INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the original text directly from the copy submitted. Thus, some dissertation copies are in typewriter face, while others may be from a computer printer.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyrighted material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is available as one exposure on a standard 35 mm slide or as a 17" × 23" black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. 35 mm slides or 6" × 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Molecular analysis of human myosin alkali light chain genes

Fodor, William Lester, Ph.D.
The Ohio State University, 1988

Copyright ©1989 by Fodor, William Lester. All rights reserved.
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark √.

1. Glossy photographs or pages ______
2. Colored illustrations, paper or print ______
3. Photographs with dark background ______
4. Illustrations are poor copy ______
5. Pages with black marks, not original copy √
6. Print shows through as there is text on both sides of page ______
7. Indistinct, broken or small print on several pages ______
8. Print exceeds margin requirements ______
9. Tightly bound copy with print lost in spine ______
10. Computer printout pages with indistinct print ______
11. Page(s) _______lacking when material received, and not available from school or author.
12. Page(s) _______seem to be missing in numbering only as text follows.
13. Two pages numbered ______. Text follows.
14. Curling and wrinkled pages ______
15. Dissertation contains pages with print at a slant, filmed as received √
16. Other ____________________________
   ___________________________________
   ___________________________________
MOLECULAR ANALYSIS OF HUMAN MYOSIN ALKALI LIGHT CHAIN GENES

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of the Ohio State University

By

William Lester Fodor, B.S.

The Ohio State University
1988

Dissertation Committee:
C.A. Breitenberger
P.A. Fuerst
G.A. Marzluf

Approved by
Advisor
Department of Molecular Genetics
To My Wife Elizabeth
ACKNOWLEDGMENTS

I express sincere appreciation to my advisor, Elio, for all his understanding in the past few months and for his insight and guidance throughout my research. The technical assistance of Doug Fletcher, Susan Morrow, and Jegatheesan Seharaseyon is greatly acknowledged for their help in purifying bacteriophage clones. Thanks to all the other members of the lab, Rene', Vidya, Piri, and Fadia for being excellent lab mates. To my wife Elizabeth and my son Ryan, I wish to express my gratitude for their patience and understanding, especially in the past year. I would also like to thank Mom and Dad Given and my parents for the emotional and the financial support when times got difficult.
VITA

DATE OF BIRTH: April 19, 1958

PLACE OF BIRTH: Bethesda, Md.

NATIONALITY: U.S.A.

ADDRESS: Department of Genetics
Room 746 Biological Sciences
484 W. 12th Avenue
Ohio State University
Columbus, Ohio 43210

Telephone: (614) 292-1139

EDUCATION:
1976-1977 Utah State University, Logan, Utah
1977-1981 Ohio State University, Columbus, Ohio
   Bachelor of Science Degree in Genetics
1982-present Ohio State University, Columbus, Ohio
   Genetics Graduate Doctoral Program

ACADEMIC APPOINTMENTS:
1982-1985 Graduate Teaching Assistant
1985-1986 Graduate Research Assistant
1985-1986 President Genetics Graduate Students

AWARDS:
1987 Graduate Fellow, Procter & Gamble Fellowship Award

TEACHING EXPERIENCE
1982-1985 Graduate Teaching Assistant
   Department of Genetics
   Ohio State University, Columbus, Ohio
   Assisted in teaching undergraduate and graduate level
courses in biology and genetics.
1985 Wi.Qtr. Genetics 140: Introductory Genetics for Non-Majors
   Department of Genetics
   Ohio State University, Columbus, Ohio
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEDICATION</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>VITA</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
<tr>
<td>ABBREVIATIONS</td>
<td>xv</td>
</tr>
</tbody>
</table>

## CHAPTER

**I. INTRODUCTION**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>B. Goals, Hypothesis and Rationale</td>
<td>1</td>
</tr>
<tr>
<td>C. Research Summary</td>
<td>5</td>
</tr>
</tbody>
</table>

**II. MYOSIN ALKALI LIGHT CHAIN LITERATURE REVIEW**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Introduction</td>
<td>6</td>
</tr>
<tr>
<td>B. Protein Function: The Myosin Hexamer</td>
<td>6</td>
</tr>
<tr>
<td>C. MLC-ALK Protein Characteristics</td>
<td>10</td>
</tr>
<tr>
<td>D. MLC-ALK Isoforms and Their Genes</td>
<td>15</td>
</tr>
<tr>
<td>1. Fast Twitch Skeletal Muscle Isoforms</td>
<td>18</td>
</tr>
<tr>
<td>2. Slow Twitch Skeletal Muscle Isoforms</td>
<td>22</td>
</tr>
<tr>
<td>3. Cardiac Muscle Isoforms</td>
<td>23</td>
</tr>
<tr>
<td>4. Smooth Muscle and Nonmuscle Isoforms</td>
<td>25</td>
</tr>
<tr>
<td>5. Invertebrate Isoforms</td>
<td>26</td>
</tr>
<tr>
<td>E. MLC-ALK Expression</td>
<td>29</td>
</tr>
<tr>
<td>1. Development</td>
<td>29</td>
</tr>
<tr>
<td>2. Disease Related Effects</td>
<td>32</td>
</tr>
<tr>
<td>3. Manipulated Muscle: Innervation Studies</td>
<td>33</td>
</tr>
<tr>
<td>4. Manipulated Muscle: Tissue Culture</td>
<td>34</td>
</tr>
<tr>
<td>F. Chromosomal Distribution of MLC-ALK Genes</td>
<td>35</td>
</tr>
<tr>
<td>G. Literature Review Summary</td>
<td>38</td>
</tr>
</tbody>
</table>
### III. GENERAL EXPERIMENTAL PROCEDURES

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. DNA Preparation</td>
<td>39</td>
</tr>
<tr>
<td>B. Restriction Enzyme Analysis</td>
<td>39</td>
</tr>
<tr>
<td>C. Restriction Fragment Isolation</td>
<td>40</td>
</tr>
<tr>
<td>D. Plasmid Subcloning</td>
<td>40</td>
</tr>
<tr>
<td>E. Oligonucleotide Probes</td>
<td>41</td>
</tr>
<tr>
<td>F. Probe Synthesis</td>
<td>41</td>
</tr>
<tr>
<td>G. Genomic Southern Analysis</td>
<td>41</td>
</tr>
<tr>
<td>H. Genomic Library Screening</td>
<td>42</td>
</tr>
<tr>
<td>I. Exon Identification &amp; DNA Sequencing</td>
<td>42</td>
</tr>
</tbody>
</table>

### IV. HETEROLOGOUS GENOMIC SOUTHERN ANALYSIS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Results</td>
<td>44</td>
</tr>
</tbody>
</table>

### V. ISOLATION AND CHARACTERIZATION OF THE HUMAN SLOW TWITCH/VENTRICULAR MYOSIN ALKALI LIGHT CHAIN GENE, hMLC1V/Sb

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Results</td>
<td>49</td>
</tr>
<tr>
<td>1. Isolation and Preliminary Analysis</td>
<td>49</td>
</tr>
<tr>
<td>of hMLC.1</td>
<td>49</td>
</tr>
<tr>
<td>2. Genomic Southern Analysis</td>
<td>53</td>
</tr>
<tr>
<td>3. Isolation and Preliminary Analysis</td>
<td>59</td>
</tr>
<tr>
<td>of hMLC.2</td>
<td>59</td>
</tr>
<tr>
<td>4. Gene Structure Analysis of hMLC1V/Sb</td>
<td>61</td>
</tr>
<tr>
<td>a. Gene Orientation</td>
<td>64</td>
</tr>
<tr>
<td>b. Identification of Exons 3-6</td>
<td>64</td>
</tr>
<tr>
<td>c. Identification of Exons 1 and 2</td>
<td>77</td>
</tr>
<tr>
<td>d. Identification of Exon 7</td>
<td>94</td>
</tr>
<tr>
<td>5. Chromosomal Localization</td>
<td>101</td>
</tr>
<tr>
<td>B. Discussion</td>
<td>103</td>
</tr>
</tbody>
</table>

### VI. ISOLATION OF hMLC.3 and hMLC.4

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Results</td>
<td>107</td>
</tr>
</tbody>
</table>
VII. PRELIMINARY ANALYSIS OF hMLC.3 ................. 109
   A. Results .................................. 109
       1. Preliminary Restriction Enzyme
           Analysis.............................. 109
       2. Repetitive DNA Analysis............ 109
       3. Genomic Southern Analysis....... 113
       4. Exon Identification................ 113
   B. Discussion................................ 123

VIII. PRELIMINARY ANALYSIS OF hMLC.4 ................. 126
   A. Results .................................. 126
       1. Preliminary Restriction Enzyme
           Analysis.............................. 126
       2. Repetitive DNA Analysis............ 126
       3. Genomic Southern Analysis....... 130
       4. Localization of MLC-ALK Homology .. 135
       5. DNA Sequence Analysis............. 141
   B. Discussion................................ 141

IX. GENERAL DISCUSSION............................... 144

APPENDIX

   A. Contractile Systems.................... 146
       1. Striated Muscle..................... 146
       2. Smooth Muscle...................... 152
       3. Nonmuscle............................ 154

   LIST OF REFERENCES.......................... 157
## LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Myosin Alkali Light Chain Protein Homology</td>
<td>4</td>
</tr>
<tr>
<td>2. Chromosomal Distribution of Myosin Alkali Light Chain Genes</td>
<td>36</td>
</tr>
<tr>
<td>3. Correlation of Human MYL1 Sequences with Human Chromosomes in Rodent x Human Somatic Cell Hybrids</td>
<td>102</td>
</tr>
<tr>
<td>4. Nucleic Acid Sequence Homology of hMLC1V/Sb to Other MLC-ALK Genes</td>
<td>106</td>
</tr>
<tr>
<td>5. Nucleotide and Amino Acid Homology of Exon 4 from hMLC.3</td>
<td>124</td>
</tr>
<tr>
<td>6. Hybridizing Fragments Observed in the Genomic Southern Analysis of hMLC.4-PstI0.42</td>
<td>139</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Schematic Diagram of the Functional Myosin Hexamer</td>
<td>7</td>
</tr>
<tr>
<td>2. Diagram of a Right Hand Depicting the Relative Positions of the E and F Helices and the Divalent Cation Binding Loop in an E-F Hand Protein Domain</td>
<td>12</td>
</tr>
<tr>
<td>3. Schematic Diagram of an E-F Hand Protein Domain</td>
<td>13</td>
</tr>
<tr>
<td>4. Diagram of a One Dimensional Protein Gel Indicating the Different MLC-ALK Protein Isoforms</td>
<td>16</td>
</tr>
<tr>
<td>5. Diagram of the Vertebrate Skeletal MLC-ALK Gene</td>
<td>20</td>
</tr>
<tr>
<td>6. Nucleotide Sequence of the Rat MLC-ALK cDNA Fragment to be used as a Heterologous Gene Probe</td>
<td>45</td>
</tr>
<tr>
<td>7. Autoradiogram Results of the Southern Blot Hybridization Experiment with Restricted Human Genomic DNA Probed with the Rat MLC-ALK cDNA Fragment, pMLC-84/R10.431</td>
<td>46</td>
</tr>
<tr>
<td>8. Autoradiogram Results of the Southern Blot Hybridization Experiment with Restricted hMLC.1 DNA probed with pMLC-84/0.431</td>
<td>51</td>
</tr>
<tr>
<td>9. Restriction Map of hMLC.1</td>
<td>52</td>
</tr>
<tr>
<td>10. Autoradiogram Results of the Southern Blot Hybridization Experiment with Restricted hMLC.1 DNA probed with Human Genomic DNA</td>
<td>55</td>
</tr>
<tr>
<td>11. Autoradiogram Results of the Southern Blot Hybridization Experiment with Restricted hMLC.1 DNA Probed with an AluI Family Repeat Clone</td>
<td>56</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>12. Autoradiogram Results of the Southern Blot Hybridization Experiment with Restricted Human Genomic DNA Probed with XRI1.0</td>
<td>58</td>
</tr>
<tr>
<td>13. Autoradiogram Results of the Southern Blot Hybridization Experiment with Restricted Human Genomic DNA Probed with BII1.1</td>
<td>60</td>
</tr>
<tr>
<td>14. Autoradiogram Results of the Southern Blot Hybridization Experiment with Restricted Clones Isolated from the 3' Walk Experiment Probed with pMLC-84/RI0.431</td>
<td>62</td>
</tr>
<tr>
<td>15. Partial Restriction Map of the DNA Contained in the Two Clones, hMLC.l and hMLC.2</td>
<td>63</td>
</tr>
<tr>
<td>16. Autoradiogram Results of the Southern Blot Hybridization Experiment with Restricted hMLC.l DNA Probed with pMLC-84/RIA10.121</td>
<td>66</td>
</tr>
<tr>
<td>17. Partial Restriction Map of hMLC.l Indicating the Region Exhibiting Homology to Exon 3 from the Rat cDNA Clone</td>
<td>67</td>
</tr>
<tr>
<td>18. Autoradiogram Results of the Southern Blot Hybridization Experiment with Restricted pBI1.2 DNA probed with pMLC-84/RIA10.121</td>
<td>68</td>
</tr>
<tr>
<td>19. Partial Restriction Map of pBI1.2 indicating the position of Exon 3</td>
<td>70</td>
</tr>
<tr>
<td>20. DNA Sequence and the Predicted Amino Acid Sequence for Exon 3</td>
<td>71</td>
</tr>
<tr>
<td>21. Autoradiogram Results of the Southern Blot Hybridization Experiment with Restricted pBI1.1 DNA Probed with pMLC-84/RIA0.431</td>
<td>73</td>
</tr>
<tr>
<td>22. Partial Restriction Map of pBI1.1 indicating the position of Exon 4</td>
<td>74</td>
</tr>
<tr>
<td>23. DNA Sequence and the predicted Amino Acid Sequence for Exon 4</td>
<td>75</td>
</tr>
<tr>
<td>24. Partial Restriction Map of pXRI1.0 indicating the position of Exon 5</td>
<td>76</td>
</tr>
<tr>
<td>25. DNA Sequence and the Predicted Amino Acid Sequence of Exon 5</td>
<td>78</td>
</tr>
</tbody>
</table>
26. Autoradiogram Results of the Southern Blot 
Hybridization Experiment with restricted 
hMLC.2 DNA probed with pMLC-84/HRI0.038........... 79

27. DNA Sequence and the Predicted Amino Acid 
Sequence for Exon 6................................ 80

28. Partial Restriction Map of pXR11.0 indicating 
the position of Exon 6............................... 81

29. Nucleotide Sequences and the predicted Amino 
Acid Sequences for Exon 1 and Exon 2 
Oligonucleotide Probes, Oligo1 and Oligo2............ 83

30. Autoradiogram Results of the Southern Blot 
Hybridization Experiment with restricted hMLC.1 
DNA probed with Oligo1................................ 85

31. Partial Restriction Map of hMLC.1 indicating 
relative position of Exon 1........................... 86

32. Autoradiogram Results of the Southern Blot 
Hybridization Experiment with restricted 
pSXII.0 DNA probed with Oligo1....................... 88

33. Partial Restriction Map of pSXII.0 indicating 
the position of Exon 1............................... 89

34. DNA Sequence and the Predicted Amino Acid 
Sequence of Exon 1...................................... 90

35. Autoradiogram Results of the Southern Blot 
Hybridization Experiment with restricted hMLC.1 
DNA probed with Oligo2............................... 91

36. DNA Sequence and the Predicted Amino Acid 
Sequence of Exon 2...................................... 92

37. Partial restriction map of pBI1.1 indicating 
the position of Exon 2................................. 93

38. Partial Restriction Map of the DNA contained in 
hMLC.1 and hMLC.2 indicating the positions of 
Exons 1-6.................................................. 95

39. Autoradiogram Results of the Southern Blot 
Hybridization Experiment with restricted hMLC.2 
DNA probed with pHMLC1-Sb/AII-PvuII0.200............ 97
40. DNA Sequence of the 3' Untranslated Region
Identified with the pHMLCl-Sb probe............... 98

41. Partial Restriction Map of hMLC.1 indicating
the positions of Exons 1-7.................. 99

42. The complete DNA Sequence and the Predicted
Amino Acid Sequence for the Human Ventricular/
Slow Twitch MLC-ALK gene.................... 100

43. Autoradiogram Results of the Southern Blot
Hybridization Experiment with Restriction
Enzyme Digestion Products of Clones 3.3,
3.4, 3.7a, and 3.7b Probed with pMLC-84/RI0.431.. 108

44. Restriction Map of hMLC.3......................... 110

45. Autoradiogram Results of the Southern Blot
Hybridization Experiment with Restricted
hMLC.1 DNA probed with Human Genomic DNA...... 111

46. Autoradiogram Results of the Southern Blot
Hybridization Experiment with Restricted
hMLC.3 DNA probed with an AluI Repeat Clone.... 112

47. Autoradiogram Results of the Southern Blot
Hybridization Experiment with Restricted Human
Genomic DNA probed with hMLC.3-SK0.7............. 114

48. Autoradiogram Results of the Southern Blot
Hybridization Experiment with Restricted hMLC.1 DNA probed with pMLC-84/RIA10.121....... 115

49. Restriction Map of hMLC.3 Showing the Region
of Hybridization to pMLC-84/RIA10.121............ 117

50. Autoradiogram Results of the Southern Blot
Hybridization Experiment with Restricted
hMLC.3-RI6.0 DNA probed with pMLC-84/AIR10.308... 118

51. Autoradiogram Results of the Southern Blot
Hybridization Experiment with Restricted
hMLC.3-RI6.0 DNA probed with pMLC-84/AIR10.308
at increased Stringency ........................ 119

52. DNA Sequence Data Determined for Exon 4 from
hMLC.3............................................. 121
53. Partial Restriction Map of hMLC.3-RI6.0
   Indicating the regions of Exons 3, 4, and 5...... 122

54. Partial Restriction Map of hMLC.4............... 127

55. Autoradiogram Results of the Southern Blot
   Hybridization Experiment with Restricted
   hMLC.4 DNA probed with Human Genomic DNA...... 128

56. Autoradiogram Results of the Southern Blot
   Hybridization Experiment with Restricted
   hMLC.4 DNA probed with an AluI Repeat Clone..... 129

57. Autoradiogram Results of the Southern Blot
   Hybridization Experiment with Restricted
   hMLC.4-RI3.5 DNA probed with Human Genomic DNA... 131

58. Partial Restriction Map of hMLC.4-RI3.5
   Indicating the Position of the Repeat Element.... 132

59. Autoradiogram Results of the Southern Blot
   Hybridization Experiment with Restricted
   hMLC.4-RI3.5 DNA probed with pMLC-84/AIRI0.308... 133

60. Partial Restriction Map of hMLC.4-RI3.5
   Indicating the Region of Homology to the
   Rat cDNA clone.................................... 134

61. Autoradiogram Results of the Southern Blot
   Hybridization Experiment with Restricted
   Human Genomic DNA probed with hMLC.4-PstI0.42.... 136

62. Autoradiogram Results of the Southern Blot
   Hybridization Experiment with Restricted
   Human Genomic DNA probed with hMLC.4-PstI0.42
   at increased Stringency........................... 137

63. Autoradiogram Results of the Southern Blot
   Hybridization Experiment with Restricted
   Human Genomic DNA probed with hMLC.4-PstI0.42
   at Maximum Stringency............................. 138

64. Autoradiogram Results of the Southern Blot
   Hybridization Experiment with Restricted
   hMLC.4-RI3.5 DNA and Single Strand DNA
   Preparations from Subclones pHcPI0.35,
   pPstI0.42, and pPICIO.7............................ 140

65. Partial Restriction map of hMLC.4-RI3.5
   indicating the region of homology to the
   rat cDNA clone, pMLC-84............................. 142
66. DNA Sequence Data obtained from pHcPI0.35, pPstI0.42 and pPIC10.7........................... 143
67. Schematic Diagram of Striated Muscle.............. 147
68. Schematic Diagram of the Myosin Hexamer .......... 149
69. Schematic Diagram of the Calcium Regulatory Complex in Striated Muscle......................... 150
ABBREVIATIONS

A: Angstrom
ATP: Adenosine Triphosphate
BSA: Bovine Serum Albumin
cpm: counts per minute
dCTP: Deoxycytosine Triphosphate
DNA: Deoxyribonucleic Acid
DTNB: 5,5’ dithio-2 bis nitrobenzoic acid
EDTA: Ethylenediaminetetraacetate Dihydrate
kbp: kilobase pair
ml: milliliter
MLC: Myosin Light Chain (number designations following this abbreviation refer to specific isoforms of the protein)
MLC-ALK: Myosin Alkali Light Chain
MLC-DTNB: Myosin DTNB Light Chain
MHC: Myosin Heavy Chain
MW: Molecular Weight
pfu: plaque forming unit
PVP: Polyvinyl Pirrolidone
RNA: Ribonucleic Acid
SDS: Sodium Dodecylsulfate
SSC: Sodium chloride, Sodium Citrate solution
ug: microgram
UTR: Untranslated Region (of a transcript)
UV: Ultraviolet light
A. Introduction

Contractile proteins are highly conserved throughout the evolution of eukaryotic cells. Virtually all eukaryotic cells possess contractile properties which are essential to the maintenance of cell shape and cell motility. Actin and myosin are the principal structural and functional proteins in the process of contraction from single cell eukaryotic organisms to the more recently evolved muscular contractions in metazoan muscle tissue. Structural analysis on contractile proteins in vertebrates has shown that they exist as multiple isoforms with different primary sequences. This has led to the discovery of highly conserved families of proteins which are encoded by multigene families (Buckingham and Minty, 1983).

