Several other measures might have helped. These include immunoprecipitation of NCAM prior to gel electrophoresis or purification of a "plasma membrane fraction". Comparison of immunoblots of whole cell lysates and membrane extracts might also clarify the "status" of minor isoforms - e.g. whether or not they are precursors of the major isoforms.
The intensity of signal generated by adult brain is greater than that generated by cell lines. It is possible that this is due to differences in glycosylation, splicing, other differences which generate epitope diversity, or a greater concentration of NCAM per mg of total cellular protein. However, the assumption that intensity of signal correlates with quantity of antigen has not been verified experimentally in this study.

Extensive smears detected by Western blotting extracts of undigested cell lines suggest that all the samples analysed here exhibit the sialylated form of the molecule. These results were corroborated by a similar staining pattern obtained on blotting SK-N-BE(2C) lysate with a monoclonal antibody recognising polysialic acid (data not displayed). The size of the smear obtained with the anti-PSA antibody suggests that the majority of PSA detected here is in association with NCAM and not the α sub-unit of sodium channels - the latter association has been implicated in rat brain (Zuber et al., 1992).

Heavy polysialylation of NCAM expressed by neuroblastoma tissues and the IMR-32 cell line has been observed by other investigators (Lipinski et al., 1987; Figarella-Branger, et al. 1990). Other tumours have been shown to exhibit heavy polysialylation of NCAM (Figarella-Branger et al., 1990; Zuber and Roth, 1990; Rygaard et al., 1992). The heterogeneous isoform complement is consistent with mRNA profiles obtained for these cell lines (K. Patel, personal communication), all of which have been found to include species of 6.7 and 5.4 kb, believed to encode NCAM-140 and NCAM-120 isoforms respectively (see Table 3.3). Those cell lines found to express NCAM-180 also harbour an mRNA transcript of 7.4 kb. An additional weak band of 5.0 kb is observed in some of the cell RNA extracts, although its significance remains unclear. Its presence does not correlate with the minor isoforms observed on Western blotting.
Table 3.3

NCAM mRNA identified in human neuroblastoma cell lines
The data presented in this table was obtained by Northern blotting RNA extracted from the cell lines indicated. The experiments were performed by others.
### Table 3.3

<table>
<thead>
<tr>
<th>Cell Lines</th>
<th>NCAM RNA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7.4 kb</td>
</tr>
<tr>
<td>NB100</td>
<td>✓</td>
</tr>
<tr>
<td>NB1</td>
<td>✓</td>
</tr>
<tr>
<td>Kelly</td>
<td>✓</td>
</tr>
<tr>
<td>IMR-32</td>
<td>✓</td>
</tr>
<tr>
<td>SK-N-SH</td>
<td>✓</td>
</tr>
<tr>
<td>SK-N-BE(2C)</td>
<td>✓</td>
</tr>
<tr>
<td>GOTO</td>
<td>✓</td>
</tr>
</tbody>
</table>
While the molecular weight of NCAM isoforms as expressed by undigested and digested cell lines are consistent with the assumption that the neuraminidase "type X" sequentially hydrolyses α2-8 homopolymers which characterise "embryonic" NCAM, the lack of positive control containing α2-8 linked sialic acid residues necessitates a degree of caution in the interpretation of data presented in Figures 3.2 - 3.10. A more specific neuraminidase is "endo N" - a bacterial enzyme which selectively cleaves long chain length α2-8 linked PSA leaving approximately five sialic residues (Vimr et al., 1984). Tang et al. conducted a comparative study of endo N and an exoneuraminidase (of unspecified origin) using embryonic chick spinal cord as a substrate and monoclonal antibodies specific for NCAM and α2-8 linked PSA. Their results support the assumption that isoforms of 120, 140 and 180 kDa observed in Figures 3.2 - 3.10 are visualised as a consequence of desialylation, and suggest that this is largely, but not entirely due to cleavage of α2-8 linked sialic acid residues.

Collectively, neuroblastoma cell lines are seen to be capable of expressing all isoforms expressed by adult brain with the exception of NCAM-170; perhaps this isoform is restricted to brain or the CNS. The data presented here suggests that NCAM-140 is the predominant isoform found in neuroblastoma. This is consistent with the fact that NCAM-140 is the most primitive NCAM isoform - it is detected at the earliest stages of development in the human. It is also consistent with characterisation of NCAM isoforms in murine neuroblastoma (Pollerberg et al., 1985) and other tumour tissues and cell lines in which NCAM is expressed (Figarella-Branger et al., 1990; Lanier et al., 1991; Møller et al., 1992).

The isoform profile of IMR-32 is consistent with that observed by Lipinski et al., 1987, but contrasts with that determined by Molenaar et al., 1991 who documented isoforms of approximately 130, 160 and 200 kDa while characterising a panel of antibodies raised against small cell lung carcinoma. It may be the higher m. wts. recorded by this group arise from incomplete desialylation. Indeed it appears that the authors made no attempt to enzymatically digest the cells. It is therefore odd that the three isoforms should be distinctly visible. It may be that the antibody used in this assay only detects lightly sialylated isoforms; that its epitope is masked by heavier sialylation. Figarella-Branger et al., (1990) immunoblotted a panel of neuroblastoma tumours and also observed extensive polysialylation. Unfortunately incomplete desialylation obscured underlying isoform expression.

The rhabdomyosarcoma cell lines in this study express isoforms of similar size to the established neural complement. On the basis of previous
work, one would expect major isoforms of 125 and 145 KDa instead of isoforms of 120 and 140 kDa. However, the resolution of proteins which differ by 5 kDa by a 7% polyacrylamide gel can be imprecise. It would thus be unwise to conclude that the major muscle NCAM isoforms observed by others are not expressed by rhabdomyosarcoma cells. Other factors might give rise to the detection of NCAM-120 and NCAM-140 in rhabdomyosarcoma cell lines and tissue. Different glycosylation patterns between normal and transformed cells may account for the difference in molecular weights of their respective NCAM isoforms. Furthermore, the discovery of NCAM isoforms of novel size (i.e. deviating from the "classical" muscle profile as described by Moore et al., 1987) in tissue derived from muscle would not be surprising as the size of isoforms characterised in muscle tissue does vary (Al-Mahdawi et al., 1990; Tassin et al., 1991; Figarella-Branger et al., 1992).

The finding that the Kg1a myeloid leukaemia cell line expresses a predominant isoform of 140 kDa is consistent with a contemporaneous study by Lanier et al., 1991.

Of the neuroblastoma cell lines which express NCAM, two groups are apparent with respect to the major isoforms - those that express NCAM-180, and those that do not. Previous work suggests that the former group may be more highly differentiated than those lacking NCAM-180. Immunoblots revealing NCAM protein expression by the latter group appear to be indistinguishable from those of rhabdomyosarcoma cell lines, rendering their differentiation impossible by this means.

(c) Western blot analysis of neuroblastoma and rhabdomyosarcoma tissue extracts

Generally, tissue samples appear to have a similar NCAM profile to that of cell lines with respect to sialylation and expression of major isoforms. The neuroblastoma tissue samples exhibit minor variations. Firstly, in 2/3 samples, bands of lower molecular weight could be discerned in native samples, suggesting a lesser degree of sialylation than that of cell lines. The disappearance of these bands on digestion with neuraminidase suggests that they are sialylated species of NCAM. The different techniques
used in lysate and homogenate preparation for cells and tissues respectively cannot be excluded as the cause of this variation.

Secondly, assuming that intensity of signal is proportional to the quantity of isoforms present, the neuroblastoma tissues appear to contain a higher amount of the isoforms of lower molecular weight: NCAM-120 and NCAM-110. It is difficult to comment on this difference without knowing more about the source and characteristics of NCAM-110. If a secreted product, it is probably prevalent in the extracellular matrix and may therefore be better represented in tissue sections than in cell lines. This may also be true of NCAM-120, some of which may be present as a soluble extracellular matrix component. If NCAM-110 is a degradation product - possibly arising from proteolysis (Cunningham et al., 1987) - its levels may be affected by handling after excision of tumour tissue.

That NCAM appears to be expressed in a similar if not identical fashion by tissue and cell lines, is reassuring to those who seek to characterise tumour biology by investigating the characteristics and behaviour of cell lines. These are often assumed to be representative of the tumour from which they are derived. My results accord with those of Rygaard et al. (1992) who investigated NCAM expression on small cell lung carcinoma cell lines and derived xenografts. They observed no qualitative difference between cell lines and xenografts.

NCAM expression on 3/3 rhabdomyosarcoma cell lines and 2/2 tissues examined is consistent with the concept that NCAM is an oncodevelopmental antigen and the fact that it is found on fetal muscle cells (Al-Mahdawi et al., 1990). In accordance with this finding, Mechtersheimer et al. (1991) noted NCAM in all rhabdomyosarcoma tissues examined.
(d) Glycosyl-phosphatidylinositol-anchored NCAM isoforms on the JR-1 rhabdomyosarcoma cell line as revealed by Western blotting

It is difficult to rationalise some of the results of this experiment. It comes as no surprise that NCAM is both spontaneously and specifically released into the extracellular environment; these capacities have been documented previously (He et al., 1987; Theveniau et al., 1991; Figarella-Branger et al., 1992). The finding that released NCAM is polysialylated, however, is controversial. Figarella-Branger et al., (1992) have concluded that GPI-anchored NCAM is never sialylated while studies by Knudsen et al. (1989, 1990) suggest that it is.

That PIPLC should release isoforms of approximately 125 and 135 kDa is surprising considering that the myoblast expresses a predominant transmembrane isoform of 140 kDa and an isoform of 120/125 kDa, believed to be GPI-linked (Walsh et al., 1989; Moore et al., 1987). Perhaps the GPI-anchored bands of approximately 125 and 135 kDa revealed in this experiment derive from the minor and major isoforms of approximately 135 and 140 kDa respectively. This hypothesis assumes that at least a subpopulation of NCAM-135 and 140 observed in extracts of native JR-1 cells is linked via a PI-anchor. Previous experiments have failed to identify such PI-linked isoforms (Moore et al., 1987; Knudsen et al., 1990). Alternatively, the species of approximately 125 kDa may arise from the degradation of the 135 kDa species.

(e) NCAM expression and cell density

NCAM expression was investigated on two cultures of the same cell line grown to different densities. IMR-32 was chosen as the cell line with which to investigate a possible change in expression, as it expresses all three major isoforms.

NCAM is a cell surface molecule which is subject to intricate developmental regulation and appears to affect development accordingly.
It would therefore seem likely that its expression might be modulated in response to environmental cues - e.g. cell density. If this is the case, it is not apparent from these results. It may be that this assay is insufficiently sensitive to detect such changes. For instance, small changes in isoform density might not be detected without sensitive quantitation of signal arising from the different isoforms. In addition, caution must be exercised in the interpretation of these results, as no replicate experiment was performed.

These results appear to contrast with those obtained by Mayerhofer et al., 1992a, who observed an increase of NCAM expression and de novo synthesis of NCAM-140 and NCAM-180 after four days' culture of mouse Leydig cells. In contrast, NCAM protein was not detected after two days, at which the cell cultures were presumably less dense. That Leydig cells are "normal", in contrast with the transformed cells analysed here, may account for this difference.

It is possible that changes in alternative splicing within the extracellular domain of NCAM may arise in response to changes in environment. However, the slight changes in molecular weight effected by alternative splicing in the extracellular domain are inconsistent with detection by Western blotting.

### 3.5 Conclusion

Where the major isoforms of NCAM are concerned, the cell lines of rhabdomyosarcoma, neuroblastoma, neuroepithelioma, and myeloid leukaemia demonstrated no marked difference. All were found to express heavily sialylated NCAM. Neuraminidase digestion did not reveal tumour-specific patterns of isoform expression - i.e. alternative splicing patterns of exons 15 and 18 did not show a consistent difference between the cell lines. However, the diversity of NCAM isoform expression is potentially greater than indicated above. Two patterns of alternative splicing in the
extracellular portion of the molecule have been described as being specific either to neural or muscle tissue (Cunningham et al., 1987; Moore et al., 1987). Investigation of the splicing patterns at these junctions is appropriate in the quest for a molecular means of differentiation.
CHAPTER 4

EXPRESSION OF THE VARIABLE ALTERNATIVELY SPliced EXON (VASE) IN NEUROBLASTOMA AND RHABDOMYOSARCOMA CELL LINES

4.1 Introduction

VASE, a 30 bp exon, is alternatively-spliced at the splice junction of exons 7 and 8 (see Figure 4.1). The inclusion of VASE in NCAM mRNA has been demonstrated to down-regulate neurite outgrowth (Doherty et al., 1992b). It must therefore influence the interaction of NCAM with either itself or other molecules. Several characteristics of the VASE-encoded polypeptide recommend it as a modulator of binding strength or specificity. The VASE peptide sequence is hydrophilic in character, suggesting that it protrudes from the surface of the Ig-fold into the aqueous environment. Secondly when NCAM and immunoglobulin sequences are aligned, the position of the exon 7/8 splice site in NCAM partly coincides with a hypervariable region in antibodies (Small et al., 1988). As the latter is thought to be involved in specific antigen binding, it may be that the VASE-encoded polypeptide, by virtue of its analogous position within NCAM, has the potential to influence intermolecular contact.

VASE was originally characterised in a cDNA clone from rat brain (Small et al., 1987) and later demonstrated to lie within 10.5 kb of sequence which separates exons 7 and 8 in NCAM DNA (Small et al., 1988). Its presence has also been demonstrated in mouse brain RNA (Santoni et al., 1989b), and its sequence found to coincide with that found in rat NCAM. Initially, its inclusion was thought to be restricted to the nervous system, and as such, subject to rigid tissue-specific control. Since its discovery, a variety of neural and muscle cell lines and tissues have been analysed to
Figure 4.1

a) Rodent sequence of VASE with flanking sequences and corresponding amino acids
The underlined sequence is that of VASE (Small et al., 1987; Santoni et al., 1989).

b) Schematic representation of a 140 kDa polypeptide encoded by RNA containing VASE
The five immunoglobulin-like domains are in the amino-terminal half of the molecule, as illustrated. Two fibronectin type III-like repeats are found in the extracellular portion, proximal to the plasma membrane. VASE is situated in the fourth domain.

Key:
- = plasma membrane
- = constitutively-spliced NCAM polypeptide
- = VASE-encoded polypeptide
- = fibronectin type III like repeat
- = immunoglobulin-like domain
Figure 4.1

a)  

<table>
<thead>
<tr>
<th></th>
<th>exon 7</th>
<th>exon 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>5'</td>
<td>GAA AAG CCA TCG TGG ACT CGA CCA GAG AAG CAA GAG ACT CTA</td>
<td>3'</td>
</tr>
</tbody>
</table>

Glu Lys Ala Ser Trp Thr Arg Pro Glu Lys Gin Glu Thr Leu

b)
test this hypothesis (see Table 4.1), and VASE is now known to be present in NCAM isoforms found in rat heart tissue (Reyes et al., 1991). However, little if any was detected in myoblast and neural cell lines or tissues of the peripheral nervous system (Small and Akeson, 1990). Furthermore, its varying levels of inclusion in different neural tissues and at different stages of development, suggest spatial and temporal control of alternative splicing. However, with respect to the major size classes of neuronal NCAM, (i.e. NCAM-120, -140, and -180) alternative splicing of VASE is thought to be stochastic (Andersson et al., 1990b, Reyes et al., 1993).

It was therefore of interest to examine the splicing patterns at the exon 7/8 splice junction in neuroblastoma and rhabdomyosarcoma cell lines. Differences in splicing patterns between the two types of cell line might assist in their differential diagnosis.

### 4.2 Southern Blotting of PCR-amplified cDNA Spanning the Exon 7/8 Splice Junction

Alternative splicing, by virtue of its tissue-specific regulation, may be a means by which to distinguish tumours of different lineages. The alternative splicing pattern at the exon 7/8 splice junction of eight neuroblastoma cell lines and one rhabdomyosarcoma cell line was investigated by PCR amplification of DNA complementary to RNA spanning this region.

(a) Methods

RNA was extracted from cell lines using guanidium isothiocyanate. Two extractions were made of each cell line growing exponentially. Each line was resuscitated from two different stock cultures.

Complementary DNA was generated by reverse transcription of RNA template, using an oligonucleotide primer complementary to a region in exon 8, downstream from the exon 7/8 splice junction (Figure 4.2.1). The primer was designed according to recommendations (Innis and Gelfand,
Table 4.1

VASE expression in human cell lines and distinct rat brain regions
The data presented summarizes results obtained by Small and Akeson, 1990. Extent of VASE inclusion was assessed according to the intensity of signal and is a comparative measure of VASE-containing PCR products, measured against VASE-lacking PCR products.
Table 4.1

<table>
<thead>
<tr>
<th>Cell Line</th>
<th>Tissue Type</th>
<th>VASE Inclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>Basal Ganglia</td>
<td>+</td>
</tr>
<tr>
<td>-</td>
<td>Brain Stem</td>
<td>+</td>
</tr>
<tr>
<td>-</td>
<td>Cerebellum</td>
<td>++</td>
</tr>
<tr>
<td>-</td>
<td>Cortex</td>
<td>++</td>
</tr>
<tr>
<td>-</td>
<td>Mid-brain</td>
<td>++</td>
</tr>
<tr>
<td>-</td>
<td>Hippocampus</td>
<td>++</td>
</tr>
<tr>
<td>-</td>
<td>Thalamus</td>
<td>+++</td>
</tr>
<tr>
<td>-</td>
<td>Olfactory epithelium</td>
<td>+/-</td>
</tr>
<tr>
<td>-</td>
<td>Spinal cord</td>
<td>+</td>
</tr>
<tr>
<td>-</td>
<td>Dorsal root ganglia</td>
<td>+</td>
</tr>
<tr>
<td>-</td>
<td>Adrenal</td>
<td>+/-</td>
</tr>
<tr>
<td>-</td>
<td>Embryonic heart</td>
<td>+/-</td>
</tr>
<tr>
<td>-</td>
<td>Adult Heart</td>
<td>+++</td>
</tr>
<tr>
<td>B35</td>
<td>Neuronal</td>
<td>+/-</td>
</tr>
<tr>
<td>B104</td>
<td>Neuronal</td>
<td>+/-</td>
</tr>
<tr>
<td>RN22</td>
<td>Schwanoma</td>
<td>+/-</td>
</tr>
<tr>
<td>C6</td>
<td>Glial</td>
<td>+/-</td>
</tr>
<tr>
<td>B12</td>
<td>Glial</td>
<td>+/-</td>
</tr>
<tr>
<td>L6</td>
<td>Muscle</td>
<td>+/-</td>
</tr>
<tr>
<td>PC12</td>
<td>Phaeochromocytoma</td>
<td>+/-</td>
</tr>
</tbody>
</table>
Figure 4.2.1

Exons 7 and 8 and the positions of primers and the non-VASE probe
Partial sequence of exons 7 and 8; the respective 5' and 3' primers used in cDNA synthesis and amplification by PCR, are framed. The non-VASE oligonucleotide is complementary to that part of exon 7 which is underlined.
Figure 4.2.1

5' Exon 7
AAA CCC AAA ATC ACA TAT GTA GAG AAC CAG ACT

5' primer
GCC ATG GAA TTA GAG GAG CAG GTC ACT CTT ACC TGT

5' Exon 8
GAA GCC TCC GGA GAC CCC ATT CCC TCC ATC ACC TGG

AGG ACT TCT ACC CGG AAC ATC AGC AGC GAA GAA AAG

Exon 8
ACT CTG GAT GGG CAC ATG GTG GTG CGT AGC CAT GCC

CGT GTG TCG TCG CTG ACC CTG AAG AGC ATC CAG TAC

ACT GAT GCC GGA GAG TAC ATC TGC ACC GCC AGC 3'

CTC ATG TAG ACG TGG CGG

← 3' Primer
which maximise the probability of efficient, yet specific hybridisation with template. The sequence of mouse NCAM was also taken into account, as the positive control used in these experiments is a clone derived from mouse brain cDNA (Barthels et al., 1988).

The position of primers used in PCR amplification across the exon 7/8 splice junction are indicated in Figure 4.2.1; the 3' primer is identical to that used in the synthesis of cDNA by reverse transcription. Optimal conditions of amplification by PCR were previously established by another member of the laboratory.

An acute lymphoblastic leukaemia cell line, GH1, which had previously been shown to lack NCAM expression by Northern blotting (Patel, K; personal communication) and fluorescence-activated flow cytometry (see Chapter 3) was used as a negative control. An additional negative control consisted of all PCR reagents to which water, instead of cDNA, was added. 20 ng of N1, a cDNA clone in which mouse VASE was originally characterised (Barthels et al., 1988), was used as a positive control.

A 10 µl aliquot of each PCR amplification was electrophoresed through a 2% agarose mini-gel in the presence of 0.5 µg/ml ethidium bromide, to ensure that the reaction had been successful in generating products. On visualisation of bands of appropriate sizes, 10 µl aliquots of each PCR amplification were electrophoresed through a large 2% agarose gel (20 x 20cm), and transferred to a nylon membrane.

As polyacrylamide gels were latterly used as a means to isolate PCR products, the former were also used to prepare Southern blots. 50 µl aliquots of PCR products were electrophoresed through an 8% polyacrylamide gel, and transferred to a nylon membrane by horizontal, semi-dry electrophoresis.
Southern blots were obtained on hybridising immobilised PCR products with a $^{32}$P labelled oligonucleotide complementary to rodent VASE, referred to in the rest of the text as the "VASE oligonucleotide". The melting temperature of the VASE oligonucleotide was calculated to be 62°C (Wahl et al., 1987) and so a hybridisation temperature of 55°C was chosen. Filters were washed to a final stringency of 0.1% SDS, 0.1 x SSC at 55°C, and exposed to film. After exposure, filters were stripped of the VASE oligonucleotide and exposed to film for an appropriate length of time, to ensure that any residual oligonucleotide would not bias interpretation of data obtained on subsequent hybridisations.

Filters were rehybridised with the non-VASE oligonucleotide, so called because it hybridises PCR products which lack VASE as well as those which contain it, and is complementary to a region in exon 7 (see Figure 4.2.1). The non-VASE oligonucleotide was hybridised at 60°C; 7°C below its calculated melting temperature. Washes were carried out to a final stringency of 0.1% SDS, 0.1 x SSC, at 60°C. Blots were drained, wrapped in cling film, and exposed to film for 4 - 6 hrs.

(b) Results

Southern blotting of PCR products generated across the exon 7/8 junction, suggests that RNA species, some of which contain or lack VASE, are present in all neuroblastoma and rhabdomyosarcoma cell lines tested.

(i) Agarose gel electrophoresis

On Southern blotting PCR products with the VASE oligonucleotide (Figure 4.2.2) a similar pattern is observed in lanes containing PCR products from the neuroblastoma cell lines (lanes 1 to 8) and the rhabdomyosarcoma cell line JR-1 (lane 10). All revealed a product of approximately 240 bp which hybridised VASE. Fainter bands of approximately 194 and 224 bp were observed in most cell lines; their intensity varied between cell lines and from blot to blot. Other minor bands of approximately 300, 660 and 900

135.
**Figure 4.2.2**

Southern blot of PCR products spanning the exon 7/8 splice site from a selection of cell lines; probed with the VASE oligonucleotide and prepared by agarose gel electrophoresis

**Key:**

|----------------|---------|----------|----------|-------|---------|----------|-------|-------|---------|------------|

1-8 = neuroblastoma cell lines  
9 = T-cell acute lymphoblastic leukaemia cell line; negative control  
10 = rhabdomyosarcoma cell line  
11 = N1 clone containing rodent VASE; positive control
Figure 4.2.2
900 were also observed in some lanes. The positive control, N1 (lane 11), yielded a single detectable PCR product of 224 bases, consistent with the calculated size of a PCR product containing VASE. No bands were observed in either of the negative controls containing water (not shown) or cDNA generated from GH1 (lane 9).