B. Goals, Hypothesis and Rationale

The myosin alkali light chain is part of the functional myosin hexamer in contractile systems and recently has been investigated at the level of the gene. At the time when this project was initiated the myosin alkali light chain gene from Drosophila (Falkenthal et al., 1984) and the gene which encodes the vertebrate skeletal muscle isoforms, MLC1/3, from chicken, mouse and rat had
been reported (Nabeshima et al., 1984; Robert et al., 1984; Periasamy et al., 1984; respectively). Human myosin alkali light chain genes had not been investigated, therefore it was hypothesized that a heterologous gene would be useful in identifying human MLC-ALK genes. A rat skeletal MLC-ALK cDNA clone, pMLC-84, was kindly supplied by Drs. L. Garfinkel and B. Nadal-Ginard and was the only available clone to use as a gene probe at the time this study was initiated. The rationale that provided the basis for the hypothesis is as follows: a number of the major myofibrillar structural protein genes have been cloned and characterized (actin; myosin heavy chain; tropomyosin; the troponins I, C, and T; myosin alkali light chain; myosin DTNB light chain) from a variety of species (reviewed by Buckingham and Minty, 1983). Many of these genes have been isolated by probing genomic and/or cDNA libraries with previously cloned heterologous genes (Drosophila actins, Fyrberg et al. 1980, 1981; human actins, Engel et al. 1981 and Humphries et al. 1981; Drosophila myosin heavy chain, Bernstein et al. 1983, etc.). The high degree of sequence conservation between the myofibrillar protein genes in a variety of diverse organisms permits the isolation of similar sequences from distantly related genomes.

Initially, the extent of sequence conservation between all known vertebrate MLC-ALK proteins (actual amino acid sequences or predicted from the gene sequence) was investigated. Sequence
comparisons were performed to determine the extent of homology between all MLC-ALK proteins. The results of those comparisons are presented in Table 1. Evident from the data presented in Table 1 is the high degree of sequence conservation between all MLC-ALK proteins. The region of the protein which exhibits the highest degree of homology between all isoforms of the protein is E-F hand Domain 3, i.e. 90% or greater and which is represented in the above comparison by amino acids residues 130-159 (data not shown). Interestingly, it was also apparent from these initial comparisons that the extent of homology for similar isoforms between species is greater than the homology between isoforms within a species, e.g. the extent of homology between the chicken skeletal fast twitch isoform and the mouse, rabbit and rat skeletal fast twitch isoforms is greater than the homology between the chicken skeletal sequence and the chicken cardiac and gizzard (smooth muscle) isoforms. Based on the results in Table 1 and the rationale presented above the following goals were set to analyze human myosin alkali light chain genes.
**TABLE 1:** The extent of amino acid sequence homology for the 141-COOH terminal amino acids of MLC-ALK isoforms.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Homology (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>chicken cardiac/chicken smooth</td>
<td>72%</td>
</tr>
<tr>
<td>chicken cardiac/chicken skeletal</td>
<td>73%</td>
</tr>
<tr>
<td>chicken smooth/chicken skeletal</td>
<td>73%</td>
</tr>
<tr>
<td>chicken cardiac/rabbit skeletal</td>
<td>76%</td>
</tr>
<tr>
<td>chicken smooth/rabbit skeletal</td>
<td>77%</td>
</tr>
<tr>
<td>chicken smooth/rat skeletal</td>
<td>75%</td>
</tr>
<tr>
<td>chicken cardiac/rat skeletal</td>
<td>78%</td>
</tr>
<tr>
<td>chicken smooth/mouse skeletal</td>
<td>75%</td>
</tr>
<tr>
<td>chicken cardiac/mouse skeletal</td>
<td>78%</td>
</tr>
<tr>
<td>chicken skeletal/rabbit skeletal</td>
<td>83%</td>
</tr>
<tr>
<td>chicken skeletal/rat skeletal</td>
<td>86%</td>
</tr>
<tr>
<td>chicken skeletal/mouse skeletal</td>
<td>86%</td>
</tr>
<tr>
<td>rabbit skeletal/rat skeletal</td>
<td>91%</td>
</tr>
<tr>
<td>rabbit skeletal/mouse skeletal</td>
<td>91%</td>
</tr>
<tr>
<td>mouse skeletal/rat skeletal</td>
<td>99%</td>
</tr>
</tbody>
</table>

The homology data presented was compiled from Parisamy et al., 1984; Robert et al., 1984; Nabeshima et al., 1984; Matsuda et al., 1981; and Maita et al., 1980. The skeletal muscle isoforms represent the common region between MLC1 and MLC3. Current amino acid sequence data for the rabbit, mouse and rat cardiac and smooth MLC-ALK isoforms is unavailable. Only the C-terminal 141 amino acids were used in this homology comparison, given that the probe to be used for the analysis of human MLC-ALK genes only contains coding sequences for these amino acids.
The goals of this thesis project were to:

1. isolate members of the human myosin alkali light gene family from human genomic libraries

2. characterize the gene structure of isolated genes, i.e. determine the position and boundaries of the individual exons and introns

3. determine the coding potential of isolated genes

4. determine the chromosomal localization of isolated genes in collaboration with Uta Francke’s laboratory at Yale University.

C. Research Summary

The following report describes the genomic characterization, DNA sequence analysis, and chromosomal localization of a human MLC-ALK gene, $\text{MLC}_{V/Sb}^1$, that was isolated from a human partial EcoRI genomic library cloned into lambda Charon 4A. Preliminary analysis and the partial characterization of two addition clones, hMLC.3 and hMLC.4, that were isolated from a human partial Sau3AI library cloned into lambda Charon 28, is also presented. All human clones were isolated by using a 431bp EcoRI fragment from a rat skeletal MLC-ALK cDNA clone (Garfinkel et.al., 1982) as a heterologous gene probe.
CHAPTER II

A REVIEW OF THE MYOSIN ALKALI LIGHT CHAIN LITERATURE

A. Introduction

The following literature review will focus on a single component of the contractile apparatus, the myosin alkali light chain (MLC-ALK), and will be devoted to a review of the current knowledge obtained on the protein, its function, the different isoforms, and the genes which encode the various isoforms. A synopsis of the different contractile systems present in higher eukaryotic organisms is discussed in Appendix A. Appendix A is not meant to be exhaustive, but will only serve to introduce the main components of the contractile apparatus and familiarize the reader with each of the following systems; striated muscle, smooth muscle and non-muscle.

B. Protein Function: The Myosin Hexamer

The functional myosin molecule exists as a hexamer and is depicted in Figure 1. The molecule is composed of two myosin heavy chain monomers (MW = 220,000) and two pairs of distinct myosin light monomers (Weeds & Pope, 1977). Each myosin heavy chain contains a globular head domain and a long rod-like alpha-helical tail domain. The myosin light chain components of the hexamer are
Figure 1. Schematic diagram of the functional myosin hexamer indicating the myosin heavy chain (MHC) and myosin light chain components (MLC-ALK and MLC-DTNB).
characterized based upon their chemical dissociation from the myosin complex (Weeds, 1969; Weeds and Lowey, 1971; Gazith et al., 1970), the myosin DTNB light chain (MW = 20,000-30,000, DTNB = 5,5' dithio-2 bis nitrobenzoic acid) and the myosin alkali light chain (MW = 17,000-20,000). One myosin alkali light chain monomer and one myosin DTNB light chain monomer associate with each of the myosin heavy chain globular domains in the proximity of the head-tail junction domain (Flicker et al. 1983; Hardwicke and Szent-Gyorgi, 1984; Winklemann and Lowey, 1983; Sellers and Harvey, 1984). The position of the MLC-ALK protein in the functional myosin molecule has been investigated using two different experimental approaches. Immunoelectron microscopy with purified myosin molecules and monoclonal antibodies raised against the myosin alkali light chain identified the relative position of the MLC-ALK protein in the hexamer. Increased density in the electron micrographs was observed in the region between the globular head and the rod-like tail of the myosin heavy chain (Flicker et al. 1983). In vitro reassociation experiments with electrophoresed myosin heavy chain protease digestion products and labeled $^{14}$C alkali light chain identified the myosin heavy chain protein fragment that binds the myosin alkali light chain. Myosin heavy chain proteins were isolated from skeletal and smooth muscle, proteolytically digested with trypsin, chymotrypsin and papain, electrophoresed on polyacrylamide gels and then incubated with $^{14}$C labeled myosin alkali light chain protein. In both muscle
types, the alkali light chain associated with the COOH terminal portion of the myosin heavy chain globular head region, and corresponds to the head-tail junction domain of the heavy chain (Sellers and Harvey, 1984).

The function of the myosin hexamer in the contractile apparatus is discussed in Appendix A, however it is important to mention here that the hexamer possesses the ATPase activity involved in breaking the actomyosin crossbridge. Loss of the ATPase results in a loss of contractility and produces a static state of rigor. Initial studies on isolated myosin from rabbit skeletal muscle indicated that removal of the alkali light chain (i.e. pH 11.0) resulted in a loss of ATPase activity (Weeds and Lowey, 1971). However, it has been shown that the enzymatic activity of the myosin ATPase is functional with purified myosin heavy chain and therefore can function in the absence of the MLC-ALK protein (Wagner and Giniger, 1981). Photoaffinity labeling experiments with ATP analogs and rabbit striated skeletal muscle have determined that on the myosin heavy chain contains the photoincorporated analog, specifically at residue 130 (i.e. tryptophan 130). Therefore, the ATPase active site is localized to the myosin heavy chain globular head and functions without MLC-ALK amino acid contributions (Okamoto and Yount, 1985). Recent photoaffinity labelling experiments on chicken gizzard smooth muscle produced different results. Okamoto et al. (1986) found 17% (+2%) of the photoincorporated analog was associated with the
alkali light chain, which may give some indication as to the function and location of the alkali light chain in smooth muscle myosin, i.e. close association with the myosin heavy chain Mg ATPase. The marked difference in the labelling pattern observed for smooth muscle myosin compared with that seen with striated muscle myosins may reflect variation in the heavy chain sequences between smooth and striated heavy chain isoforms or it may reflect differences in the regulation of the contractile apparatus between these different tissues (see Appendix A; striated muscle is regulated by the troponin complex, whereas smooth muscle regulation is myosin-linked). The actual function of the MLC-ALK protein is still undetermined, but experiments have shown that they influence the actin-activated ATPase (Wagner and Weeds, 1977; Trayer and Trayer, 1985) and indicates a possible role in the actin-myosin crossbridge interaction.

C. MLC-ALK Protein Characteristics

The MLC-ALK protein is a member of the troponin C superfamily of proteins. These are a related group of proteins characterized by the presence of E-F hand calcium binding domains. Proteins in this family include troponin I, troponin C, troponin T, calmodulin, parvalbumin and the myosin DTNB light chain (Kretsinger 1980).

The E-F hand structural domain of calcium binding proteins was originally described by Moews and Kretsinger (1975). Based on the crystal structure of carp parvalbumin, isotype 4.25, as determined
by X-ray crystalography at 1.9 Å, they identified distinct structural features of the protein. Fourier analysis of the X-ray diffraction data indicated the presence of six alpha-helical regions, denoted A-F. The helices are bordered by a non-helical NH₂-terminus and five loops which occur between each of the six helices. The E helix, loop EF (the loop located between the E helix and the F helix), and the F helix are described by Kretsinger (1980) as resembling a right hand (see Figure 2), where the forefinger represents the E helix and the thumb represents the F helix. The middle finger represents a loop which is capable of coordinating a calcium ion. The term "EF hand domain" has originated from this description. Amino acid sequence data comparisons have identified homologous domains to the parvalbumin EF hand in each of the members of the Troponin C superfamily of proteins (Kretsinger 1980, Collins et.al. 1973, Weeds and Mclachlan 1974, Tufty and Kretsinger 1975).

The EF hand domain as described by Moews a.l Kretsinger has the following features (see Figure 3); it consists of 29 amino acids which are organized into an E alpha-helix of 9 amino acids, a loop of 12 amino acids, and an F alpha-helix of 8 amino acids. The structure is denoted n(2), n(5), n(6), n(9), X(10), Y(12), Z(14), -Y(16), -X(18), -Z(21), n(25), n(26), n(28), n(29). The numbers represent the relative positions of residues in the domain. Residues denoted "n" represent the hydrophobic or inner aspects of helix E (residues 2, 5, 6, and 9) and of helix F (residues 25, 26,
Figure 2. Diagram of a right hand depicting the relative positions of the E and F helices and the divalent cation binding loop in an E-F Hand protein domain.
FIGURE 3. A diagrammatic representation of an E-F hand domain. Important residues are numbered as described above. The E helix and F helix are indicated with arrows. The highly conserved amino acid residues glycine 15 and isoleucine 17 are indicated by G and I, respectively. A discussion of these residues is presented in the text.
The calcium binding ligands within the 12 amino acid loop can be assigned to the vertices of an octahedron and are indicated by X, Y, Z, -X, -Y, -Z. In order to bind calcium, the 12 amino acid loop of the E-F hand must possess residues with oxygen containing side chains at 5 vertices; X(10), Y(12), Z(14), -X(18) and -Z(21); the sixth, -Y (16), is contributed by the oxygen of the alpha-carbonyl carbon atom. Any one of the five vertices X, Y, Z, -X, or -Z, can lack an oxygen containing side chain if glycine is present at that position (glycine is capable of coordinating calcium through the oxygen of a hydrogen bonded water molecule (ibid)). Functionally established calcium binding loops possess a number of characteristics. Asparagine or aspartic acid are usually found at the X vertex. Glutamic acid is usually found at the -Z vertex. Serine or threonine may be present at any single vertex in the loop, but they never occur twice in the same loop (serine and threonine are the shortest ligands with oxygen bonded to the beta-carbon atom, which implies a spacing requirement). The sum of the carboxylate groups in the loop is usually 3 or 4 and all calcium binding loops contain valine or isoleucine at position 17, which implies the necessity for a hydrophobic residue to stabilize the loop in the hydrophobic core of the protein. Functional E-F hand domains also possess a glycine at position 15, which would indicate some change in loop structure associated with any substitution for glycine (Kretsinger, 1980). Deletions and/or insertions are not tolerated in the loop if the domain is to remain functional.
Myosin alkali light chains contain four putative E-F hand domains, but the protein fails to bind any calcium either as an isolated protein or when associated with the myosin hexamer (Kendrick-Jones et al., 1970; Szent-Gyorgyi et al., 1973). One of the four domains (i.e. domain 3) has retained four acidic groups among its six ligating residues, however this site appears distorted since it lacks both the glycine and isoleucine residues characteristic of E-F hand domains as well as possessing bulky apolar residues which may make this site inaccessible to divalent cations. Interestingly, domain 3 has been highly conserved between all MLC-ALK proteins and therefore argues that it must provide some essential function, although this is still undetermined. All of the other putative E-F hand domains have diverged significantly and are unable to bind divalent cations.

D. MLC-ALK Isoforms and Their Genes

Vertebrate isoforms of the myosin alkali light chain protein range in molecular weight from 17,000 to 21,000 daltons depending on the species and tissue type. Figure 4 is a schematic representation of a one dimensional protein gel depicting the different isoforms of the MLC-ALK proteins (reproduced from Barton and Buckingham 1985). The different vertebrate isoforms identified to date include: an embryonic isoform; fast-twitch skeletal muscle isoforms MLC1_F and MLC3_F; slow twitch isoforms MLC1_Sa (mammalian
Figure 4. Diagram of a one dimensional protein gel indicating the different MLC-ALK protein isoforms. The MLC-ALK proteins are denoted by the number 1 in the skeletal and cardiac panels and by G2 in the smooth muscle panel. The proteins indicated by the number 2 and G1 represent the DTNB light chain.
specific) and MLC_{sb}; two cardiac isoforms, a ventricular and an atrial; a smooth muscle; and a non-muscle isoform. An initial observation that can be made from Figure 4 is that protein variability exists between the vertebrate isoforms within the MLC-ALK family (a discussion of the various isoforms is presented in the following sections). Two features of the protein which vary and distinguish certain isoforms in the family are:

(1) the size of the NH_{2} terminus and

(2) the presence of tri-methylalanine at the NH_{2} terminus.

The vertebrate isoforms MLC_{F}, MLC_{A}, MLC_{V}, and MLC_{S} possess additional NH_{2} terminal amino acids and can be distinguished from the MLC_{F} skeletal isoform, the smooth muscle, non-muscle, as well as invertebrate proteins on 1-dimensional protein electrophoresis gels by their reduced migration (see Figure 4). The vertebrate isoforms MLC_{F}, MLC_{A}, MLC_{V}, and MLC_{S} examined to date also possess the unusual NH_{2} terminal alanine residue, tri-methylalanine (Henry et al. 1985). The vertebrate skeletal isoform MLC_{F}, the smooth muscle isoform from vertebrates, and all invertebrate species' proteins examined to date do not contain this modification (ibid). Empirically, this NH_{2} terminal modification would increase the hydrophobicity of this residue, however functional implications of this have not been examined.
The following sections describe the different isoforms found in different vertebrate tissue types and the proteins identified in invertebrates. Where available, the genes which encode these different proteins will also be discussed.

1. Fast twitch skeletal muscle isoforms. Initial protein analysis on rabbit fast twitch skeletal muscle identified two myosin alkali light chain isoforms $\text{MLC}_{1F}$ ($MW = 21,000$ daltons) and $\text{MLC}_{3F}$ ($MW = 17,000$ daltons; Frank and Weeds, 1974). Amino acid sequence data for fast-twitch isoforms $\text{MLC}_{1F}$ and $\text{MLC}_{3F}$ from rabbit and chicken (Frank and Weeds, 1974; Matsuda et. al. 1981a; respectively) showed that these isoforms contain an identical 141 C-terminal amino acids, but differ in the N-terminal amino acids both in number and sequence composition. The $\text{MLC}_{1F}$ isoform contains 46 or 49 amino acids (rabbit and chicken, respectively), whereas the $\text{MLC}_{3F}$ isoform contains a distinct 8 amino acids at the amino terminus in both the rabbit and the rat. Matsuda et. al. (1981a) also made the observation that amino acid differences that occurred in the common region between rabbit and chicken were conserved for both $\text{MLC}_{1F}$ and $\text{MLC}_{3F}$. For example, the COOH-terminal amino acid in the rabbit $\text{MLC}_{1F}$ and $\text{MLC}_{3F}$ proteins is an isoleucine, whereas in chicken a valine residue is present at the COOH-terminus in both proteins.

Subsequently, it has been shown that these two isoforms are encoded by the same gene.
The vertebrate striated skeletal fast twitch MLC-ALK gene from three different species, chicken (Nabeshima et al., 1984), mouse (Robert et al., 1984), and rat (Periasamy et al., 1984) have been cloned and sequenced. In all cases the vertebrate fast twitch gene has been shown to encode the two different isoforms, MLC1\_F and MLC3\_F, via differential promoter utilization and differential splicing of the primary transcripts (ibid). In the following description of the vertebrate fast twitch gene structure, the numbering convention used to describe the exons in the mouse gene (Robert et al., 1984) will also be used here. This numbering scheme provides a more logical way of describing the transcripts that are generated from the gene. Analysis of skeletal fast twitch cDNA clones and the corresponding genomic DNA has revealed that a total of 9 exons (numbered 1, 1’, 2, 2’ and 3-7) are present in the MLC1\_F/MLC3\_F gene, however only seven exons are used to encode either the MLC1\_F or MLC3\_F isoform, see Figure 5. Exons 1, 2, 3-7 are utilized to encode the MLC1\_F isoform, and exons 1’, 2’, 3-7 are utilized to encode the MLC3\_F isoform. Figure 5 is a diagramatic representation of the vertebrate fast twitch gene and shows the mechanism by which these two transcripts are generated from a single gene. The MLC1\_F transcript is initiated at the 5’ most promoter and proceeds 3’ through all nine exons. This transcript is then processed to produce a mature mRNA that contains exons 1, 2, and 3-7 and as a result excludes exons 1’ and 2’. The MLC3\_F transcript is initiated at the 3’ most promoter located in intron 1, and
Figure 5. Diagram of the vertebrate skeletal MLC-ALK gene. Exons are numbered as described in the text and are represented by the open boxes and hatched boxes. The hatched areas represent the 5' and 3' untranslated regions of the transcripts. Intron sizes are indicated in base pairs between the exons.
proceeds 3' through exons 1', 2', and 3-7. This transcript is
processed to produce a mature mRNA which contains exons 1', 2', and
3-7 and results in a transcript that excludes exon 2. Exons 3-7
are present in both transcripts and represent a common coding
region and a common 3' untranslated region (3' UTR). As a result
the MLC1_F isoform and the MLC3_F isoform share an identical 141-COOH
terminal amino acids but differ in their amino termini. The
mechanism responsible for generating these mRNAs must be under
strict regulation, since the expression of the two proteins is
modulated during fetal development, where MLC1_F accumulates first
and not until later in development does MLC3_F appear (Streter
et.al. 1975, Roy et.al. 1979, Gauthier et.al. 1982, Barton et.al.
1985). A more detailed discussion of MLC-ALK expression during
development is presented in a following section (see E.1.
Development).