On hybridising the same blot with the non-VASE oligonucleotide, major bands of approximately 194, 224 and 240 bp were observed in all neuroblastoma and rhabdomyosarcoma cell lines analysed. (Figure 4.2.3; lanes 1-8 and 10) Minor bands of approximately 300 and 900 bp were also observed in some cell lines. Inexplicably, no product was observed in the lane containing the positive control (lane 11), and both negative controls remained free of bands (lane 9; water control not shown).

In total, seven different blots were obtained in analysing both sets of PCR products. Each was hybridised with both the VASE and the non-VASE oligonucleotide probes. Results were reproducible, although faint traces of signal were observed in the negative controls of two blots. As a consequence, all PCR reagents were renewed, and stored in a separate building to that in which the reactions were carried out, in an effort to reduce the risk of contamination. This proved to be effective.

(ii) Polyacrylamide gel electrophoresis

PCR products from representative cell lines were also electrophoresed through a polyacrylamide gel, transferred to a nylon membrane, and Southern blotted. Here, the 240 bp band was not detected in the Kelly and NB1 neuroblastoma cell lines nor the JR-1 rhabdomyosarcoma cell line when hybridised with the VASE probe, although a band of 900 bp was observed (Figure 4.2.4, lanes 2-4). When the same blot was hybridised with the non-VASE probe, this pattern was repeated, with additional bands of 660 bp in all cell lines, and one of approximately 240 bp in JR-1 (Figure 4.2.5). In both blots, the positive
Figure 4.2.3

Southern blot of PCR products spanning the exon 7/8 splice site from a selection of cell lines; probed with the non-VASE oligonucleotide and prepared by agarose gel electrophoresis

Key:

1. SK-N-BE (2C)  4. IMR-32  8. PCF
2. GOTO  5. NB1  9. GH1
3. Kelly  6. NB100  10. JR-1

1-8 = neuroblastoma cell lines
9 = T-cell acute lymphoblastic leukaemia cell line; negative control
10 = rhabdomyosarcoma cell line
11 = N1 clone containing rodent VASE; positive control
Figure 4.2.3
Figure 4.2.4

Southern blot of PCR products spanning the exon 7/8 splice site from a selection of cell lines; probed with the VASE oligonucleotide and prepared by polyacrylamide gel electrophoresis

Key:

1. H₂O 3. NB1 5. N1 clone
2. Kelly 4. JR-1 6. m. wt. markers

1 = negative control
2 & 3 = neuroblastoma cell lines
4 = rhabdomyosarcoma cell line
5 = N1 clone containing rodent VASE; positive control
Figure 4.2.5

Southern blot of PCR products spanning the exon 7/8 splice site from a selection of cell lines; probed with non-VASE oligonucleotide and prepared by polyacrylamide gel electrophoresis

Key:

1. H₂O
2. Kelly
3. NB1
4. JR-1
5. N1 clone
6. m. wt. markers

1 = negative control
2 & 3 = neuroblastoma cell lines
4 = rhabdomyosarcoma cell line
5 = N1 clone containing rodent VASE; positive control
control, N1 (lane 5), yielded a single detectable PCR product of 224 bases, a size consistent with a PCR product containing VASE. No bands were observed in the negative control containing water (lane 1).

In total, three different blots derived from polyacrylamide gels were obtained; each successive blot was obtained using increasing quantities of PCR product. 50 µl aliquots of PCR product were loaded into each well of the gel which gave rise to the blot presented in Figures 4.2.4 and 4.2.5. Each blot was probed with the VASE oligonucleotide, and one, with the non-VASE oligonucleotide. Those probed with the VASE oligonucleotide demonstrated minor differences in that the consistency of the bands themselves was variable. Their appearance suggested irregularities in transfer from gel to Genescreen™ by horizontal electrophoresis - a technical problem countered in this instance by increasing the quantity of PCR products applied to the gel.

(c) Discussion

The data presented here demonstrates that alternative splicing of VASE occurs in the RNA of all neuroblastoma and rhabdomyosarcoma cell lines examined. The data is summarised in Table 4.2. Southern blots of seven other rhabdomyosarcoma cell lines and one small cell lung cancer cell line suggested an identical complement of alternatively spliced RNA (A. Hancock; personal communication).

By virtue of its size and blotting profile, the PCR product of 194 bases is believed to comprise parts of exons 7 and 8 as defined by the primers, with no intervening sequence at the splice junction. This conclusion assumes that the minor band of 194 bp, observed on hybridisation of the PCR products of some cell lines with the VASE oligonucleotide (for example, lane 6, Figure 4.2.2) is non-specific. Aberrant migration of a subset of VASE-containing products could also explain the presence of this band. The size and blotting profile of the PCR product of
Table 4.2

PCR products putatively spanning the exon 7/8 splice junction: their sizes and blotting profiles
The data presented below is obtained from Figures 4.2.2 - 4.2.5. Cell types in column (a) are either rhabdomyosarcoma (RMS) or neuroblastoma (NBL). Cell lines which were analysed are listed below. Gel type as listed in column b specifies the type of gel used for electrophoretic separation.

<table>
<thead>
<tr>
<th>cell type</th>
<th>cell line</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBL</td>
<td>SK-N-BE (2C)</td>
</tr>
<tr>
<td></td>
<td>GOTO</td>
</tr>
<tr>
<td></td>
<td>Kelly</td>
</tr>
<tr>
<td></td>
<td>IMR-32</td>
</tr>
<tr>
<td></td>
<td>NB1</td>
</tr>
<tr>
<td></td>
<td>NB100</td>
</tr>
<tr>
<td></td>
<td>SK-N-SH</td>
</tr>
<tr>
<td></td>
<td>PCF</td>
</tr>
<tr>
<td>RMS</td>
<td>JR-1</td>
</tr>
<tr>
<td>a. Cell Type</td>
<td>b. Gel Type</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>NBL &amp; RMS</td>
<td>“</td>
</tr>
<tr>
<td></td>
<td>“</td>
</tr>
<tr>
<td></td>
<td>“</td>
</tr>
<tr>
<td>NBL &amp; RMS</td>
<td>Poly-acrylamide</td>
</tr>
<tr>
<td></td>
<td>“</td>
</tr>
<tr>
<td></td>
<td>“</td>
</tr>
<tr>
<td></td>
<td>“</td>
</tr>
<tr>
<td>RMS</td>
<td>“</td>
</tr>
</tbody>
</table>
224 bases suggests that it contains VASE. Sequence analysis of these bands confirmed their suspected identities (see section 4.3).

The identity of the 240 and 300 bp products observed in agarose gels, and those of 900 and 660 bp observed in both agarose and acrylamide gels, is less easy to assign. If they are assumed to be PCR products which "honestly" reflect RNA, it may be that a novel exon, hitherto uncharacterised, is spliced into the exon 7/8 splice junction alongside VASE. Alternatively, such bands may represent artifacts.

The status of the 240 bp species was rendered questionable by its apparent absence in neuroblastoma cell lines, when selected samples were analysed by means of a polyacrylamide gel (Figure 4.2.4). The 240 bp species seen in the rhabdomyosarcoma cell line JR-1 (Figure 4.2.5, lane 4) suggests novel alternatively-spliced sequence at the exon 7/8 splice junction which is not associated with VASE. Its identical size to the 240 bp band seen in blots derived from agarose gels would hinder its detection therein. Alternatively, it may be analogous to the 240 bp species observed when using an agarose gel although this is inconsistent with their different blotting profiles. The relevance of this band should be viewed with caution, as it has not been demonstrated to be reproducible.

The appearance of what appear to be high molecular weight NCAM PCR products of 660 and 900 bp on hybridising PCR products with the non-VASE oligonucleotide (Figure 4.2.5) is consistent with novel alternatively-spliced sequence. It was of interest to investigate these species by sequence analysis, as blotting data obtained using both polyacrylamide and agarose gels provide reproducible results for the 194 and 224 bp PCR products only. A Southern blot obtained by PAGE (Figure 4.2.5) suggested that the JR-1 rhabdomyosarcoma cell line may be distinguished from neuroblastoma cell lines by virtue of a 240 bp band, revealed on hybridisation with the non-
VASE oligonucleotide probe. It remains a possibility however, that this band is artifactual; sequence analysis should clarify its identity.

That NCAM-positive PCR products derived from PCF neuroblastoma cell line are easily visualised, bears testimony to the sensitive detection of the polymerase chain reaction: NCAM was not detected by FACs analysis on this cell line (see Chapter 3).

All Southern blots described here would suggest the presence of VASE-containing and VASE-lacking PCR products, in RNA from both the neuroblastoma and the representative rhabdomyosarcoma cell line. That such RNA may comprise ribosomal and/or transfer RNA cannot be excluded; an improved experimental protocol would include isolation of poly A mRNA for analysis. Some blots suggest the presence of artifacts or PCR products with novel inserts at the exon 7/8 splice junction, either on their own, or alongside VASE (See Table 4.2). While gel electrophoresis and Southern blotting provide valuable information about the likely identity of PCR products, their true identity can only be ascertained by sequence analysis.

4.3 Direct Sequence Analysis of PCR Products

(a) Methods

(i) Establishing a method

Two aliquots from a sample derived from the neuroblastoma cell line NB1 and containing products obtained from PCR amplification across the exon 7/8 splice junction, were electrophoresed through a 2% low melting agarose gel. Bands of 194 bp were excised from each sample. One gel slice was boiled in TE buffer for 5 min, followed by extraction with phenol/chloroform. Elution of DNA from the other gel slice was achieved by two cycles of freezing, thawing and vortexing in TE buffer.
Eluates were assayed by DNA dipstick, and diluted to 4 ng/μl, whereupon they were re-amplified by PCR using one of four mastermixes. The ratio of 5' primer to 3' primer (1:1, 1:50, 1:75 and 1:100) varied between the mastermixes. A negative control containing water only, was performed with each mastermix.

Resultant "first generation" products were visualised by electrophoresing an aliquot through a 2% agarose gel in the presence of 0.5 μg/ml ethidium bromide. If visible under U.V. light, a 20 μl aliquot of remaining sample was subjected to a second amplification, giving rise to "second generation" PCR products. The latter were obtained using two mastermixes with different primer ratios (1:1 and 1:50), the latter with an excess of 3' primer. A negative control was performed for each mastermix.

"Second generation" PCR products were visualised by electrophoresing 10 μl aliquots through a 2% agarose gel in the presence of 0.5 μg/ml ethidium bromide. On visualising PCR product, remaining volumes were purified and assayed by DNA dipstick analysis. Samples were diluted to 50 pmole/ml, and sequenced using the 5' primer as lead sequence (Winship, 1989).

(ii) Sequence analysis of PCR products

PCR products (20μl) were separated by electrophoresis in an identical manner to that used for Southern blotting. Initially, agarose gels were used to separate products for sequencing. Bands were excised, and the DNA eluted. In some instances, products were isolated from a polyacrylamide gel which permits sharper resolution and is therefore more likely to yield a pure product. In such instances, DNA was obtained from excised gel slices by the "crush and soak" method (Sambrook et al., 1989).

Aliquots of excised, eluted products were electrophoresed through a 2% agarose gel, and Southern blotted using the non-VASE oligonucleotide probe, to confirm their size. Remaining eluate was asymmetrically
amplified by two PCRs with an excess of 3' primer. Thus, predominantly anti-sense PCR products were generated from a single visible band. The products were purified and sequenced (Winship, 1989) using the 5' (sense) primer.

(b) Results

(i) Establishing a method

A method of sequencing PCR products was successfully established by comparing two methods of isolating products from an agarose gel; by boiling and by eluting.

Only eluates gave rise to PCR products which could be visualised on electrophoresis (Figure 4.3.1, lanes 2, 5, and 8). The intensity of observed bands decreased with an increase in 3':5' primer ratio; PCR products arising by amplification with a 50:1 ratio were visible (lane 5) while those generated by a 100:1 ratio could not be seen (lane 11). Extracts obtained by boiling did not give rise to visible products (lanes 1, 4, 7 and 10), nor did the negative controls (lanes 3, 6, 9 and 12).

PCR products arising from eluates amplified symmetrically, with a 1:1 primer ratio (Figure 4.3.1, lane 2), and asymmetrically, with a 3':5' primer ratio of 50:1 (lane 5) were selected for a second amplification, on the basis of their appearance in the ethidium bromide-stained gel. Each sample was amplified in the manner by which it arose; asymmetric products were re-amplified asymmetrically and symmetric products, by symmetric PCR. Products were clearly visible on electrophoresis in the presence of 0.5 μg/ml ethidium bromide. No product was observed in the negative control (data not displayed).

On sequencing two aliquots from each purified product, asymmetric product was found to yield clear sequence, while products generated by symmetric PCR yielded unreadable sequence (Figure 4.3.2). Consequently,
Amplified extracts of the 194bp PCR product; extraction procedures and primer ratios compared
Two different extraction procedures were used to prepare PCR products for sequencing. PCR products were separated by electrophoresis and excised. A suspension of isolated PCR product was obtained, either by boiling the gel slice in the presence of TE buffer, or by elution. Equal quantities of resultant suspension were amplified using varying primer ratios from 1:1 to 1:100, in the latter of which, the 3' primer was in excess.

Key:

1. boiled extract
2. eluate
3. dH₂O
4. boiled extract
5. eluate
6. dH₂O
7. boiled extract
8. eluate
9. dH₂O
10. boiled extract
11. eluate
12. dH₂O
13. 1 kb m. wt. markers

1-3; samples amplified with a 3':5' primer ratio of 1:1
4-6; samples amplified with a 3':5' primer ratio of 50:1
7-9; samples amplified with a 3':5' primer ratio of 75:1
10-12; samples amplified with a 3':5' primer ratio of 100:1
Figure 4.3.2

Sequence of symmetrically and asymmetrically-amplified PCR products compared
Eluates of the 194 bp product generated from primers spanning the exon 7/8 splice junction were amplified, either by two rounds of symmetric amplification of two rounds of asymmetric amplification (see text). Products were purified and sequenced using Winship's method (Winship, 1989).

Key:

A: sequence obtained from symmetrically-amplified PCR products

B: sequence obtained from asymmetrically-amplified PCR products

1 products of sequencing reaction in which ddATP served as the terminating nucleotide

2 products of sequencing reaction in which ddGTP served as the terminating nucleotide

3 products of sequencing reaction in which ddCTP served as the terminating nucleotide

4 products of sequencing reaction in which ddTTP served as the terminating nucleotide
Figure 4.3.2
asymmetric PCR amplification with a 3':5' primer ratio of 50:1 was used in subsequent preparations of PCR products for sequence analysis.

(ii) Sequence analysis of PCR products

Eluates derived from gel slices containing the three major PCR products of 240, 224 and 194 bp migrated as expected (Figure 4.3.3, lanes 5 - 7). Sequencing results generally confirmed the conclusions drawn from Southern blotting data. Each sample was sequenced once, unless specified otherwise. The 194 bp PCR product from two samples derived from the Kelly neuroblastoma cell line was found to comprise exons 7 and 8 juxtaposed; no intervening sequence was found at the exon junction (Figure 4.3.4). The 224 bp PCR product of the Kelly and NB1 neuroblastoma cell lines was found to consist of exons 7 and 8, separated by 30 bases of sequence with 97% homology to rodent VASE (Figures 4.3.5 and 4.3.6). VASE in Kelly and NB1 was found to differ from rodent VASE by one nucleotide. At the third nucleotide position, a thymidine residue replaces an adenine residue. Four samples from the Kelly cell line and one sample from the NB1 cell line were found to contain reproducible VASE sequence.

All sequences presented here were obtained by preparing PCR products with an excess of 3' primer, and sequenced using the 5' primer. Lack of time prevented confirmation of the data by reversing the primers - i.e. generating PCR products with an excess of 5' primer, and using the 3' primer to initiate the sequencing reaction.

In contrast with the equivalent PCR products in the neuroblastoma cell lines, the 224 bp band from JR-1 cell lines yielded sequence which could only be read several bases into the VASE component (Figure 4.3.7). Sequence beyond this point could not be read, due to heavy background on the sequencing film which obscured reading. This problem recurred
Figure 4.3.3

Southern blot of isolated PCR products
Individual PCR products, generated by amplification across the exon 7/8 splice junction, were excised from a gel, as a preliminary step in their preparation for direct sequencing. Eluates were Southern blotted with the non-VASE oligonucleotide probe to confirm their electrophoretic identities.

Key:

1. N1 clone                    5. Kelly; 240 bp product
2. dH2O (negative control)    6. Kelly; 224 bp product
3-4. Kelly                     7. Kelly; 194 bp product
Figure 4.3.3

- 194 bp
- 224 bp
- 240 bp
Figure 4.3.4

Sequence of 194 bp PCR product of the Kelly neuroblastoma cell line

(i) **Sequence**
The 194 bp PCR product was found to lack VASE. The heavy line represents the exon 7/8 splice junction.

**Key:**

A = adenine residue  
C = cytosine residue  
G = guanidine residue  
T = thymidine residue

(ii) **Schematic representation of sequence**

**Key:**

□ = constitutively-spliced NCAM sequence exon as indicated below
Figure 4.3.4

(i)

5'
A GAC CCC ATT CCC TCC ATC ACC TGG AGG ACT
TCT ACC CGG AAC ATC AGC AGC GAA GAA AAG | ACT
CTG GAT GGG CAC ATG GTG GTG CGT AGC CAT GCC
CGT GTG TCG

3'

(ii)

5' ------------ 3'

7  8

151.
Figure 4.3.5

Sequence of 224 bp PCR product of the Kelly and NB1 neuroblastoma cell lines

i) **Sequence**
The 224 bp PCR product was found to contain VASE. The heavy lines represent the exon 7/8 splice junction, while the bold, italic print represents the VASE sequence. See Figure 4.3.4 (i) for Key.

(ii) **Schematic representation of sequence**

**Key:**

- □ = constitutively-spliced NCAM sequence
- □ = exon as indicated below
- ■ = VASE

Figure 4.3.6

**Human and rodent VASE sequences compared**
Comparison of the human and rodent VASE sequences reveals that the third nucleotide differs; this is framed in the figure, and does not effect a change in amino acid sequence. See Figure 4.3.4 (i) for key.
Figure 4.3.5

(i)

5'
CCC TCC ATC ACC TGG AGG ACT TCT ACC CGG AAC
ATC AGC AGC GAA GAA AAG GCT TCG TGG ACT CGA
CCA GAG AAG CAA GAG ACT CTG GAT GGG CAC ATG
GAT GTG CGT AGC XAT GCC CGT GTG TCG TCG CTG
ACC CTG AAG AGC ATC CAG TAC ACT GAT GCC GGA
GAG TAC ATC TGC ACC
3'

(ii)

5' 7 VASE 8 3'

Figure 4.3.6

5'
Human: GCT TCG TGG ACT CGA CCA GAG AAG CAA GAG
Rodent: GCA TCG TGG ACT CGA CCA GAG AAG CAA GAG

Amino acid Sequence: Ala Ser Trp Thr Arg Pro Glu Lys Gln Glu
Figure 4.3.7

**Sequence of 224 bp PCR product of the JR-1 rhabdomyosarcoma cell line**

(i) **Sequence**
The 224 bp PCR product was found to contain the first 5 bases of VASE (V = VASE); sequence beyond this point could not be elucidated. The exon 7/8 splice junction is indicated by a bold vertical line, and bold, italic print represents VASE sequence. See Figure 4.3.4 (i) for Key.

(ii) **Schematic representation of sequence**

**Key:** As previously
Figure 4.3.7

(i)

5'
GAC CCC ATT CCC TCC ATC ACC TGG AGG ACT TCT
ACC CGG AAC ATC AGC AGC GAA GAA AAG | GCT TCG
TGG
3'

(ii)

5' 3'

←7→ V
each time the sequencing reaction was repeated with fresh sample (three times).

The 240 bp band observed in agarose gels (Figure 4.3.8) and the 660 and 900 bp bands observed in polyacrylamide gels (Figure 4.3.9) appeared to consist of at least two products - the 194 bp product and the 224 bp product. This was concluded on examination of the sequence data. Clear sequence was observed which extended to the splice junction, after which two sequencing bands were sometimes observed at the same nucleotide position. Such bands were always consistent with superimposition of the VASE-containing (224 bp) and the VASE-lacking (194 bp) products. Single sequencing bands were observed at positions where the VASE-containing and VASE-lacking products shared the same base. Three samples of the 240 bp PCR product were sequenced. These were derived from the Kelly and NB1 neuroblastoma cell lines and from the JR-1 rhabdomyosarcoma cell line. PCR products of 660 and 900 bp derived from both the Kelly and NB1 neuroblastoma cell lines were sequenced. All sequences derived from PCR products of the same size were found to concur with slight variations in the length of sequence obtained. All sequence data presented here have been obtained from at least two samples.

(c) Discussion

The data presented here, confirms that obtained from Southern blots; that the 194 bp product lacks VASE, and the 224 bp product contains it. Furthermore, the sequencing data has yielded the human VASE sequence, which had not been characterised prior to the commencement of this study. In addition, the data suggests that other bands observed on Southern blots are artifactual, and arise from either heteroduplex formation or aggregation.

Human VASE, as sequenced from the Kelly and NB1 neuroblastoma cell lines (Figure 4.3.4) differs from rodent VASE sequence at the third nucleotide position (Figure 4.3.5). Here, thymidine replaces adenine - a
Figure 4.3.8

Sequence of 240 bp PCR product of the Kelly and NB1 neuroblastoma cell lines and the JR-1 rhabdomyosarcoma cell line

(i) Sequence
The 240 bp PCR product was found to contain exon 7. At the exon 7/8 splice junction, indicated by a bold vertical line, two sequences were apparent, each of which is represented as one part of a fraction. The numerator comprises a part of exon 8 (plain type) while the denominator comprises part of VASE (bold type). See Figure 4.3.4 (i) for Key.

Human VASE =
GCT TCG TGG ACT CGA CCA GAG AAG CAA GAG

(ii) Schematic representation of sequence

Key:

☐ = constitutively spliced sequence
exon as indicated below

☒ = VASE (V) and constitutively spliced sequence

Figure 4.3.9

Sequence of 660 and 900 bp PCR products of the Kelly neuroblastoma cell line

(i) Sequence
The 660 and 900 bp PCR products were found to contain exon 7. At the exon 7/8 splice junction, indicated by a bold vertical line, two sequences were apparent, each of which is represented as one part of a fraction. The numerator comprises a part of exon 8 (plain type) while the denominator comprises part of VASE (bold type). See Figure 4.3.4 (i) for Key.

Key: as in Figure 4.3.8 (ii)
Figure 4.3.8

(i)

5'
CC TGG AGG ACT TCT ACC CGG AAC ATC AGC AGC
GAA GAA AAG

\[ \begin{array}{c}
\text{G} \\
\text{A} \\
\text{C} \\
\text{T} \\
\text{C}
\end{array} \]

GAT TGG

3'

(ii)

5' --------- 3'
<- 7 -> | V+8

Figure 4.3.9

(i)

5'
AC CCC ATT CCC TCC ATC ACC TGG AGG ACT TCT

ACC CGG AAG ATC AGG AGC GAA GAA

\[ \begin{array}{c}
\text{G} \\
\text{G} \\
\text{A} \\
\text{C} \\
\text{T}
\end{array} \]

GAT

TGG ACT

AC ATG

GAG AA

3'

(ii)

5' --------- 3'
<- 7 -> | V+8
difference which does not alter the amino acid sequence. These findings are in agreement with those of Hemperly et al., 1990, who published the human VASE sequence obtained from a cDNA clone of NCAM derived from the Kelly neuroblastoma cell line, during the course of this work. The high degree of homology between rodent and human NCAM VASE sequences - 97% of DNA sequence, and 100% of amino acid sequence - suggests that this region may be of functional importance.

Unlike that of the neuroblastoma cell lines, the JR-1 rhabdomyosarcoma cell line PCR product of 224 bp proved difficult to purify sufficiently for sequencing beyond the first 9 bases of VASE (Figure 4.3.7).