Human cDNA clones for the MLC1_F (Seidel et.al. 1987) and MLC3_F
(Seidel et.al. 1987, Wade et.al. in press) isoforms have also been
reported. Nucleotide sequence analysis revealed 100% homology
between the two clones except for 201 and 79 nucleotides at the 5'
end of MLC1_F and MLC3_F respectively. A similar gene structure to
that of the other vertebrate fast twitch skeletal muscle isoform
genes is proposed, but the actual gene structure has not been
reported.

A mouse processed pseudogene, corresponding to the skeletal
fast twitch common coding region, has been identified in the
species *Mus musculus* (Robert et al. 1984). DNA sequence analysis of this pseudogene identified regions which are characteristic of a processed-type pseudogene (Vanin 1980, Nishioka 1980, Hollis 1982). The sequence contains no intron, contains a stretch of 10 adenine residues at the end of the sequence representing the 3' UTR. A putative insertion site represented by a direct repeat flanking the pseudogene was also identified (ibid). Interestingly, the homology between the 5' end of the pseudogene and the fast-twitch common coding region stops precisely at the 5' end of the common coding region, i.e. at the 5' splice site of exon 3.

2. Slow Twitch Skeletal Muscle Isoforms. The myosin alkali light chain content in vertebrate slow twitch striated skeletal muscle exhibits variation between species and fiber type. In mouse, a single myosin alkali light chain isoform (MLC1_S) is present in slow twitch muscle (Butler-Brown & Whalen, 1984). Rat slow twitch skeletal muscle contains two distinct isoforms, MLC1_S and MLC1_G, however the MLC1_S isoform appears only as a minor component (Whalen et al., 1978). Rabbit soleus muscle also contains two slow twitch alkali light chains, MLC1_Sa and MLC1_Sb (which are synonomous to the MLC1_S and MLC1_G isoforms from rat, respectively), in relatively equal proportions (Weeds 1976). Peptide mapping indicates that the MLC1_Sa and MLC1_Sb isoforms contain regions of similar sequence, however partial amino acid sequence data on the thiol peptides from these proteins indicates
that that these isoforms are distinct (Weeds, 1976). In addition, monoclonal antibodies recognize distinct epitopes on each of these proteins (Margreth et al. 1980).

Recently a human slow twitch skeletal muscle MLC-ALK cDNA clone (pHMLC1-Sb) has been isolated from human cDNA library constructed from the gastrocnemius muscle (Wade et al., in press). The DNA sequence of this clone was determined and shown to contain 888 nucleotides, which were distinct from the human fast muscle cDNA clones. The coding potential of pHMLC1-Sb predicted a protein of 195 amino acid residues and exhibited the greatest degree of homology to the chicken cardiac isoform. This clone was also shown to be expressed in adult ventrical (Wade et al. in press). Recent evidence indicates that the MLC1\(_S\) isoform in mouse is also expressed in adult ventricle (see Cardiac Isoforms).

3. Cardiac Muscle Isoforms. Mammalian cardiac muscle contains two myosin alkali light chain isoforms, a ventricular (MLC1\(_V\)) and an atrial (MLC1\(_A\)), (Syrovy et al., 1979; Wikman-Coffelt and Srivastava, 1979). Chicken, however only contains a single cardiac isoform for which the complete amino acid sequence is known (Maita, 1980). The mammalian MLC1\(_V\) isoform and the chicken cardiac isoform have been shown to exhibit the same electrophoretic mobility and based on immunological data also share common epitopes (Lowey and Risby, 1971; Sakar et al., 1971; Weeds and Pope, 1971; Obinata et al., 1979; Sartore, 1981). It was mentioned
in the previous section that the human cDNA clone, pHMLC1-Sb, is expressed in both slow twitch skeletal muscle and adult ventricle. This also appears to be the situation in mouse. A partial MLC1\textsubscript{v} isoform cDNA clone from mouse ventricle has been isolated and used to test the relationship between MLC1\textsubscript{Sb} and MLC1\textsubscript{v} mRNAs. Both isoforms are encoded by mRNAs that are indistinguishable on the basis of size and sequence homology to the MLC1\textsubscript{v} cDNA clone (Barton et al. 1985). Two-dimensional gel analysis and partial peptide mapping of MLC1\textsubscript{Sb} and MLC1\textsubscript{v} indicate that these isoforms are indistinguishable (Margreth et al., 1980; Weeds, 1976; respectively).

Mammalian atria and the Purkinje fibers in the heart contain a specific isoform, MLC1\textsubscript{A} (Long et al., 1977; Syrovy et al., 1979) which has an electrophoretic mobility similar to the MLC1\textsubscript{F} isoform. This isoform is tissue specific in adult heart muscle, but 2-D gel analysis (Whalen et al. 1978; Whalen et al. 1982; Dalla-Libera 1981) and peptide mapping (Whalen & Sell 1980, Cummins 1982) have found that this isoform is indistinguishable from an MLC1 fetal isoform, MLC1\textsubscript{emb}, which is found in fetal striated muscle (Whalen 1978) as well as in fetal atrial and ventricular muscle (Price 1980, Cummins 1980). Nucleotide analysis of the MLC1\textsubscript{A} and MLC1\textsubscript{emb} isoforms in mouse support the theory that these isoforms are identical and suggests that they are encoded by the same gene. A mouse atrial MLC1\textsubscript{A} cDNA clone has been isolated from an adult atrial cDNA library. Partial nucleotide analysis confirmed that it
was generated from a myosin alkali light chain transcript (Barton et.al., 1985). Hybridization and R-loop analysis of atrial and fetal mRNAs has shown that both molecules share a common size and sequence homology to the atrial cDNA clone (Barton et.al. 1985).

A chicken fetal muscle myosin alkali light chain cDNA clone has recently been isolated from embryonic gizzard muscle and sequenced. The chicken fetal MLC-ALK cDNA clone (L_{23}) contains an insert of 832 nucleotides and predicted a polypeptide of 185 amino acids. Genomic southern blot analysis revealed a single hybridizing fragment different from the skeletal and smooth/non-muscle genes, indicating that the gene which encodes this transcript is unique (Kawashima et.al. 1987). RNA blot analysis with the 3' untranslated region of the cDNA clone identified transient expression of the transcript in embryonic skeletal, cardiac and smooth muscles. The transcript was also identified in brain tissue during all stages of embryonic development as well as in one day old chicks (ibid). Seven day old chick brain RNA also exhibited very faint hybridization to the 3' UTR of L_{23}.

4. Smooth Muscle and Nonmuscle Isoforms. Smooth muscle and nonmuscle contain a single myosin alkali light chain isoform which has a similar electrophoretic mobility to the skeletal MLC_{3p} isoform (Burridge 1974). The chicken gizzard isoform has been sequenced and contains 150 amino acids (Matsuda et.al., 1981).

Most recently, cDNA clones for the chicken smooth and chicken
nonmuscle myosin alkali light chains have been sequenced (Nabeshima et.al. 1987 and Kawashima et.al. 1987, respectively). The non-muscle and smooth muscle clones have been predicted to encode proteins of 151 amino acids in length. The two cDNA clones are identical in the 5’ untranslated region and for 427 nucleotides extending 3’ from the translational initiation codon. The two sequences diverge for 39 nucleotides representing the COOH-terminal coding region, but then are again identical in their 3’ untranslated region. Genomic southern analysis with the two clones as probes has revealed a common restriction fragment banding pattern (Nabeshima et.al., 1987). These results suggest that the two isoforms are encoded by the same gene and possibly undergo a tissue specific splicing mechanism (the Drosophila MLC-ALK gene employs this strategy and will be discussed in the following section).

5. Invertebrate isoforms. Myosin light chain proteins have been identified in Acanthameoba (Pollard and Korn, 1973; Maruta and Korn, 1977), Dictyostelium (Clarke and Spudich, 1974; Mockrin and Spudich, 1976), and Physarum (Jacobsen et.al., 1976; Hatano and Owaribe, 1976). The following invertebrate species so far examined; scallop(Kendrick-Jones et.al., 1976), squid (Kendrick-Jones et.al., 1976), crayfish and lobster(Ko et.al., 1979), the nematode Ascaris lumbricoides (Nakamura et.al., 1975), Lethocerus codofanus (water bug, Bullard et.al. 1973), Heliocopris japetus (dung beetle, ibid), all contain a single alkali-type myosin light
chain in the 15,000-17,000 dalton size range. Drosophila contains two 17,000 dalton MLC-ALK isoforms (Takano-Ohmura et al. 1983b; Falkenthal et al., 1985).

The myosin alkali light chain gene from Drosophila has been completely characterized (Falkenthal et al., 1984; Falkenthal et al., 1985). The Drosophila myosin alkali light chain gene is present in single copy per haploid genome and encodes two different isoforms, the IFM (indirect flight muscle, i.e. striated muscle) specific and the tubular muscle specific isoforms (ibid).

Differential splicing of the primary transcript from this gene generates these two isoforms in a developmental and tissue specific manner. The genomic copy of the gene contains a total of 6 exons. In larva, the mature transcript contains all 6 exons and encodes a protein of 155 amino acids. During pupariation and in the adult fly an additional transcript is generated from the gene. S1 nuclease analysis of this transcript has shown that exon 5 is absent from the mature transcript and consequently only contains exons 1, 2, 3, 4, and 6. Sequence analysis of the 3' splice site in intron 4 revealed a non-consensus polypurine stretch of nucleotides preceding the 3' acceptor site. The resultant IFM protein also contains 155 amino acids, but is unique for the 13 amino acids at the C-terminus (ibid). Sequence analysis of exon 6 reveals the dual function of this exon. In the larval transcript, exon 6 serves as the 3'UTR, whereas in the IFM transcript exon 6 provides coding sequence for the 13 COOH-terminal amino acids as well as the
3' UTR. The difference resulting from a shift in the reading frame of exon 6 created by the 4-6 splice junction. S1 nuclease analysis also revealed transcripts which differ in size due to alternative polyadenylation. DNA sequence analysis of the 3' UTR of the gene identified 4 putative polyadenylation signals (i.e. the consensus sequence, aataaa), at least 2 of which are utilized as determined by S1 nuclease experiments (ibid).

The only other invertebrate MLC-ALK sequence data comes from the adductor muscle from the species *Aequipectin irradians* (scallop). Protein sequence data reveals that it contains 156 amino acids and is similar in size to the Drosophila isoforms (Collins et.al. 1986). Most recently, the nucleotide sequence (generated from two overlapping cDNA clones; Goodwin et.al. 1987) has been published and predicts an amino acid sequence which corresponds exactly to the protein sequence. Interestingly, northern blot analysis of poly(A)^+ RNA from the adductor muscle probed with a mixture of the cDNA clones resulted in hybridization to multiple transcripts (1.6kb, 1.3kb, and 0.65kb), however the relationship between these transcripts was not investigated (ibid). Genomic southern analysis, again using both cDNAs as probe, revealed a single hybridizing fragment (ibid). These results seem to suggest that the scallop gene may also utilize differential splicing, differential promoter utilization, and/or differential polyadenylation to generate multiple transcripts from a single gene. Analyzing the genomic DNA should reveal if any of these
mechanisms result in the various transcripts or if in fact the multiple transcripts are generated by similar genes organized as a small gene cluster.

E. MLC-ALK Expression

1. Development. Developmental transitions occur in MLC-ALK isoform expression and have been documented in Drosophila and vertebrates. A switch in gene expression from non-muscle to muscle-type isoforms when mononucleated myoblasts fuse to form multinucleate muscle fibers in the developing organism (Buckingham, 1985). Maturation of the muscle fiber is characterized by the expression of a distinct isoform(s) which is(are) stably maintained under normal physiological conditions. Development of striated muscle in vertebrates generally exhibits the following pattern of MLC-ALK expression:

myoblasts ----> fiber formation ----> fiber maturation
(non-muscle) (embryonic) (adult isoforms)

Analysis of myoblast cultures provides the only evidence for what is believed to occur in vivo for the switch from nonmuscle to muscle specific isoforms. Myoblast cultures grown as monolayers of mononucleate cells has shown that nonmuscle contractile protein isoforms are expressed until fusion and fiber formation occur where upon muscle specific isoforms accumulate (Buckingham 1985). The
different muscle types within an organism also exhibit slightly different modes of MLC-ALK expression and therefore a discussion of the development of each muscle type will be presented. In addition, avian muscle development differs from mammalian muscle development and will be discussed separately.

Developing mammalian fast twitch muscle fibers predominantly contain the MLC1\textsubscript{emb} isoform (Whalen 1978, Whalen & Sell 1980, Price 1980, Barton 1985a). However, in humans the MLC1\textsubscript{F} isoform already begins to replace the MLC1\textsubscript{emb} isoform in the second trimester and is the predominant isoform at birth (Strohman 1983, Biral 1984). The MLC1\textsubscript{Sa} isoform has also been detected in vivo in human 3-4 month fetal muscle (Biral et.al. 1984). In rodents, the MLC1\textsubscript{emb} persists until birth along with the MLC1\textsubscript{F} isoform. MLC3\textsubscript{F} only accumulates in significant amounts after birth (Sreter 1985, Roy 1979b, Gauthier 1982).

Fetal slow twitch skeletal muscle fibers contain the MLC1\textsubscript{emb}, MLC1\textsubscript{F} and MLC1\textsubscript{Sa} isoforms, but at birth the MLC1\textsubscript{Sa} isoform begins to accumulate significantly and eventually predominates (Gauthier 1982). The MLC1\textsubscript{Sa} isoform accumulates later in these fibers (ibid). Developing mammalian ventricular muscle fibers co-accumulate MLC1\textsubscript{emb} and MLC1\textsubscript{V} (Cummins 1980, Whalen & Sell 1980) whereas the developing atria only contains the MLC1\textsubscript{A/emb} isoform which persists and is maintained in the adult atria (Cummins 1980, Whalen & Sell 1982).
Avian muscle exhibits a slightly different mode of MLC-ALK expression. A fetal isoform is expressed during the early stages of development in cardiac, skeletal and smooth muscle (Takano-Ohmuro 1985). This isoform has distinct electrophoretic and immunological properties from the adult striated isoforms (Kato & Kubo 1978, Takano-Ohmuro 1985). Avian fetal skeletal muscle at 5-7 days in ovo only expresses the MLC1_p isoform in potentially fast or slow twitch fibers (Obinata et al. 1980; Stockdale et al. 1981; Lowey et al. 1983; Matsuda et al. 1983). MLC1_s accumulates from days 8-12 in both fast and slow twitch muscle (ibid). The adult pattern of expression begins to establish itself by day 12 and by day 16 fast twitch fibers begin to accumulate the MLC3_p isoform. By day 19 only the MLC1_p and MLC3_p isoforms are detectable (Crow 1983, Crow and Stockdale 1984). Adult slow twitch skeletal muscle predominantly contains the MLC1_s, although MLC1_p is seen as a minor component and ultimately depends on the extent of a fast fiber component in the characteristically slow twitch muscle group. MLC3_p is never detected in slow muscle regardless of the amount of the fast fiber component (Crow 1983). Avian cardiac muscle only expresses the MLC1_v isoform both in the developing heart (5 Day) as well as in the adult, however Obinata et al. (1983) have reported an MLC1-like isoform (i.e. the fetal isoform), distinct from MLC1_v, in chicken cardiac tissue which is present in trace amounts before day 10 in ovo. Avian gizzard muscle (smooth muscle tissue) has been examined during development and the fetal isoform present in other
muscle types is also present in the developing gizzard (Katoh & Kubo 1978, Takano-Ohmuro 1985).

2. Disease related effects. Different physiological factors can affect the expression, the accumulation and/or the abundance of contractile protein isoforms within the various muscle types. Thyroid hormone levels influence the expression of $MLC_{1F}$ and $MLC_{3F}$ in fast twitch skeletal muscle. Salviati et. al. (1985) observed that the $MLC_{1Sb}$ isoform predominates in the biceps brachialis muscle (fast twitch) in hypothyroid humans, which under normal hormone conditions is known to contain the $MLC_{1F}$ and $MLC_{3F}$ isoforms. Thyroidectomy experiments in rats have shown that the soleus muscle, which normally is of mixed fiber type and expresses both fast and slow twitch $MLC$-ALK isoforms, predominantly expresses the $MLC_{1s}$ isoform (Johnson et.al. 1980). Hypertrophy in human atrial muscle, resulting from coronary artery disease and consequent pressure overload, exhibits a shift in $MLC$-ALK expression from the $MLC_{1A}$ isoform to the $MLC_{1V}$ isoform (Cummins et.al. 1982). Muscular dystrophy in vertebrates also results in altered $MLC$-ALK expression. Dystrophic chicken skeletal muscle exhibits a $MLC_{1F}/MLC_{3F}$ to $MLC_{1S}$ shift in normally fast twitch muscle (Bandman 1984). Reduced levels of $MLC_{3F}$ in the gastrocnemius muscle, predominant expression of $MLC_{1Sb}$ in normally mixed fiber muscles (e.g. soleus) and traces of $MLC_{1Sb}$ in normally pure fast twitch muscle have all been detected in dystrophic mice (Fitzsimmons & Hon 1983, John 1974). Duchenne muscular dystrophy
in humans is characterized by the presence of fetal myosin, which contains the MLC_{emb} isoform, in mature muscle (Margrath et al. 1984). Human nemaline myopathy also exhibits altered patterns of MLC-ALK expression; for example, in the gastrocnemius muscle (mixed fiber type) of affected individuals, MLC_{Sb} is the only isoform observed (Volpe et al. 1982).

3. Manipulated Muscle; Innervation Studies. The type of innervation that a particular muscle receives has a direct influence on the expression of myosin isoforms. Slow twitch skeletal muscle appears to be more dependant on nerve contact for the maintenance of its isoforms than fast twitch skeletal muscle. Denervation of chicken fast twitch muscle (pectoralis) does result in an altered MLC_{F}/MLC3_{F} ratio (Matsuda et al. 1984), however, denervation of developing and adult rat soleus muscle (slow twitch) resulted in the partial loss of MLC_{Sb} and consequent replacement with MLC1_{F} and MLC3_{F} (Rubinstein and Kelly 1978). A change in MLC-ALK expression, from fast to slow, can also be induced. In experiments where chronic low frequency stimulation is supplied to rabbit fast twitch muscle a change in phenotype is observed (Roy et al. 1979) and a corresponding increase in MLC_{Sb} mRNA is detected (Heilig & Pette 1983). Cross-innervation experiments also reveal a change in MLC-ALK expression. When the extensor digitorum longus (fast) and soleus muscle (slow) nerves from rats are surgically dissected and used to cross-innervate the respective muscles, a change in alkali light chain phenotype is observed (Sreter et al. 1974, Weeds et al.)
1974, Gauthier et.al. 1983). The conversion of fast to slow is complete, but only partial conversion from slow to fast is observed (Gauthier et.al. 1983).

4. Manipulated Muscle; Tissue Culture Analyses. Cultured mammalian skeletal myoblasts can be isolated and grown as monolayers of mononucleated cells, which will fuse and form multinucleated fibers. Upon formation of the fiber, muscle isoforms of contractile proteins accumulate in sarcomeric structures (Buckingham 1985). Analysis of MLC-ALK expression in a variety muscle cell lines indicates variation in the MLC-ALK isoforms which accumulate. In primary cultures of bovine skeletal muscle the MLC1<sub>emb</sub>, MLC1<sub>F</sub>, and MLC1<sub>Sa</sub> isoforms are expressed (Daubas 1981, Whalen 1982). Mouse cell lines, C2 and T984-C1 10, express MLC1<sub>emb</sub>, MLC1<sub>F</sub>, and MLC3<sub>F</sub> (Caravatti 1982), in addition the C2 line also expresses MLC1<sub>Sa</sub> at low levels (Blau 1983).