Sequence analysis of the 240 bp band observed in Southern blots derived from agarose gels (Figures 4.2.3 and 4.2.4), and bands of 660 and 900 bp in blots derived from polyacrylamide gels (Figure 4.2.6), suggests that these bands contain the 194 bp VASE-lacking, and the 224 bp VASE-containing PCR products. This was concluded by the superimposition of the beginning of exon 8 and VASE at the exon 7/8 splice junction (Figures 4.3.7 and 4.3.8). Lack of clarity made it impossible to read through the entire VASE sequence on each film where it was apparent that at least two templates were present in the sequencing reaction.

It is curious that while both the 660 and 900 bp PCR products appear to be of hybrid identity according to the sequencing data, only the 900 bp species was seen to hybridise both probes (Figures 4.2.4 and 4.2.5). This discrepancy between the blotting data and the sequencing data is difficult to explain, and suggests that Southern blotting may not always reveal species which are present.

It could be argued that bands of 240, 660, and 900 bp may contain species which harbour a novel exon between VASE and exon 8. However, the Southern blotting data and sequencing data suggests that they may be
artifacts, generated by the formation of heteroduplices or aggregates which
migrate anomalously through a gel. Hybridisation of the VASE-containing
and VASE-lacking products might give rise to a looped structure (Figure
4.3.10) which would migrate more slowly than homoduplices of its
constituent strands.

Previous work suggests that PCR products can give rise to artifacts
when electrophoresed. This is observed in contemporaneous investigations
of alternative splicing in NCAM using PCR amplification; electrophoresing
products through a denaturing polyacrylamide gel has been observed to
eliminate artifact(s) (Small and Akeson, 1990; Zorn and Krieg, 1991).

While direct sequencing has confirmed and clarified much of the
data obtained from Southern blots, it has failed to yield conclusive data
regarding the representative rhabdomyosarcoma cell line JR-1. As high
background observed on sequencing films suggested a technical difficulty
in template purification, an alternative strategy was decided upon. This
involved cloning the PCR products, to enable the preparation of “clean”
template and consequent facilitation of effective sequencing.

4.4 Cloning of PCR Products from the JR-1 Rhabdomyosarcoma
Cell Line: Generation of Clone sets 1 and 2

The efficiency of direct sequencing depends on a homogeneous
template; a requisite which presents a technical challenge when presented
with a mixed population of PCR products. JR-1 PCR products were cloned
into a plasmid vector, in attempt to produce homogeneous template and
thus facilitate sequencing.
Figure 4.3.10

Potential structure of a VASE-related PCR-generated artifact
Partially single-stranded species may result from hybridisation between VASE-containing and VASE-lacking PCR products; such products would migrate more slowly through a gel, in comparison with homoduplices of its constituent strands.

Key:

--- = Constitutively spliced exons 7 and 8

----- = VASE
Figure 4.3.10
(a) Methods

Two populations of PCR products were cloned. The first comprised random, "parental generation" PCR products - i.e. products generated directly from cDNA resulting from reverse transcription. Two molar ratios of plasmid vector : PCR product were used: 1:1 and 1:3. Resultant clones were termed clone set 1, so called because they were the first set of clones constructed for the purpose of examining the splicing patterns at the exon 7/8 splice site. A negative control containing ligation reagents and plasmid only was incubated alongside the other reactions, providing a reference by which to estimate the frequency of false positive bacterial colonies.

The second population comprised non-random "first generation" PCR products and gave rise to clone set 2. These were generated by extracting the 240 bp product from a (polyacrylamide) gel slice by the "crush and soak" method, and re-amplifying by 30 cycles of PCR under the same conditions as those used to generate "parental generation" PCR products. Products were cloned at plasmid : PCR product molar ratios of 1:1, 1:3 and 1:10. Again, a negative control for the cloning reaction containing ligation reagents and plasmid only, was incubated alongside the other reactions.

A cloning kit specifically designed for the amplification of PCR products was used. The kit includes a plasmid, engineered in such a way as to specifically favour ligation with PCR products (Mead et al., 1991). This was achieved by constructing a plasmid containing an origin of replication, a gene encoding kanamycin resistance, and a gene encoding β-galactosidase, with two inverted HphI restriction enzyme sites positioned in the latter (Figure 4.4.1). The HphI sites are engineered in-phase with the β-galactosidase gene (lacZ') and oriented so that a single 3' deoxythymidylate extension remains after restriction digestion. It has been observed that Taq polymerase catalyses the addition of a single dATP residue to the 3' termini of a blunt-ended DNA duplex (Clark, 1988); the
Figure 4.4.1

Structure of the pCR1000 plasmid
(a) = the origin of replication, (b) = gene encoding X-galactosidase, and (c) = gene conferring resistance to kanamycin. Asterisks indicate sites at which the restriction enzyme Not-1 digests; complete digestion yields fragments of 150, 350 and 2400 bp. Sequence flanking the insertion site, with two overhanging adenylate residues, is depicted with reference to its position in the plasmid. "Universal" and "Reverse" primers used for sequencing are indicated above the approximate region of the plasmid to which they are complementary.
plasmid described above was designed with this in mind. Thus, PCR products which have deoxyadenylate extensions from the 3' ends of each terminus should complement the insertion site of the plasmid vector.

Transformation of *E. coli* was effected by heat shock in the presence of rubidium salts, according to instructions supplied by the manufacturer. Positive transformants were selected on the basis of colour. White colonies, being unable to metabolise X-galactose, theoretically harbour plasmids in which the X-galactosidase gene has been disrupted. This can occur by insertion of a PCR product, giving rise to true positive colonies. Alternatively, re-ligation of the plasmid ends resulting in an altered reading frame, gives rise to false positive colonies.

Positive and negative clones (the latter served as negative controls) were isolated and cultured. Episomal DNA was extracted, and digested with the Not-1 restriction enzyme. Not-1 digests the native plasmid at three sites, yielding fragments of 2400, 350 and 150 bp. The fragment of 350 bp contains the insertion site, and so the size of inserted PCR product can be estimated by gauging the size of restriction fragments on electrophoretic separation through a 1% agarose gel in the presence of ethidium bromide. Restriction enzyme digests were observed under U.V. light. A further step was taken to ensure that samples chosen for sequencing were suitable for analysis. The products were transferred to a nylon membrane and Southern blotted with VASE and non-VASE probes.

(b) Results

Initially, random PCR products derived directly from RNA of JR-1 rhabdomyosarcoma cells were used for cloning, giving rise to clone set 1. The cloning efficiency (5 and 6% for reactions using a vector : insert ratio of 1:1 and 1:3 respectively) did not vary significantly between the two ligation reactions, nor between these and the negative control (see Table 4.4.1).
Table 4.4.1

Colonies arising from the cloning of random PCR products; clone set 1
Random, parental generation PCR products generated from RNA extracted from the JR-1 rhabdomyosarcoma cell line were ligated with the pCR™1000 vector at 3 different vector : insert molar ratios: 1:0 (negative control), 1:1, and 1:3. The number of positive and negative colonies arising from transformation of E. coli by products of each ligation reaction are indicated.
Table 4.4.1

<table>
<thead>
<tr>
<th>Plasmid : PCR product ratio (molar)</th>
<th>1:0 (negative control)</th>
<th>1:1</th>
<th>1:3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>negative</td>
<td>positive</td>
<td>negative</td>
</tr>
<tr>
<td>number of colonies</td>
<td>181</td>
<td>11</td>
<td>159</td>
</tr>
<tr>
<td>mean</td>
<td>115</td>
<td>6</td>
<td>161</td>
</tr>
<tr>
<td>cloning efficiency</td>
<td>6%</td>
<td>5%</td>
<td>6%</td>
</tr>
</tbody>
</table>
None of the positive clones appeared to contain a PCR product including VASE (see Figure 4.4.2), although the positive control gave rise to a band of 224 bp, while the negative control remained blank. On rehybridising the same blot with the non-VASE probe, five clones appeared to contain PCR products of NCAM sequence (Figure 4.4.3; samples 15, 20, 30, 45 and 46). Again, the positive control gave rise to a band of 224 bp, while the negative control remained blank.

The molecular weight of the smallest restriction fragment arising from digestion with Not-1, and which hybridised the non-VASE probe, varied considerably from clone to clone (see Table 4.4.2). Fragments from 530 to 1100 bp were observed; these correspond to insert sizes of 180 to 750 bp. Larger species were also visible.

Clone set 2 was generated by plasmid: PCR product ratios of 1:1, 1:3 and 1:10. In this case, the PCR products had been previously selected to bias the frequency of VASE-containing products. The 240 bp band arising from PCR amplification of the exon 7/8 splice junction of JR-1 was excised and re-amplified, resulting in a "first generation" population of PCR products.

Again, the cloning efficiency of all three ligations was unexpectedly low and did not vary significantly from the negative control (see Table 4.4.3). An increase in the concentration of PCR product did not affect cloning efficiency, as has been observed in other instances (Mead et al., 1991). However, Southern blotting suggested clone set 2 to contain two clones (samples 28 and 38) which include VASE (Figure 4.4.4). Visible but fainter bands were also observed in lanes containing RE products derived from other plasmid samples. By error, a positive control was not included in this blot, although the negative control remained blank.

Like the NCAM-containing samples in clone set 1, the sizes of restriction fragments were difficult to rationalise (Table 4.4.4). Insert sizes
Figure 4.4.2

Southern blot of selected samples from clone set 1; hybridised with the VASE oligonucleotide
Episomal DNA from selected colonies of clone set 1 was extracted, digested with the Not-1 restriction enzyme, and Southern blotted with the VASE oligonucleotide probe.

Key:

N1 = positive control
samples are as indicated
NC = negative control
Figure 4.4.2

224 bp
Figure 4.4.3

- 3050 bp
- 1850 bp
- 1100 bp
- 600 bp
- 530 bp
- 224 bp
Table 4.4.2

Clone set 1; products from the JR-1 rhabdomyosarcoma cell line, cloned and digested with (Not-1) restriction enzyme

Column a: Sample number corresponds directly to those displayed in Figure 4.4.3. Column b: the molecular weights of restriction enzyme fragments observed in Figure 4.4.3. These were calculated by comparing with a standard molecular weight ladder, electrophoresed alongside the digested samples. Visualisation was achieved by staining with ethidium bromide, prior to transferral to a nylon membrane. Column c: the calculated size of inserted PCR product (lowest value in column b minus 350 bp). Column d: theoretical sizes of restriction enzyme fragments on partial digestion of cloned PCR product. These were calculated on the assumption that Not-1 may digest only two of the three sites available for digestion (see Figure 4.4.6).
Table 4.4.2

<table>
<thead>
<tr>
<th>a. sample</th>
<th>b. fragments binding non-VASE (bp)</th>
<th>c. size of insert (bp)</th>
<th>d. Theoretical sizes (approx.) of RE fragments on partial digestion (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>530 (1600)</td>
<td>180</td>
<td>700, 3000</td>
</tr>
<tr>
<td>20</td>
<td>600</td>
<td>250</td>
<td>750, 3000</td>
</tr>
<tr>
<td>30</td>
<td>1100</td>
<td>750</td>
<td>1,250, 3500</td>
</tr>
<tr>
<td>45</td>
<td>575</td>
<td>225</td>
<td>725, 3000</td>
</tr>
<tr>
<td>46</td>
<td>1100</td>
<td>750</td>
<td>1250, 3500</td>
</tr>
</tbody>
</table>
Table 4.4.3

Colonies arising from the cloning of selected PCR products; clone set 2
Selected PCR products generated from RNA extracted from the JR-1 rhabdomyosarcoma cell line were ligated
with the pCR™1000 vector at 4 different vector : insert molar ratios: 1:0 (negative control), 1:1, 1:3 and 1:10.
The number of positive and negative colonies arising from transformation of E. coli by products of each ligation
reaction are indicated.
<table>
<thead>
<tr>
<th>Number of colonies</th>
<th>Plasmid : PCR product ratio (molar)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1:0 (negative control)</td>
<td>1:1</td>
</tr>
<tr>
<td></td>
<td>negative  positive</td>
<td>negative  positive</td>
</tr>
<tr>
<td></td>
<td>61  3</td>
<td>129  6</td>
</tr>
<tr>
<td></td>
<td>75  6</td>
<td>120  7</td>
</tr>
<tr>
<td></td>
<td>38  2</td>
<td>188  9</td>
</tr>
<tr>
<td></td>
<td>32  3</td>
<td>202  16</td>
</tr>
<tr>
<td></td>
<td>100  4</td>
<td>51.5  3.5</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>4%</td>
</tr>
</tbody>
</table>

Cloning efficiency
Figure 4.4.4

Southern blot of selected samples from clone set 2; hybridised with the VASE oligonucleotide
Episomal DNA from selected colonies of clone set 1 was extracted, digested with the Not-1 restriction enzyme, and Southern blotted with the VASE oligonucleotide probe.

Key:

samples are as indicated

NC = negative control
Figure 4.4.4
Table 4.4.4

Clone set 2; products from the JR-1 rhabdomyosarcoma cell line, cloned and digested with (Not-1) restriction enzyme

**Column a:** Sample number corresponds directly to those displayed in Figure 4.4.5  **Column b:** the molecular weights of restriction enzyme fragments observed in Figure 4.4.5 (with the exception of sample 9). These were calculated by comparing with a standard molecular weight ladder, electrophoresed alongside the digested samples. Visualisation was achieved by staining with ethidium bromide, prior to transferral to a nylon membrane. Molecular weights in brackets indicate bands of weak intensity.  **Column c:** the calculated size of inserted PCR product (lowest value in column b minus 350 bp).  **Column d:** theoretical sizes of restriction enzyme fragments on partial digestion of cloned plasmid. These were calculated on the assumption that Not-1 may digest only two of the three sites available for digestion (see Figure 4.4.6).
Table 4.4.4

<table>
<thead>
<tr>
<th>a. sample</th>
<th>b. fragments observed upon electrophoretic separation</th>
<th>c. size of insert (bp)</th>
<th>d. Theoretical sizes (approx.) of RE fragments on partial digestion (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>510</td>
<td>160</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>1700</td>
<td>1350</td>
<td>1850</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td></td>
<td>3100</td>
</tr>
<tr>
<td>32</td>
<td>530</td>
<td>180</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(700)</td>
<td></td>
<td>680</td>
</tr>
<tr>
<td></td>
<td>1700</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2700</td>
<td></td>
<td>2900</td>
</tr>
<tr>
<td>38</td>
<td>880</td>
<td>530</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td></td>
<td>1000</td>
</tr>
<tr>
<td></td>
<td>(1700)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2400)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2700</td>
<td></td>
<td>3300</td>
</tr>
</tbody>
</table>
of VASE-containing samples 28 and 38 did not correspond to products of an expected size (i.e. 224 or 240 bp). Clones containing NCAM positive, VASE-lacking PCR products were also generated (Figure 4.4.5, samples 20, 22, 30, 32, 33 and 36). Again, the negative control remained blank.

(c) Discussion

Southern blot analysis of clone sets 1 and 2 suggests that VASE-containing and VASE-lacking, NCAM-positive PCR products had been successfully cloned, but constitute a small fraction of all the clones examined. The different sizes of inserted PCR products, as calculated from the blot shown in Figure 4.4.3 (summarised in Table 4.4.2) are at variance with the insertion of five VASE-lacking PCR products as characterised by direct sequencing. In this case one would expect to see a band of 545 bp in each sample - approximately 195 bp of non-VASE PCR product, plus 350 bp of digested plasmid. The comparatively large size of restriction fragments from samples 20, 30, 45 and 46 is consistent with the inclusion of a novel exon at the exon 7/8 splice junction, although the corresponding insert size of the latter two samples corresponds poorly with PCR products observed in Table 4.2.

Bands of high molecular weight (1.6 - 3.5 kb) observed in the lanes containing samples 15, 20, 30, 45 and 46 could be explained by incomplete digestion of the plasmid by Not-1 restriction enzyme (Figure 4.4.6). However, the size of these bands, even though approximate, is inconsistent with some of the theoretical weights of such products (See Table 4.4.2). Similarly, restriction fragments observed on digestion of selected samples from clone set 2 (Table 4.4.4) only partially coincide with theoretical fragments, calculated on the basis of insert size.

It is also difficult to explain the presence of clones with inserts, yet apparently lacking in NCAM sequence. These may relate to non-specific products which arise during the PCR reaction, although the intensity of the
Figure 4.4.5

Southern blot of selected samples from clone set 2; hybridised with the non-VASE oligonucleotide
The Southern blot presented in Figure 4.4.3, was stripped of VASE oligonucleotide probe, and rehybridised with the non-VASE oligonucleotide probe.

Key:
samples are as indicated
NC = negative control
latter on gels stained with ethidium bromide, suggests that they should not seriously hinder the cloning of specific bands. It may be that non-specific PCR products are generated alongside specific PCR products and that these are of similar size; thus in cloning specific products, one inadvertently clones a population of non-specific products as well. This hypothesis is supported by the observation that non-specific bands are seen in agarose gels containing PCR products derived from the negative control GH1 cell line; bands of corresponding sizes are also observed in neuroblastoma and rhabdomyosarcoma cell lines (data not displayed). One means by which this possibility might be explored is to sequence the major products observed in the GH1 cell, and then to reprobe the nylon filters used to generate the blots above with a suitable oligonucleotide.

There are several means of minimising the generation of non-specific products. These include the empirical determination of stringent annealing conditions and optimum reagent concentrations (see section 4.5 (c)), and the strategic choice of primers for the RT-cDNA and PCR reactions. Three strategies are generally used. These are summarised below:

1. An exon-specific 3' primer is used to prime the RT-cDNA reaction. The same primer, together with an exon-specific 5' primer, is used for PCR amplification.

2. An exon-specific 3' primer is used to prime the RT-cDNA reaction. An internal "nested" 3' primer and an exon-specific 5' primer are used for PCR amplification.

3. Oligo dT or random primers are used to prime the RT-cDNA reaction. Two exon-specific primers are used for PCR amplification.

The second method has an obvious advantage over the first in generating specific products; an improved experimental protocol would adhere to this strategy. A comparative study of these methods suggests that method (2) may also result in more specific products than method (3) (Baier \textit{et al.}, 1993). This may be due to the fact that an "unselected" population of
cDNA species is generated by the RT-PCR of method (3). The use of oligo dT, however, minimises potential "distortions" due to a possible excess 3' primer carried over into the PCR reaction.

The success of all of these methods depends on the ability of primers to specifically anneal. It seems that some primers simply do not anneal to certain templates. This may be due to secondary structure of RNA or cDNA (for examples see McCall et al., 1992 and Cuddy et al., 1993). Exon-"specific" primers may be promiscuous in that they bind non-specifically at various points on available templates - in most instances, this may be overcome by establishing stringent annealing conditions and/or by reducing the primer concentration.

A reduction in PCR cycles and/or cDNA generated by RT-PCR may also result in a more homogeneous population of specific product; optimum number of cycles or quantity of cDNA can only be determined empirically. Most primers bind preferentially to their targeted sequence when compared with "non-specific" sequences, with which they have reduced complementarity. As the rate of product formation is exponential, it is best to halt the reaction as soon as a sufficient quantity of specific product is generated.

The technical task of obtaining NCAM-positive clones was further hindered by very low cloning efficiencies. Transformation of *E. coli* by the products of five cloning reactions gave rise to a total of 1399 colonies, of which 119 were positive (see Tables 4.4.1 and 4.4.3). It is unlikely that this was a result of inappropriate reaction conditions, as chemical components of the ligation reaction were supplied in the cloning kit, and as such, had been standardised by the manufacturer. In addition, the results were reproducible. Irregularities in the vector, inefficient deoxyadenylation of PCR products by *Taq* polymerase, or inadvertent modification of the PCR product during isolation and purification are all potential factors which could affect cloning efficiency, although it is impossible to ascertain if any of them have contributed to the results presented here.

The apparent cloning of irrelevant PCR products also contributed to the paucity of NCAM-positive clones - of the 48 "positive" clones analysed in the Southern blots presented here, only 13 appeared to harbour NCAM sequence. It may be that the remaining clones were in fact false positives, although sizes of electrophoresed products arising from restriction enzyme

172a.
digestion with Not-1 suggest that many do contain an insert (data not displayed).

Sequencing relevant clones from clone sets 1 and 2, should clarify the unexpected sizes of inserts presented in Tables 4.4.2 and 4.4.4, as calculated from restriction fragments observed in Figures 4.4.3 and 4.4.4.

4.5 Sequence Analysis of Samples from Clone sets 1 and 2

Samples 15, 20, 30, 45 and 46 from clone set 1 were sequenced. These hybridised the non-VASE probe on Southern blotting (Figure 4.4.3), and gave rise to unexpected insert sizes (see Table 4.4.1); sequence analysis was performed to elucidate the insert sequence. For similar reasons, samples 28 and 38 of clone set 2 were sequenced; these hybridised VASE in addition to non-VASE, and also gave rise to restriction fragments of unexpected sizes. Samples 9 and 32 of clone set 2, neither of which hybridised VASE, were sequenced for comparison's sake.

(a) Methods

Clones were sequenced using the Sequenase 2 kit (USB), according to instructions supplied by the manufacturer. A variety of primers were used: the 3' and 5' primers used in PCR amplification, and two which were complementary to plasmid sequence adjacent to the insertion site. The latter two were designated the "Universal" and "Reverse" primers, and complementary to sequence 3' and 5' of the insertion site respectively (see Figure 4.4.6). Radiolabelled products of the sequencing reaction were analysed by denaturing polyacrylamide gel electrophoresis, as described in Chapter 2.

(b) Results

Cloned PCR products generally yielded clear, unambiguous sequence (see Figure 4.5.1 for example). Sequence analysis of the samples in clone set 1 confirmed their suspected identities - i.e. that they contain NCAM sequence,
Figure 4.5.1

Sequence film of a cloned PCR product
PCR product generated from primers spanning the exon 7/8 splice junction were cloned into the pCR1000TM vector and sequenced using the Universal primer. The sequence presented here is derived from sample 38, clone set 2, and is complementary to data presented in Figure 4.5.3 (b).

Key:
A: sequencing products electrophoresed for 4.5 hr
B: sequencing products electrophoresed for 3 hr
1 products of sequencing reaction in which ddATP served as the terminating nucleotide
2 products of sequencing reaction in which ddGTP served as the terminating nucleotide
3 products of sequencing reaction in which ddCTP served as the terminating nucleotide
4 products of sequencing reaction in which ddTTP served as the terminating nucleotide
but lack VASE. Results obtained from clone set 2 were also consistent with data arising from Southern blots; samples 28 and 38 contain VASE, while samples 9 and 32 lack it. This represents conclusive evidence that VASE is expressed by the rhabdomyosarcoma cell line JR-1, and in conjunction with data presented in section 4.3, reveals it to be identical with that found in neuroblastoma cell lines.

Data arising from clone set 1 revealed that PCR products are more heterogeneous than originally perceived. Sample 45 alone (Figure 4.5.2a) yielded sequence corresponding to the orthodox VASE-lacking PCR product. Sample 15 (Figure 4.5.2b) was found to contain a similar insert, although the 3' primer was truncated.

Three other samples from clone set 1 also contain PCR products which appear to have been generated from the 5' primer, as their sequences extend beyond the 3' primer and further into exon 8. Such is the case with sample 30 (Figure 4.5.2c). Sample 20 yielded a similar sequence, although a curious feature was observed at its 3' end (Figure 4.5.2d). A second 3' primer, followed by a short unidentifiable sequence, appears to have fused to the end of the first sequence. Sample 46 (Figure 4.5.2e) also contains NCAM sequence which extends beyond the 3' primer. In this case it appears to extend into exon 12 although this can only be confirmed by obtaining sequence which bridges those obtained with the 5' and Universal primers. The primary purpose of the exercise described here was to characterise alternative splicing patterns at the exon 7/8 junction by sequencing PCR products. Thus, complete sequence analysis of the entire PCR insert is considered desirable but not imperative; some inserts have not been sequenced from end to end.

Sequencing data from samples of clone set 2 also confirm the blotting data. Samples which bind both VASE and non-VASE probes (samples 28 and 38; Figures 4.5.3a and 4.5.3b respectively) are seen to
Figure 4.5.2 (a)

**Sequence obtained from sample 45, clone set 1**
Digestion with Not-1 restriction enzyme, followed by Southern blotting (Figure 4.4.2 and 4.4.3; Table 4.4.2), suggested that sample 45 has an insert of approximately 225 bp, containing part of exon 7, but lacking VASE.