Differentiated cultures of rat L6 and L8 myogenic cell lines only express the MLC1<sub>emb</sub> isoform (Whalen et.al. 1978, Garfinkel et.al. 1982, Periasamy et.al. 1984, and Wade et.al. 1988). A human rhabdomyosarcoma has also been shown to express both MLC1<sub>F</sub> and MLC3<sub>F</sub> in approximately equimolar ratios, but no expression of the MLC1<sub>Sb</sub> was detected (Wade 1988). Human differentiated primary myoblast cultures provides an interesting deviation from what is observed in development and other mammalian cell cultures. The only MLC-ALK isoform detected in this cell culture, as determined by northern blot analysis, is the MLC3<sub>F</sub> isoform (Wade et.al. 1988).
This is an interesting result given that it is the only case where MLC3p is the only isoform expressed. MLC3p in all other tissues and cell cultures examined is expressed after MLC1f has accumulated.

Avian cell cultures exhibit an MLC-ALK pattern of expression that mimics muscle development; co-accumulation of MLC1s and MLC1f (Keller and Emerson 1980, Stockdale et.al. 1981, Monterras & Fiszman 1983, Matsuda et.al. 1983). Clonal analysis of myoblasts from different muscles at different ages shows that all cells express MLC1s and MLC1f with variable low levels of MLC3p (Keller and Emerson 1980).

F. Chromosomal Distribution of MLC-ALK Genes

Localization of the MLC-ALK genes in mouse (see Table 2) have been determined using cDNA clones corresponding to MLC1emb/MLC1A, MLC1b/MLC1v, MLC1f/MLC3f and the MLC1p/MLC3p pseudogene. These genes have been mapped with respect to the segregation pattern of restriction fragment length polymorphisms (RFLPs) in comparison with the segregation pattern of a series of chromosomal markers observed between the mouse species Mus spreus and Mus musculus. An F1 backcross between these species has shown that all of the members of the MLC-ALK gene family are unlinked and are represented on different chromosomes (Robert et.al. 1985, Barton et.al. 1985, Robert et.al 1984). The following chromosomal assignments are
Table 2: Chromosominal Distribution of Vertebrate Myosin Alkali Light Chain Genes

<table>
<thead>
<tr>
<th>MLC-ALK Genes</th>
<th>Chromosome #</th>
</tr>
</thead>
<tbody>
<tr>
<td>mouse</td>
<td></td>
</tr>
<tr>
<td>MLC1&lt;sub&gt;1A/emb&lt;/sub&gt;</td>
<td>11</td>
</tr>
<tr>
<td>MLC1&lt;sub&gt;F/3F&lt;/sub&gt;</td>
<td>1</td>
</tr>
<tr>
<td>MLC1&lt;sub&gt;V/S&lt;/sub&gt;</td>
<td>9</td>
</tr>
<tr>
<td>MLC pseudogene</td>
<td>12</td>
</tr>
<tr>
<td>human</td>
<td></td>
</tr>
<tr>
<td>MLC1&lt;sub&gt;1A/emb&lt;/sub&gt;</td>
<td>17</td>
</tr>
<tr>
<td>MLC1&lt;sub&gt;F/3F&lt;/sub&gt;</td>
<td>2</td>
</tr>
<tr>
<td>MLC1&lt;sub&gt;V/Sb&lt;/sub&gt;</td>
<td>3</td>
</tr>
</tbody>
</table>
given to the mouse MLC-ALK gene family: MLC1_p/MLC3_p, chromosome 1; MLC1_emb/MLC1_A, chromosome 11; MLC1_Sb/MLC1_v, chromosome 9; and pseudo-MLC1/3 chromosome 12.

Chromosomal assignments have been given to 3 human myosin alkali light chain genes (see Table 2); the fast skeletal MLC-ALK gene, MLC1_F/MLC3_F, the atrial/fetal muscle gene (MLC1_A) and the ventricular/slow skeletal muscle gene (MLC1_v). The MLC1_F/MLC3_F gene was localized using the mouse processed pseudogene as a probe against genomic DNA from a panel of 25 independent human-rodent somatic cell hybrids. The results of these experiments localized this gene to chromosome 2 (Serero et.al., 1987). The MLC1_A gene was localized using a mouse MLC1_A cDNA clone as a probe against a similar panel of human-rodent somatic cell hybrids and shown to reside on chromosome 17 (Cohen-Haguenauer et.al, 1987). The MLC1_v gene was mapped as described above using a human MLC1_v cDNA clone as the hybridization probe. This gene was localized to chromosome 3 (ibid, also see Chapter IV: Isolation and Characterization of hMLC1_v/Sb).

The Drosophila MLC-ALK gene has been localized, utilizing in situ hybridization to salivary gland chromosomes, to the 98B region of chromosome 3 (Falkenthal et.al. 1984).
G. Review Summary

The myosin alkali light chain is conserved throughout eukaryotic evolution and is essential for contractile processes. The function of the protein is still undetermined although its position in the functional myosin hexamer has been localized to the myosin heavy chain head-tail junction. In the vertebrate species so far examined, the protein is represented by several isoforms which are encoded by a multigene family. The various isoforms are developmentally regulated and exhibit tissue specificity. Nucleic acid analysis of the genes which encode these isoforms has led to the discovery that more than one isoform may be encoded by a single gene. In addition, a single isoform may be expressed in two different tissue types.
CHAPTER III
GENERAL EXPERIMENTAL PROCEDURES

A. DNA Preparation. Human genomic DNA was prepared from fresh placenta and rat genomic DNA was prepared from fresh liver, as described previously (Blattner et al. 1978).

Recombinant bacteriophage DNA was prepared from primary lysates by the rapid isolation of bacteriophage lambda DNA protocol (Maniatis et al., 1982).

Plasmid DNA was prepared by the alkaline lysis procedure ( ), and purified in CsCl gradients.

B. Restriction Enzyme Analysis. Genomic, bacteriophage, and plasmid DNAs were digested with various restriction enzymes using the recommended reaction conditions provided by the suppliers (Bethesda Research Labs, BRL; International Biotechnologies Inc., IBI). Restriction enzyme digestion products were electrophoresed on agarose and/or polyacrylamide gels, stained with ethidium bromide (1ug/ml) and photographed under UV illumination.

Restriction fragment sizes were determined by the distance migrated in the gel with respect to DNA fragments of known size plotted on semi-logarithmic graph paper. Restriction maps were constructed based on single and double digest restriction fragment patterns.
C. Restriction Fragment Isolation. Isolated restriction fragments were prepared by digesting the DNA with the appropriate restriction enzyme(s) and electrophoresing the products on a 5% acrylamide gel. The gel was stained with ethidium bromide (1ug/ml final concentration) and the desired restriction fragment was cut from the gel while under ultraviolet illumination. The gel slice was then incubated overnight at 37° C in a sealed 1 ml. pipet tip with 400ul of Elution Solution (500mM NH₄OAc + 10mM MgOAc + 1mM EDTA + 1% SDS). The eluted DNA was recovered by centrifuging the Elution solution through siliconized glass wool and the DNA was subsequently precipitated by adding 2 volumes of 95% ethanol. Precipitated DNA was recovered by centrifugation, washed with 70% ethanol and resuspended in 1X TE buffer (10mM Tris-HCl + 1mM EDTA).

D. Plasmid Subclones. Various subclones were generated from isolated bacteriophage clones by ligating desired restriction fragments into plasmid vectors. Plasmid DNA was prepared for ligation by digesting it with the appropriate restriction enzyme(s) and subsequently dephosphorylating it with CIAP (calf intestinal alkaline phosphotase) to prevent it from religation. Typical dephosphorylation reaction conditions were carried out using the recommended reaction conditions provided by the supplier (Beoringer Manheim Biochemicals). Ligation reactions were typically carried out in 30ul reaction volumes with bacteriophage T₄ ligase using the reaction buffer and conditions provided by the supplier (BRL or IBI). Equimolar ratios of plasmid and insert DNA were calculated
by weight and subsequently used in the reaction to generate the desired subclone. Ligation reaction products were then used to transform an appropriate bacterial host using the CaCl transformation procedure (Mandel and Higa 1970).

E. Oligonucleotides. The two oligonucleotides that were used as exon specific probes (see Chapter V; Results, section 4c.) were synthesized on an Applied Biosystems Model 380B at the Biochemical Instrument Center (College Biological Sciences, OSU) using alpha-cyanoethylphosphoramidite chemistry.

F. Probe Synthesis. All restriction fragments used as probes were radiolabeled with alpha-\(^{32}\)P-dCTP (3000Ci/mmol) from Amersham Corp. using the oligo-labeling procedure (Feinberg and Vogelstein, 1983). Oligonucleotide probes were radiolabeled with gamma-\(^{32}\)P-ATP (800Ci/mmol, ICN) using Polynucleotide Kinase and the recommended reaction conditions provided by the supplier (Bethesda Research Laboratories). Oligonucleotide probes were also radiolabeled using the oligo-labeling procedure (ibid). Radiolabeled oligonucleotide probes were purified on Nensorb columns before being used as hybridization probes.

G. Genomic Southern Analysis. Human genomic DNA was digested with EcoRI, BamHI, HindIII, and XbaI (where indicated on individual southern blots) rat genomic DNA was digested with EcoRI. Fragments were electrophoresed on 0.8% or 1% agarose, stained, photographed and blotted to nitrocellulose (Southern 1975). Heterologous hybridizations were performed by incubating a P-32 labeled 431bp
EcoRI fragment from the rat cDNA clone (Garfinkel et al. 1982) in 6X SSC + Denhardt's (1X SSC = 0.3 M sodium chloride, 0.03 M sodium citrate, pH : Denhardt's solution = 0.02% ficol, 0.02% PVP, 0.02% BSA, 0.5% SDS) at 55°C overnight. Filters were then washed under low stringency, 3X SSC + 0.5% SDS at 65°C. Homologous hybridizations were performed as above, with the exception of incubation temperature; 68°C, and stringency washes; 0.1X SSC + 0.5% SDS.

H. Genomic Library Screening. A partial EcoRI human genomic library cloned into lambda Charon 4A and a partial Sau3AI human genomic library cloned into lambda Charon 28 (Slightom et al., 1980) were screened as described by Blattner et al. (1978; except that large 25x25 cm plates, Gibco Laboratories, were used for the initial screenings), using the 431bp rat cDNA fragment as probe and applying the hybridization conditions determined for heterologous genomic hybridizations. The 3' walk experiment was performed by screening the human partial Sau3AI genomic library as described above except a 1.0 kbp XbaI-EcoRI restriction fragment from the 3' end of hMLC.1 was used as the probe.

I. Exon Identification and DNA Sequence Analysis.
Heterologous exons, corresponding to the common coding region, were isolated from the rat skeletal cDNA using convenient restriction sites (see Figure 6). Hybridizations to restricted hMLC.1 DNA with specific exon restriction fragments were performed under low stringency as described above. Oligonucleotide probes,
corresponding to the 5' coding region, were used to identify exons 1 and 2. The nucleotide sequences for both oligonucleotides were determined by predicting the appropriate bases, including redundancies, to encode polypeptides that corresponded to a known MLC-ALK protein sequence (see Results and Discussion for a detailed explanation). To identify exons 1 and 2, the appropriate oligonucleotide probe was used in low stringency hybridization experiments: 37° C in 6X SSC + Denhardt's for 16 hours, washed with 3X SSC at 42° C for 30 minutes. DNA sequence analysis was performed using the method of Maxam and Gilbert (1980) with the modifications of Slightom et al. (1980) and the Sanger dideoxy method (1977) employing the Bluescript phagemid vector (Stratagene) in conjunction with the Sequenase system (United States Biochemicals).
CHAPTER IV
HETERLOGOUS GENOMIC SOUTHERN ANALYSIS

Results

Genomic Southern Analysis. The first step in the analysis of human MLC-ALK genes was to test the rat MLC-ALK cDNA clone (pMLC-84 isolated from rat skeletal muscle L6E9 myotubes; Garfinkel et al. 1982) for its ability to identify human MLC-ALK gene sequences. This test involved isolating a 431bp EcoRI fragment (MLC-84/RI0.431) which contains E-F hand domain 3 and represents the common coding region (exons 3-6, see Figure 6) shared by the MLC1 and MLC3 isoforms. This restriction fragment was isolated from the plasmid clone to avoid a repeat element that was shown to exist in the 3' untranslated region (Periasamy et al., 1984). The 431bp EcoRI fragment was band isolated, labeled with P-32 and used as a probe in a human genomic southern blot experiment. Several low stringency hybridization and wash conditions were used to maximize the extent of hybridization, however only the experimental conditions described in Figure 7 proved to be useful. As an internal control, rat genomic DNA was included in the analysis. The 431bp rat cDNA fragment hybridizes to a number of bands in both the rat and human digests (sizes are indicated in kbp). In the rat EcoRI digest, 5 bands are present. The strongly hybridizing 2.0
EcoRI
\[ 5(3) \]

\[
\text{GAATTCAAGGAGGCATTTCTTCTTGACAGACGACCCGCTGGCTGCTGATGTCCGCTGAAGACCCCTAGGGAATCGAGGCTGCTGTGGGCACCAATCCCACAAATGCAGAAGTCAAGAAG} \\
\text{GlnValGlyAspValLeuArgAlaLeuGlyThrAsnProThrAlaAsnGluValLysLys}
\]

\[
\text{AvaI} \quad 5(3) \quad 6(4)
\]

\[
\text{GTTCTCGGGAACCTAGCAATGAAGATGAATGCTAAGAAAATCGAGTGTGAAACATGGTT}
\]

\[
\text{ValLeuGlyAsnProSerAsnGluGluMetAsnAlaLysLysIleGluPheGluGlnPhe}
\]

\[
\text{CTGCCCATGATGGAACCTCAACACCAACCAGCAGGAGGTGCTGATGCTGCTGCTGCTGCT}
\]

\[
\text{LeuProMetMetGinAlaIleSerAsnAsnLysLysIleGlyTyrGluAspPheVal}
\]

\[
\text{GAGGATCTCCTGCGCTCTGCAAGAGCCAGGGCATGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCG}
\]

\[
\text{GluGlyLeuArgValPheAspLysGluGlyAsnGlyThrValMetGlyAlaGluLeuArg}
\]

\[
\text{6(4) \quad 7(5)}
\]

\[
\text{CATGCTCGCCACTTGGAGAGAGATGAAGAGGAGGAGGTAGAGCATTGTGGCG}
\]

\[
\text{HisValLeuAlaThrLeuGlyGluLysMetLysGluGluValGlAlaLeuLeuAla}
\]

\[
\text{HindIII} \quad 7(5) \quad 8(6)
\]

\[
\text{GGCCAGAGGACTTGAATGGCTGATCAACTGATGACGGCTGCAAGAGCAGCATCTGCT}
\]

\[
\text{GlyGlnGluAspSerAsnGlyCysIleAsnTyrGluAlaPheValLysHisIleMetSer}
\]

\[
\text{EcoRI}
\]

\[
\text{GTCTAAACGAGAATTC}
\]

\[
\text{Val}^{* * *}
\]

Figure 6. Nucleotide sequence of the rat MLC-ALK cDNA fragment to be used as a heterologous gene probe. Restriction enzyme sites used to isolate exon specific fragments are indicated. Exons are numbered with the numbers in parentheses indicating the numbers used to number the mouse exons and also correspond to the exon numbers in the human gene, hMLC1v/Sb.
Figure 7. Autoradiogram results of the southern blot hybridization experiment. Restriction enzyme digested human genomic DNA was probed with the rat MLC-ALK cDNA fragment, pMLC-84/RI0.431. The filter was hybridized with P-32 labeled pMLC-84/RI0.431 (2 X 10^6 cpm/ml) at 55°C in 6X SSC + Denhardt's solution for 16hrs. The filter was then washed in 3X SSC + 0.5% SDS at 65°C for 2 hrs. Sizes of hybridizing restriction fragments are indicated in kbp.
kbp band corresponds to the genomic fragment that contains exons 4-6. Exon 3 is contained on the 0.5 kbp hybridizing band. In addition a number of other bands hybridize to the probe and probably represent the other members of the rat MLC-ALK gene family. Several bands of varying intensity are evident in both human digests, 7 bands in the BamHI digest and 4 bands in the EcoRI digest. Equation 1 was used to determine the extent of homology that the above hybridization experiment would detect.

\[
\text{Equation 1: } T_M = 69.3 + 0.41(G+C\%) - \frac{650}{L}
\]

Where \( T_M \) equals the temperature at which the double stranded DNA molecule will dissociate into single strands, \( G+C\% \) equals the percentage of guanine and cytosine residues in the DNA, and \( L \) equals the length of the DNA in question (Marmur and Doty, 1962; Wetmer and Davidson, 1968). Also, the percent mismatch between any two hybrid DNA molecules has been determined to reduce the \( T_M \) of that hybrid molecule by \( 1^\circ \) C for every 1% mismatch (Bonner et.al., 1973). Therefore employing Equation 1, it was determined that the \( T_M \) of the probe equaled \( 89^\circ \) C (where \( G+C\% = 51; \ L = 431 \)) and that the wash condition (i.e. \( 65^\circ \) C) used in the hybridization experiment presented in Figure 7 was \( 24^\circ \) C below the \( T_M \). These conditions would then theoretically identify sequences which are
76% homologous or greater to the rat cDNA probe. The result of this experiment indicated the utility of the rat cDNA fragment for isolating heterologous MLC-ALK sequences. The above conditions were then applied to screening two different human genomic libraries, a partial EcoRI library cloned into lambda Charon 4A and a partial Sau 3AI library cloned into lambda Charon 28.
CHAPTER V
ISOLATION AND CHARACTERIZATION OF THE HUMAN hMLC1V/Sb GENE

A. Results

1. Isolation and Preliminary Analysis of hMLC1. A human partial EcoRI library, cloned into Charon 4A, was screened with the rat cDNA probe (pMLC-84/RI0.431) using the conditions determined for the genomic southern (i.e. 5X SSC + Denhardt's, incubated at 55° C for 16 hours and washed with 3X SSC at 55° C). Employing these hybridization conditions it was expected that clones which contained at least 70% homology to the probe would be to be able to be isolated. The human partial EcoRI library was plated on 4 megaplates (Gibco Laboratories) at a density of 150,000 pfu per plate. Plating at this density and making the assumption that 15kbp represented the average size of the DNA ligated into the library, it was calculated, using Equation 2, that the total number of phage represented $9 \times 10^9$ bp of human DNA. This represents approximately a 3n compliment of the human genome, given the current estimate of $3 \times 10^6$ kbp of DNA/haploid cell.