(i) **Sequence**
Sample 45 was found to contain an insert of 194 bp, comprising NCAM-positive sequence which lacked VASE. The primers are framed, and bold vertical lines indicate the exon 7/8 splice junction.

**Key:** See Figure 4.3.4 (i) for key to sequence

--- = sequence obtained using universal primer
■■■ = sequence obtained using 5' primer

(ii) **Representation of Sequence**
Sequence from sample 45 is represented in terms of its constituents.

**Key:**

□□□ = plasmid sequence; sequence on the 3' and 5' sides of the insertion site are as indicated

□ = constitutively-spliced sequence; exon is indicated below

3' = 3' primer

5' = 5' primer
Figure 4.5.2 (a)

(i)

5' TGAGAACCGTTT A3 CAG GTC ACT CTT ACC TGT GAA
GCC TCC GGA GAC CCC ATT CCC TCC ATC ACC TGG
AGG ACT TCT ACC CCG AAC ATC AGC AGC GAA GAA
AAG ACT CTG GAT GGG CAC ATG GTG GTG CGT AGC
CAT GCC CGT GTG TCG TCG CTG ACC CTG AAG AGC
ATC CAG TAC ACT GAT GCC GGA GAG TAC ATC TGC
ACC GCC AAACCTT

3'

(ii)

3'   5'  

7  8  

176.
Figure 4.5.2 (b)

Sequence obtained from sample 15, clone set 1
Digestion with Not-1 restriction enzyme, followed by Southern blotting (Figure 4.4.2 and 4.4.3; Table 4.4.2), suggested that sample 15 has an insert of approximately 180 bp, containing part of exon 7, but lacking VASE.

(i) Sequence
Sample 15 was found to contain an insert of 187 bp, comprising NCAM-positive sequence which lacked VASE. Primers and primer-derived sequence are framed; the 3' primer was found to be truncated. Plasmid sequence is in bold type.

Key:
A = adenine residue
C = cytosine residue
G = guanidine residue
T = thymidine residue
X = unidentified residue
— = sequence obtained using universal primer
■■■ = sequence obtained using 5' primer

(ii) Representation of Sequence
Sequence from sample 15 is represented in terms of its constituents.

Key:
■■■ = plasmid sequence. Sequence on the 3' and 5' sides of the insertion site are as indicated
□ = constitutively-spliced sequence; exon is indicated below
3' = truncated 3' primer
5' = 5' primer
Figure 4.5.2 (b)

(i)  

```
5' GAGAACCGTTT A|G  CAG GTC ACT CTT ACC T G t| GAA
GCC TCC GGA GAC CCC ATT CCC TCC ATC ACC TGG
AGG ACT TCT ACC CGG AAC ATC AGC AGC GAA GAA
AAG ACT CTG GAT GGG CAC ATG GTG GTG CGT AGC
CAT GCC CGT GTG TCG TXX CTG ACC CTG AAG AGC
ATC CAG TAC ACT GAT XXC GGA GAG TAC ATC TG
AAACCCTTCTCAGATAAG
3'
```

(ii)

```
3' 5' 7 8 3'
```

177.
Figure 4.5.2 (c)

**Sequence obtained from sample 30, clone set 1**

Digestion with Not-1 restriction enzyme, followed by Southern blotting (Figure 4.4.2 and 4.4.3; Table 4.4.2), suggested that sample 30 has an insert of approximately 750bp, containing part of exon 7, but lacking VASE.

(i) **Sequence**

Sample 30 was found to contain an insert exceeding 222 bp, comprising NCAM-positive sequence and lacking VASE. The single 3' primer is framed, and a bold vertical line indicates the exon 7/8 splice junction.

**Key:**

■■■■ = sequence obtained using 5' primer

(ii) **Representation of Sequence**

Sequence from sample 30 is represented in terms of its constituents.

□ = constitutively-spliced sequence; exon is indicated below

3' = 3' primer
Figure 4.5.2 (c)

(i)

5'

CC CGG AAC ATC AGC AGC GAA GAA AAG ACTCTG
GAT GGG CAC ATG GTG GTG CGT AGC CAT GCC CGT
GTG TCG TCX CTG ACC CTG AAG AGC ATC CAG TAC
ACT GAT GCC GGA GAG TAC AXX XXX ACC GCC AGC
AAC ACC ATC GGC CAG G A GTCC CAG TCC ATG TAC
CTT GAA GTG CAA TAT GCC CCA AAG CTA CAG GGC
CCT G TG GCT GTG TAC ACT TGG GAG G

3'

(ii)

| 3' |

← 7 | 8 →
Sequence obtained from sample 20, clone set 1

Digestion with Not-1 restriction enzyme, followed by Southern blotting (Figure 4.4.2 and 4.4.3; Table 4.4.2), suggested that sample 20 has an insert of approximately 250p, containing part of exon 7, but lacking VASE.

(i) **Sequence**

Sample 20 was found to contain an insert exceeding 197 bp, comprising NCAM-positive sequence and lacking VASE. The sequence extends into exon 8, beyond the 3' primer. Two 3' primers (framed) were found to be present; the second was followed by a short sequence of unknown identity. Bold vertical lines indicate exon borders. Plasmid sequence is in bold type.

**Key:**

■ ■■ = sequence obtained using 5' primer

(ii) **Representation of Sequence**

Sequence from sample 20 is represented in terms of its constituents.

**Key:**

■ = plasmid sequence; sequence on the 3' and 5' sides of the insertion site are as indicated

□ = constitutively-spliced sequence; exon is indicated below

3' = 3' primer

□ = unidentified sequence
Figure 4.5.2 (d)

(i)

5' CC CGG AAC ATC AGC AGC GAA GAA AAG ACT CTG

GAT GGG CAC ATG GTG GTG CGT AGC CAT GCC CGT

GTG TCG TCG CTG ACC CTG AAG AGC ATC CAG TAC

ACT GAT GCC GGA [GAG TAC ATC TGC ACC GCC] AGC

AAC AAC ATC GGC CAG CAC GCC ATG CAG TAC

CTT GXX [GAG TAC ATC TGC ACC GCC] AAC CCT

ACTCGCGGA

3'

(ii)

5' 3' 3'

←7 ←8 ←8 →
Figure 4.5.2 (e)

**Sequence obtained from sample 46, clone set 1**

Digestion with Not-1 restriction enzyme, followed by Southern blotting (Figure 4.4.2 and 4.4.3; Table 4.4.2), suggested that sample 46 has an insert of approximately 750 bp, containing part of exon 7, but lacking VASE.

(i) **Sequence**

Sample 46 was found to contain an insert exceeding 329 bp, comprising NCAM-positive sequence and lacking VASE. The primer is framed, and bold vertical lines indicate the exon borders.

**Key:**

— — = sequence obtained using universal primer

■ ■ ■ = sequence obtained using 5' primer

(ii) **Representation of Sequence**

Sequence from sample 46 is represented in terms of its constituents.

**Key:**

□ = constitutively-spliced sequence; exon is indicated below

3' = 3' primer
Figure 4.5.2 (e)

(i)

5'

ACG TGG AGG ACT TCT ACC CGG AAC ATC AGC AGC
GAA GAA AAG ACT CTG GAT GGG CAC ATG GTG GTG
CGT AGC CAT GCC GTG TCG TCG CTG ACC CTG
AAG AGC ATC CAG TAC ACT GAT GCC GGA [GAG TAC
TX XXX ACC GCC] AGC AAC ACC ATC GGC CAG GAC
TCC CAG TCC CAG ATG TAC CTG AAA GTG CAA TAT
GCC CCA AAG CTA CAG GG ...

...TT CTT GGG GGA AGA AGT GGT TAG
CTG GTC CAC CTC TGT ATG AGG TCG TGT CTC GTC
CAC GTC AAA CTA CTT GGT CTC CGG TGT CCA CCC
CAC GGG TAG TAC CTT ATG TTT CTA CTC ACC TCT
CGT CAA CCA CTT CTT CAT ACC GTC ACG TAC ACC
ATA CTA CGG TTC CTT CTT GGG TCG TAC CTC ATG ...

3'

(ii)

3'

←7 || 8 →

10 | 11 | 12
Figure 4.5.3 (a)

Sequence obtained from sample 28, clone set 2
Digestion with Not-1 restriction enzyme, followed by Southern blotting (Figures 4.4.4 and 4.4.5; Table 4.4.4), suggested that sample 28 has an insert of approximately 1350 bp, containing part of exon 7 and VASE.

(i) Sequence
Sample 28 was found to contain an insert of 300 bp, comprising NCAM-positive sequence including VASE, followed by unidentifiable sequence. Bold vertical lines indicate exon borders. Plasmid sequence is in bold type and VASE is in bold italics. Asterisks indicate nucleotides which deviate from established sequence.

Key: For key to all sequences in Figures (a) - (d) see Figure 4.5.2 (b).

— — = sequence obtained using universal primer

• • • = sequence obtained using reverse primer

(ii) Representation of Sequence
Sequence from sample 28 is represented in terms of its constituents.

Key:

= plasmid sequence; sequence on the 3' and 5' sides of the insertion site are as indicated

square = constitutively-spliced sequence; exon is indicated below

= VASE (V)

= 3' primer

= 5' primer

= unidentified sequence
Figure 4.5.3 (a)

(i)

5'  
GAGAACCAGTTT [AG CAG GTC ACT CTT ACC TGT] AG  
CAG GTC ACT CTT ACC TGT [AG CAG GTC ACT CTT]  
ACC TGT [GAA GCC TCC GGA GAC CCC ATT CCC TCC]  
ATC ACC TGG AGG ACT TCT ACC CGG AAC ATC AGC  
[ACC GAA GAA AGG] TCG TGG ACT CGA CCA GAG  
AAG CAA GAG [ACT CTG GAT GGG CAC ATG GTG GTG]  
CGT AGC CAT GCC CGT GTG TCG TCG CTG ACC CTG  
AAG AGC ATC CAG TAC ACT GAT GCC GGA GAG TAC  
ATC TGC ACC GCC AGC  
3'

(ii)

3'  
[CGT AGC CAT GCC CGT GTG TCG TCG CTG ACC CTG]  
AAG AGC ATC CAG TAC ACT GAT GCC GGA [GAG TAC]  
ATC TGC ACC GCC AGCCACTTGCTCCCCTCTA [GAG TAC]  
ATC TGC ACC GCC AACCTTCTCACCTCGGCCGATAAGC  
3'  

*site of aberrant sequence

181.
Figure 4.5.3 (b)

Sequence obtained from sample 38, clone set 2
Digestion with Not-1 restriction enzyme, followed by Southern blotting (Figures 4.4.4 and 4.4.5; Table 4.4.4), suggested that sample 38 has an insert of approximately 530 bp, containing part of exon 7 and VASE.

(i) Sequence
Sample 38 was found to contain an insert of exceeding 416 bp. It comprises NCAM-positive sequence including VASE (bold italics), three 5' primers, one 3' primer, and unidentifiable sequence. Bold vertical lines indicate exon borders. Primers are framed, and plasmid sequence is in bold type.

Key:

--- = sequence obtained using universal primer

• • • = sequence obtained using reverse primer

(ii) Representation of Sequence
Sequence from sample 38 is represented in terms of its constituents.

Key:

 بلاك = plasmid sequence; sequence on the 3' and 5' sides of the insertion site are as indicated

□□□□□□□□□ = constitutively-spliced sequence; exon is indicated below

[Black square] = VASE (V)

[3'] = 3' primer

[5'] = 5' primer

[Black diamond] = unidentified sequence
Figure 4.5.3 (b)

(i)  

5'  
ATCGGCGAAGTGAAGAAGGT 
  ACC TGT GAA GCC TCC  
GGA GAC CCC ATT CCC TCC ATC ACC TGG AGG ACT  
TCT ACC CGG AAC ATC AGC AGC GAA GAA AAG GCT  
TCG TGG ACT CGA CCA GAG AAG CAA GAG ACT CTG  
GAT GGG CAC ATG GTG GTG CGT AGC CAT GCC CGT  
GTG TCG TCG CTG ACC CTG AAG AGC ATC CAG TAC  
ACT GAT GCC GGA GAG TAC . . .  

3' ...
TGAGTACATCTGCCAACGCAAGGACCTGTA

(ii)  

5'  

|    7 | V | 8 |

3'
contain an NCAM PCR product containing VASE, although sequence from sample 28 deviates at two positions from previously established sequence. This region was sequenced three times to confirm the sequence aberrations. Sample 38 yielded obscured sequence at the exon 7/VASE border in 5 sequencing reactions. Addition of Mn⁺ to the sequencing reaction restored clear sequence (obtained once and in the first instance).

As some of the other samples of clone set 2 yielded faint signals on hybridisation with VASE (Figure 4.4.3) two (samples 9 and 32) were selected for sequence analysis. Sample 9 was found to contain a string of whole and partial 3' primers followed by sequence irrelevant to both NCAM and the pCR™1000 plasmid vector (Figure 4.5.3c). Sample 32, which hybridised the non-VASE probe on Southern blotting (Figure 4.4.5) is similar to products of clone set 1 - i.e. it contains NCAM sequence but not VASE (Figure 4.5.3d).

(c) Discussion

Data obtained from sequencing clones containing PCR products generated from NCAM demonstrates that VASE is spliced into some RNA species synthesised by the rhabdomyosarcoma cell line JR-1.

While sequencing data of clones generally confirmed conclusions drawn from Southern blots presented in Section 4.4, it suggested that several events take place during PCR amplification which had not been foreseen. These give rise to PCR products of unexpected size, thus explaining the unexpected insert sizes of clones in Tables 4.4.2 and 4.4.4. Amplification of sequence beyond the 3' primer, non-specific hybridisation of primers, and fusion of oligonucleotides all appear to take place.

Samples 20, 30 and 46 of clone set 1 all contain sequence from exon 8 which lies beyond the 3' primer, which was used to generate cDNA by reverse transcription. The 3' primer should therefore define the 3' terminus of PCR products subsequently generated. In these clones, it does not,
Figure 4.5.3 (c)

Sequence obtained from sample 9, clone set 2
Digestion with Not-1 restriction enzyme, followed by Southern blotting (Figures 4.4.4 and 4.4.5; Table 4.4.4), suggested that sample 9 has an insert of approximately 160 bp, containing no NCAM positive sequence as detected by the non-VASE and VASE oligonucleotide probes.

(i) **Sequence**
Sample 9 was found to contain an insert of 158 bp, containing NCAM-positive sequence which lacked VASE. The insert was found to comprise a series of fused 5' primers, some of which were whole, and others of which were partial. Primers and primer-derived sequence are framed; plasmid sequence is in bold type.

**Key:**

= sequence obtained using universal primer

(ii) **Representation of Sequence**
Sequence from sample 9 is represented in terms of its constituents.

**Key:**

= plasmid sequence. Sequence on the 3' and 5' sides of the insertion site are as indicated

= 5' primer; n.b. some are incomplete. The 5' primer sequence is AGCAGGTCACCTCCTACCTGT

= unidentified sequence
Figure 4.5.3 (c)

(i)  

\[
\begin{align*}
5' & \quad \text{GGTGAGAGGGTTT} \quad \text{ACAGGTCACCTTACCTGTAACAGGTC} \\
& \quad \text{CTCTTACCTGTGCTGCAGGTCA} \\
& \quad \text{ACTCTTACCTTACTCTCTGATCTCAGTTG} \\
& \quad \text{AAGTCGGGAAACGTATTCACCGTAGGCTACGTATCTAGGTA} \\
& \quad \text{AACGGTTCTC}
\end{align*}
\]

(ii)  

\[
\begin{align*}
3' & \quad \text{5' 5' 5' 5' 5' 5' 5'}
\end{align*}
\]
Figure 4.5.3 (d)

**Sequence obtained from sample 32, clone set 2**
Digestion with Not-1 restriction enzyme, followed by Southern blotting (Figures 4.4.4 and 4.4.5; Table 4.4.4), suggested that sample 32 has an insert of approximately 180 bp, containing part of exon 7, but lacking VASE.

(i) **Sequence**
Sample 32 was found to contain an insert of 208 bp, comprising NCAM-positive sequence but not VASE. Primers are framed, and plasmid sequence is in bold type. A bold vertical line indicates the exon 7/8 splice junction.

**Key:**

--- = sequence obtained using universal primer

• • • = sequence obtained using reverse primer

(ii) **Representation of Sequence**
Sequence from sample 32 is represented in terms of its constituents.

**Key:**

[ ] = plasmid sequence. Sequence on the 3' and 5' sides of the insertion site are as indicated

[ ] = constitutively-spliced sequence; exon is indicated below

[ ] = truncated 3' primer

[ ] = 5' primer
Figure 4.5.3 (d)

(i)

5'  5'  
GAGAACCCTT G CAG GTC ACT CTT ACC TGT G CAG
GTC ACT CTT ACC TGT GAA GCC TCC GGA GAC CCC
ATT CCC TCC ATC ACC TGG AGG ACT TCT ACC CGG
AAC ATC AGC AGC GAA GAA AAG ACT CTG GAT GGG
CAC ATG GTG GTG CGT AGC CAT GCC CGT GTG TCG
TCG CTG ACC CTG AAG AGC ATC CAG TAC ACT GAT
GCC GGA GAG TAC ATC TGC AC GAA CTCTTCACCTCG

(ii)

3'  5'  

| 7 | 8 | 3' |
suggesting that RNA may act as a template in the PCR amplification reaction, alongside complementary DNA generated by reverse transcription.

Non-specific hybridisation of primers is suggested by the presence of unidentified sequence juxtaposed with primer, as seen in sample 20 of clone set 1, and samples 9, 28 and 38 of clone set 2. Unidentified sequence has been analysed by a DNA alignment program (DNAstar Inc. Wilbur-Lipman) k-tuple = 3; gap penalty = 3, and conforms neither to NCAM sequence nor to that of the pCRT™ 1000 vector. One can only conclude that both the 3' and the 5' primers bind non-specifically and that non-specific hybridisation can give rise to the unidentified sequences observed.

This conclusion is supported by the appearance of visible bands on PCR amplification of cDNA generated from RNA derived from the negative control GH1 cell line. While these are clearly visible when electrophoresed through an ethidium bromide-stained gel (data not shown), Southern blots suggest that they do not contain NCAM-positive sequence, and are hence non-specific. Analogous bands are also observed in NCAM-positive neuroblastoma and rhabdomyosarcoma cell lines. It would be of interest to rehybridise the samples presented in Figures 4.4.1 and 4.4.3 using first one and then the other primer as a probe, to elucidate the extent to which non-specific annealing during the PCR reaction contributes to the large number of non-NCAM clones.

While reagent concentrations used in these experiments correspond to those of recommended protocols (Sambrook et al., 1989; Kawasaki, 1990) it is possible that "carry-over" of dNTPs, Mg2+ and/or primers may reduce the specificity of the PCR reaction and contribute to generation of artefacts. There are numerous ways to reduce non-specific annealing. An empirical determination of the highest permissible annealing temperature is one. Reduction of Taq polymerase, Mg2+, and dNTP concentration to minimal permissible levels, reduces the frequency of mispairing and the chances that a non-specifically annealed primer will be extended. Reducing the primer concentration to minimal permissible levels increases the resulting ratio of specific : non-specific product.

Sequencing data suggest that the artifactual bands of 240 bp, 660 bp and 900 bp observed in Figures 4.2.2 - 4.2.5 are probably not due to non-specific priming. While it is conceivable that non-specific binding may play a role, it appears that aggregation and/or heteroduplex formation contribute(s). Indeed, other investigators using RT-PCR to investigate splicing patterns of NCAM RNA have also detected consistent artifactual bands which resolve upon electrophoresis through a denaturing gel (Small and Akeson, 1990; Zorn and Krieg, 1991).
Sequence analysis of samples 20, 30 and 46 of clone set 1 suggest that RNA may be amplified during the initial PCR reaction. Elimination of RNA after cDNA synthesis and prior to amplification by PCR, might reduce the number and heterogeneity of templates to which primers might "non-specifically" anneal. 'The use of nested primers in the generation of first and second' generation PCR products might reduce the number of irrelevant products generated. Variations in sequence or contamination of oligonucleotide primer may contribute to the production of irrelevant products. However, this seems unlikely, as the stepwise efficiency of synthesis of both primers exceeded 98%.

Deoxyoligonucleotide fusion would appear to be a feature of the reaction which complicates matters. It is thought that this may take place when polymerase terminates synthesis on one DNA strand and continues synthesis after priming occurs on another (Innis et al., 1990). However, it is difficult to envisage how this behaviour could give rise to a product such as sample 9 of clone set 1, which is largely composed of primer and fragments of primer.

The fusion of oligonucleotides during the PCR reaction has been described as the cause of "primer-dimers" (Mullis, 1991) - short fragments of approximately two primers in length, which can interfere with the amplification of target sequence - but has yet to be described in the context of longer amplified products. Samples 9, 28, 32 and 38 of clone set 2, all show evidence of "primer-dimer" formation, but not in the usual sense, where both the 3' and 5' primers are observed. The primer fusion observed in these samples invariably occurs between two or more primers which are either 3', or 5', but not both.

Another factor which did not facilitate the attainment of clones of interest was the low cloning frequency. Increasing the vector : plasmid molar ratio as recommended by the manufacturer, did not affect the efficiency. Perhaps the vector : plasmid ratio was not sufficiently increased. Alternatively, different brands of \textit{Taq} polymerase may adenylate PCR products to different extents; the polymerase used by the manufacturer of
the TA-cloning kit was obtained by cloning, while that used in these experiments was extracted from *T. aquaticus* by biochemical methods.

Alternative strategies may result in higher cloning efficiencies of specific PCR product. One could generate one's own "TA" vector using the single stranded bacteriophage M13mp18 (Pharmacia LKB Biotechnology) and a series of primers (Zimmer, 1993).

Blunt-ended ligation could be achieved by "polishing" the termini of PCR products using Klenow fragment. A disadvantage with this method is that the cloning reaction competes with self-ligation of the plasmid. In contrast, vectors which utilise "sticky ends" generated by RE digestion, are not prone to self-ligation. PCR products can be generated with complementary "sticky ends" by a variety of methods. Relevant sequence should be screened for naturally-occuring RE sites as a first step. Naturally-occuring partial RE-sites can also be exploited by using a specially-designed primer which converts them into full sites upon amplification. This kind of primer deviates from its "normal" counterpart in that it is not entirely complementary to its targeted sequence as it contains sequence which results in a full RE site on amplification. For this reason, it is advisable to design the primers so that the "non-matching" sequence is near the 5' end of the primer. The shorter the "non-matching" sequence (or the longer the naturally-occuring partial RE site), the more likely it is that relevant products will be generated.

If no conveniently-placed RE sites are present, sequences which result in full RE-sites on PCR amplification can be simply added on to the 5' end of each primer (Mullis *et al.*, 1986; for example, see Hamshere *et al.*, 1991). After amplification, the products are purified and digested with (an) appropriate enzyme(s) and ligated into a vector with complementary ends.

Sequencing clones of PCR products is not as "honest" an indication of true sequence as direct sequencing, in which a population of PCR products is sequenced. This is due to the error rate of *Taq* polymerase in the PCR reaction. It is estimated that under the conditions used in these experiments, erroneous nucleotides are incorporated at a rate of 0.25% in a 30 cycle reaction (Innis and Gelfand, 1990).

Two 30-cycle reactions were employed in the second preparation of PCR products for cloning (giving rise to clone set 2): the first reaction resulted in the initial generation of PCR products, while the second was used to amplify DNA eluted from excised bands. Thus products of approximately 224 bases (the theoretical length of the VASE-containing product generated by the primers indicated in Figure 4.2.1) which result from a total of sixty cycles of PCR amplification may contain nucleotides which do not correspond to
those of the original template. Such errors are concealed when sequencing a heterogeneous population of PCR products, as an aberration in one product will be masked by "native" sequence of others. In contrast, cloning effectively amplifies a single PCR product and any erroneous bases it may contain. When using the latter method, it is thus advisable to obtain identical sequence from at least two clones derived from different PCR amplifications to compensate for this procedural weakness.

Alignment of sequences obtained from samples 28, 38 and direct sequence (Figure 4.5.4) suggests that two nucleotide misalignments leading to base pair substitutions, have occurred during the amplification of VASE and its 5' flanking sequence in sample 28. "G" replaces "A" at the second position in the triplet immediately preceding the exon 7/8 splice junction, and another "G" replaces "T" at the third position in the first triplet in VASE.
Comparison of VASE sequences obtained from the JR-1 rhabdomyosarcoma cell line
Sequence from samples 28 and 38 of clone set 2 are compared with that obtained from direct sequencing (see Figure 4.3.4). VASE sequence is in bold type, and bases which vary from previously established sequence are framed.
Figure 4.5.4

5'   | 3'
---|---
Sample 28: GAA AAG GCC TCG TGG ACT CGA CCA GAG AAG CAA GAGACT CTG
Sample 38: GAA AAG GCT TCG TGG ACT CGA CCA GAG AAG CAAGAG ACT CTG
Direct Sequence: GAA AAG GCT TCG TGG
The proximity of these errors suggests that this region is a "hot spot" - where reading errors by *Taq* polymerase are likely to occur (Eckert and Kunkel, 1991). In this respect, it is notable that difficulties were encountered in sequencing sample 38 across these two triplets - clear sequence could not be resolved without the addition of Mn+ to the reaction. Finally, by virtue of its function, it is not surprising that a splice site might have an unusual topography which could challenge the fidelity of *Taq* polymerase, and/or the efficiency of Sequenase 2.

Data obtained from sequencing selected samples from clone sets 1 and 2 confirms the (Southern) blotting profile of these samples (Table 4.5) and suggests that Southern blotting is a reliable means by which to identify samples of interest. Generally, the size of insert as estimated from Southern blots, roughly corresponds with the number of bases sequenced, in cases where the full insert has been characterised.

### 4.6 Conclusion

The data presented here demonstrates that NCAM RNA in neuroblastoma and rhabdomyosarcoma cell lines is alternatively spliced at the exon 7/8 junction. Some species include VASE, while others lack it. The considerable level of VASE-containing species suggested by these data, is in contrast to that observed in a similar analysis in which neural and muscle cell lines were found to contain very little (if any) VASE. It could be that transformation and/or a reversion to embryonic phenotype is responsible for increased levels of VASE, although this runs counter to the finding that VASE is upregulated with development and is generally associated with static, differentiated tissue.

Direct sequence analysis implied that some bands, visible both on ethidium bromide-stained gels and on hybridisation, were artifacts. It remains a possibility that species containing novel, alternatively-spliced
<table>
<thead>
<tr>
<th>Column (A)</th>
<th>Column (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary of data presented in Figures 4.4.2 - 4.4.5</td>
<td>Summary of data presented in Figures 4.5.2 and 4.5.3</td>
</tr>
</tbody>
</table>
Table 4.5

<table>
<thead>
<tr>
<th>clone set</th>
<th>Sample</th>
<th>A. Data from Southern Blots</th>
<th>B. Sequencing Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>size of insert (bp)</td>
<td>blotting profile</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VASE</td>
<td>non-VASE</td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>180</td>
<td>-</td>
</tr>
<tr>
<td>&quot;</td>
<td>20</td>
<td>250</td>
<td>-</td>
</tr>
<tr>
<td>&quot;</td>
<td>30</td>
<td>750</td>
<td>-</td>
</tr>
<tr>
<td>&quot;</td>
<td>45</td>
<td>225</td>
<td>-</td>
</tr>
<tr>
<td>&quot;</td>
<td>46</td>
<td>750</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>160</td>
<td>-</td>
</tr>
<tr>
<td>&quot;</td>
<td>28</td>
<td>1350</td>
<td>√</td>
</tr>
<tr>
<td>&quot;</td>
<td>32</td>
<td>180</td>
<td>-</td>
</tr>
<tr>
<td>&quot;</td>
<td>38</td>
<td>530</td>
<td>√</td>
</tr>
</tbody>
</table>
exons exist at levels undetected by the methods used here. It is also possible that the 240 bp product of the JR-1 rhabdomyosarcoma cell line, observed in lane 4, Figure 4.2.5, contains novel sequence. Electrophoretic separation of JR-1 PCR products through a denaturing polyacrylamide gel would clarify the status of this band. However, its identical size with a band observed on agarose gel electrophoresis and demonstrated to be artifactual, suggests that it does not reflect a genuine alternatively-spliced product. It may be that the ratio of RNA which contains VASE to that which lacks VASE, differs between the rhabdomyosarcoma cell line and neuroblastoma cell lines. This might result in a 240 bp product which does not resolve on electrophoresis through a polyacrylamide gel. Such variance might also explain the difficulty encountered in the direct sequencing of the 224 bp product of JR-1 (Figure 4.3.6).

Seven neuroblastoma and one rhabdomyosarcoma cell lines were found to have similar profiles on electrophoretic separation of PCR products through an agarose gel, followed by Southern blotting. An additional six rhabdomyosarcoma cell lines were found to have an identical profile to that obtained with JR-1 (A. Hancock, personal communication). Thus, from a diagnostic perspective, alternative splicing patterns at the exon 7/8 splice junction are not sufficiently distinct to lend themselves to tumour differentiation, as judged by the splicing patterns in the cell lines examined here.
CHAPTER 5

EXPRESSION OF THE MUSCLE-SPECIFIC DOMAIN (MSD1) IN NEOUROBLASTOMA AND RHABDOMYOSARCOMA CELL LINES

5.1 Introduction

The exon 12/13 junction, at which MSD1 is alternatively-spliced, lies between the two fibronectin-like repeats found in the "stem" of the molecule - i.e. between the immunoglobulin-like domains and the site of membrane attachment (Figure 5.1). In this respect, NCAM is similar to fibronectin - both molecules contain an alternatively-spliced domain at the nexus of two fibronectin type III (-like) repeats. As its name suggests, the expression of MSD1 is believed to be restricted to muscle tissue.

Analysis of an NCAM human genomic clone indicated that MSD1 comprises four constituent exons of 15, 48, 42 and 3 bp (Thompson et al., 1989). The first three exons were designated MSD1a, MSD1b and MSD1c. The 3 bp exon which has only gradually acquired the status of an exon since its discovery, is known as the "AAG triplet". It has been referred to as MSD1d (Prediger et al., 1988); a more appropriate term implying the true exon nature of the AAG triplet. Another exon, SEC, found immediately 3' of the MSD1 exons a, b and c, was found to be alternatively spliced in cDNA derived from a skeletal muscle mRNA (See Chapter 1).

Since the discovery of MSD1, analogues have been identified in chicken cardiac and skeletal muscle (Prediger et al., 1988), rat heart (Reyes et al., 1991) and mouse muscle (Hamshere et al., 1991). However, it cannot be assumed that an MSD1 analogue will be found in skeletal or cardiac muscle across species, as sensitive assays involving PCR amplification have revealed only an analogue of MSD1a in such tissues of the frog Xenopus

193.
Figure 5.1

a) Human sequence of MSD1 with flanking sequences and corresponding amino acids
The underlined sequence is that of MSD-1; its constituent exons are as indicated above the sequence (Dickson et al., 1987; Thompson et al., 1989).

b) Schematic representation of a 125 kDa polypeptide encoded by RNA containing MSD1
The five immunoglobulin-like domains are in the amino-terminal half of the molecule, as illustrated. MSD1 is found between two fibronectin type III-like repeats found in the extracellular portion, proximal to the plasma membrane.

Key:

= protein encoded by MSD1 domain
= plasma membrane
= glycosyl phosphatidylinositol anchor
= Ig-like domain
= fibronectin type III-like repeat
= polypeptide encoded by constitutively-splited RNA
Figure 5.1

a)

**Exon 12**

<table>
<thead>
<tr>
<th>CAG CCA GTC C AT</th>
<th>AGC CCT CCT CCA CCG GCA TCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gln Pro Val His Ser Pro Pro Pro Pro Ala Ser</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GCT AGC TCG TCT ACC CCT GTT CCA TTG TCT CCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala Ser Ser Ser Thr Pro Val Pro Leu Ser Pro</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CCA GAT ACA ACT TGG CCT CTT CCT GCC CTT GCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pro Asp Thr Thr Trp Pro Leu Pro Ala Leu Ala</td>
</tr>
</tbody>
</table>

b)

**Exon 13**

<table>
<thead>
<tr>
<th>ACC ACA GAA CCA GCT AAA G GG GAA CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thr Thr Glu Pro Ala Lys Gly Glu Pro</td>
</tr>
</tbody>
</table>

b)
laevis (Zorn and Krieg, 1991). Neither can it be assumed that MSD1 and its constituents are confined exclusively to muscle tissue. Even before Joanne Thompson established that MSD1 comprises several smaller exons, S1 nuclease experiments had suggested the presence of a 15 bp alternatively-spliced exon between exons 12 and 13 in mouse brain (Barthels et al., 1987), later confirmed to be MSD1a (Santoni et al., 1989b). More recently, PCR analysis of chicken feather filament tissue, derived from both the epidermis and dermis, suggests the presence of mRNA containing the avian 93 bp analogue of MSD1 (Marsh and Gallin, 1992).

It should be noted that the 5' half of chicken MSD1 does not bear the same high degree of homology with human MSD1 as those of mouse and rat. Like human MSD1, chicken MSD1 comprises four exons: MSD1a (15 bp), MSD1b (33 bp), MSD1c (42 bp) and MSD1d (3 bp). Unlike human MSD1, Northern blotting suggests that the MSD1 domain is expressed in transmembrane as well as GPI-linked isoforms in chicken cardiac and skeletal muscle. Exons MSD1c and MSD1d are similar to those found in human skeletal muscle with a nucleotide homology of 76%, but MSD1a and MSD1b have only a 37% nucleotide homology (and 19% amino acid homology) with their "counterparts" in human MSD1 (Prediger et al., 1988). This is in contrast with the mouse, rat and human, where MSD1a is perfectly conserved across these species.

As exons, MSD1a, MSD1b, MSD1c and MSD1d have the potential for independent splicing, thereby generating further diversity in NCAM isoform structure. Indeed, recent research has revealed that they are found in different combinations in developing skeletal and cardiac muscle in the mouse and rat respectively (Hamshere et al., 1991; Reyes et al., 1991; see Table 5.1). Furthermore, certain combinations appear to be subject to developmental regulation. For instance, in mouse muscle, inclusion of MSD1a, b and c as a unit, increases upon denervation and during
Alternative splicing patterns of the MSD1 domain

Alternative splicing patterns of the MSD1 domain have been characterised in different tissues of several species. Some have been discovered by restriction and sequence analysis of cDNA clones, while others have been detected by PCR amplification, followed by sequencing. The presence of yet others were detected by the size of PCR products generated across a splice junction; confirmatory sequence analysis had not been performed.

References:
1. Dickson et al., 1987
2. Thompson et al., 1989
3. Carbone et al., 1991
4. Hamshere et al., 1991
5. Santoni et al., 1989
6. Reyes et al., 1991
7. Reyes et al., 1993
8. Prediger et al., 1988
9. Marsh and Gallin, 1992
Table 5.1

<table>
<thead>
<tr>
<th>Species</th>
<th>Tissue</th>
<th>Splice variant</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>Skeletal Muscle</td>
<td>abcd</td>
<td>1, 2</td>
</tr>
<tr>
<td></td>
<td>SCLC</td>
<td>abc</td>
<td>3</td>
</tr>
<tr>
<td>Mouse</td>
<td>Neuroblastoma</td>
<td>d</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Muscle and/or</td>
<td>a</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Developing</td>
<td>abc</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Developing Muscle</td>
<td>abcd</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ac</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ad</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cd</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Brain</td>
<td>a</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ad</td>
<td>4, 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d</td>
<td>4, 5</td>
</tr>
<tr>
<td>Rat</td>
<td>Heart and/or</td>
<td>a</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Developing</td>
<td>abcd</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Heart</td>
<td>abc</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>abd</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>acd</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ad</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>bcd</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>bd</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Brain</td>
<td>a</td>
<td>6 &amp; 7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ad</td>
<td>6 &amp; 7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d</td>
<td>6 &amp; 7</td>
</tr>
<tr>
<td></td>
<td>Phaeochromocytoma</td>
<td>d</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Schwannoma</td>
<td>d</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Glioma</td>
<td>d</td>
<td>4</td>
</tr>
<tr>
<td>Chicken</td>
<td>Skeletal muscle</td>
<td>abcd</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Feather filament</td>
<td>abcd</td>
<td>9</td>
</tr>
</tbody>
</table>
development, with a concomitant reduction in mRNA containing MSD1d alone. Similarly, low levels of mRNA containing MSD1a and MSD1a plus d are detected upon fusion of myoblasts and in myotube tissue only (Hamshere et al., 1991).

Alternative splicing of the constituent exons comprising MSD1 implies that each exon influences NCAM function. Exactly how these exons influence function remains to be elucidated. Current hypotheses are based largely upon the nature of the corresponding amino acid sequence and its context position in the molecule. MSD1a encodes three consecutive proline residues; four when in tandem with MSD1b in the mouse and human. It may be that this is sufficient to introduce a kink or "hinge" in the tertiary structure of NCAM, and thereby affect the orientation of the molecule. Supporting this theory are two facts: (1) MSD1 bears some homology to the hinge region of Igs (Walsh, 1988) and (2) under the electron microscope, NCAM appears as a rod with a bend at a position near MSD1; i.e. between the immunoglobulin-like domains and the site of cell surface attachment (Hall and Rutishauser, 1987). The insertion of MSD1 or its constituents in the proximity of the hinge may influence its angle.

The MSD1 sequence is rich in serine and threonine - amino acids which can be O-glycosylated - and evidence strongly suggests that the MSD1 sequence is O-glycosylated (Walsh et al., 1989). However, the functional consequences of O-glycosylation also remain unclear, although it is postulated to play a role in stabilising membrane molecules, thereby increasing surface expression.

Temporal regulation of MSD1-containing mRNA suggests that the corresponding proteins of such species may differ in their capacity for mediating adhesion, when compared with those encoded by mRNA which lacks MSD1 (See Chapter 1). It was hypothesised that the 145 kDa transmembrane isoform, lacking MSD1 and observed in myoblast extracts,
mediates fusion, implying a greater capacity for homophilic adhesion than the GPI-linked MSD1-containing isoforms. It was suggested that excision of MSD1 might confer this capacity. However, transfectants containing transcripts encoding GPI-linked NCAM both containing and lacking MSD1 showed no difference in intercellular adhesion (Pizzey et al., 1989). Similar experiments revealed no difference in their capacity for supporting neurite outgrowth (Doherty et al., 1989).

Other proposed functions of MSD1 and its associated carbohydrate, include the cis-binding of NCAM molecules within the plane of the cell membrane (G. Dickson, personal communication), and influencing ligand properties of the molecule (Walsh et al., 1989).

Published data on the inclusion of MSD1 and its constituents in tumour tissue is sparse. Rat tumour cell lines have been analysed by PCR analysis followed by sequencing, and found to contain exon d (Hamshere et al. 1991). The analysis of human small cell lung carcinoma cell lines by PCR suggests that MSD1 is present in NCAM expressed by one small cell lung carcinoma cell line, although data was not presented in support of this claim (Carbone et al., 1991).

Cumulative data clearly demonstrates that alternative splicing of the MSD1 exons is subject to complex spatial and temporal regulation. It was therefore of interest to examine the splicing patterns at the exon 12/13 splice junction in neuroblastoma and rhabdomyosarcoma cell lines. Differences in splicing patterns between the two types of cell line might assist in distinguishing one from the other.

5.2 Generation of PCR Products which Span the Exon 12/13 Splice Junction and their Analysis by Southern Blotting

The alternative splicing pattern at the exon 12/13 splice junction of eight neuroblastoma cell lines and one rhabdomyosarcoma cell line was investigated by PCR amplification of DNA complementary to RNA spanning
this region. Six additional rhabdomyosarcoma cell lines were analysed by another member of the laboratory.

(a) **Methods**

RNA was extracted from cell lines using guanidium isothiocyanate. Two extractions were made of each cell line growing at an exponential rate, and resuscitated from two different stocks.

Complementary DNA was generated by reverse transcription of RNA template, using an oligonucleotide primer complementary to a region in exon 13, downstream from the exon 12/13 splice junction (Figure 5.2.1). The primer was designed according to recommendations (Innis and Gelfand, 1990) which maximise the probability of efficient, yet specific hybridisation with template.

The position of primers used in PCR amplification across the exon 12/13 splice junction are indicated in Figure 5.2.1; the 3' primer is identical to that used in the synthesis of cDNA by reverse transcription. Conditions for amplification by PCR were identical to those used to amplify a region spanning the exon 7/8 splice junction.

PCR products derived from a T-ALL cell line GH1 were used as a negative control. An additional negative control consisted of all PCR reagents to which water, instead of cDNA, was added. 20 ng of λ9.5 (Dickson *et al.*, 1987) was used as a positive control.

Products of the PCR reaction were electrophoresed through a 2% agarose mini-gel in the presence of 0.5 μg/ml ethidium bromide, to ensure that the reaction had been successful. On visualisation of bands of appropriate sizes, products were electrophoresed through a large 2% agarose gel (20 x 20cm), and transferred to a nylon membrane.

Southern blots were obtained on hybridising immobilised PCR products with a 32P labelled oligonucleotide complementary to the entire length of MSD1a. The melting temperature of the oligonucleotide probe
Figure 5.2.1

Exons 12 and 13 and the positions of primers used in PCR amplification and sequencing
Partial sequence of exons 12 and 13; the respective 5' and 3' primers used in cDNA synthesis and amplification by PCR are framed. *Primer 1, used in sequence analysis only, is underlined. The splice site is indicated by a heavy, black vertical line.

* sequence complementary to
Figure 5.2.1

Exon 12
CTG AAG CCC GAA ACA ACG

5' primer
TAC GCC GTA AGG CTG GCC GCG CT C AAT GGC AAA

GGG CTG GGT GAG ATC AGC GCG GCC TCC GAG TTC

Exon 13
AAG ACG CAG CCA GTC GGG GAA CCC AGT GCA CCT

AAG CTC GAA GGG CAG ATG GGA GAG G AT GGA AAC

TCT ATT AAA GTG AAC CTG ATC AAG CAG GAT GAC

GGC GGC TCC CCC ATC AGA CAC TAT CTG GTC AGG

TAC CGA GCG CTC TCC TCC GAG TGG AAA CCA GAG

ATC AGG CTC CCG TCT GGC AGT GAC CAC GTC ATG
G TCC GAG GCC AGA CCG TCA CT

← 3' primer
was calculated to be 52°C (Wahl et al., 1987) and so a hybridisation temperature of 45°C was chosen. Filters were washed to a final stringency of 0.1% SDS, 0.1 x SSC at 45°C, and exposed to film for 2-3 days. After exposure, filters were stripped of MSD1a probe and exposed to film for an appropriate length of time, to ensure that any residual oligonucleotide would not bias interpretation of data obtained on subsequent hybridisations.

Immobilised PCR products were re-hybridised with the MSD1b oligonucleotide, so called because it is complementary to the MSD1b sequence (see Figure 5.1). Immobilised PCR products were hybridised at 62°C; 7°C below the theoretical melting temperature of the probe (Wahl et al., 1987). Washes were carried out to a final stringency of 0.1% SDS, 0.1 x SSC, at 62°C. After exposing the blot for an adequate length of time (approx. 2 days), it was stripped of MSD1b oligonucleotide and exposed to film to ensure adequate removal of probe. The same blot was re-hybridised with oligonucleotides complementary to the MSD1c sequence, the entire MSD1 domain, and then with λ9.5, the latter of which was labelled with 32P using random hexamers. Hybridisation with the MSD1c probe was carried out at 60°C, and hybridisation with the MSD1 and λ9.5 probes at 70°C. Average exposure times for blots hybridised with these oligonucleotide probes were 1.5 days, 1 day and 1 hour respectively. Prior to each hybridisation, the blot was stripped of previous probe as above. It is impossible to detect the presence of MSD1d using this method. A 2% agarose gel does not permit adequate resolution to differentiate MSD1d-containing products on the basis of size, and indiscriminate binding of a complementary triplet probe would obscure its detection.

Three blots were obtained; two of which contained PCR products generated from RNA obtained from different stock cultures. Each blot was hybridised with all five probes, although the order in which each blot was hybridised with different probes varied. This was to ensure that loss of immobilised PCR products, which occurs upon hybridisation did not bias the results of subsequent hybridisations. The λ9.5 "probe", however, was
applied ultimately, as it proved especially difficult to remove from the blot. The data presented here is representative of all three blots, except where specified.

(b) Results

Southern blotting of PCR products generated across the exon 12/13 splice junction, suggests that neuroblastoma cell lines produce RNA in which constituents of the MSD1 domain are present. Three of the eight neuroblastoma cell lines yield PCR products in which MSD1a is detected, while all eight yield PCR products in which yield bands on hybridisation with the MSD1b oligonucleotide probe. MSD1c is not detected in PCR products arising from any of the neuroblastoma cell lines, although it is readily detectable in the JR-1 rhabdomyosarcoma cell line, alongside MSD1a and MSD1b.

(i) Southern blotting using the MSD1a oligonucleotide probe

On Southern blotting PCR products with the MSD1a oligonucleotide (Figure 5.2.2) a similar pattern is observed in lanes containing PCR products from the neuroblastoma cell lines GOTO, IMR-32 and NB1 (lanes 2, 3 and 5). All revealed a major product of approximately 270 bp. A minor band of approximately 300 bp was also observed in the latter two cell lines. No signal was seen in lanes containing PCR products from the other neuroblastoma cell lines, but a major band of 360 bp was noted in the lane containing PCR products generated from the rhabdomyosarcoma cell line JR-1 (lane 12). Fainter bands of approximately 270 and 300 bp were also observed in this lane. The positive control, λ9.5 (lane 10) yielded a single detectable PCR product of 360 bp, consistent with the calculated size of a PCR product containing the entire MSD1 domain. No bands were observed in either of the negative controls containing water (lane 11) or cDNA generated from GH1 (lane 9).
Figure 5.2.2

Southern blot of PCR products spanning the exon 12/13 splice site from a selection of cell lines; hybridised with the MSD1a oligonucleotide

1. SK-N-BE (2C)  
2. GOTO  
3. IMR-32  
4. Kelly  
5. NB1  
6. NB100  
7. SK-N-SH  
8. PCF  
9. GH1  
10. λ9.5  
11. H2O  
12. JR-1

1-8 = neuroblastoma cell lines  
9 = T-cell acute lymphoblastic leukaemia cell line; negative control  
10 = λ9.5, PCR'd; positive control  
11 = H2O; additional negative control  
12 = rhabdomyosarcoma cell line
Figure 5.2.2

- 360bp
- 300bp
- 270bp
(ii) **Southern blotting using the MSD1b oligonucleotide probe**

On Southern blotting PCR products with the MSD1b oligonucleotide (Figure 5.2.3) a similar pattern comprising bands of approximately 250 and 300 bp is observed in lanes containing PCR products from 7/8 neuroblastoma cell lines (lanes 1 to 7). In the blot displayed, the IMR-32 NBL cell line lacks the 250 bp band, although this deficit was not reproducible. The intensity of signal however, varies from sample to sample, and between bands; the 300 bp band is stronger in the majority of lanes containing PCR products from neuroblastoma cell lines.

A major band of 360 bp was observed in the lane containing PCR products generated from the rhabdomyosarcoma cell line JR-1 (lane 12). Shorter exposures of this blot reveals a minor band of approximately 300 bp (data not displayed). The positive control, λ9.5 (lane 10) yielded a single detectable PCR product of 360 bp, consistent with the calculated size of a PCR product containing the entire MSD1 domain. No bands were observed in either of the negative controls containing water (lane 11) or cDNA generated from GH1 (lane 9). It must be noted, however, that the intensity of signal generated by JR-1 in lane 12 might obscure bands of a similar size arising in lane 11, although shorter exposures of the blot revealed no such bands.