Equation 2

$($# pfu/plate$)(15.0$kbp/pfu$)(4$ plates$) = $Total DNA (in kbp)$
Initially, 72 putative positive clones were isolated from the phage megaplates. Each of these was then individually rescreened for hybridization to pMLC-84/RI0.431. From these initial clones, 5 consistently continued to hybridize and were subsequently plaque purified. DNA from each of the clones was isolated (as described in GENERAL EXPERIMENTAL PROCEDURES) and analyzed with various restriction enzymes. Based on initial restriction analysis and southern blot hybridization data, all 5 clones were determined to be identical. This isolate was designated hMLC.1 and was shown to contain a 15kbp EcoRI fragment. Initial restriction enzyme analysis and southern blot hybridization data with pMLC-84/RI0.431 identified 2 hybridizing regions of different intensity, see Figure 8. The strongly hybridizing band mapped to an internal 1.1kbp BamHI fragment (designated hMLC.1-B11.1). The weaker signal mapped to a 1.0kbp XbaI-EcoRI fragment (designated hMLC.1-XRI1.0). The 15kbp EcoRI fragment was subsequently subcloned into pAT153 and designated phMLC.1. The plasmid clone was then analyzed with a variety of restriction enzymes and used to construct the restriction map given in Figure 9, where the shaded regions represent the regions of hybridization. The initial observation from this preliminary analysis is that the 2 regions of hybridization are separated by at least 500 bp and probably represent 2 adjacent exons in pMLC-84/RI0.431. A more detailed discussion of exon localization is presented in a following section (Gene Structure Analysis of hMLC.1).
<table>
<thead>
<tr>
<th>Lane</th>
<th>Enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A. AvaI</td>
</tr>
<tr>
<td>B</td>
<td>B. AvaI-EcoRI</td>
</tr>
<tr>
<td>C</td>
<td>C. BamHI</td>
</tr>
<tr>
<td>D</td>
<td>D. BamHI-EcoRI</td>
</tr>
<tr>
<td>E</td>
<td>E. lambda marker</td>
</tr>
<tr>
<td>F</td>
<td>F. EcoRI</td>
</tr>
<tr>
<td>G</td>
<td>G. HindIII</td>
</tr>
<tr>
<td>H</td>
<td>H. HindIII-EcoRI</td>
</tr>
<tr>
<td>I</td>
<td>I. HpaI</td>
</tr>
<tr>
<td>J</td>
<td>J. HpaI-EcoRI</td>
</tr>
<tr>
<td>K</td>
<td>K. lambda marker</td>
</tr>
<tr>
<td>L</td>
<td>L. KpnI</td>
</tr>
<tr>
<td>M</td>
<td>M. KpnI-EcoRI</td>
</tr>
<tr>
<td>N</td>
<td>N. SacI</td>
</tr>
<tr>
<td>O</td>
<td>O. SacI-EcoRI</td>
</tr>
<tr>
<td>P</td>
<td>P. XbaI</td>
</tr>
<tr>
<td>Q</td>
<td>Q. XbaI-EcoRI</td>
</tr>
<tr>
<td>R</td>
<td>R. EcoRI</td>
</tr>
<tr>
<td>S</td>
<td>S. lambda marker</td>
</tr>
</tbody>
</table>

Figure 8. Autoradiogram results of the southern blot hybridization experiment with restricted hMLC.1 DNA probed with pMLC-84/RI0.431. Restriction enzymes used to digest hMLC.1 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled pMLC-84/RI0.431 (2 X 10^5 cpm/ml) at 55° C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 3X SSC at 65° C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80° C.
Figure 9. Restriction map of hMLC.1. Restriction enzymes are indicated above the horizontal bar. Shaded regions indicate the regions of homology to pMLC-84/R10.431.
2. Genomic Southern Analysis. Genomic southern blot analysis with specific restriction fragments from hMLC.1 was performed for three reasons. First, in the previous section a discussion of the two regions in hMLC.1 which hybridized to the rat cDNA clone was presented and based on restriction analysis it was determined that one region of hybridization (hMLC.1-XRI1.0) represented the 3' most region of the 15kbp insert. It was concluded, at this time in the study, that hMLC.1 may not contain the entire gene (remember that the rat cDNA fragment used as the probe only contains sequences representing the COOH terminal coding region common to the two skeletal isoforms and does not contain the 3' untranslated region). Consequently the genomic DNA 3' to hMLC.1 should be isolated (i.e. a 3' walk experiment should be performed) to ensure that the entire gene was available for further analysis. Performing a walk experiment involves isolating a single copy restriction fragment from the end of hMLC.1 for use as a hybridization probe in a subsequent human genomic library screening. Secondly, in order for Dr. Francke's laboratory to localize this gene to its respective human chromosome, a single copy restriction fragment had to be isolated from hMLC.1 so that it could be used as a probe in genomic southern blot hybridization experiments against a panel of rodent x human somatic cell hybrid DNAs. Finally, the fragment that exhibits the greatest homology to the rat cDNA clone and produces the more intense signal seen in the autoradiogram presented in Figure 8 (i.e. hMLC.1-BII1.1), probably represents the sequences
responsible for encoding E-F hand domain 3, the most conserved domain between all MLC-ALK proteins. It was concluded at this time that this fragment may also be useful as a probe for isolating other human MLC-ALK genes from human genomic libraries under low stringency hybridization conditions.

The first step in identifying single copy DNA in hMLC.1 involved identifying repetitive DNA in the 15kbp insert. Analyzing the DNA for the presence of repetitive elements was performed as follows: restricted hMLC.1 phage DNA was probed with P-32 labeled human genomic DNA, an AluI family repeat clone and 2 KpnI repeat clones, each representing half of the KpnI repeat element. These repeat element clones were isolated from the human beta-globin gene cluster in the region between the epsilon-globin gene and the gamma-globin gene. The AluI and KpnI repeat elements are also referred to as SINEs (Short INTerspersed repeat Elements) and LINEs (Long INTerspersed repeat Elements), respectively and have been shown to be dispersed throughout the human genome (reviewed by Weiner et al., 1986). The KpnI repeat element has also been shown to exist in the genome in one of two states, as a complete element or as a partial element where half of the sequences representing the repeat are not present (ibid). Therefore it was necessary to analyze hMLC.1 with both halves of the KpnI repeat element in order to accurately assess the presence of these sequences in hMLC.1. The results of these analyses are presented in Figures 10 and 11. A single hybridizing region was identified with both the genomic
Figure 10. Autoradiogram results of the southern blot hybridization experiment with restriction enzyme digested hMLC.1 DNA probed with human genomic DNA. Restriction enzymes used to digest hMLC.1 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled human genomic DNA that was first digested with EcoRI (1 X 10^3 cpm/ml) at 65° C for 16hrs in 6X - SSC + Denhardt solution. The filter was subsequently washed with 1X SSC at 68° C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80° C.
<table>
<thead>
<tr>
<th>Lane</th>
<th>Restriction Enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. KpnI</td>
<td></td>
</tr>
<tr>
<td>B. KpnI-SmaI</td>
<td></td>
</tr>
<tr>
<td>C. SmaI</td>
<td></td>
</tr>
<tr>
<td>D. lambda marker</td>
<td></td>
</tr>
<tr>
<td>E. SacI-XhoI</td>
<td></td>
</tr>
<tr>
<td>F. SacI-XbaI</td>
<td></td>
</tr>
<tr>
<td>G. SacI</td>
<td></td>
</tr>
<tr>
<td>H. lambda marker</td>
<td></td>
</tr>
<tr>
<td>I. BamHI-EcoRI</td>
<td></td>
</tr>
<tr>
<td>J. BamHI</td>
<td></td>
</tr>
</tbody>
</table>

Figure 11. Autoradiogram results of the southern blot hybridization experiment with restricted hMLC.1 DNA probed with an AluI family repeat clone. Restriction enzymes used to digest hMLC.1 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled AluI DNA (1 X 10^5 cpm/ml) at 65° C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 1X SSC at 68° C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80° C.
DNA probe and with the AluI probe (Figures 10 and 11, respectively), no hybridization was observed with either of the KpnI clones (data not shown). The region of the repeat element is indicated by the horizontal solid bar in Figure 9 and was localized to the DNA between two SacI sites. The intensity of hybridization in this region was weak when compared to similar blots of other phage clones with known AluI repeat elements (compare Figures 10 and 11 to Figures 46, 55, and 56). and probably represents a evolutionarily divergent copy of the AluI repeat. From these data, hMLC.1 was determined to be almost entirely single copy, however it is important to realize that this analysis only identifies repetative DNA that is present in the genome in approximately $10^4$ copies (estimates on the copy number of the AluI and KpnI repeat elements are $10^5$ and $10^4$, respectively) and does not detect repetative DNA present in fewer copies, i.e. low level repeats. Therefore it is also necessary to analyze any fragment that is suspected to be single copy by using it as a probe in a genomic southern blot experiment. The fragment, hMLC.1-XRI1.0, represented a fragment that was suspected to be unique DNA in the genome and also represented the 3' most fragment from hMLC.1, thus fitting the criteria required for a 3' walk experiment. Therefore this fragment was band isolated, labeled with P-32 and used as a probe in a genomic southern blot hybridization experiment. The results of that experiment are given in Figure 12. The 1.0kbp XbaI-EcoRI hybridizes to the expected 15kbp EcoRI fragment and to a 6.5 kbp
Figure 12. Autoradiogram results of the southern blot hybridization experiment with restricted human genomic DNA probed with XRI1.0. Restriction enzymes used to digest genomic DNA are indicated above the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled XRI1.0 (2 x 10^6 cpm/ml) at 65°C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 1X SSC at 68°C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80°C.
BamHI fragment. Interestingly, a 6.5kbp fragment is also observed in the BamHI digest in Figure 7, the heterologous genomic southern. The results of the above experiments proved that the 1.0kbp XbaI-EcoRI fragment would be useful for both the chromosomal localization experiments and for the 3' walk experiment, it was therefore subcloned into the plasmid vector, pUC19 (as described in GENERAL EXPERIMENTAL PROCEDURES) for future experiments.

The fragment, hMLC.1-BI1.1, was isolated and also used to probe a human genomic southern blot. The rationale behind this experiment was to test it as a probe for identifying other human MLC-ALK gene sequences (as previously discussed). The results of this hybridization experiment are given in Figure 13. The 1.1 BamHI fragment under high stringency wash conditions exhibited hybridization to the 15kbp EcoRI fragment and to a 1.0kbp BamHI fragment, as expected. However, low stringency hybridization and wash conditions resulted in a smear (data not shown), and therefore in its entirety, hMLC.1-BI1.1, would not be useful as an additional gene probe. The BamHI 1.1 probe required additional washings at increased stringency, (see figure legend for conditions) presumably due to homology to the AluI family sequence as indicated on Figure 9.

3. Isolation of hMLC.2. A 3' walk experiment was performed to ensure that the DNA encoding the entire gene was available for analysis. From the genomic southern data obtained in Figure 9, it was known that hMLC.1-XRI1.0 is single copy and that it identifies
Figure 13. Autoradiogram results of the southern blot hybridization experiment with restricted human genomic DNA probed with BI1.1. Restriction enzymes used to digest genomic DNA are indicated above the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled BI1.1 DNA (2 X 10^6 cpm/ml) at 65°C for 16 hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 0.1X SSC at 68°C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80°C.
Therefore a human partial Sau3A1 library cloned into lambda Charon 28 at the BamHI sites, was screened using the hMLC.1-XRI1.0 fragment as a probe. An initial 64 clones were isolated, and from these 64, 2 clones continued to consistently hybridize through plaque purification. Initial restriction analysis and southern blot data (see Figure 14) indicated that these 2 clones were identical. The clones were pooled and designated hMLC.2. A partial restriction map of the entire region is shown in Figure 15. The region of overlap between hMLC.1 and hMLC.2 extends 3' from the 3' most BamHI site in hMLC.1 to the EcoRI site at the 3' end of hMLC.1. Therefore hMLC.2 also contains the entire 1.0kbp XbaI-EcoRI as well as approximately 0.5kbp of DNA 5' to the XbaI site. The additional DNA present in hMLC.2 extends from the EcoRI site 3' for approximately 8.0kbp.

4. Gene Structure Analysis of hMLC1 вспи. The following section is divided into four parts (a.- d.). The first part describes the orientation of the gene in the DNA represented by hMLC.1 and hMLC.2. The second part describes how exons 3-6 were identified, localized, and sequenced. The third part describes how oligonucleotide probes were used to identify and localize exons 1 and 2. Part 3 also presents the DNA sequence data from exons 1 and 2. The fourth part of this section describes how exon 7 was identified, localized and sequenced. It is also important to mention that the numbering scheme used in describing the exons of
### Figure 14. Autoradiogram results of the southern blot hybridization experiment with restricted clones isolated from the 3' walk experiment probed with pMLC-4/RI0.431. Lanes B, C, and D correspond to clone 2.7; lanes E, F, and G correspond to clone 2.10. Restriction enzymes used to digest the bacteriophage DNA are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled XRI1.0 DNA (1 X 10^5 cpm/ml) at 65° C for 16hrs in 6X SSC + Denhardt solution. The filter was subsequently washed with 1X SSC at 68° C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80° C.

<table>
<thead>
<tr>
<th>Lane</th>
<th>Digestion Product</th>
<th>Size (kb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. lambda marker</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. BamHI</td>
<td></td>
<td>21</td>
</tr>
<tr>
<td>C. BamHI-EcoRI</td>
<td></td>
<td>13.7</td>
</tr>
<tr>
<td>D. BamHI-XbaI</td>
<td></td>
<td>9.5</td>
</tr>
<tr>
<td>E. BamHI</td>
<td></td>
<td>7.5</td>
</tr>
<tr>
<td>F. BamHI-EcoRI</td>
<td></td>
<td>6.7</td>
</tr>
<tr>
<td>G. BamHI-XbaI</td>
<td></td>
<td>5.7</td>
</tr>
<tr>
<td>H. lambda marker</td>
<td></td>
<td>4.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.1</td>
</tr>
</tbody>
</table>
Figure 15. Partial restriction map of the DNA contained in the two clones, hMLC.1 and hMLC.2. The solid horizontal bars below the map represent the respective DNA contained in hMLC.1 and hMLC.2. The region in hMLC.1 that hybridized to the AluI clone is indicated by the horizontal hatched bar above the map. The regions of homology to pMLC-84 are indicated by the shaded regions. Restriction enzymes that were used in mapping both clones are extended from the map, all other enzymes sites were only determined for hMLC.1. The region of the genomic repeat element is indicated by the solid horizontal bar above the map, labeled AluI.
hMLC1\textsubscript{v/Sb} (i.e. a total of 7 exons were discovered in hMLC1\textsubscript{v/Sb}) is the same as that used for the mouse MLC1/3\textsubscript{p} gene, however the rat cDNA maintains the numbering system that the authors originally defined (Periasamy et.al., 1984). Please keep this in mind and hopefully it will limit any confusion.

a. During the 3' walk experiment, initial DNA sequence analysis was also being performed on fragments, XRI1.0 and BI1.0, to try and determine the regions of homology to the rat cDNA clone. Based on initial sequence data extending from the XbaI site towards the EcoRI site, a region homologous to the 5' half of the rat exon 7 was identified. It was therefore concluded that the strong hybridization of pMLC-84/RI0.431 to hMLC.1-BI1.1 must be homology to the rat exon 6, the E-F hand coding sequence. The initial sequence data in conjunction with the hybridization data provided orientation to the clone with respect to the coding sequences and also indicated that exons 6 and 7 from pMLC-84 need not be separated to co-hybridize in low stringency hybridization experiments. It was also concluded at this time that hMLC.2 would be the logical clone to screen and would hopefully provide sufficient DNA in the search for exons homologous to rat exons 8 and 9.

b. A description of how exons 3-6 were identified is presented in this section and includes a more detailed discussion of the human exon 5. The discussion of each exon is presented in numerical order starting with exon 3 and proceeding through exon 6.
This does not represent the actual order of discovery but provides a more organized way of describing the results.

Restriction fragments from the rat cDNA clone that represented exons 5 and 8 were separately isolated from pMLC-84 using convenient restriction sites (see Figure 6) and were subsequently used as exon specific probes. Previous hybridization analysis and initial DNA sequence had determined the relative positions of exons 4 and 5 in hMLC.1, therefore these exons need only be further localized (remember that exon 4, i.e. rat exon 6 homology, was located on a 1.1kbp BamHI fragment and therefore should be localized further to assist DNA sequence analysis) and sequenced.

Homology to exon 5 from pMLC-84 was identified in hMLC.1 by a southern blot hybridization experiment using the rat exon 5 specific probe; pMLC-84/RIAI0.121 (see Figure 6). The result of that hybridization experiment is indicated in Figure 16. The homologous region was localized to a 1.2kbp BamHI fragment 5' to exons 4 and 5 (see Figure 17). Further localization of the homologous region was performed by generating a clone of the BI1.2 fragment and subsequently analyzing it for the presence of an exon. A BamHI 1.2 subclone (pBI1.2) was constructed in pUC19, prepared in large scale, and analyzed with a variety of restriction enzymes. The digested DNA was probed with pMLC-84/RIAI0.121 under low stringency hybridization conditions. Figure 18 is the result of that experiment. The homology to the rat exon 5 was localized to a 450bp HinfI fragment. In addition, compiling the restriction
Lane
A. lambda + Drosophila MLC-ALK marker  
B. HincII  
C. HincII-HindIII  
D. HindIII  
E. BamHI-HindIII  
F. BamHI  
G. marker  
H. BamHI  
I. BamHI-XbaI  
J. HindIII-XbaI  
K. EcoRI-XbaI  
L. SphI  
M. SphI-XbaI  
N. marker  

Figure 16. Autoradiogram results of the southern blot hybridization experiment. Restriction enzyme digested hMLC.1 DNA probed with pMLC-84/RIA10.121. Restriction enzymes used to digest hMLC.1 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled pMLC-84/RIA10.121 (2 X 10^5 cpm/ml) at 50°C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 3X SSC at 57°C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80°C.
Figure 17. Partial restriction map of hMLC.1 indicating the region exhibiting homology to exons 5, 6 and 7 from the rat cDNA clone. Enzyme sites are indicated as in previous figures.
Figure 18. Autoradiogram results of the southern blot hybridization experiment with restricted pBI1.2 (3'-5' clone) DNA probed with pMLC-84/RIAIO.121. Restriction enzymes used to digest pBI1.2 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.5% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled pMLC-84/RIAIO.121 (2 x 10^5 cpm/ml) at 50° C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 3X SSC at 57° C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80° C.
data and corresponding hybridizing fragments, this exon was positioned in the 250bp AvaI-AvaII fragment in the 3' half of pBI1.2 as indicated in Figure 19. Figure 19 is a partial restriction map determined for pBI1.2. The shaded region indicates exon 3. The actual size and DNA sequence of exon 3 was determined by sequencing the hybridizing region and then comparing this sequence to the rat exon 5. That sequence data is presented in Figure 20 along with the comparison to the rat exon. An interesting distinction between exon 3 from hMLC.1 and exon 5 from pMLC-84 is the presence of an apparent 2 amino acid insertion. Considering this insertion when aligning the two sequences, it was observed that the two sequences exhibit 65% homology and appear to be defined by similar boundaries. Consensus splice site junctions and putative intron branch point sequences are also indicated in Figure 20 and support the conclusion that this represents a functional exon. The significance of the identified insertion will be discussed in following sections (see Oligonucleotide Probes and Discussion).

Exon 4 was first identified in a 1.1kbp BamHI fragment, as previously discussed. Further localization of the actual exon was accomplished by generating a subclone of hMLC.1-BII.1 (pBI1.1) and subsequently analyzing it for the presence of exon 4. The BamHI 1.1 kbp fragment was ligated into pUC19 at the BamHI site located in the polylinker of the vector. The plasmid was prepared on a large scale, purified and digested with a variety of restriction enzymes.
Figure 19. Partial restriction map of pBI1.2 indicating the position of Exon 3. Enzyme sites are indicated and the shaded region represents the homology to exon 5 from pMLC-84.
NUCLEOTIDE HOMOLOGY TO pMLC-84 EXON 5 = 65%

Figure 20. DNA sequence and the predicted amino acid sequence for Exon 3. Putative coding sequences are represented by capital letters and intron sequences are indicated by lower case letters. Identical nucleotides between exon 3 and pMLC-84 are indicated by the vertical lines. The dashes in the pMLC-84 sequence indicate the position of the 2 amino acid insertion in the human sequence. The overall nucleotide homology is also given. Consensus dinucleotide splice sites are underlined and the vertical arrows indicate the exon-intron boundaries.
enzymes. The digestion products were electrophoresed on a 2.5% agarose gel blotted to nitrocellulose, and probed with the rat cDNA clone in a southern blot hybridization experiment. Figure 21 is the result of that experiment. A single region of hybridization was identified with pMLC-84/RI0.431 and determined to reside within a 430bp AvaII-Hinfl fragment. Strong hybridization is also observed with one or more DdeI fragments, but due to the complicated restriction pattern observed for the DdeI digest these fragments were not mapped. Figure 22 is a partial restriction map of pBI1.1. Homology to the rat exon 6, as determined by the experiment presented in Figure 21, is indicated by the horizontal hatched bar. The shaded region represents the exact size of the exon as determined by DNA sequence analysis. Additional subclones of this fragment are also indicated. The DNA sequence data from pBI1.1 and the 4 additional subclones is given in Figure 23. Again, a comparison is made between the human sequence and pMLC-84. Important intron splice sequences and the putative coding sequence are indicated as described previously. The overall homology to the rat exon is 79%, however the 3' half of this exon exhibits 88% homology to the rat cDNA clone and explains the strong hybridization observed in the initial hybridization experiments.

Exon 5 was identified by initial DNA sequence analysis, as previously described in section 4a. Gene Orientation. The subclone of hMLC.1-XRI1.0 (pXRI1.0) that was generated was mapped with additional restriction enzymes. Figure 24 represents a map of
Figure 21. Autoradiogram results of the southern blot hybridization experiment with restricted pBI1.1 DNA probed with pMLC-84/RIO.431. Restriction enzymes used to digest pBI1.1 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 2.5% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled pMLC-84/RIO.431 (2 X 10^5 cpm/ml) at 55°C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 3X SSC at 65°C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80°C.
Figure 22. Partial restriction map of pBI1.1 indicating the position of Exon 4. The shaded region indicates the position of exon 4. Restriction enzyme sites are indicated and additional subclones from this region are indicated below the map.
Figure 23. DNA sequence and the predicted amino acid sequence for Exon 4. Coding sequences are capitalized and intron sequences are indicated by lower case letters. Identical bases between exon four and exon 6 from pMLC-84 are indicated by the vertical lines between the two sequences. Consensus splice site dinucleotides are underlined. Vertical arrows indicate the exon-intron boundaries. The overall nucleotide homology between the two sequences is also indicated.
Figure 24. Partial restriction map of pXRI1.0 indicating the position of Exon 5. Subclones of this region that were cloned into BS+ are also indicated by the solid bars below the map.
pXRII.0. The shaded region represents exon 5. The DNA sequence from this region is given in Figure 25. A comparison is made with exon 7 from pMLC-84 and is shown to exhibit 78% homology with hMLC.1. Interestingly, the 3' region of this exon also exhibits a high degree of sequence conservation to the rat cDNA exon 7, whereas the 5' region is more divergent. The 3' 45 nucleotides of exon 5 exhibit 91% homology to the rat exon 7 and therefore explains why these exons co-hybridize to the rat cDNA clone.