(iii) **Southern blotting using the MSD1c oligonucleotide probe**

On Southern blotting PCR products with the MSD1c oligonucleotide (Figure 5.2.4) no product is observed in any of the neuroblastoma cell lines (lanes 1 to 8).

A major band of 360 bp and a minor band of approximately 270 bp was observed in the lane containing PCR products generated from the rhabdomyosarcoma cell line JR-1 (lane 12). The positive control, λ9.5 (lane 10) yielded a single detectable PCR product of 360 bp, consistent with the
Figure 5.2.3

Southern blot of PCR products spanning the exon 12/13 splice site from a selection of cell lines; hybridised with the MSD-1b oligonucleotide

1. SK-N-BE (2C)  5. NB1  9. GH1
2. GOTO  6. NB100  10. λ9.5
4. Kelly  8. PCF  12. JR-1

1-8 = neuroblastoma cell lines
9 = T-cell acute lymphoblastic leukaemia cell line; negative control
10 = λ9.5, PCR'd; positive control
11 = H2O; additional negative control
12 = rhabdomyosarcoma cell line
Figure 5.2.4

Southern blot of PCR products spanning the exon 12/13 splice site from a selection of cell lines; probed with MSD1c oligonucleotide

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SK-N-BE (2C)</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>GOTO</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>IMR-32</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>Kelly</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>GH1</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>$\lambda 9.5$</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>H$_2$O</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>JR-1</td>
<td></td>
</tr>
</tbody>
</table>

1-8 = neuroblastoma cell lines  
9 = T-cell acute lymphoblastic leukaemia cell line; negative control  
10 = $\lambda 9.5$, PCR'd; positive control  
11 = H$_2$O; additional negative control  
12 = rhabdomyosarcoma cell line
Figure 5.2.4

- 360bp
calculated size of a PCR product containing the entire MSD1 domain. No bands were observed in either of the negative controls containing water (lane 11) or cDNA generated from GH1 (lane 9).

(iv) Southern blotting using an oligonucleotide probe complementary to the MSD1 domain

On Southern blotting PCR products with an oligonucleotide complementary to the entire MSD1 domain (Figure 5.2.5), a pattern was revealed which was very similar to that observed on hybridisation with MSD1b. Bands of approximately 250 and 300 bp and varying intensity were observed in lanes containing products derived from neuroblastoma cell lines (lanes 1 to 7). Bands of approximately 250, 300 and 360 bp were observed in the lane containing products from the JR-1 rhabdomyosarcoma cell line (lane 12). The positive control, λ9.5 (lane 10) yielded a major PCR product of 360 bp, and a minor band of approximately 250 bp although the latter was not found to be reproducible. The neuroblastoma cell line PCF and the two negative controls: H2O (lane 11) and the GH1 cell line (lane 9) yielded no PCR products detected by hybridisation with MSD1.

(v) Southern blotting using cDNAs generated by random priming of λ9.5, as a collective probe

On Southern blotting PCR products with the "λ9.5" probe (Figure 5.2.6), a pattern was revealed which was again similar to that observed on hybridisation with MSD1b. A strong band of approximately 250 bp was observed in all the neuroblastoma cell lines (lane 1-7) save PCF (lane 8), in which it was seen to be present, but fainter. Strong bands of approximately 270 bp were observed as products from three of the neuroblastoma cell lines (GOTO in lane 2; IMR-32 in lane 3; and NB1 in lane 5), while weak bands of the same size were observed in lanes containing products from 2 other neuroblastoma cell lines (NB100, lane 6 and PCF, lane 8). Weak bands of approximately 300 bp were observed in five neuroblastoma cell lines (lanes 1 to 5).
Southern blot of PCR products spanning the exon 12/13 splice site from a selection of cell lines; hybridised with the MSD-1 oligonucleotide

1. SK-N-BE (2C)  
2. GOTO  
3. IMR-32  
4. Kelly  
5. NB1  
6. NB100  
7. SK-N-SH  
8. PCF  
9. GH1  
10. λ9.5  
11. H2O  
12. JR-1

1-8 = neuroblastoma cell lines  
9 = T-cell acute lymphoblastic leukaemia cell line; negative control  
10 = λ9.5, PCR'd; positive control  
11 = H2O; additional negative control  
12 = rhabdomyosarcoma cell line
Bands of approximately 250, 300 and 360 bp were revealed in the lane containing products derived from the rhabdomyosarcoma cell line JR-1 (lane 12). The positive control, λ9.5 yielded a band of 360 bp, consistent with a PCR product containing the entire MSD1 domain. No product was seen in the lane containing the negative control of water (lane 9), but a faint band of approximately 270 bp is observed in the negative control cell line GH1 (lane 11).

(c) Discussion

The data presented here suggests that alternative splicing of the MSD1 domain occurs in the RNA of neuroblastoma and rhabdomyosarcoma cell lines examined. Furthermore, it is possible to distinguish the two types of cell by the alternatively spliced exons found at the exon 12/13 splice site. Southern blots of six other rhabdomyosarcoma cell lines (A. Hancock; personal communication) suggest that the data obtained from the JR-1 rhabdomyosarcoma cell line is representative of the larger group.

The theoretical sizes of possible PCR products generated by the primers indicated in Figure 5.2.1, are summarised in Table 5.2.

Figure 5.2.2 suggests that 3 of the 8 neuroblastoma cell lines analysed, synthesise RNA in which the MSD1a exon is present. This is consistent with the presence of a strong band of approximately 270 bp (lanes 2, 3 and 5). JR-1, the representative rhabdomyosarcoma cell line also appears to synthesise this species, in addition to a species of 360 bp, consistent with the presence of MSD1a, b, c and possibly d exons (lane 12). The latter band co-migrates with that arising from the amplification of λ9.5, the positive control, in which the entire MSD1 domain is known to be present. The faint bands of approximately 300 bp are less easy to rationalise. Two different events might give rise to "non-specific" binding (defined below), and the aberrant migration of MSD1a-containing PCR products.
Table 5.2

**Theoretical sizes of potential PCR products generated by amplification across the exon 12/13 splice junction**

Potential PCR products generated by amplification across the exon 12/13 splice junction contain exons MSD1a (15 bp), MSD1b (48 bp), MSD1c (42 bp) and MSD1d (3 bp) or any combination of these four. The theoretical sizes of some products are presented here. The theoretical size of products in which more than one MSD1 exon is spliced, can be calculated by adding the sum of the sizes (given above) to 255 bp.
Table 5.2

<table>
<thead>
<tr>
<th>MSD1 exons found at exon 12/13 splice junction</th>
<th>Theoretical size of PCR product (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>255</td>
</tr>
<tr>
<td>a</td>
<td>270</td>
</tr>
<tr>
<td>b</td>
<td>303</td>
</tr>
<tr>
<td>c</td>
<td>297</td>
</tr>
<tr>
<td>abc</td>
<td>360</td>
</tr>
</tbody>
</table>
"Non-specific" binding is a convenient but misleading term. It implies that an oligonucleotide might bind to another which has no apparently relevant sequence to itself. In fact, hybridisation only occurs between two oligonucleotides if their sequences are sufficiently complementary to permit annealing. Thus, it may take place if a partial homology exists - the likelihood of successful hybridisation depends on the extent and distribution of non-complementary sequences, the constitution of complementary sequence and reaction conditions. For instance, a stretch of non-complementary sequence in the middle of two otherwise aligned oligonucleotides, is more likely to interfere with their hybridisation than an identical sequence at their termini. Effects of sequence constitution reflect the fact that base pairs composed of guanidine and cytosine confer greater stability than those composed of adenine and thymidine residues.

For the purpose of discussion, hybridisation between oligonucleotides with partial homology is referred to as being "non-specific". It may be that the faint band of 300 bp observed on hybridisation of the blot with the MSD1a oligonucleotide is "non-specific". If this were the case, one would expect to observe such a band in other cell lines; a major PCR product of approximately 300 bp appears to be present in the PCR product mixtures arising from all neuroblastoma cells examined here (see below).

The alternative hypothesis seems a more likely explanation. It may be that a sub-population of PCR products containing MSD1a migrate at a slower rate than those which constitute the band of 270 bp. The formation of heteroduplces and oligonucleotide fusion during the PCR reaction could give rise to such products.

Figure 5.2.3 suggests that 7/8 neuroblastoma cell lines (lanes 1-7) and the representative JR-1 rhabdomyosarcoma cell line (lane 10) examined here, synthesise RNA which includes MSD1b. The band of approximately
300 bp seen in the NBL and RMS cell lines (although this is not apparent in the case of JR-1 from the photograph displayed here) is consistent with a PCR product in which MSD1b is present - either alone or in concert with MSD1d. The presence of MSD1b in the major 360 bp band of JR-1 (lane 12) corroborates its identity as an analogue of that seen in the positive control, λ9.5 (lane 10).

Figure 5.2.4 suggests that MSD1c is absent from all neuroblastoma cell lines examined here (lanes 1-8), but present in the 360 bp product derived from the rhabdomyosarcoma cell line JR-1 (lane 12). This cell line therefore appears to synthesise RNA in which MSD1a, MSD1b and MSD1c are spliced in tandem at the exon 12/13 splice junction.

Hybridisation of the immobilised PCR products with an oligonucleotide complementary to the entire MSD1 domain (Figure 5.2.5) made visible all species detected by previous probes. No species was detected which had not been identified by prior hybridisation except the faint non-reproducible band of 250 bp observed in the positive control (lane 10). It may be that the alternatively-spliced SEC exon of 239 bp is absent from RNA of the neuroblastoma cell lines; this exon has yet to be found spliced independently of MSD1. In addition, hybridisation of two of the three sets of PCR products with an oligonucleotide complementary to a region in the SEC exon, failed to identify any species. These results, however, should be interpreted with caution, as no positive control was available.

cDNA generated from λ9.5 by random hexamers was used as a collective probe to identify any NCAM-positive PCR products, regardless of their MSD1 content. While most species previously identified were observed, the band of 300 bp, so strongly visible on hybridisation with MSD1b and MSD1, was very faint in lanes containing products from five of
the neuroblastoma cell lines (lanes 1-5) and not visible in NB100 or SK-N-SH (lanes 6 and 7). In contrast, the band of 250 bp hybridised strongly. Its size and intensity of signal (with reference to previous blots) suggest that this band is derived from RNA in which exons 12 and 13 are either flush, or separated by the triplet MSD1d.

The presence of a faint band at approximately 270 bp in the lane containing products from the negative control cell line GH1 suggests that one of two possible events has taken place. Either the "9.5 probe" has bound a non-specific product of 270 bp (the heterogeneous nature of the probe might contribute to the likelihood of such an event) or low levels of NCAM are in fact present in the GH1 cell line - sufficiently low to go undetected by immunofluorescence or Northern blotting. However the 270 bp band does not appear to hybridise oligonucleotide probes which are complementary to any of the MSD1 exons; a finding which is inconsistent with its size, were it derived from NCAM. If the former hypothesis is correct, then bands of identical weight observed in lanes 6 and 8 should be interpreted with caution.

The fact that the neuroblastoma cell line PCF would appear to have little or no RNA encoding NCAM, accords with the FACs results presented in Chapter 3, but contrasts with the levels of NCAM-positive PCR product obtained using primers which span the exon 7/8 junction (see data presented in Chapter 4, section 2). This discrepancy could be explained by the fact that different primers have different binding affinities with template and will therefore reverse-transcribe and/or amplify with varying efficiencies (for example, see Garson et al., 1992).

The length of time required to expose film varied between replicate blots. For instance, film exposed to blots hybridised with the VASE oligonucleotide could take between 1.5 hrs and 4 hrs to develop.
Variations in batch-to-batch PCR efficiency or the efficiency of transfer from gel to GeneScreen could account for these differences.

More marked temporal variations were observed in the required exposure of film to blots hybridised with different oligonucleotides. For example, products generated by PCR amplification across the exon 7/8 splice junction and hybridised with the VASE probe (Figure 4.2.2) required approximately two hours to develop, while those spanning the exon 12/13 splice junction and hybridised with the MSD1b probe (Figure 5.2.3) required an average of 2 days. These time requirements are broadly representative of replicate experiments, and demonstrate no correlation with labeling efficiencies. It could be that such variations are due to the relative proportion of each splice variant in its respective pool of PCR products. Blots of the same PCR products hybridised either with a probe complementary to constitutively-spliced exon 7 (Figure 4.2.3) or randomly-primed constitutive sequence (Figure 5.2.6) support this hypothesis. These suggest that VASE-containing products constitute a much greater proportion of "exon 7/8" products, than do MSDb-containing products in a population of "exon 12/13" products. Consistent with this was the intensity of ethidium-stained bands in gels (not displayed) giving rise to the blots - a band of approximately 224 bp, consistent in size with the VASE-containing product, was always clearly observed. In contrast, the level of MSD1b splice variant was probably too low to be visualised; a band of the appropriate size (approximately 300 bp) was not observed on any gel.

While "non-specific" and/or aberrant migration may obscure an easy interpretation of the blotting data, two clear conclusions can be drawn from the data presented in this section. Firstly, some neuroblastoma cell lines synthesise RNA which includes MSD1a; others do not. Secondly, splicing
patterns at the exon 12/13 junction of the neuroblastoma and rhabdomyosarcoma cell lines examined here, demonstrate a reproducible difference in the alternative splicing patterns of the two types of cell. The rhabdomyosarcoma cell line JR-1, and six others (A. Hancock; personal communication) contain a variety of NCAM RNA species but in all the cell lines MSD1c was detectable. In contrast, all eight neuroblastoma cell lines were found to lack MSD1c. This may prove to be of diagnostic use in the differentiation of these two "small round cell" tumours.

Sequence analysis of PCR products detected by Southern blotting is desirable. As data in Chapter 4, section 3 illustrates, it is possible that some bands may be artifactual. Sequencing would confirm the suspected identities of the bands observed, and also detect the presence or absence of MSD1d.

5.3 Cloning of PCR Products from the NB-1 and Kelly Neuroblastoma Cell Lines: Generation of Clone sets a and b

Direct sequencing requires a homogeneous template. Results from Southern blotting suggest that some of the PCR products resulting from amplification across the exon 12/13 splice junction are similar in size. For instance, three of the cell lines - IMR-32, GOTO, and NB1 - yield products which are thought to contain MSD1a. These cell lines also give rise to a product which appears to lack the three larger alternatively-spliced exons; the size difference between these products is approximately 15 bp, enabling them to be visibly separated on a 2% agarose gel albeit by a small margin. It is unlikely that these products could be sufficiently separated by gel electrophoresis to allow for sequence analysis. PCR products were cloned into a plasmid vector, in an attempt to produce homogeneous template and thus facilitate sequencing. Positive clones were subjected to Southern blot analysis to identify clones of interest.
(a) Methods

PCR products spanning the exon 12/13 splice junction were cloned contemporaneously with those spanning the exon 7/8 splice junction. Consequently, it was not possible to implement changes suggested in Chapter 4 for optimising the acquisition of NCAM-positive clones.

Three populations of PCR products were cloned. The first comprised random, "parental generation" PCR products - i.e. products generated directly from cDNA resulting from reverse transcription. Four molar ratios of plasmid vector : PCR product were used: 1:1, 1:3, 1:10 and 1:30. Resultant clones were termed clone set a. A negative control containing ligation reagents and plasmid only was incubated alongside the other reactions, providing a reference by which to estimate the frequency of false positive bacterial colonies.

The second population comprised non-random "first generation" PCR products and gave rise to clone set b. These were generated by eluting products of approximately 270 bp from a gel slice containing "parental generation" DNA and re-amplifying by PCR amplification. Clone set b was derived from PCR products derived from both Kelly and NB1 neuroblastoma cell lines. NB1 gave rise to clone set b₁ and Kelly gave rise to clone set b₂. Products from each cell line were cloned at plasmid : PCR product ratios of 1:1, 1:3 and 1:10. Again, a negative control for the cloning reaction containing ligation reagents and plasmid only, was incubated alongside the other reactions.

The pCR™1000 plasmid, of the TA cloning kit (Invitrogen) was employed as a vector with which to clone PCR products. The molecular features which render the pCR™1000 plasmid an appropriate vehicle for this task are discussed in Chapter 4. Details of the ligation reaction, transformation of E. coli, and the extraction and purification of episomal
DNA were identical to the procedures used in the analysis of PCR products in Chapter 4, section 4.

Positive and negative clones (the latter served as negative controls) were isolated and cultured. Episomal DNA was extracted, and digested with the Not-1 restriction enzyme. As a fragment of 350 bp contains the insertion site, the size of inserted PCR product is estimated by gauging the size of restriction fragments on electrophoretic separation. Digests were Southern blotted using the MSD1 oligonucleotide probe, to determine whether complementary inserted sequence was present. An aliquot of product arising from the PCR amplification of λ9.5 was used as a positive control.

Selected samples from clone sets b₁ and b₂, most of which were previously observed to hybridise the MSD1 oligonucleotide, were Southern blotted with oligonucleotides complementary to MSD1a, MSD1b and MSD1c and with the "λ9.5" probe. Appropriate positive and negative controls were included. Methods of Southern blotting were identical to those described in section 2 of this chapter, and with the exception of that probed with the "λ9.5" probe, each hybridised blot was exposed to film for the same length of time. The latter set of blots was reproduced once.

(b) Results

Different concentrations of random PCR products from the NB1 neuroblastoma cell line were ligated with the pCR™1000 plasmid, giving rise to clone set a. Ligation products were used to transform INVαF' E. coli. Resultant cloning efficiencies have been calculated from data presented in Table 5.3.1. A PCR product : vector ratio of 1:1, 1:3, 1:10 and 1:30 gave corresponding cloning efficiencies of 3%, 3.4%, 2.5% and 8%. No false positives were observed in the negative control.

Southern blot analysis of digested plasmid DNA (Figure 5.3.1) suggests that at least 6 of these samples contain sequence which hybridises
Table 5.3.1

**Colonies arising from the cloning of random PCR products; clone set a**

Random parental generation PCR products designed to span the exon 12/13 splice junction and generated from RNA extracted from the NB-1 neuroblastoma cell line were ligated with the pCR™1000 vector at 5 different vector:insert molar ratios: 1:0 (negative control), 1:1, 1:3, 1:10, and 1:30. The number of positive and negative colonies arising from transformation of *E. coli* by products of each ligation reaction are indicated.
<table>
<thead>
<tr>
<th></th>
<th>1:0 (negative control)</th>
<th>1:1</th>
<th>1:3</th>
<th>1:10</th>
<th>1:30</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>number of colonies</strong></td>
<td>negative</td>
<td>positive</td>
<td>negative</td>
<td>positive</td>
<td>negative</td>
</tr>
<tr>
<td>102</td>
<td>-</td>
<td>321</td>
<td>6</td>
<td>137</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>116</td>
<td>116</td>
<td>5</td>
<td>63</td>
<td>2</td>
</tr>
<tr>
<td><strong>mean</strong></td>
<td>218.5</td>
<td>7</td>
<td>126.5</td>
<td>4.5</td>
<td>76</td>
</tr>
<tr>
<td><strong>cloning efficiency (%)</strong></td>
<td>-</td>
<td>3</td>
<td>3.4</td>
<td>2.5</td>
<td>8</td>
</tr>
</tbody>
</table>
Figure 5.3.1

Southern blot of selected samples from clone set a; hybridised with MSD1 oligonucleotide
Episomal DNA from selected colonies of clone set a was extracted, digested with the Not-1 restriction enzyme, and Southern blotted with the MSD1 oligonucleotide probe. Samples are as indicated and molecular weights are indicated in kb on the right side of the blot.

Key:

+ = positive control; λ9.5, amplified by PCR
samples are as indicated
NC = negative control
the MSD1 oligonucleotide probe. Bands are observed in lanes containing samples 8 (approximately 2850 bp), 19, 20, 27, 34 and 35 (the 5 latter samples are approximately 620 bp in length). Other samples yielded signal of lesser strength; for example samples 1, 10, 11 and 28. The negative control (NC) remained blank, while the four positive controls of λ9.5, observed in each corner, gave rise to a strong band of 360 bp.

PCR products which had been selected on the basis of size were also ligated with pCR™1000 vector, and the ligate cloned. Two sets of reactions were performed; clone set b1 was obtained using PCR products derived from the NB1 neuroblastoma cell line, and clone set b2, with PCR products derived from the Kelly neuroblastoma cell line. Varying the ratio of PCR product : vector resulted in little variation of cloning efficiency (Tables 5.3.2 and 5.3.3).

Southern blot analysis suggests that samples 8 and 13 of clone set b1 (Figure 5.3.2) may contain sequence which is homologous with MSD1. Major bands of approximately 1.7 - >4 kb are observed in the lane containing sample 8, and bands of approximately 0.6 - >4 kb are observed in the lane containing sample 13. Bands of similar sizes to those seen in lanes 8 and 13 are observed in many other lanes, although they are of lesser intensity, and comparable to those observed in the negative control (NC). A strong band of approximately 360 bp is observed in the two lanes containing amplified λ9.5 at the right of the film.

A Southern blot of clone set b2 (Figure 5.3.3) suggested a similar complement to that of clone set b1. Samples 24, 27, 28 and 37 gave rise to major bands, but bands of lesser intensity could be observed in most other lanes. Faint bands were observed in the negative control (NC), although what appeared to be artifact partially obscured this lane. λ9.5, the positive control, gave rise to a strong band of 360 bp although this is not wholly included on the photograph.
Table 5.3.2

*Colonies arising from the cloning of selected PCR products; clone set b1*
Selected parental generation PCR products designed to span the exon 12/13 splice junction and generated from RNA extracted from the NB-1 neuroblastoma cell line were ligated with the pCR™1000 vector at 5 different vector : insert molar ratios: 1:0 (negative control), 1:3, 1:10, and 1:30. The number of positive and negative colonies arising from transformation of *E. coli* by products of each ligation reaction are indicated.
Table 5.3.2

<table>
<thead>
<tr>
<th>Plasmid : PCR product ratio (molar)</th>
<th>1:0 (negative control)</th>
<th>1:3</th>
<th>1:10</th>
<th>1:30</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of colonies</td>
<td>negative</td>
<td>positive</td>
<td>negative</td>
<td>positive</td>
</tr>
<tr>
<td>116</td>
<td>3</td>
<td>49</td>
<td>5</td>
<td>49</td>
</tr>
<tr>
<td>83</td>
<td>4</td>
<td>50</td>
<td>8</td>
<td>40</td>
</tr>
<tr>
<td>mean</td>
<td>99.5</td>
<td>3.5</td>
<td>49.5</td>
<td>6.5</td>
</tr>
<tr>
<td>cloning efficiency (%)</td>
<td>3.3</td>
<td>11.6</td>
<td>11</td>
<td>10</td>
</tr>
</tbody>
</table>


Table 5.3.3

Colonies arising from the cloning of selected PCR products; clone set b2
Selected parental generation PCR products designed to span the exon 12/13 splice junction and generated from RNA extracted from the Kelly neuroblastoma cell line were ligated with the pCR™1000 vector at 4 different vector : insert molar ratios: 1:0 (negative control), 1:3, 1:10, and 1:30. The number of positive and negative colonies arising from transformation of *E. coli* by products of each ligation reaction are indicated.
Table 5.3.3

<table>
<thead>
<tr>
<th>Plasmid : PCR product ratio (molar)</th>
<th>1:0 (negative control)</th>
<th>1:3</th>
<th>1:10</th>
<th>1:30</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>negative</td>
<td>positive</td>
<td>negative</td>
<td>positive</td>
</tr>
<tr>
<td>number of colonies</td>
<td>116</td>
<td>3</td>
<td>65</td>
<td>11</td>
</tr>
<tr>
<td>mean</td>
<td>83</td>
<td>4</td>
<td>79</td>
<td>6</td>
</tr>
<tr>
<td>cloning efficiency (%)</td>
<td>99.5</td>
<td>3.5</td>
<td>72</td>
<td>8.5</td>
</tr>
<tr>
<td></td>
<td>3.3</td>
<td></td>
<td>10.5</td>
<td></td>
</tr>
</tbody>
</table>
Figure 5.3.2

Southern blot of selected samples from clone set b₁; hybridised with MSD1 oligonucleotide
Episomal DNA from selected colonies of clone set b₁ was extracted, digested with the Not-1 restriction enzyme, and Southern blotted with the MSD1 oligonucleotide probe. Samples are as indicated and molecular weights are indicated in kb on the right side of the blot.