Exon 6 was identified in hMLC.2 by using a 37bp HindIII-EcoRI fragment (pMLC-84/H3-RI0.037) from pMLC-84 (see Figure 6) as a probe against restricted hMLC.2 DNA under low stringency hybridization and wash conditions. The rationale for using hMLC.2 in this analysis is congruent with the argument presented for the 3' walk experiment and the fact that hMLC.2 contains sufficient overlap to span the region between the exon 5 homolog and the end of hMLC.1. The result of that experiment is given in Figure 26. Homology to exon 6 from the rat cDNA was located in the 3' XbaI-EcoRI 1.0 fragment. Extensive DNA sequencing of pXRII.0 revealed a sequence which exhibits 83% homology with exon 8 from the rat skeletal gene. The sequence and a comparison with rat exon 6 is given in Figure 27. A partial restriction map of pXRII.0 indicating the position of this homologous sequence is presented in Figure 28.


Exons 1 and 2 were identified by oligonucleotide probes. The
hMLC.1  attcagtgggtccccccaccccatcatgcctctgcacgcacggtgtgctgactcagcc
rMLC1/3

TlyGluArgLeuThrGluAspGluValGluLysLeuMetAla

hMLC.1  tcccactcctgcag GTGAGAGGCTGACAGAAGACGAAGTGGAGAAGTTGATGGCT
rMLC1/3  GAGAGAATGAAGGAGGAGGAGGGTATGGCTGCG

rMLC1/3  GGCCAGGAGACTCCAATGGCTGCATCAACTATGAAG
hMLC.1  GGGCAAGGAGACTCCAATGGCTGCATCAACTATGAAG tgggctcagcaggggcag

GlyGlnGluAspSerAsnGlyCysIleAsnTyrGluA

NUCLEOTIDE HOMOLOGY = 78%

Figure 25. DNA sequence and the predicted amino acid sequence of Exon 5. Coding sequences are capitalized and intron sequences are indicated by lower case letters. Identical bases between exon 5 and exon 7 from pMLC-84 are indicated by the vertical lines between the two sequences. Consensus splice site dinucleotides are underlined. Vertical arrows indicate the exon-intron boundaries. The overall nucleotide homology between the two sequences is also indicated.
Figure 26. Autoradiogram results of the southern blot hybridization experiment with restricted hMLC.2 DNA probed with pMLC-84/HRI0.037. Restriction enzymes used to digest hMLC.2 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled pMLC-84/HRI0.037 (2 x 10^5 cpm/ml) at 37° C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 3X SSC at 42° C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80° C.
Figure 27. DNA sequence and the predicted amino acid sequence for Exon 6. Coding sequences are capitalized and intron sequences are indicated by lower case letters. Identical bases between exon 6 and exon 8 from pMLC-84 are indicated by the vertical lines between the two sequences. Consensus splice site dinucleotides are underlined. Vertical arrows indicate the exon-intron boundaries. The overall nucleotide homology between the two sequences is also indicated.
Figure 28. Partial restriction map of pXRI1.0 indicating the position of Exon 6.
sequences of these oligonucleotides are given in Figure 29. These sequences were deduced using the following rationale: up to this point in the analysis of hMLC.1, four regions have been identified which exhibit a high degree of homology to pMLC-84. All regions have characteristics representative of functional exons, in addition exon boundaries are conserved with respect to the rat, mouse and chicken skeletal genes (i.e. the point at which coding sequences are interrupted by introns). Exact exon sizes are also conserved for the regions homologous to exons 4 and 5. The four "exons" are homologous to the common coding region between skeletal isoforms 1 and 3 and constitute >70% of the coding sequences. Interestingly, the amino acid sequence predicted from this DNA sequence data exhibited the greatest degree of homology to the amino acid sequence from the chicken ventricular isoform. In addition, an identical two amino acid insertion is present in both the chicken cardiac isoform and pMLC.1, which is not present in any other isoform. Therefore, it was concluded that this human gene probably encodes the MLCl V/Sb isoform and that the amino acid sequence from the chicken cardiac isoform should be used to predict the nucleotide sequences to be synthesized. Another assumption that was made in the determination of the probe sequences was that exon boundaries would be conserved with respect to the skeletal isoform from the three species mentioned. Therefore, two oligonucleotides were synthesized for the following reason; the 50 NH₂-terminal amino acids of the chicken ventricular isoform are
NH2 - Terminal Amino Acid Sequence of the Chicken Cardiac MLC-ALK Isoform:

MET - ALA - PRO - LYS - LYS - PRO - GLU - PRO - LYS - LYS

OLIGO1:

ATG-GC(T,A)-CC(C,T,A)-AAG-AAG-CC(C,T,A)-GA(G,A)-CC(C,T,A)-AAG-AAG

Amino Acid Sequence of the Chicken Cardiac MLC-ALK Isoform "Exon 2":

VAL - GLU - PHE - THR - PRO - ASP - GLN - ILE - GLU

OLIGO2:

ag-GT(C,G)-GA(G,A)-TT(C,T)-AC(C,T,A)-CC(C,T,A)-GA(C,T)-CA(G,A)-
AT(T,C)-GA(G,A)

Figure 29. Nucleotide sequences and the predicted amino acid sequences for Exon 1 and Exon 2 oligonucleotide probes, Oligo1 and Oligo2.
most homologous to exons 1 and 4 in the skeletal gene and therefore may be encoded by two exons and consequently the possibility existed that hMLC.1 would also employ two exons to encode these amino acids. Based on these arguments, the amino acid sequence from the chicken cardiac isoform was used to predict the nucleotide sequences, including redundancies and taking into account the preferred codon usage from pMLC.1 DNA sequence (as discussed in Figure 29). Each oligonucleotide was labeled and used in southern blot hybridization experiments with restricted phMLC.1 DNA. Figure 30 is the result obtained when Oligol was used as a hybridization probe against phMLC.1. The oligonucleotide exhibited hybridization to a single band in all the digests, in particular the region of hybridization could be localized to a 820bp SmaI-BamHI fragment. Figure 31 is the restriction map for hMLC.1 and shows the region of hybridization to Oligol. Subsequently, a 1.0kbp SmaI-XbaI fragment which contained the region of hybridization plus an additional 185bp 3' to the region was subcloned into pUC19 (pSX11.0). The reason for cloning this fragment vs. the SmaI-BamHI fragment was to ensure that the entire exon was present (i.e. the additional DNA present in this clone, 3' to the region of hybridization, would be sufficient to contain the entire exon if by chance the hybridizing region was very close to the BamHI site). The clone, pSX11.0, was analyzed with various restriction enzymes. Restriction enzyme digestion products were blotted to nitrocellulose and reprobed with Oligol. The result of that hybridization experiment is indicated
Figure 30. Autoradiogram results of the southern blot hybridization experiment with restricted hMLC.1 DNA probed with Oligol. Restriction enzymes used to digest hMLC.1 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled Oligol (2 X 10^3 cpm/ml) at 37° C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 3X SSC at 42° C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80° C.
Figure 31. Partial restriction map of hMLC.1 indicating relative position of Exon 1. Positions of exons 3, 4, 5, and 6 are also indicated by the shaded regions.
in Figure 32. From this experiment, exon 1 could be further localized to a 400bp AvaII fragment. Figure 33 is a map of pSX1.0 and shows the region of hybridization to Oligo1. This region was sequenced and shown to contain sequences which exhibit a high degree of homology to exon 1 from the skeletal MLC1f gene, and as expected the predicted amino acid sequence exhibited a high degree of homology to the chicken cardiac sequence. Figure 34 is the sequence obtained from this region. Included in Figure 34 is a comparison between the first 50 amino acids of the chicken cardiac protein and the predicted amino acid sequence from hMLC.1-pSX1.0. As predicted, the amino-terminal coding exon did not encode the entire 50 amino acids observed to exist in the chicken cardiac protein, a total of 9 amino acids remained unidentified. It was therefore necessary to employ Oligo2 in a similar analysis of hMLC.1. The result of the hybridization experiment, using Oligo2 as a probe, is given in Figure 35. Figure 35 is the result obtained when restricted hMLC.1 was analysed. Oligo2 hybridized to the BamHI 1.2 kbp fragment, which also contains the exon 3. Sequence data previously generated from pBI1.2 was searched for homology to Oligo2 before further localization experiments were performed. This analysis proved to be beneficial. Homology to Oligo2 was found 127bp from exon 3 and that sequence is shown in Figure 36. Figure 37 is the restriction map of pBI1.2 and shows the position of Oligo2 homology.
Figure 32. Autoradiogram results of the southern blot hybridization experiment with restricted pSX1.0 DNA probed with Oligol. Restriction enzymes used to digest pSX1.0 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled Oligol (2 x 10^5 cpm/ml) at 37°C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 3X SSC at 42°C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80°C.
Figure 33. Partial restriction map of pSX1.0 indicating the position of Exon 1. Subclones from this region that were constructed in BS+ are indicated by the horizontal solid bars below the map.
Figure 34. DNA sequence and the predicted amino acid sequence of Exon 1. Coding sequences and the putative 5' UTR are capitalized and intron sequences are indicated by lower case letters. Identical amino acids between exon 1 and the chicken cardiac isoform (ccMLC1) are indicated by asterisks, deletions are indicated by dashes. The consensus splice site dinucleotide is underlined. The vertical arrow indicates the exon-intron boundary. The overall amino acid homology between the two sequences is also indicated.
Figure 35. Autoradiogram results of the southern blot hybridization experiment with restricted hMLC.1 DNA probed with Oligo2. Restriction enzymes used to digest hMLC.1 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled Oligo2 (2 X 10^5 cpm/ml) at 37°C for 16 hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 3X SSC at 42°C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80°C.
AMINO ACID HOMOLOGY OF EXON 2 TO THE CHICKEN CARDIAC ISOFORM = 78%

Figure 36. DNA sequence and the predicted amino acid sequence of Exon 2. Coding sequences are capitalized and intron sequences are indicated by lower case letters. Identical amino acids between exon 2 and the chicken cardiac isoform (ccMLC1) are indicated by asterisks. The consensus splice site dinucleotides are underlined. The vertical arrow indicates the exon-intron boundary. The overall amino acid homology between the two sequences is also indicated.
Figure 37. Partial restriction map of pBI1.2 indicating the position of Exon 2.
d. Identification of Exon 7. A total of six exons were identified in hMLC.1 (see Figure 38) and it was concluded at this time in the study that hMLC.1 probably represented the human MLC\textsubscript{V/Sb} gene. This conclusion was based upon the homology to the chicken cardiac isoform (see Table 1); the results obtained by Wade et al., 1988, which showed that the human slow twitch skeletal muscle MLC-ALK cDNA clone, pMLC1-Sb, hybridizes to human adult ventricular RNA; and the data obtained by Barton et al., 1985, which showed that the isoforms MLC\textsubscript{L} and MLC\textsubscript{V} are encoded by the same gene. Compiling these facts, it seemed likely that hMLC.1 may be the gene that encodes the transcript represented by pMLC1-Sb. The sequences from pHMLC1-Sb and hMLC.1 were compared to determine if in fact this was the case. The result of that comparison showed that pMLC1-Sb and hMLC.1 exhibited 100% homology and supported the conclusion that hMLC.1 is the ventricular/slow twitch gene, however sequences representing the the 3' UTR were still unidentified in hMLC.1. If hMLC.1 and pMLC1-Sb do represent the gene and corresponding transcript, respectively; then the 3'UTR region from pHMLC1-Sb should be able to identify the 3' UTR exon in hMLC.1. Therefore a southern blot hybridization experiment with a 200bp AvaII-PvuII fragment isolated from pHMLC1-Sb (kindly supplied by Drs Wade and Kedes) was performed. The fragment represents the 3' half of the 3' UTR and was used as a probe in a hybridization experiment with a southern blot that contained restriction fragments of hMLC.2 DNA. The results of that experiment are given
Figure 38. Partial restriction map of the DNA contained in hMLC.1 and hMLC.2 indicating the positions of Exons 1-6.
in Figure 39. The 3’ UTR probe exhibited hybridization to the 1.0kb XbaI-EcoRI fragment. Therefore pXRII.0 contained exon 5 and exon 6 homology as well as the 3’ UTR, i.e. exon 7. Sequence data previously generated from pXRII.0 was subsequently searched for homology to the 3’UTR from pHMLCl-Sb. Sequence comparisons to pHMLCl-Sb positively identified a region approximately 120bp from exon 6 which exhibited 100% homology to the 3’ UTR of pHMLCl-Sb. The sequence of this region is given in Figure 40.

The complete DNA sequence and the predicted amino acid sequence of hMLCl\textsuperscript{V/Sb} is given in Figure 42. Sequence comparisons were made to pHMLCl-Sb and confirmed all exon splice junctions. Exon sequences are indicated by bold capital letters and intron sequences are given in lower case letters. Important intron and putative RNA processing sites are indicated as described in the legend. Putative transcriptional promotor sequences (i.e. ATA) are also indicated. Important residues characteristic of E-F hand domains are indicated in domain 3. A discussion of the expression of pHMLCl-Sb was presented previously and it was shown that the gene was expressed in adult ventricle (Wade et.al., in press). Unpublished results from our lab have also confirmed this data (data not shown). Based on the results presented here it was concluded that the entire MLC\textsubscript{V/Sb} isoform gene sequence and structure had been elucidated. Continued analysis of the 5’ transcriptional start site is currently underway.
Figure 39. Autoradiogram results of the southern blot hybridization experiment with restricted hMLC.2 DNA probed with pHMLC1-Sb/AII-PvuII0.200. Restriction enzymes used to digest hMLC.1 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled pHMLC-Sb/AII-PvuII0.200 (2 X 10^3 cpm/ml) at 65° C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 3X SSC at 68° C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80° C.
Figure 40. DNA sequence of the 3' Untranslated Region identified with the pHMLCl-Sb probe. Coding sequences are capitalized and intron sequences are indicated by lower case letters. The consensus splice site dinucleotide is underlined. Vertical arrows indicate the exon-intron boundaries. A putative poly(A)$^+$ consensus signal is indicated by the asterisks. The C residue marked with a vertical line indicates the 3' end of the cDNA clone pMLCl-Sb.
Figure 41. Partial restriction map of hMLC.1 indicating the positions of Exons 1-7.
Figure 42. The complete DNA sequence and the predicted amino acid sequence for the Human Ventricular/Slow Twitch MLC-ALK gene, hMLC1v/Sb. Coding sequences are capitalized and the predicted amino acid sequence is indicated below the nucleotide sequence. Intron sequences are indicated by lower case letters and intron sizes are indicated in parentheses. Consensus splice site dinucleotides are underlined. Vertical arrows indicate the putative 5' and 3' ends of the transcript. Asterisks above the sequence indicate putative transcriptional control signals.
5. Chromosomal Localization. Determining the chromosomal locus of hgMLC.1 was a collaborative effort between Uta Franke's laboratory at Yale University and our laboratory. Chromosomal localization was determined by hybridizing nick-translated pXRI-1.0 to southern blots of 12 rodent x human somatic cell hybrid DNAs digested with BglII or HindIII. Human-specific restriction fragments of 13.3 kbp (BglII) and 11.3 kbp (HindIII) were detected only in hybrids with human chromosome 3. Possible sites on other chromosomes were excluded by three or more discordant clones (see Table 3). Table 3 is the compiled data of informative hybrids and discordant hybrids. Analysis of this data revealed that hybrids that contained chromosome 3 had a discordance of 0, whereas all the other chromosomes produced a discordance of 3 or more, for example chromosome 1 exhibited a discordance of 6. In other words, whenever the MYL1 signal was present in the hybrid DNA (i.e. hybridization to pXRI1.0) chromosome 3 was also present. This was not the case for any other chromosome. All human-specific fragments were also present in 2 hybrids that had retained the short arm of chromosome 3 and were absent in 3 hybrids that had retained only the long arm.

These results are consistent with the existence of a single locus for this polypeptide for which the designation MYL1 is proposed.
TABLE 3. Correlation of human MYL1 sequences with human chromosomes in rodent x human somatic cell hybrids.

<table>
<thead>
<tr>
<th>Hybridization/ Chromosome</th>
<th>Human Chromosomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1    2    3    4    5    6    7    8    9    10   11   12   13   14   15   16   17   18   19   20   21   22   X</td>
</tr>
<tr>
<td>+/+</td>
<td>3    1    8    2    2    7    3    4    1    3    2    4    2    4    5    3    3    0    4    3    1    6    6    3</td>
</tr>
<tr>
<td>-/-</td>
<td>5    6    6    4    5    3    5    3    8    3    5    3    2    3    5    5    3    2    4    2    0    2</td>
</tr>
<tr>
<td>+/-</td>
<td>6    8    0    6    8    3    7    6    9    7    7    6    6    4    6    5    9    4    7    9    4    4    2</td>
</tr>
<tr>
<td>-/+</td>
<td>3    0    0    3    2    4    1    4    3    0    3    3    5    5    5    3    3    5    6    4    4    6    2</td>
</tr>
</tbody>
</table>

Discordant hybrids: 9 8 0 9 10 7 8 10 12 7 10 9 11 9 11 8 12 9 13 13 8 10 4
Informative hybrids: 17 15 14 15 17 16 17 16 18 15 18 16 17 16 17 16 18 16 16 16 9

The numbers of hybrids that are concordant (+/+ or -/-) and discordant (+/- or -/+), with the human MYL1 sequence are given for each chromosome. Hybrids in which a particular chromosome was structurally rearranged or present in fewer than 10% of cells were excluded.
B. Discussion

The results of the above experiments have led to the conclusion that the gene contained in the lambda phage clone, hMLC.1, is the gene that encodes the myosin alkali light chain ventricular/slow twitch isoform. This is the first MLC-ALK gene that has been sequenced in humans (although several cDNA MLC-ALK sequences have been deduced) and also represents the first MLC-ALK gene other than the skeletal gene for which the exon-intron structure has been determined. The gene contains 7 exons and 6 introns. The first six exons were found to encode a protein of 195 amino acids, exon 7 contained the sequences representing the 3' UTR of the mature transcript. Interestingly, the location of introns 1-6 are in analogous positions to the introns of the skeletal fast twitch MLC1p isoform. Analyzing the intron exon boundaries it was observed that the consensus RNA splicing control sequences are present in all the introns except intron 3. The 3' acceptor site in intron 3 contains the consensus AG dinucleotide, however this acceptor site is not preceded by the consensus polypyrimidine stretch of nucleotides that usually precedes the 3' acceptor site. Several purine residues are present in this sequence that are not observed in the other vertebrate MLC-ALK gene introns which have been sequenced, however a polypurine stretch of nucleotides is observed in intron 4 of the Drosophila MLC-ALK gene. It has been proposed that this polypurine stretch of nucleotides is somehow functionally involved in the differential splicing observed for the
Drosophila gene transcript (Falkenthal et al., 1985). The functional significance, if any, of the nonconsensus sequence observed in intron 3 of hMLC1_Sb is not clear. The transcriptional start site of hMLC1_Sb is currently under investigation by the laboratory and preliminary results have indicated that exon 1 is 183 nucleotides in length and therefore maps the transcriptional start site at the beginning of the cDNA sequence determined for pMLC1-Sb (Wade et al., 1988). The putative consensus transcription control sequence ATA is observed 26 nucleotides 5' to the start of the cDNA and seems to suggest that the cDNA may indeed be representative of a full length transcript.

The derived amino acid of hMLC1_Sb was compared to the other MLC-ALK amino acid sequences (either predicted from a DNA sequence or actual protein sequence) from vertebrates. These comparisons indicated that the predicted protein sequence of hMLC1_Sb is most homologous to the chicken cardiac isoform and indicates that analogous tissue isoforms between species are more homologous than isoforms present within a species. These results suggest that specific functional constraints for a particular tissue type reduce the extent of divergent evolution. It was also observed that hMLC1_Sb and the chicken cardiac/slow twitch isoform contain a two amino acid insertion (lysine and cysteine, amino acids 66 and 67) when compared to the amino acids of other known isoforms. These results suggest that the gene duplication event that gave rise to this isoform must have occurred before mammals and avians diverged, which is estimated at 270 million years ago.
A comparison of the coding region exons from hMLCl/V/Sb and the appropriate exons or regions (cDNA sequences where the exon boundaries are unknown) from other MLC-ALK sequences is presented in Table 4. Evident from this data is the fact that exons 4, 5, and 6 exhibit a greater degree of homology than the homology between exons 1, 2, and 3 and the appropriate regions from other MLC-ALK sequences. The combined homology for exons 4, 5, and 6 between hMLCl/V/Sb and the various MLC-ALK sequences ranges from 76% to 80%, whereas the combined homology for exons 1, 2, and 3 between hMLCl/V/Sb and the other MLC-ALK isoforms ranges from 54% to 61%. These results suggest that the amino acids required for common MLC-ALK function reside in the C-terminal region of the protein and that tissue specificity is conferred by the N-terminal region.