Key:

+ = positive control; λ9.5, amplified by PCR
samples are as indicated
NC = negative control
Figure 5.3.2
Figure 5.3.3

Southern blot of selected samples from clone set b$_2$; hybridised with MSD1 oligonucleotide
Episomal DNA from selected colonies of clone set b$_2$ was extracted, digested with the Not-1 restriction enzyme, and Southern blotted with the MSD1 oligonucleotide probe. Samples are as indicated and molecular weights are indicated in kb on the right side of the blot.

Key:

+ = positive control; λ9.5, amplified by PCR
samples are as indicated
NC = negative control
Figure 5.3.3
Samples which gave rise to strong bands on hybridisation with the MSD1 oligonucleotide (samples 8 and 13 of clone set b1 and samples 24, 27, 28 and 37 of clone set b2), were also found to hybridise the MSD1a, MSD1b and MSD1c oligonucleotide probes more strongly than samples 36 and 38 of clone set b2 or the negative control (Figure 5.3.4). Varying intensities of signal were observed; signal obtained using the oligonucleotide complementary to MSD1c was much stronger than that obtained using the other probes. Predictably, the "λ9.5" probe gave rise to a similar binding pattern. The negative controls (NC) were visible, and of an intensity comparable to those observed in samples 36 and 38. Inexplicably, a band of approximately 430 bp - a size which is inconsistent with products arising from an unligated plasmid - was observed in the lanes containing the negative controls. All of the positive controls gave rise to a strong band of approximately 360 bp.

(c) Discussion

The cloning efficiencies obtained in the generation of clone sets a, b1, and b2 did not vary considerably from those presented in the last chapter. There did not appear to be a consistent correlation between PCR product : vector ratio and cloning efficiency (see Tables 5.3.1 - 5.3.3). An average cloning efficiency of 6.8% was obtained. The reasons behind these poor efficiencies remain obscure: possible explanations are presented in Chapter 4.

The Southern blot data suggests that PCR products which contain at least part of the MSD1 sequence have been successfully cloned. In total, twelve samples appear to hybridise the MSD1 probe on Southern blotting: samples 8, 19, 20, 27, 34 and 35 of clone set a (Figure 5.3.1), samples 8 and 13 of clone set b1 (Figure 5.3.2), and samples 24, 27, 28 and 37 (Figure 5.3.3) of clone set b2. However, many other samples also give rise to visible bands albeit of weaker intensity.
Figure 5.3.4

Southern blot analysis of selected samples from clone sets b1 and b2; hybridised with MSD1a, MSD1b, MSD1c and "λ9.5" oligonucleotides

Samples 8 and 13 of clone set b1, and samples 24, 27, 28, 36, 37 and 38 of clone set b2 were digested with Not-I restriction enzyme and Southern blotted with various probes, alongside appropriate positive and negative controls. Samples are as indicated and molecular weights are indicated in kb on the right side of each blot.

Key:

a. blot hybridised with MSD1a probe
b. blot hybridised with MSD1b probe
c. blot hybridised with MSD1c probe
d. blot hybridised with "λ9.5" probe

M = molecular weight markers
NC = negative control
+ = positive control, PCR'd λ9.5
The majority of the samples which give rise to major bands on hybridisation of MSD1, yielded restriction enzyme fragments which are of a size consistent with those seen in Southern blots in section 5.2. For instance samples 19, 20, 27, 34 and 35 of clone set a give rise to an RE fragment of 620 bp. Theoretically, 350 bp of this fragment comprises plasmid sequence (see Figure 4.4.1). Thus an insert of approximately 270 bp appears to be present. Bands of this size were observed on PCR amplification of cDNA arising from reverse transcription of neuroblastoma RNA - the 270 bp species is believed to contain MSD1a alone (see Figure 5.2.2).

Blotting patterns generated by some samples are difficult to rationalise. For instance, samples 24 and 37 of clone set b give rise to major bands of approximately 600 bp, corresponding with an insert of roughly 250 bp. Species of this size were detected in the initial characterisation of cDNAs generated from neuroblastoma cell lines (see Figure 5.2.6) but as a consequence of size and blotting profile, are thought not to contain MSD1 exons a b and c. The fact that these species hybridise the MSD1 oligonucleotide, is inconsistent with their predicted insert size of approximately 250 bp.

Each clone set contains a sample which yields a band of greater size than that expected: sample 8 of clone set a (insert = approx. 2.5 kb), sample 8 of clone set b (insert = approx. 1.35 kb), and sample 28 of clone set b (insert = approx. 1.1 kb). It may be that non-specific hybridisation of the primers during PCR amplifications and/or oligonucleotide fusion has generated these fragments of unpredictable size - sequencing data presented in Chapter 4 has demonstrated that such events can take place.
As clones sets a, b1, and b2 have not been Southern blotted with an oligonucleotide designed to hybridise all NCAM-positive products irrespective of alternatively-spliced sequence at the exon 12/13 splice junction (e.g. the λ9.5 probe), it is impossible to ascertain the proportion of NCAM-positive clones with respect to the total number of clones analysed.

However, the generation of non-NCAM products alongside NCAM-positive species would not be surprising. "Non-specific" bands are seen on gel electrophoresis of PCR products arising from negative control cell line GH1, but they are not observed on hybridisation with any of the oligonucleotides complementary to all or part of the MSD1 sequence. If their analogues are generated alongside NCAM-positive sequences in the neuroblastoma cell lines, a mixed population of PCR products would result. One means by which this possibility might be explored is to sequence the major products observed in the GH1 cell, and then to re-probe the nylon filters used to generate the blots above with a suitable oligonucleotide. The inadvertent amplification of "non-specific" products from whole cell preparations and the appearance of products with no relevant sequence on cloning PCR products into the pCR1000™ vector has been observed by others (G. Middleton and D. M. Hunt; personal communications).

Sequencing relevant clones from clone sets a and b should clarify the unexpected sizes of inserts of the samples 8, 8, and 28 of clone sets a, b1 and b2 respectively. Similarly, sequence analysis should clarify the inconsistent size and blotting profiles of samples 24 and 37 of clone set b2.

5.4 Sequence Analysis of Selected Samples in Clone sets a and b

Samples 8, 19, 20, 27 and 34 from clone set a were sequenced. These hybridised the MSD1 probe on Southern blotting (Figure 5.3.1); sample 8 giving rise to a fragment of unexpected size. Sample 10 of clone set a was also sequenced, as the signal to which it gave rise on Southern
blotting was intermediate in intensity, and therefore difficult to identify as either positive or negative.

Samples 8 and 13 of clone set b₁ were sequenced. These also hybridised MSD1, and sample 8 gave rise to a restriction fragment of unexpected size. Samples 24, 27 and 28 of clone set b₂ were sequenced to verify their suspected composition and samples 1 and 2 of clone set b₂ - neither of which gave strong signal on hybridisation with MSD1 - were sequenced in the hope that consequent data might shed some light on their origin.

(a) Methods

Clones were sequenced using the Sequenase 2 kit (USB), according to instructions supplied by the manufacturer. A variety of primers were used: the 5' primer used in PCR amplification, primer 1, designed to hybridise to a region in exon 13, 3' of the exon 12/13 splice junction (see Figure 5.2.1), and two which were complementary to plasmid sequence, flanking the insertion site. The latter two were designated the "Universal" and "Reverse" primers, and complementary to sequence 3' and 5' of the insertion site respectively (see Figure 4.4.1). Radiolabelled products of the sequencing reaction were analysed by denaturing polyacrylamide gel electrophoresis, as described in Chapter 2.

Sequence analysis was performed using the Genebank database in conjunction with the "Align" (Wilbur-Lipman) programme (DNAStar Inc.).

(b) Results

Sequence analysis of selected clones from clone sets a and b yielded results which contrasted with the conclusions drawn from their Southern blotting profiles presented in section 5.3. Data presented here, demonstrates that clones which strongly hybridise the MSD1 oligonucleotide on Southern blotting all contain PCR products in which exons 12 and 13 are either flush, or separated by the MSD1d exon.
Clone set a was generated from a random pool of "parental generation" PCR products from the NB1 neuroblastoma cell line. Sample 8 was found to contain a PCR product in which the exon 12/13 splice junction is intact - no intervening sequence is found at the nexus of exons 12 and 13 (Figure 5.4.1a). In contrast, samples 19 (Figure 5.4.1b), 27 (Figure 5.4.1c), 34 (Figure 5.4.1d), and 35 (Figure 5.4.1e) were all found to contain MSD1d at the exon 12/13 splice junction. It should be noted however, that none of these analyses spanned the entire PCR insert, from the 5' side of the plasmid insertion site to the 3' portion of the same. Sample 10 (sequence not displayed) was sequenced using both the Universal and Reverse primers; 189 bp of unidentifiable sequence was obtained.

Clone sets b1 and b2, generated by amplification of selected PCR products, were found to comprise a similar complement to clone set a. The former were generated from a 270 bp band excised from the NB1 and Kelly neuroblastoma cell lines: NB1 giving rise to clone set b1, and Kelly, to clone set b2.

Samples 8 (Figure 5.4.2a) and 13 (Figure 5.4.2b) of clone set b1 (see Figure 5.3.2) were found to contain PCR products with MSD1d. In contrast samples 24 (Figure 5.4.2c), 27 (Figure 5.4.2d) and 28 (Figure 5.4.2e) of clone set b2 (see Figure 5.3.3) were found to contain PCR products with no exon between exons 12 and 13. Samples 1 and 2 were analysed, to confirm that they did not contain any NCAM sequence. Identical and unidentifiable sequences of approximately 150 bases were obtained for each, using both the Universal and Reverse primers (data not displayed).

(c) Discussion

The discovery that clones which hybridise the MSD1 oligonucleotide do not appear to contain MSD1a, b or c is surprising. MSD1d was found to be present in 6 of the 10 clones analysed. Even if this triplet was sufficient
Figure 5.4.1 (a)

**Sequence obtained from sample 8, clone set b**
Digestion with Not-1 restriction enzyme, followed by Southern blotting (Figure 5.3.1), suggested that sample 8 has an insert of approximately 2.5 kb, containing at least part of the MSD1 domain.

(1) **Sequence**
Sample 8 was found to contain an insert of at least 186 bp, comprising NCAM-positive sequence which lacked any insert between exons 12 and 13. The exon border is indicated by a bold, vertical line.

**Key:** See Figure 4.5.2 (b) for key to all sequences presented in Figures 5.4.1 (a) - (e).

- - - = sequence obtained using 5' primer

(ii) **Representation of Sequence**
Sequence from sample 15 is represented in terms of its constituents.

**Key:**

□ = constitutively-spliced sequence;
exon is indicated below
Figure 5.4.1 (a)

(i)

GCC TCC GAG TTC AAG ACG CAG CCA GTC CSG GAA
CCC AGT GCA CCT AAG CTC GAA GGG CAG ATG GGA
GAG GAT GGA AAC TCT ATT AAA GTG AAG CTG ATC
AAG CAG GAT GAC GGC GGC TCC CCC ATC AGA CAC
TAT CTG GTC AGG TAC CGA GCX CTC TCC TCC GAG
TGG AAA CCA GAG ATC AGG CTC

(ii)

[Diagram with annotations indicating positions -12 and -13]
Figure 5.4.1 (b)

**Sequence obtained from sample 19, clone set a**
Digestion with Not-1 restriction enzyme, followed by Southern blotting (Figure 5.3.1), suggested that sample 19 has an insert of approximately 270 bp containing at least part of the MSD1 domain. Exon borders are indicated by bold vertical lines, the start of the 3' primer is framed, and MSD1d is in bold italics.

(i) **Sequence**
Sample 19 was found to contain an insert exceeding 185 bp, comprising NCAM-positive sequence and MSD1d.

**Key:**
- ■ ■ ■ = sequence obtained using 5' primer

(ii) **Representation of Sequence**
Sequence from sample 19 is represented in terms of its constituents.

**Key:**
- □ = constitutively-spliced sequence; exon is indicated below
- 3' = 3' primer
- ■ = MSD1d
Figure 5.4.1 (b)

(i)

 GCC TCC GAG TTC AAG ACG CAG CCA GTC C\textbf{AA} GGG
g
 GAA CCC AGT GCA CCT AAG CTC GAA GGG CAG ATG
g
 GGA GAG GAT GGA AAC TCT ATT AAA GTG AAC CTG
g
 ATC AAG CAG GAT GAC GGC XXX TCC CCC ATC AGA
g
 CAC TAT CTG GTC AGG TAC CGA XXX XTC TCC TCC
g
 GAG TGG AAA CCA GAG \textbf{ATC AG}

(ii)

\[\text{3'}\]
\[\text{32}\]
\[\text{13}\]
Figure 5.4.1 (c)

**Sequence obtained from sample 27, clone set a**

Digestion with Not-1 restriction enzyme, followed by Southern blotting (Figure 5.3.1), suggested that sample 19 has an insert of approximately 270 bp containing at least part of the MSD1 domain. Exon borders are indicated by bold vertical lines, and MSD1d is in bold italics.

(i) **Sequence**

Sample 27 was found to contain an insert exceeding X bp, comprising NCAM-positive sequence and MSD1d.

**Key:**

- = sequence obtained using 5' primer

(ii) **Representation of Sequence**

Sequence from sample 27 is represented in terms of its constituents.

**Key:**

= constitutively-spliced sequence; exon is indicated below

= MSD1d
Figure 5.4.1 (c)

(i)

CC TCC GAG XXC AAG AGC CAG CCA GTC CAA GGG
GAA CCC AGT GCA CCT AAG CTC GAA GGG CAG ATG
GGA GAG GAT GGA AAC TCT ATT AAA GTG AAC CTG
ATC AAG CAG GAT GA

(ii)

12 \[ \text{|d|} \] 13
Figure 5.4.1 (d)

**Sequence obtained from sample 34, clone set a**
Digestion with Not-1 restriction enzyme, followed by Southern blotting (Figure 5.3.1), suggested that sample 34 has an insert of approximately 270 bp containing at least part of the MSD1 domain. Exon borders are indicated by bold vertical lines, and MSD1d is in bold italics.

(i) **Sequence**
Sample 34 was found to contain an insert exceeding 135 bp, comprising NCAM-positive sequence and MSD1d.

**Key:**

- ■ ■ ■ = sequence obtained using 5' primer

(ii) **Representation of Sequence**
Sequence from sample 34 is represented in terms of its constituents.

**Key:**

- □ = constitutively-spliced sequence; exon is indicated below

- † = MSD1d
Figure 5.4.1 (d)

(i)

GCC TCC GAG TTC AAG ACG CAG CCA GTC |AA| GSG
GAA CCC AGT GCA CCT AAG CTC GAA GGG CAG ATG
GGA GAG GAT GGA AAC TCT ATT AAA GTG AAC CTX
ATC AXG CAG GAT GAC GGX XXX XXC CCC ATC AGA
CAC

(ii)

\[ -12 \quad \text{|d|} \quad -13 \]
Sequence obtained from sample 35, clone set a
Digestion with Not-1 restriction enzyme, followed by Southern blotting (Figure 5.3.1), suggested that sample 35 has an insert of approximately 270 bp containing at least part of the MSD1 domain. Exon borders are indicated by bold vertical lines, and MSD1d is in bold italics.

(i) Sequence
Sample 35 was found to contain an insert exceeding 147 bp, comprising NCAM-positive sequence and MSD1d.

Key:

### = sequence obtained using 5' primer

(ii) Representation of Sequence
Sequence from sample 35 is represented in terms of its constituents.

Key:

□ = constitutively-spliced sequence;
   exon is indicated below

□□ = MSD1d
Figure 5.4.1 (e)

(i)

GCC TCC GAG TTC AAG ACG CAG CCA GTC CAA GGG
GAA CCC AGT GCA CCT AAG CTC GAA GGG CAG ATX
XXX XAG GAT GGA AAC TCT ATT AAA GTG AAC CTG
ATC AAG CAG GAT GAC GGX XXX XCC CCC ATC AGA
CAC TAT CTG GTC AGG

(ii)

\[ -12 \ \ |d| \ \ -13 \]
**Figure 5.4.2 (a)**

**Sequence obtained from sample 8, clone set b1**

Digestion with Not-1 restriction enzyme, followed by Southern blotting (Figure 5.3.2), suggested that sample 8 has an insert of approximately 1.35 kb, containing at least part of the MSD1 domain.

(i) **Sequence**

Sample 8 was found to contain an insert of at least 122 bp, comprising NCAM-positive sequence which contained MSD1d between exons 12 and 13. Plasmid sequence is in bold type, the 5' primer is framed, the exon border is indicated by a bold, vertical line, and MSD1d is in bold italics.

**Key:** See Figure 4.5.2 (b) for key to all sequences presented in Figures 5.4.2 (a) - (e).

---

(ii) **Representation of Sequence**

Sequence from sample 8 is represented in terms of its constituents.

**Key:**

- □ = constitutively-spliced sequence; exon is indicated below
- 5' = 5' primer
- ■ = MSD1d
- □ = plasmid sequence on the 3' side of the insertion site is indicated above
Figure 5.4.2 (a)

(i)

\[
\text{GGTGAGAACGTT } \text{TC GCC GTA AGG CTG GCG GCG}
\]
\[
\text{CTC AAT GGC AAA GGG CTG GGT GAG ATC AGC GCG}
\]
\[
\text{GCC TCC GAG TTC AAG ACG CAG CCA GTC cjaA}
\]
\[
\text{GAA CCC AGT GCA CCT AAG CTC GAA GGG CAG ATG}
\]
\[
\text{GGA}
\]

(ii)

3' \[\begin{array}{c}
\text{5'}
\end{array}\] 12 13
**Figure 5.4.2 (b)**

**Sequence obtained from sample 13, clone set b₁**

Digestion with Not-1 restriction enzyme, followed by Southern blotting (Figure 5.3.2), suggested that sample 13 has an insert of approximately 270 bp, containing at least part of the MSD1 domain.

(1) **Sequence**

Sample 13 was found to contain an insert of 249 bp, comprising NCAM-positive sequence which contained MSD1d between exons 12 and 13. Plasmid sequence is in bold type, the 3' and truncated 5' primer are framed, the exon border is indicated by a bold, vertical line, and MSD1d is in bold italics.

**Key:**

- — — = sequence obtained using universal primer

- • • • = sequence obtained using reverse primer

(ii) **Representation of Sequence**

Sequence from sample 13 is represented in terms of its constituents.

**Key:**

- = constitutively-spliced sequence; exon is indicated below

- 5' = truncated 5' primer

- 3' = 3' primer

- = MSD1d

- = plasmid sequence; that found on the 3' and 5' sides of the insertion site are as indicated above
Figure 5.4.2 (b)

(i)

\[
\begin{align*}
\text{GGTGAGAACCCTTT T} & \text{AGG CTG GCG GCG CT} \ldots \\
\text{AAT GGC AAA GGG CTG GGT GAG ATC AGC GCXGCC} & \\
\text{TCC GAG TTC AAG ACG CAG CCA GTC} & \text{GAA} \\
\text{CCC AGT GCA CCT AAG CTC GAA GGG CAG ATG GGA} & \\
\text{GAG GAT GGA AAC TCT ATT AAA GTG AAC CTX XXX} & \\
\text{TAT CTG GTC AGG TAC CGA} & \text{GCG CTC TCC TCC GAG} \\
\text{TGG AAA CCA GAG ATC AGG CTC CCG TCT GGC AGT} & \\
\text{GA} & \text{ACCCCTTCTCACCTCGG} \\
\end{align*}
\]

(ii)

![Diagram showing the sequence and its components with distances labeled as 12 and 13]
Figure 5.4.2 (c)

Sequence obtained from sample 24, clone set b2
Digestion with Not-1 restriction enzyme, followed by Southern blotting (Figure 5.3.3), suggested that sample 24 has an insert of approximately 250 bp, containing at least part of the MSD1 domain.

(i) Sequence
Sample 24 was found to contain an insert of 249 bp, comprising NCAM-positive sequence which contained no insert between exons 12 and 13. Plasmid sequence is in bold type, the 3' and truncated 5' primers are framed, and the exon border is indicated by a bold, vertical line.

Key:

• • • = sequence obtained using reverse primer

—— = sequence obtained using primer 1

(ii) Representation of Sequence
Sequence from sample 24 is represented in terms of its constituents.

Key:

□ = constitutively-spliced sequence; exon is indicated below

5' = truncated 5' primer

3' = 3' primer

■■ = plasmid sequence; that found on the 3' and 5' sides of the insertion site are as indicated above
Figure 5.4.2 (c)

(i)

\[ \text{GGTGAGAACGTT} \]

\[ ^{5'} \]

\[ \text{GTA AGG CTG GCC GCG CTC} \]

\[ \text{AAT GGC AAA GGG CTG GGT GAG ATC AGC GCCGCC} \]

\[ \text{TCC GAG TTC AAG ACG CAG CCA GTC GGG GAA CCC} \]

\[ \text{AGT GCA CCT AAG CTC GAA GAA CAG ATG GGA GAG} \]

\[ \text{GAT GGA AAC TCT ATT AAA GTG AAC CTG ATC AAG} \]

\[ \text{CAG GAT GAC GCC GCC TCC CCC ATC AGA CAC TAT} \]

\[ \text{CTG GTC AGG TAC CGA GCG CTC TCC TCC GAG TGG} \]

\[ \text{AAA CCA GAG ATC AGG CTC CCG TCT GGC AGT GA} \]

\[ \text{ACCCCTTCTCACCTCGG} \]

(ii)

\[ \text{3'} \]

\[ \text{5'} \]

\[ \text{12} \]

\[ \text{13} \]
Figure 5.4.2 (d)

**Sequence obtained from sample 27, clone set b<sub>2</sub>**

Digestion with Not-1 restriction enzyme, followed by Southern blotting (Figure 5.3.3), suggested that sample 27 has an insert of approximately 300 bp, containing at least part of the MSDI domain.

(i) **Sequence**

Sample 27 was found to contain an insert of approximately 305 bp, comprising NCAM-positive sequence which contains no insert between exons 12 and 13. Plasmid sequence is in bold type, the 3' primers are framed, and the exon border is indicated by a bold, vertical line.

**Key:**

- • • • = sequence obtained using reverse primer
- ___ = sequence obtained using primer 1

(ii) **Representation of Sequence**

Sequence from sample 27 is represented in terms of its constituents.

**Key:**

- □ = constitutively-spliced sequence; exon is indicated below
- 3' = 3' primer
- □□ = plasmid sequence; that found on the 3' and 5' sides of the insertion site are as indicated above
- □□□ = unidentified sequence
Figure 5.4.2 (d)

(i)  

5'  
A AGG CTG GCX  

GXG CTC AAT CXX XTT GGG CTG GGT GAG ATC AGC  
GCG GCX TCC GAG XXC AAG AGC CAG CCA GTC GGG  
GAA CCC AGT GCA CCT AAG CTC GAA GGG CAG ATG  
GGG GAG GAT GGA AAC TCT ATT AAA GTG AAC CTG  
A . . . . . . . GGC TCC CCC ATC AGA CAG TAT  
CTG GTC AGG TAC CGA GCG CTC TCC TCC GAG TGG  
3'  
AAA CCA GAG ATC AGG CTC CCG TCT GCC AGT GGC  
CGTTAACGGCTGGCCTGCGCTGAGAAAGACGGCTGAACTTGGACTATC  
TAGAGGAAGTAAAGTGTAA C AGG CTC CCG TCT GCC  
AGT GA AACCCTTCTCACC  

(ii)  

5'  

3'  

12  3'  

13  3'  

13
Figure 5.4.2 (e)

**Sequence obtained from sample 28, clone set b2**
Digestion with Not-1 restriction enzyme, followed by Southern blotting (Figure 5.3.3), suggested that sample 28 has an insert of approximately 1.1 kb, containing at least part of the MSD1 domain.