Interestingly, the chromosomal "linkage group" also seems to be conserved between species for this MLC-ALK gene. In the mouse, 3 myosin alkali light chain genes have been assigned to chromosomes 1, 9, and 11 (Robert et al., 1985). The Myl-3 gene on distal MMU 9 encodes the ventricular/slow twitch skeletal muscle isoforms. Other loci on distal MMU 9, such as Bgl and Acv are homologous to genes on human chromosome 3p, and no other known homologies exist between 3p and MMU 1 or MMU 11. Therefore it is possible that MYL1 is the homologue of mouse Myl-3 and may be more appropriately named MYL3.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Human Fast Twitch</td>
<td>57(74)</td>
<td>61(17)</td>
<td>63(95)</td>
<td>80(140)</td>
<td>79(62)</td>
<td>83(24)</td>
<td>61(186)</td>
<td>80(226)</td>
<td>70(412)</td>
</tr>
<tr>
<td>Mouse Fast Twitch</td>
<td>43(56)</td>
<td>61(17)</td>
<td>65(98)</td>
<td>76(133)</td>
<td>78(61)</td>
<td>79(23)</td>
<td>56(171)</td>
<td>77(217)</td>
<td>66(388)</td>
</tr>
<tr>
<td>Rat Fast Twitch</td>
<td>50(64)</td>
<td>57(16)</td>
<td>65(97)</td>
<td>79(137)</td>
<td>78(61)</td>
<td>79(23)</td>
<td>58(177)</td>
<td>79(221)</td>
<td>68(398)</td>
</tr>
<tr>
<td>Chicken Fast Twitch</td>
<td>40(52)</td>
<td>64(18)</td>
<td>63(95)</td>
<td>74(128)</td>
<td>77(60)</td>
<td>86(25)</td>
<td>54(165)</td>
<td>76(213)</td>
<td>64(378)</td>
</tr>
<tr>
<td>Chicken Smooth Muscle</td>
<td>-</td>
<td>-</td>
<td>65(98)</td>
<td>76(132)</td>
<td>82(64)</td>
<td>45(13)</td>
<td>65(98)*</td>
<td>74(209)</td>
<td>71(307)</td>
</tr>
<tr>
<td>Chicken Non-Muscle</td>
<td>-</td>
<td>-</td>
<td>65(98)</td>
<td>76(132)</td>
<td>82(64)</td>
<td>69(20)</td>
<td>65(98)*</td>
<td>77(216)</td>
<td>73(314)</td>
</tr>
</tbody>
</table>
CHAPTER VI
ISOLATION OF hMLC.3 AND hMLC.4

A. Results

1. Isolation of hMLC.3 and hMLC.4. A partial Sau3AI human library cloned into lambda Charon 28 was screened with pMLC-84/R10.431 under low stringency hybridization conditions as previously described. An initial 46 putative positive clones were isolated and individually rescreened for hybridization to the rat cDNA sequence. Four clones; 3.3, 3.4, 3.7a and 3.7b continued to exhibit consistent hybridization to the probe and were subsequently plaque purified. Clone 3.4 consistently hybridized with less intensity than the other 3 clones. DNA from each of these phage clones was isolated and analyzed by restriction enzyme digestion. Initial digests of the clones revealed two distinct restriction patterns for each of the following enzymes; BamHI, EcoRI, and HindIII. A Southern blot hybridization experiment with this gel probed with P-32 labeled pMLC-84/R10.431 revealed a single hybridizing band in each of the digests, again exhibiting 2 distinct banding patterns. Figure 43 is the autoradiogram produced from this experiment. Clones 3.3, 3.7a and 3.7b were concluded to be identical from this analysis, pooled and renamed hMLC.3. Clone 3.4 was concluded to be unique and subsequently renamed hMLC.4.
Figure A3. Autoradiogram results of the southern blot hybridization experiment with restriction enzyme digestion products of clones 3.3, 3.4, 3.7a, and 3.7b probed with pMLC-8/RI0.431. Restriction enzymes used to digest bacteriophage DNA are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled pMLC-8/RI0.431 (2 \times 10^5 \text{ cpm/ml}) at 55^\circ C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 3X SSC at 65^\circ C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80^\circ C.
CHAPTER VII
PRELIMINARY ANALYSIS OF hMLC.3

A. Results

1. Restriction Analysis of hMLC.3. Restriction enzyme analysis of hMLC.3 was performed using a variety of restriction enzymes. A partial restriction map of hMLC.3 is given in Figure 44.

2. Repeat Analysis. The analysis of hMLC.3 for repeated DNA was performed for the same reason presented in the analysis of hMLC1V/Sb, i.e., in order for Dr. Francke's laboratory to determine the chromosomal localization for this clone, a single copy restriction fragment needed to be isolated for use as a probe in the somatic cell hybrid analysis. Restriction digests of hMLC.3 were probed with labeled human genomic DNA and the repeat element clones AluI0.55, KpnI1.55 and KpnI1.65. The results of these experiments are discussed in this section. Figure 45 is the result of hMLC.3 probed with P-32 labeled genomic DNA and Figure 46 is the result of hMLC.3 probed with P-32 labeled AluI DNA. Comparing these two autoradiograms it is obvious that the two probes are hybridizing to identical fragments. The region of strong hybridization is localized to a 1.75kbp SmaI fragment and also seems to extend into the adjacent 3.0kbp SmaI fragment 5' to it in the clone (see Figure 44, hMLC.3 restriction map). The slightly
Figure 44. Partial restriction Map of hMLC.3. Enzyme sites are indicated above the map and the region of hybridization to the rat MLC-ALK cDNA clone is indicated by the shaded region. The position of the genomic repeat element is indicated by the solid bar above the map.
Figure 45. Autoradiogram results of the southern blot hybridization experiment with restricted hMLC.3 DNA probed with human genomic DNA. Restriction enzymes used to digest hMLC.3 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled human genomic DNA (1 X 10^5 cpm/ml) at 65° C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 1X SSC at 68° C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80° C.
Figure 46. Autoradiogram results of the southern blot hybridization experiment with restricted hMLC.3 DNA probed with an AluI repeat clone. Restriction enzymes used to digest hMLC.3 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled AluI DNA (1 X 10^5 cpm/ml) at 65° C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 1X SSC at 68° C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80° C.
hybridizing band seen in lanes 4, 6, 11, 12, and 14 represents the 3.0kbp SmaI fragment and gives support to this conclusion. No KpnI family repeat elements were observed to hybridize to hMLC.3.

3. Genomic Southern Analysis. The repeat analysis of hMLC.3 presented above indicated that sequences flanking the DNA homologous to the AluI repeat would be useful in a genomic southern hybridization experiment. A unique 0.7kbp SmaI-KpnI fragment located at near the 3' end of hMLC.3, and contained within hMLC.3-RI6.0, was band isolated, labeled with P-32 and used to probe a genomic southern blot of human DNA. The result of that experiment is given in Figure 47. The approximate sizes of the hybridizing bands are indicated. Evident from this analysis is the hybridization to the 6.0kbp EcoRI band, as expected, and to a 14kbp band in the BamHI digest which also corresponds to the BamHI 14kbp band seen in Figure 7. A single hybridizing band is also seen in the HindIII digest. This fragment was subsequently sent to Dr. Francke for the purpose of localizing this gene to its respective chromosome.

4. Exon Identification. Exons homologous to the rat exons 5, 6, and 7 from pMLC-84 were identified as described in Chapter III. The results of exon 5, 6, and 7 homology are discussed in the following sections.

Restriction enzyme digests from hMLC.3 were probed with pMLC-84/RIAI0.121 (i.e. the rat cDNA fragment containing exon 5). The result of that hybridization experiment is given in Figure 48. The
Figure 47. Autoradiogram results of the southern blot hybridization experiment with restricted human genomic DNA probed with SmaI-KpnI0.7 (SK0.7). Restriction enzymes used to digest genomic DNA are indicated above the corresponding lanes (BI = BamHI, RI = EcoRI, HIII = HindIII). Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled SK0.7 DNA (2 X 10^6 cpm/ml) at 65° C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 0.1X SSC at 68° C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80° C.
<table>
<thead>
<tr>
<th>Lane</th>
<th>Enzyme Combination</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>BglIII</td>
</tr>
<tr>
<td>B</td>
<td>BglIII-RamHI</td>
</tr>
<tr>
<td>C</td>
<td>BglIII-EcoRI</td>
</tr>
<tr>
<td>D</td>
<td>BglIII-Smal</td>
</tr>
<tr>
<td>E</td>
<td>ClaI</td>
</tr>
<tr>
<td>F</td>
<td>ClaI-Smal</td>
</tr>
<tr>
<td>G</td>
<td>lambda marker</td>
</tr>
<tr>
<td>H</td>
<td>EcoRI</td>
</tr>
<tr>
<td>I</td>
<td>EcoRI-Smal</td>
</tr>
<tr>
<td>J</td>
<td>PstI</td>
</tr>
<tr>
<td>K</td>
<td>PstI-Smal</td>
</tr>
<tr>
<td>L</td>
<td>lambda marker</td>
</tr>
<tr>
<td>M</td>
<td>SacI</td>
</tr>
<tr>
<td>N</td>
<td>SacI-Smal</td>
</tr>
<tr>
<td>O</td>
<td>HindIII</td>
</tr>
<tr>
<td>P</td>
<td>HindIII-Smal</td>
</tr>
<tr>
<td>Q</td>
<td>XbaI</td>
</tr>
<tr>
<td>R</td>
<td>XbaI-Smal</td>
</tr>
<tr>
<td>S</td>
<td>XbaI-BglIII</td>
</tr>
</tbody>
</table>

**Figure 48.** Autoradiogram results of the southern blot hybridization experiment with restricted hMLC.3 DNA probed with the rat cDNA clone. Restriction enzymes used to digest hMLC.3 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled RIA10.121 (1 X 10^5 cpm/ml) at 65° C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 1X SSC at 68° C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80° C.
homology to this probe was localized to the region between the two SacI fragments as indicated by the hybridization observed in lanes 6 and 7 in Figure 48. Figure 49 shows the restriction map of hMLC.3 and the region of hybridization to the rat exon 5.

Initial hybridization experiments, using pMLC-84/RI0.431 as the probe, identified a single EcoRI fragment (hMLC.3-RI6.0) that exhibited strong hybridization to the rat cDNA fragment (see Figure 43). This fragment was subsequently subcloned into the Bluescript plus (BS+) vector. Restriction enzyme digestion products were electrophoresed on a 1.5% agarose gel, blotted to nitrocellulose and probed with a 308bp AvaI-EcoRI fragment from pMLC-84, which contains exons 5, 6 and 7. Figure 50 is an autoradiogram of that hybridization experiment and shows that two fragments exhibit hybridization. Increasing the wash stringency with this blot resulted in a single banding pattern in each of the digests. The result of that experiment is given in Figure 51. Two possible conclusions can be drawn from the results presented in Figures 50 and 51; (1) one exon is represented by the hybridization data where the KpnI site bisects it and increasing the wash stringency only removed the segment of the exon exhibiting the least amount of homology to the probe; or (2) at least two exons exhibit homology to the rat cDNA fragment and that the rat exon 4 homology represented the hybridization to the 700bp SmaI-KpnI fragment seen in Figure 51, which was retained even after increasing the stringency of the wash conditions (remember that exon 4 contains the sequences
Figure 49. Restriction Map of hMLC.3 Showing the Region of hybridization to pMLC-84/RIA10.121. The shaded region indicates the position of homology to exon 5 from the rat cDNA clone. Restriction enzyme sites are indicated above the map.
Figure 50. Autoradiogram results of the southern blot hybridization experiment with restricted hMLC.3-pRI6.0 DNA probed with an pMLC-84/AIR10.308. Restriction enzymes used to digest pRI6.0 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled pMLC-84/AIR10.308 DNA (1 X 10^5 cpm/ml) at 55°C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 3X SSC at 65°C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80°C.
Figure 51. Autoradiogram results of the southern blot hybridization experiment with restricted hMLC.3-pRI6.0 DNA probed with an pMLC-84/AIRI0.308. Restriction enzymes used to digest pRI6.0 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled pMLC-84/AIRI0.308 DNA (1 X 10^5 cpm/ml) at 55° C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 3X SSC at 65° C for 1 hr and then rewashed with 1X SSC at 65° C for 1 hr. The filter was then re-exposed to X-ray film overnight with an intensifying screen at -80° C.
encoding E-F hand domain 3 and represents the region of greatest homology between all MLC-ALK proteins) and that exon 5 homology resided in the region between the KpnI site and the downstream BamHI site (remember that these exons co-hybridized to hMLC.1). To determine which of the two conclusions was correct, DNA sequence analysis of the 700bp SmaI-KpnI fragment was performed. This analysis identified a region homologous to the rat exon 6 (i.e. exon 4 from hMLC1V/Sb) approximately 200bp 5' to the KpnI site (see Figure 52). The sequence data presented in Figure 52 shows that the region of hybridization detected in the preceding hybridization experiment (Figure 51), did in fact correspond to a single exon and it can therefore be concluded that the hybridization detected in the KpnI-EcoRI fragment corresponds to a separate exon. Figure 53 is a partial restriction map of the RI6.0 subclone and shows the relative positions of exons 3, 4, and 5.

Exon 8 from pMLC-84 (i.e. the HindIII-EcoRI0.037 fragment) was also used in hybridization experiments but did not exhibit any homology to the DNA present in hMLC.3.

Both oligonucleotides, Oligol and Oligo2, were also used as probes in hybridization experiments to hMLC.3 DNA. The results of those experiments were negative, both Oligol and Oligo2 failed to exhibit any hybridization to hMLC.3 DNA.
NUCLEOTIDE HOMOLOGY BETWEEN EXON 4 FROM hMLC.3 AND hMLC.1 = 82%

Figure 52. DNA sequence data determined for Exon 4 from hMLC.3. Coding sequences are capitalized and intron sequences are indicated by lower case letters. Identical bases between exon 4 and exon 4 from hMLC.1 are indicated by the vertical lines between the two sequences. Consensus splice site dinucleotide is underlined. The vertical arrow indicates the exon-intron boundary.
Figure 53. Partial restriction map of hMLC.3-RI6.0 indicating the regions of Exons 3, 4, and 5. Shaded regions represent the locations of the exons in pRI6.0. Enzyme sites are indicated above the map.
B. Discussion

The data from the above experiments on hMLC.3 indicate that this clone represents a human MLC-ALK gene. The sequence data presented in Figure 52 indicates that the residues characteristic of the MLC-ALK EF hand domain 3 are conserved and supports the conclusion that hMLC.3 encodes a human MLC-ALK isoform. The 3' end of this putative exon is followed by the consensus dinucleotide "gt" characteristic of the 5' end of a functional intron. Several attempts to produce legible sequence from the 5' most region of this exon failed repeatedly. Constructing additional subclones of this region will be required to obtain this sequence. The conclusion was made that hMLC.3 encodes an MLC-ALK isoform, however which isoform this clone encodes cannot be determined from the data obtained to date. Exon 4 from this gene does exhibit the greatest amount of homology to exon 4 from the MLC1V/Sb isoform (see Table 5) and may represent the human MLC1Sa isoform. This conclusion is based on the fact that the DNA sequence from exon 4 is different from the sequence that has been obtained from the human skeletal MLC1p, MLC3p and MLC smooth muscle (H.H. Arnold, personal communication) cDNA clones. In addition, it has been shown that the two slow twitch isoforms MLC1Sb and MLC1Sa exhibit amino acid sequence homology and share common epitopes as discussed in Chapter II. The evidence presented however, does not exclude the possibility that hMLC.3 encodes the atrial/embryonic isoform, MLC1A/emb (to date no amino acid or nucleotide sequence data is
Table 5. Nucleic Acid and Amino Acid Homology between Exon 4 from hMLC.3 and the other Human MLC-ALK Isoforms.

<table>
<thead>
<tr>
<th>MLC Isoforms</th>
<th>% Nucleic Acid</th>
<th>% Amino Acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLC1F/3F</td>
<td>78</td>
<td>75</td>
</tr>
<tr>
<td>MLC1V/Sb</td>
<td>82</td>
<td>81</td>
</tr>
<tr>
<td>MLC_smooth</td>
<td>67</td>
<td>69</td>
</tr>
</tbody>
</table>

* personal communication (unpublished results from H.H. Arnold)
available for this isoform). The chicken fetal isoform L_{23} exhibits 77% homology at the amino acid level and 80% homology at the nucleic acid level, both of these values are less than that seen in Table 5 for the hMLC_1 V/Sb isoform. Additional DNA sequence data and northern blot analysis of different human tissues needs to be performed to determine the actual isoform that this gene encodes. The chromosomal localization of this gene may also provide evidence for determining the isoform that this gene encodes. The human MLC_1 A/emb gene has been mapped to chromosome 17 (Cohen-Haguenauer et al., 1987) and if hMLC.3 also maps to this chromosome it would provide evidence supporting the conclusion that hMLC.3 encodes the atrial/embryonic isoform. The MLC_1 Sa isoform has not been mapped and therefore if a unique locus is identified then hMLC.3 may in fact encode the MLC_1 Sa isoform.
CHAPTER VIII
PRELIMINARY ANALYSIS OF hMLC.4

A. Results

1. Restriction Analysis of hMLC.4. Restriction enzyme analysis of hMLC.4 was performed using a variety of restriction enzymes. The data from this analysis was used to construct the partial restriction map presented in Figure 54. The region of hybridization to the rat cDNA clone is also indicated. Plasmid subclones from hMLC.4 were generated by cloning restriction fragments into the appropriate restriction sites located in the polylinker of the BS+ vector. These subclones are also indicated in Figure 54.

2. Repeat analysis. The following repeat analysis was performed on hMLC.4 to identify single copy DNA in this clone for the purpose of providing Dr. Francke's laboratory with a probe for the chromosomal localization analysis. Restriction digests of hMLC.4 were probed with labeled human genomic DNA and the repeat element clones AluI0.55, KpnI1.55 and KpnI1.65. The results of these experiments are discussed in this section. Figure 55 is the result of hMLC.4 probed with labeled genomic DNA. Figure 56 is the result of hMLC.4 probed with labeled AluI DNA. Comparing these two autoradiograms it is obvious that the two probes are hybridizing to identical fragments. These regions of homology to the AluI repeat
Figure 54. Partial restriction map of hMLC.4. Enzyme sites are indicated above the map. The region of this clone that hybridizes to the rat cDNA clone, pMLC-84, is indicated by the shaded region. AluI repeat elements are indicated by the solid horizontal bars above the map. Subclones constructed in BS+ are indicated below the map.
Figure 55. Autoradiogram results of the southern blot hybridization experiment with restricted hMLC.4 DNA probed with human genomic DNA. Restriction enzymes used to digest hMLC.4 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled human genomic DNA (1 X 10^5 cpm/ml) at 65°C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 1X SSC at 68°C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80°C.
Figure 56. Autoradiogram results of the southern blot hybridization experiment with restricted hMLC.4 DNA probed with AluI DNA. Restriction enzymes used to digest hMLC.4 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled AluI DNA (1 X 10^5 cpm/ml) at 65° C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 1X SSC at 68° C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80° C.
element were localized to two EcoRI fragments as indicated in Figure 54. The genomic repeat element was further localized in the 3.5 EcoRI fragment (RI3.5) by probing restriction digest fragments from the RI3.5 subclones with genomic DNA. The data from that experiment is given in Figure 57. A partial restriction map for pRI3.5 is given in Figure 58 and shows the location of the repeat element. The genomic repeat in this 3.5 EcoRI clone was further localized because this fragment of hMLC.4 also contains the region that hybridizes to the rat cDNA clone, pMLC-84. No KpnI homology was detected in hMLC.4 with either of the KpnI clones.