(i) **Sequence**
Sample 28 was found to contain an insert of at least 312 bp, comprising NCAM-positive sequence which contains no insert between exons 12 and 13. Plasmid sequence is in bold type, the primers are framed, and the exon border is indicated by a bold, vertical line.

**Key:**

• • • = sequence obtained using reverse primer

_____ = sequence obtained using primer 1

(ii) **Representation of Sequence**
Sequence from sample 28 is represented in terms of its constituents.

**Key:**

□ = constitutively-spliced sequence; exon is indicated below

5' = truncated 5' primer

3' = 3' primer

= plasmid sequence; that found on the 3' and 5' sides of the insertion site are as indicated above

= unidentified sequence
Figure 5.4.2 (e)

(i)

GGTTGGAGGTGGTCCTCTCTCTGACTGTGCTAGTCTGTATTCCAGG
CTCCGTCCTGGCAGTGAAC
AAT GXX AAA GGG CTG GGT GAG ATC AGC GCGGCC
TCC GAG TCC AAG AGC GAG CCA CTC TGG GAA CCC
AGT XXA CCT AAG CTC GAA GGG CAG ATG GCA GAG
GAT GGA AAG TCT ATT AAA GTG AAG CTG ACG AAG
CAG GAT GAC GXX GGC TCC CCC ATG AGA CAC TAT
CTG GTC AGG TAC CGA GCG CTC TCC TCC GAG TGG
AAA CCA GAG ATC AGG CTC CCG TCT GGC AGT GA
A _AACGGTTCTCA

(ii)

\[ \begin{array}{c}
\text{12} \\
\text{13}
\end{array} \]
to hybridise the MSD1 oligonucleotide, the latter hybridises with species in which no alternatively spliced exons are found at the exon 12/13 splice junction.

As most of the cloned inserts have not been sequenced from end to end, it might be argued that species which contain MSD1a, b or c could be found inserted alongside those species which have been detected. Sequence analysis of PCR product-containing clones in the quest for VASE (Chapter 4) has suggested that oligonucleotides probably undergo fusion during the PCR reaction. On this tenet, it is possible that two or more NCAM-positive PCR products might fuse to form a concatemer. However, the theoretical size of such a product would be inconsistent with the calculated molecular weights of most of these products. For instance, sample 19 has an insert of approximately 270 bp (see Figure 5.3.1). Approximately 185 bases of its sequence have been obtained through sequence analysis (Figure 5.4.1b). It seems unlikely that a PCR product containing MSD1a, b and/or c, of approximately 85 bp might be found in tandem with the sequence already characterised. In addition, two samples which hybridise the MSD1 oligonucleotide have been sequenced from one side of the plasmid insertion site to the other. Sample 13 of clone set b1 contains MSD1d and sample 24 of clone set b2 has no intervening sequence at the exon 12/13 splice junction.

It thus would appear that the heavy bands arising on Southern blot analysis of specific samples are a consequence of "non-specific" binding. A considerable degree of what is believed to be "non-specific" binding is also observed in samples other than those which hybridise strongly with the MSD1 and MSD1c oligonucleotide probes. For example, three samples (sample 10 of clone set a and samples 1 and 2 of clone set b2) which yielded faint signal on Southern blotting with the MSD1 oligonucleotide were found to contain unidentifiable sequence (sequence data not
displayed). However, this does not rule out the possibility that NCAM-
positive samples may also constitute a proportion of the samples assumed
to be "negative" for MSD1.

It may be that the MSD1 oligonucleotide has a hybridisation affinity
for the pCR™1000 vector. Indeed, faint bands could be discerned in the
lanes containing digested episomal DNA extracted from "negative" blue
colonies in Figures 5.3.2 and 5.3.3 (control lane is partly obscured by
artifact in Figure 5.5.3), although this is not true in the case of clone set a
(Figure 5.3.1). Alignment of the MSD1 sequence and the pCR1000 vector
using the DNA Walter-Lipman programme (DNASTar; k-tuple = 3; gap penalty
= 3) reveals a match of 10/11 consecutive nucleotides with a 50% GC
composition. This may be sufficient to incur "non-specific" binding. That
this region lies within the MSD1c exon, and the "positive" clones hybridised
most strongly with the MSD1c oligonucleotide probe corroborates the
hypothesis that non-specific hybridisation between oligonucleotide and
vector contributes to the appearance of strong bands in Figures 5.3.1 - 5.3.3.

The juxtaposition of pCR™1000 vector and the "exon 12/13" PCR
product may create a novel sequence to which the MSD1 and MSD1c
oligonucleotides are sufficiently complementary to generate signal on
Southern blotting. However, this seems unlikely, as sequencing data
presented in Figures 5.4.2 (c) and 5.4.2 (e) reveals that the "exon 12/13"
PCR product is inserted in a different orientation in each of these clones.

As discussed in earlier in this chapter, "non-specific" hybridisation is
influenced by a variety of factors. In this case, it is difficult to speculate
further without additional sequencing data. Assessing the exact DNA
concentration of each sample might also shed light on the mutually
contradictory nature of the blotting and sequencing data presented here;
samples of greater concentration would be more likely to give rise to non-
specific binding.
The empirical establishment of hybridisation conditions which provide maximum stringency, yet permit specific annealing of the MSD1 oligonucleotide might suppress the "non-specific" binding observed to take place in Figures 5.3.1 - 5.3.3. These results raise a question mark over the interpretation of those presented in Figures 5.2.2, 5.2.3 and 5.2.5. It could be argued that the hybridisation of the MSD1a, MSD1b and MSD1 oligonucleotide probes to "parental generation" PCR products, might be "non-specific", and that it is premature to claim that MSD1b is spliced into the RNA of neuroblastoma cell lines.

However, comparison of the blotting profiles of Figures 5.2.2 - 5.2.5, and 5.3.4 reveal two different kinds of pattern. Of all the oligonucleotide probes, MSD1c yields the strongest signal in Figure 5.3.4c, while it appears to hybirdise no product arising directly from neuroblastoma cell lines (Figure 5.2.4). Similarly, the MSD1a oligonucleotide yields faint bands on hybridisation with DNA from clones which have been demonstrated to include exons 12 and 13 (Figure 5.3.4 a), either with or without MSD1d therein. In contrast, it does not hybridise the 250 bp band arising directly from PCR amplification of neuroblastoma cDNA, and believed to comprise the same sequence.

The hybridisation of MSD1 oligonucleotides with pCR™1000 clones is positively influenced by a factor which is absent in identical hybridisations with "parental generation" PCR products. The identity of this factor - or combination of factors - remains to be elucidated, although possibilities have been discussed. While "non-specific" hybridisation by the cloned PCR products recommends a degree of caution in the interpretation of previous data obtained by Southern blotting, it does not negate it.

In summary, random and selected PCR products - the latter selected on the basis of size (270 bp) - gave rise to clones, some of which contained inserts with NCAM-positive sequence. Furthermore, these sequences
corresponded to PCR products predicted from the size and blotting profiles of PCR products of 250 bp detected in Figures 5.2.2 - 5.2.6. This may be partly due to the fact that agarose gel electrophoresis (in this instance, agarose concentrations of 2% were used) is an imprecise means by which to separate oligonucleotides of similar size.

The finding that 6/7 clones derived from the neuroblastoma NB1 cell line contained MSD1d, in contrast with 3/3 clones derived from the neuroblastoma Kelly cell line, is intriguing, considering that alternative-splicing of MSD1d is thought to be phenotypically and developmentally-regulated (Hamshere et al., 1991; Reyes et al., 1991; Reyes et al., 1993). Further experiments involving finer electrophoretic resolution and the quantification of PCR products could verify whether there is, in fact a difference in the levels of alternatively-spliced product between these two cell lines. If regulated splicing of MSD1d in neurons parallels that observed in glial cells, a higher level of MSD1d inclusion would indeed be expected in the RNA of the NB1 neuroblastoma cell line; the NCAM isoform profile of this cell line accords with that of a more differentiated cell type (see Chapter 3).

5.5 Conclusion

Southern blotting of PCR products demonstrates a reproducible difference in the alternative splicing patterns of NCAM by the neuroblastoma and rhabdomyosarcoma cell lines examined here. Southern blotting and sequencing data suggest that neuroblastoma cell lines generate RNA which can contain MSD1a, MSD1b or MSD1d at the exon 12/13 splice junction. In contrast, rhabdomyosarcoma cell lines generate RNA, some of which contains MSD1a, MSD1b, MSD1c and MSD1d in tandem. Sequence analysis of a clone containing a PCR product derived from the 360 bp PCR product of the JR-1 rhabdomyosarcoma cell line revealed the presence of the entire MSD1 domain (R. Rossel; personal communication).
Evidence suggests that different splicing patterns exist within the group of neuroblastoma cell lines tested. IMR-32, NB1 and GOTO cell lines appeared to contain RNA species containing MSD1a at the exon 12/13 splice junction while the other neuroblastoma cell lines did not. The certainty of these findings is subject to the limits of detection and specificity afforded by Southern blot analysis; sequence analysis could confirm their veracity.

Whether or not MSD1d is spliced alongside the presumed MSD1a or b exons in the neuroblastoma cell lines remains to be elucidated. Similarly, it is not possible to ascertain whether the MSD1a, b and c mini-exons can be spliced together in the absence of MSD1d, in NCAM synthesised by the JR-1 cell line. Resolution by agarose gel electrophoresis was insufficient for the identification of such products, and lack of time prevented their isolation and sequence analysis.

The alternative splicing patterns at the exon 12/13 junction in the neuroblastoma and rhabdomyosarcoma cell lines examined here, suggest that they may prove useful in distinguishing these cell types. These findings are in agreement with a recent publication which also concludes that "cell phenotype is not the dominant factor in regulating NCAM microsplicing of VASE but may be a more significant factor at the exon 12/13 junction" (Reyes et al., 1993). The results presented here suggest that rhabdomyosarcoma cell lines contain a splice variant of NCAM which contains MSD1c at the exon 12/13 splice junction, while neuroblastoma cell lines do not.
Chapter 6

SUMMARY AND SUGGESTIONS FOR FURTHER STUDIES

6.1 Summary

The principal findings of this study can be summarised as follows:

1) The association of NCAM expression with rhabdomyosarcoma and advanced disease stage (Evans Stage 4) neuroblastoma has been established. That NCAM is expressed by rhabdomyosarcoma confirms previous findings that its expression is not restricted to tissues of neuroectodermal origin.

2) All neoplastic cell lines investigated yielded a polydisperse smear on Western blotting, suggesting extensive polysialylation. On desialylation, NCAM-180 was readily detected in some neuroblastoma cell lines and one tissue sample. However, the pattern of major NCAM isoforms in most neuroblastoma and rhabdomyosarcoma cell lines and tissues does not differ.

3) NCAM expression has been demonstrated in a myeloid leukaemia cell line, a tumour type in which NCAM has recently been discovered (Lanier et al., 1991; Barker et al., 1992).

4) Alternative splicing of VASE was demonstrated to be present in all neuroblastoma and rhabdomyosarcoma cell lines analysed. The finding that VASE is present in cell lines of rhabdoid origin is consistent with a contemporaneous study (Small and Akeson, 1990) which suggests that VASE is not restricted to neuronal tissues.

5) Southern blotting of PCR products suggests that all neuroblastoma cell lines analysed contain RNA species in which the MSD1b exon is present. The inclusion of MSD1b was previously believed to be restricted to mRNA of rhabdoid tissues. These data also suggest that MSD1α exon is present in some neuroblastoma cell lines but not others.

6) The pattern of alternative splicing of NCAM RNA transcripts appears to differ between the neuroblastoma and rhabdomyosarcoma cell lines analysed. Southern blotting of PCR products indicates that the entire MSD1 domain is present in some NCAM RNA species of the rhabdomyosarcoma cell line JR-1. Seven additional rhabdomyosarcoma cell lines yielded identical blotting profiles (A. Hancock, personal communication). Sequence analysis of a clone containing a PCR product derived from
rhabdomyosarcoma has demonstrated that the entire MSD1 sequence is spliced into some RNA species of this cell line (R. Rossell, personal communication). MSD1c has not been detected in any of the neuroblastoma cell lines investigated here.

7) A sensitive technique has been developed and used to sequence PCR products as a means of verifying the Southern blotting data. However, the successful application of this technique may depend on adequate separation of PCR products of differing size. An alternative method employing the TA cloning kit™ obviates this technical obstacle, but the generation of a large number of clones containing "irrelevant" sequence impeded the acquisition of "NCAM-positive" clones.

6.2 Suggestions for Further Studies

The presence of MSD1c in rhabdomyosarcoma cell lines but not in neuroblastoma implies that the tissues from which they are derived might be distinguished by exploiting this difference. A monoclonal antibody has recently been generated, by immunising mice with a peptide corresponding in sequence to MSD1c, conjugated to Keyhole Limpet Haemocyanin. The antibody binds rhabdomyosarcoma as revealed by FACs analysis, and binding is inhibited by prior incubation with the immunising peptide (J. Hancock, personal communication). The specificity and sensitivity of this antibody and any others subsequently raised will be investigated by binding assays with transfectants, cell lines and tumour biopsies. It is hoped that an antibody will be obtained which will reliably and consistently differentiate rhabdomyosarcoma from neuroblastoma and perhaps other small round tumours.

As is seen in Figure 6.1, MSD1c of human NCAM differs from its mouse counterpart by five amino acids in close proximity: one proline, one glutamic acid, and three alanine residues. These amino acids probably create a sufficiently different tertiary structure so as to provoke an immune reaction in the mouse. MSD1b or MSD1c is probably O-glycosylated; experiments demonstrate that the MSD1 sequence is O-glycosylated at some point (Walsh et al., 1989) and both MSD1b and MSD1c have several serine and threonine residues - O-glycosylation occurs via their hydroxyl groups. Although the O-linked carbohydrate on the native protein could prevent antibody raised against a synthetic peptide from binding in an analogous fashion to that observed by Gendler et al., 1991, this appears not to have been the case, although →
Figure 6.1

Mouse and human MSD1 sequences compared: cDNA and corresponding amino acids
Data is that obtained by Hamshere et al., 1991; Moore et al., 1987; and R. Rossell - ICRF, Bristol, 1992.

Key:

M = sequence derived from mouse cDNA
H = sequence derived from human cDNA
### MSD1a

<table>
<thead>
<tr>
<th>M</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>His Ser Pro Pro Pro Gln Ala Pro Ala Asn Ser Ser</td>
<td>CAT AGC CCT CCT CCA CAG GCT CCT GCT AAC TCT TCC</td>
</tr>
<tr>
<td></td>
<td>C A T G G T Pro Ser Ser</td>
</tr>
</tbody>
</table>

### MSD1b

<table>
<thead>
<tr>
<th>M</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thr His Val Pro Leu Ser Pro Arg Ala Thr Thr Trp</td>
<td>ACC CAT GTT CCG TTG TCT CCT CGA GCT ACA ACC TGG</td>
</tr>
<tr>
<td></td>
<td>C A A C A T Pro Pro Asp</td>
</tr>
</tbody>
</table>

### MSD1c

<table>
<thead>
<tr>
<th>M</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pro Leu Pro Val Leu Pro Thr Thr Asp Leu Ser Lys</td>
<td>CCT CTT CCT GTC CTT CCA ACC ACA GAC CTG TCT AAA</td>
</tr>
<tr>
<td></td>
<td>C G A CA G Ala Ala Glu Pro Ala</td>
</tr>
</tbody>
</table>
comprehensive binding profiles of the MoAb have yet to be obtained. Extended conformation has been associated with O-glycosylated regions (Komatsu et al., 1970; Remold-O'Donnell et al., 1984; McCall et al., 1992). However, the acquisition of an MSD1c-specific antibody suggests that the MSD1c peptide is sufficiently immunogenic to elicit antibody.

Attempted perturbation of functional processes with an anti-MSD1c antibody might suggest a role for the polypeptide. Such an antibody could also be used to investigate the alternative splicing of MSD1c in other tissues by immunocytochemistry and immunohistochemistry.

Immunoprecipitation and immunoblotting experiments using an anti-MSD1c antibody could determine whether or not the exon is exclusively associated with GPI-linked isoforms, as suggested previously (Walsh et al., 1989). Alternatively, Northern blots using an oligonucleotide probe complementary to the MSD1 domain or its exon components could reveal associated alternative splicing patterns. An advantage to this approach would be its flexibility in application - the "isoform profile" of MSD1b in both rhabdomyosarcoma and neuroblastoma could be similarly established - without having to generate MSD1b-specific antibodies.

Another possible diagnostic strategy would be to subject RNA from biopsies of tumour tissue to RT-PCR, using a primer complementary to (constitutively-spliced) exon 13 to prime the RT reaction. Generation of PCR products using a 5' primer derived from MSD1c sequence, in combination with a 3' primer complementary to the 5' end of exon 13 would indicate the presence of MSD1c, and hence suggest a case of rhabdomyosarcoma. Alternatively, one could use a 5' primer identical to a region in exon 12, and a 3' primer complementary to a region in MSD1c. If these two reactions are of similar efficiency, they could be conducted simultaneously in the same mixture. If the primers are designed so that the resulting products are of discernably different sizes, the specificity of the reaction is enhanced.

Retrospective screening of all known anti-NCAM antibodies using transfectants containing different combinations of the MSD1 exons may demonstrate changes in molecular conformation with different splicing patterns.

The fact that some neuroblastoma cell lines appear to harbour RNA containing MSD1a while others do not, is intriguing. A thorough characterisation of these cell lines may reveal an association between the splicing of MSD1a, and cell phenotype. Neuroblastoma cell lines assume a range of phenotypes, described as "neuroblastic", "substrate adherent" or "intermediate" (Ciccarone et al., 1989). It is possible that the presence of
MSD1a may correlate with one of these states, in a similar manner to that of the integrin VLA-5 (Yoshihara et al., 1992). Analysis of RNA from a native cell line, and that extracted from the same cell line differentiated with an agent such as all-trans retinoic acid might also prove informative in this regard. Splicing patterns of tumour biopsies consisting of neuroblastoma tissue from different stages of the disease might also reveal a trend in splicing pattern. More generally, the extent of total NCAM expression, or the degree to which the molecule is sialylated may be altered according to stage. The fact that neuroblastoma varies dramatically in its pathology - from spontaneous regression to aggressive metastasis - invites comparative studies of tissues from patients with different grades of disease.

It is possible to generate mutated transcripts of pre-mRNA in which alternatively-spliced exons are spliced in an aberrant fashion (Tacke and Goridis, 1991). Morphogenesis of transgenic mice in which the endogenous NCAM gene is replaced by mutant NCAM DNA which prohibits the generation of specific alternatively-spliced isoforms, might suggest a functional niche for such isoforms.

A number of studies could be performed to confirm the work described here. Verification of the Western blotting data could be obtained by PCR experiments to investigate the alternative splicing which gives rise to the different major isoforms - i.e. the splicing of exons 15 and 18. Determination of experimental conditions conferring maximum stringency and yet permitting specific hybridisation of the various oligonucleotide probes, would improve the certainty of conclusions based on Southern blotting data.

PCR analysis of more neuroblastoma and rhabdomyosarcoma cell lines is desirable as a means to test further the conclusion that MSD1c is spliced into the mRNA of rhabdomyosarcoma only. All the neuroblastoma cell lines analysed here are representative of their tissue of origin in that typical ultrastructural features and neurotransmitter activity were described shortly after each cell line was established (see Chapter 2, section 1 for references). Similarly, the rhabdomyosarcoma cell lines were also deemed to be representative of their seminal tissue - again on the basis of ultrastructural features (e.g. Z bands) and by their expression of desmin as well as other muscle-specific proteins. As these cell lines were easily identified by the characteristics described, they are not representative of extremely primitive tumours, which by definition, are difficult to categorise, and with which this work is concerned. It is important that the tumour-specific splicing pattern suggested by PCR analyses presented here,
extrapolates to fresh tumour tissue and undifferentiated cell lines. Further studies would include such analyses. Ideally, these should have been carried out completed prior to attempts at raising a monoclonal antibody which recognises MSD1c.

PCR analysis of tissue samples, however, is challenging. The reaction is extremely sensitive, and obtaining sterile, pure tumour tissue, free of blood and other contaminants is very difficult if not impossible. Two additional PCR reactions can be run alongside that which amplifies across the exon 12/13 splice junction in tumour tissue to monitor contamination, such as blood which may contain leucocytes expressing NCAM. One reaction utilises mRNA derived from what is believed to be pure tumour tissue, and primers which have been previously demonstrated to specifically amplify a product in the suspected contaminant - e.g. CD45 in the case of leucocytes. Another utilises mRNA derived from the suspected contaminant, and the primers used in the "main" reaction - i.e. those hybridising to regions of exons 12 and 13 in NCAM cDNA. By comparing and contrasting PCR products (and appropriate positive controls) it should be possible to ascertain whether any of the suspected contaminant is indeed present. A limitation of this strategy is that it does not screen for unsuspected contaminants.

The previous discussion highlights one of the main advantages of using cell lines in this study. Other advantages include availability (neuroblastoma tissue is not easy to obtain) and the fact that each neuroblastoma cell line has been characterised and judged to be representative of the tumour. The disadvantage of using many different cell lines is that over time, cross-contamination is possible. In an ideal world, cell lines would be routinely subjected to karyotyping and/or analysis using restriction fragment length polymorphisms by way of monitoring. While cell lines used in this study were initially revived from seed stocks, and were immunophenotyped shortly after resuscitation, they were not profiled at regular intervals throughout the study. An improved protocol would include this measure as a means of monitoring possible contamination and/or phenotypic drift - events which probably occur more frequently than are detected (see McKeever et al., 1991).

The successful sequencing of PCR products containing MSD1a and MSD1b from neuroblastoma would corroborate the Southern blotting data. One clone has been derived from rhabdomyosarcoma which contains a PCR product in which the entire MSD1 sequence is present between exons 12 and 13 (R. Rossell, personal communication). While it is normally good
practise to obtain at least two cloned PCR products which yield the same sequence (as a safeguard against sequence errors by Taq polymerase) the fact that the sequence obtained corresponds with one previously defined in a cDNA clone derived from human muscle (Moore et al., 1987) renders this a task of low priority.

Finally, a more broadly-based strategy could be applied to identify genes/gene products which differentiate between neuroblastoma and rhabdomyosarcoma. Subtracted cDNA libraries (e.g. a cDNA library derived from a primitive RMS cell line, "subtracted" with the mRNA of a neuroblastoma cell line) could yield candidate genes. If previously unidentified, these might prove interesting in their own right, in addition to potential markers for tumour differentiation.

6.3 Conclusion

NCAM expression and alternative splicing has been investigated using a variety of different molecular and immunological techniques. PCR amplification and Southern blot analysis are appropriate techniques by which to investigate alternative splicing of VASE and the MSD1 domain, as antibodies specific to these domains have yet to be identified. While NCAM expression is not confined to tissues of neuroectodermal origin, its expression may nonetheless prove instrumental in the differential diagnosis of neuroblastoma from rhabdomyosarcoma, by virtue of its propensity for tissue-specific alternative splicing patterns. It may also provide important insights into the fundamental mechanisms involved in the processes of cellular differentiation and neoplastic progression.
REFERENCES

References by more than ten authors have been abbreviated after the tenth author.
(For additional references, see Addendum)


Schwab, M., Alitalo, K., & Klempnauer, K. K. (1983). Amplified DNA with limited homology to myc cellular oncogene is shared by human neuroblastoma cell lines and neuroblastoma tumour. 305: 245-248.


ADDENDUM

Additional References


PUBLICATIONS ARISING FROM THE THESIS


Phimister, E., Rossell, R. J., Culverwell, A., Kemshead, J. T., Patel, K. The MSD-1c mini-exon is present in neural cell adhesion molecules from rhabdomyosarcoma but not neuroblastoma (In preparation)