3. Genomic Southern Analysis. Initial hybridization data on hMLC.4 with the rat cDNA probe revealed a hybridizing 3.5 EcoRI fragment, see Figure 54. Homology to the rat cDNA clone was further localized in RI3.5 by southern blot hybridization to restriction enzyme fragments probed with pMLC-84/AIRI0.308. The results of this hybridization experiment are given in Figure 59. The region of homology was subsequently localized to the region indicated on Figure 60. Subclones of this entire region (HincII-PstI0.35 (pHcP0.35), cloned into the PstI and Smal sites; PstI0.425 (PstI0.425), cloned into the PstI site; PstI-ClaI0.7 (PC0.65), cloned into the PstI and ClaI sites) were generated and are also indicated in Figure 60. Based on the results of the genomic repeat hybridization experiment it was concluded that the PstI0.425 fragment may represent unique DNA and therefore it was subsequently used as a probe in a human genomic southern blot hybridization
Figure 57. Autoradiogram results of the southern blot hybridization experiment with restricted hMLC.4-pRI3.5 DNA probed with human genomic DNA. Restriction enzymes used to digest pRI3.5 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled human genomic DNA (1 X 10^5 cpm/ml) at 65° C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 1X SSC at 68° C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80° C.
Figure 58. Partial restriction map of hMLC.4-R13.5 indicating the region of homology to the rat cDNA clone. Restriction enzymes are indicated above the map. The region of homology to pMLC-84 is indicated by the shaded region. The AluI repeat element is indicated by the solid bar above the map. Subclones constructed in BS+ are indicated by the solid bars below the map.
Lane
A. AvaI
B. AvaI-HindIII
C. AvaII
D. AvaII-HindIII
E. pAT153 marker DNA
F. BamHI
G. BamHI-EcoRI
H. EcoRI
I. EcoRI-HindIII
J. pAT153 marker DNA
K. HindIII
L. HincII
M. HincII-HindIII
N. SacI
O. SacI-HindIII
P. pAT153 marker DNA
Q. SacII-HindIII
R. SacII

Figure 59. Autoradiogram results of the southern blot hybridization experiment with restricted hMLC.4-pRI3.5 DNA probed with the rat cDNA fragment pMLC-84/AIRI0.308 DNA. Restriction enzymes used to digest pRI3.5 are indicated in the corresponding lanes. Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled pMLC-84 DNA (2 × 10^5 cpm/ml) at 55° C for 16hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 3X SSC at 60° C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80° C.
Figure 60. Partial restriction map of hMLC.4-RI3.5 indicating the region of homology to the rat cDNA clone. Restriction enzymes are indicated above the map. The region of homology to pMLC-84 is indicated by the shaded region. The AluI repeat element is indicated by the solid bar above the map. Subclones constructed in BS⁺ are indicated by the solid bars below the map.

experiment. Figures 61-63 are the results of that experiment.

Figure 61 is the result obtained with moderate stringency washes. A number of bands are observed in all the lanes. Table 6 is a list of the approximate sizes of fragments which hybridize.

Interestingly, this fragment hybridizes to similar size fragments as the rat cDNA clone in the EcoRI digest (see Figure 9), but not in the BamHI digests. Additional bands are also observed with respect to the fragment banding pattern seen in Figure 9.

Increasing the wash stringency resulted in the banding pattern seen in Figure 62. Evident from this autoradiogram is the loss of or reduced intensity of certain bands in each of the lanes. Using high stringency washes resulted in the retention of only the expected homologous bands, see Figure 63. These results suggest that the fragment, hMLC.4-PstI0.425, contains homology to the MLC-ALK gene family and quite possibly to a related gene family (i.e. other members of the Troponin C superfamily).

4. Localization of MLC-ALK Homology. The interesting results produced by the hybridization experiments given in Figures 61-63 required that the region of homology to the rat cDNA clone be further localized to ascertain whether the PstI0.425 clone represented the homologous region. A hybridization experiment was subsequently performed on restricted RI3.5 as well as with the various subclones previously described. The result of that experiment is given in Figure 64 and indicates that the rat cDNA probe in fact does hybridize to the PstI0.425 fragment. A low
Figure 61. Autoradiogram results of the southern blot hybridization experiment with restricted human genomic DNA probed with hMLC.4-PstI.42. Restriction enzymes used to digest genomic DNA are indicated above the corresponding lanes (BI = BamHI, RI = EcoRI, HIII = HindIII, XI = XbaI). Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled PstI0.425 DNA (2 x 10^6 cpm/ml) at 65° C for 16hrs in 6X SSC + Denhardt solution. The filter was subsequently washed with 1X SSC at 55° C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80° C.
Figure 62. Autoradiogram results of the southern blot hybridization experiment with restricted human genomic DNA probed with hMLC.4-PstI0.42 at increased stringency. Restriction enzymes used to digest genomic DNA are indicated above the corresponding lanes (BI = BamHI, RI = EcoRI, HIII = HindIII, XI = XbaI). Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The same filter in Figure 61 was subsequently rewashed with 1X SSC at 65°C for 1 hr. The filter was then re-exposed to X-ray film overnight with an intensifying screen at -80°C.
Figure 63. Autoradiogram results of the southern blot hybridization experiment with restricted human genomic DNA probed with hMLC.4-PstI0.42 at maximum stringency. Restriction enzymes used to digest genomic DNA are indicated above the corresponding lanes (BI = BamHI, RI = EcoRI, HIII = HindIII, XI = XbaI). Digestion products were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The same filter in Figure 62 was subsequently rewashed with 0.1X SSC at 65°C for 1 hr. The filter was then re-exposed to X-ray film overnight with an intensifying screen at -80°C.
Table 6. Sizes of Hybridizing Restriction Fragments in the Genomic Southern Blot Analysis of pPstI0.425

<table>
<thead>
<tr>
<th>BamHI</th>
<th>EcoRI</th>
<th>HindIII</th>
<th>XbaI</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>16</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>18</td>
<td>13</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>9.5</td>
<td>9.2</td>
<td>10.5</td>
<td>11.5</td>
</tr>
<tr>
<td>8.8</td>
<td>6.2</td>
<td>6.7</td>
<td>*6.2</td>
</tr>
<tr>
<td>*8.4</td>
<td>5.3</td>
<td>6.2</td>
<td>5.9</td>
</tr>
<tr>
<td>7.2</td>
<td>5.1</td>
<td>5.7</td>
<td>5.7</td>
</tr>
<tr>
<td>5.7</td>
<td>4.3</td>
<td>5.4</td>
<td>4.6</td>
</tr>
<tr>
<td>3.5</td>
<td>4.1</td>
<td>3.9</td>
<td>3.2</td>
</tr>
<tr>
<td>3.3</td>
<td>*3.5</td>
<td>3.6</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td>3.1</td>
<td>*3.3</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>2.6</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.2</td>
<td></td>
</tr>
</tbody>
</table>

* Fragments labeled with an asterisks are the homologous fragments. All fragments other than the homologous fragments reduced in intensity or were removed from the filter after the second wash. Only the homologous fragments exhibited hybridization after the third wash. All fragment sizes are indicated in kbp.
Lane
A. A vaginal-PstI
B. PstI
C. ClaI-PstI
D. ClaI
E. HincII
F. HincII-PstI
G. Rsal
H. Rsal-PstI
I. pUC18 marker DNA
J. pCP1.0
K. pPC0.7
L. pPHcII1.0
M. pHcP0.35
N. pPstl0.425

Figure 64. Autoradiogram results of the southern blot hybridization experiment with restricted hMLC.4-RI3.5 DNA and single strand DNA preparations from subclones pPHcII1.0, pHcP0.35, pPstl0.42, pPIC10.7, and pCP1.0. Restriction enzymes used to digest pRI3.5 DNA are indicated in the corresponding lanes. Digestion products and single stranded DNA were electrophoresed in a 1.0% agarose gel and blotted to nitrocellulose. The filter was incubated with P-32 labeled pMLC-84/RI0.431 (2 X 10^5 cpm/ml) at 55° C for 16 hrs in 6X SSC + Denhardt's solution. The filter was subsequently washed with 3X SSC at 65° C for 1 hr. The filter was then exposed to X-ray film overnight with an intensifying screen at -80° C.
level of hybridization is seen to the actual fragment (see Lanes A, B, C, and F), however the PstI0.425 subclone exhibits strong hybridization. The gel represented by this figure was used for two southern blots and therefore may have resulted in a reduced amount of the PstI fragment adhering to this particular filter and as a result faint hybridization is detected to this fragment.

5. DNA Sequence Analysis. The pPI0.425 subclone and the flanking subclones were then sequenced to try and identify specific regions of homology to pMLC-84. The DNA sequence from this region is given in Figure 66. Unexpectedly, no significantly long stretches of homology are observed between this sequence and the rat cDNA clone. A small region of homology was identified using the DNAstar alignment subprogram and that region is indicated in Figure 66. The DNA sequence was subsequently entered into the Bionet System and used to search the NIH gene bank. This analysis also resulted in no significant homology to any DNA sequence published for chicken, mouse, rat, or human genes.

B. Discussion

Interestingly, hMLC.4 does not appear to be a myosin alkali light chain gene, although it does exhibit hybridization to the rat cDNA clone. If the DNA in hMLC.4 does in fact contain coding sequences for some unknown gene product, it would have to be used in a northern blot analysis of human RNA from all the major tissue types. This is the only way to determine if a transcript is produced.
Figure 65. Partial Restriction Map of hMLC.4-R13.5 indicating the region of hybridization to the rat cDNA clone. Restriction enzyme sites are indicated above the map. The shaded region between the 2 PstI sites indicates the region of homology to the rat cDNA clone.
Figure 66. DNA Sequence Data obtained from pPstI0.42. Sequence homology to exon 7 from the rat cDNA clone is underlined. The Rsal restriction site that is mentioned in the text is indicated. The PstI sites defining the ends of this fragment are also indicated.
GENERAL DISCUSSION

The myosin alkali light chain proteins are highly conserved throughout evolution and therefore must provide some essential function to the different contractile systems in eukaryotic cells. The human genes identified in this report are the first reported human myosin alkali light chain genes for which the exon-intron structure has been determined. It is also novel for being the first ventricular/slow twitch gene that has been characterized in any organism. The data presented here has also added to the accumulating data on the myosin alkali light chain protein genes and has provided knowledge on the regions of the genes and their respective proteins which have been the most conserved.

The function of the myosin alkali light chain is still undetermined, but hopefully from the accumulating data on these genes and their proteins it will be possible to elucidate the function by employing mutagenesis studies to particular domains of the protein in organism systems that are amenable to these types of genetic studies.

The regulation of the expression of the different isoforms within a species also needs to be addressed to identify the tissue specific control regions that govern the expression of an isoform in any particular tissue type.
Future experiments in our laboratory have been discussed and steps are being taken to answer specific questions concerning the regulation of hMLC1V/Sb gene in slow twitch and ventricular muscle. Deletion analysis of the 5' flanking DNA is underway to assay which sequences in the 5' region of hMLC1V/Sb are responsible for the correct transcription of the gene as well as which sequences are responsible for the tissue specific expression in both ventricular and slow twitch skeletal muscle.

The laboratory is also involved in detecting restriction fragment length polymorphisms in and surrounding the hMLC1V/Sb gene for the purpose of determining the linkage relationship between hMLC1V/Sb, aminoacylase-1, and the beta-galactosidase-A genes along the short arm of chromosome 3.
APPENDIX

A. Contractile Systems

Introduction

Contractile protein isoforms are expressed in non-muscle and in a large number of different muscle types, which can be classified as smooth muscle (e.g. stomach, intestine, uterus, veins, and arteries), cardiac muscles (atrium, ventricle, and Purkinje fibers), striated skeletal muscle which is further subdivided into fast twitch and slow twitch depending on the speed of fiber contraction. The functional contractile apparatus in these various tissues differs and can be described in three basic scenarios: striated muscle, smooth (non-striated) muscle, and microfilament associated.

1. Striated Muscle. The contractile apparatus of striated muscle is the best characterized and will be discussed in the greatest detail. Striated skeletal muscle and cardiac muscle, a specialized form of striated muscle, are composed of fibers which are multinucleate syncitia of many muscle cells. The principle structural unit of the striated muscle fiber is the sarcomere, which is defined as the region between two Z lines (see Figure 67a). Figure 67a is a longitudinal section though a striated muscle fiber. The muscle fiber is composed of myofibrils of two principle types, polymerized actin thin filaments and polymerized myosin thick filaments. Regions of the sarcomere are defined by
Figure 67. (a) Schematic diagram of striated skeletal muscle indicating the A band, the H band, and the Z line. (b) Cross sections through the different regions of the sarcomere reveal the arrangement of thick and thin filaments.
the relationship of the two fiber types with respect each other
(see Figure 67b). Figure 67b is a diagramatic representation of
cross sections though the various regions of the sarcomere. A
cross section through the I-band reveals actin thin filaments, a
cross section through the H-band reveals myosin thick filaments,
and a cross section through the A-band reveals overlapping thick
and thin filaments arranged in the following pattern; six actin
filaments surround each myosin filament and three myosin filaments
surround each actin filament. The Z-lines which define the
sarcomeric unit contain alpha-actin and play a role in anchoring
the thin filaments thus providing a base for contraction (Masaki
et.al., 1967; Stromer and Goll, 1972). Figure 68 is a schematic
representation of the functional actin thin filament and the myosin
hexamer. Movement in muscle is effected by the sliding of actin
and myosin filaments over each other. This process is driven by
the hydrolysis of ATP and is regulated by a number of calcium
sensitive myofibrillar proteins (see Figure 69). Figure 69 is a
diagramatic representation of all the components of the skeletal
muscle contractile complex. In addition to actin and myosin,
tropomyosin, troponin C, troponin I, troponin T, the myosin DTNB
light chain and the myosin alkali light chain proteins are all
intricately involved.

The functional actin filament is composed of polymerized actin
monomers (MW = 42,000 daltons) which form an alpha-helical
filament. Tropomyosin (MW = 35,000 daltons) is complexed with
Figure 69. A schematic diagram of the striated skeletal muscle regulatory complex. The troponins, T (TnT), I (TnI), and C (TnC) are associated with the actin filament and regulate contraction as described in the text. Actin the myosin heavy chain globular head and tropomyosin are also indicated. The myosin binding site is indicated by the shaded region on the actin molecule.
Figure 68. (a) A schematic diagram of the functional actin filament indicating the alpha helical polymerization of actin monomers. The relative positions of tropomyosin and the troponin complex are also indicated. (b) A diagram of the myosin hexamer indicating the myosin heavy chain and the myosin light chain components.
actin (stoichiometrically 1:7; Potter 1974) in the major groove of the alpha helix and functions both as a structural and regulatory component (Fujime 1973; Oosawa and Asakura 1975). The troponin complex (troponin C, MW = 18,000; troponin I, MW = 24,000; troponin T, MW = 37,000) is also associated with actin in a 1:7 molar ratio (ibid) and functions to regulate the actomyosin crossbridge interaction.

The functional myosin molecule exists as a hexamer (see Figure 68) composed of two myosin heavy chain monomers (MW = 220,000) each with a globular head region and a long rod-like alpha-helical tail. Two myosin DTNB light chain monomers (MW = 20,000) and two myosin alkali light chain monomers (MW = 17,000-20,000) are associated with myosin heavy chain globular head in the proximity of the head-tail junction domain (Flicker et al., 1983; Sellers and Harvey 1984; Hardwicke and Szent-Gyorgyi 1984). The globular head region of the myosin heavy chain contains the actin activated magnesium ATPase active site (Okamoto et al., 1986; Wells and Yount, 1982; Mahmood and Yount, 1984), and the rod domains of the heavy chains are involved in the association of myosin molecules into bipolar thick filaments (Burridge and Bray, 1975; Neiderman and Pollard, 1975; Pollard, 1975).

Different striated muscles (e.g. fast-twitch skeletal, slow-twitch skeletal, and cardiac) contain different isoforms of the various myofibrillar proteins just described, suggesting that these isoforms are associated with the different contractile properties of these tissues.
Initiation of the contractile response occurs upon nervous stimulation of the sarcoplasmic reticulum, calcium ions are released into the cytoplasm from the sarcoplasmic reticulum resulting in a change of the intracellular concentration of calcium from 10^(-8) M to 10^(-6) M (Ashley and Ridgway, 1970; Winegrad 1970). Binding of calcium by troponin C causes a conformational change in the troponin complex. As a result, the myosin binding site on the actin filament now becomes accessible to the myosin heavy chain globular head. Myosin then binds to the actin filament forming the actomyosin crossbridge. Hydrolysis of ATP by the myosin heavy chain actin-activated magnesium-ATPase breaks the crossbridge and permits continued contraction or reversion back to the resting state.

2. Smooth Muscle. Smooth muscle differs from striated muscle in its cellular structure and its mode of regulation of contraction. Smooth muscle is composed of single long tapered mononucleate cells that contain actin thin filaments and myosin thick filaments in a less ordered state with striations being absent. Z-lines are also absent, although alpha-actinin has been isolated from smooth muscle and is primarily associated with the long tapered ends of the cell (Lazarides and Burridge, 1975). Observations from electron micrographs suggest that the actin thin filaments are arranged in a rosette around each myosin thick filament in a molar ratio of 15:1.
The smooth muscle contractile complex includes; two actin isoforms, alpha-sm and gamma-sm (Vandekerckhove & Weber 1981); two smooth muscle tropomyosin isoforms, alpha and beta, that are distinct from other tropomyosin isoforms (Cummins & Perry, 1974); a single smooth muscle myosin heavy chain isoform (Burridge 1974); and single isoforms for each of the myosin light chain proteins (Jakes et.al. 1976). The troponins C, T, and I are not found in smooth muscle.

The mechanism of contraction in smooth muscle is similar to striated muscle, although differences do exist. Movement is effected by the sliding of thin and thick filaments over each other. This process is triggered by an influx of calcium into the cytoplasm which results in a conformational change in the myosin complex and allows the formation of the actomyosin crossbridge. Hydrolysis of ATP by the myosin heavy chain actin-activated magnesium-ATPase breaks the crossbridge and contracton proceeds or the resting state is restored.

The regulation of smooth muscle contraction differs from striated muscle. Regulation takes place in the absence of a troponin regulatory complex. A comparison of isolated contractile proteins from striated and smooth muscle indicates a similar calcium sensitivity, however the mechanism of calcium regulation of the smooth muscle contractile system is myosin-linked, not actin associated. Regulation is modulated by a calcium regulated, calmodulin-activated myosin light chain kinase (cpe 68).
Phosphorylation of the DTNB myosin light chain at a serine residue in the amino-terminal region induces the conformational change in the myosin complex which permits the myosin heavy chain globular head to bind to its actin binding site.

3. Non-Muscle Microfilament Associated. The contractile apparatus in non-muscle cells is associated with cytoplasmic microfilaments. Cytoplasmic microfilaments are composed of filamentous actin (Pollard and Weihing, 1974; Korn 1976) and are distributed primarily in the region immediately underlying the plasma membrane as highly organized bundles of filaments called stress fibers, in broad pseudopodal areas, and in cellular microprojections such as filopodia and microvilli (Pollack et al., 1975). Microfilaments also form the transitory contractile ring of dividing cells (Schroeder 1973), accumulate at the plasma membrane in regions of active phagocytosis (Korn et al., 1974), and are observed in the region between the mitotic spindle poles and the chromosomes during mitosis (Cande et al., 1977; Sanger, 1975). Two types of nonmuscle actin isoforms, beta and gamma, have been found in all nonmuscle cells examined and it has been shown that both are expressed in a single cell (Bravo et al. 1981). Myosin thick filaments analogous to those formed by myosin in skeletal muscle are absent in nonmuscle cells, however the functional myosin hexamer is present and has been identified in a variety of vertebrate cell types (Korn 1978). Immunofluorescence microscopy with anti-myosin antibodies has identified myosin
closely associated with actin microfilaments in the contractile ring of dividing cells, in stress fibers, and in the actin containing region between the chromosomes and the mitotic spindle (Painter et al., 1975; Weber and Goeschel-Stewart, 1975; Fujiwara and Pollard, 1976). Anti-myosin antibodies also diffusely stain the sub-plasma membrane (Heggeness et al., 1977; Olden et al., 1976). Two nonmuscle myosin heavy chain isoforms have been identified and are described as "brain type" and "platelet type" based on peptide differences (Burridge and Bray 1975). Fibroblasts have been shown to express both types of myosin heavy chain molecules, however it is unclear if both are expressed in the same cell (ibid). Myosin light chain isoforms have been identified in a variety of non-muscle cell types and generally consist of a 20,000 dalton DTNB light chain and a 17,000 dalton alkali light chain. Antibodies raised against smooth muscle tropomyosin also react with the stress fibers of cultured mammalian fibroblasts (Lazarides 1975; Lazarides 1976). Tropomyosin has been isolated from numerous mammalian nonmuscle cells and has a molecular weight of approximately 30,000 daltons. Two nonmuscle tropomyosins have been reported in horse platelets (Cote & Smillie 1981), however only one predominant isoform has been observed in both human fibroblast cell lines and in pig platelets (Giometti & Anderson 1981, der Terrossian et al. 1981; respectively). A troponin complex has yet to be purified from any nonmuscle cell. The actual presence of a troponin-like complex in nonmuscle is unproven. Smooth muscle
alpha-actinin antibodies also react with actin microfilaments in mammalian fibroblast stress fibers (Lazarides 1975; Lazarides 1976) and with the cytoplasmic surface of the intestinal microvillus membrane at the apparent points of attachment of the actin microfilaments (Mooseker and Tilney 1975) and may function to anchor the microfilaments to the plasma membrane (Schoenberg and Neadham 1976). Regulation of nonmuscle contraction is currently believed to be synonymous to smooth muscle and involves a myosin-linked phosphorylation event of the DTNB light chain (Kendrick-Jones et al. 1976, Kendrick-Jones et al. 1983).
LIST OF REFERENCES


157


Cohen-Haguenauer et.al. (1987)


159


