



## Durham E-Theses

---

### *Characterization of endoplasmic reticulum from castor bean and the cloning of a plant phosphatase: a basis for comprehensive plant organelle proteomics research*

Maltman, Daniel James

#### How to cite:

---

Maltman, Daniel James (2000) *Characterization of endoplasmic reticulum from castor bean and the cloning of a plant phosphatase: a basis for comprehensive plant organelle proteomics research*, Durham theses, Durham University. Available at Durham E-Theses Online: <http://etheses.dur.ac.uk/4958/>

#### Use policy

---

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

---

Academic Support Office, Durham University, University Office, Old Elvet, Durham DH1 3HP  
e-mail: [e-theses.admin@dur.ac.uk](mailto:e-theses.admin@dur.ac.uk) Tel: +44 0191 334 6107  
<http://etheses.dur.ac.uk>

CHARACTERIZATION OF ENDOPLASMIC  
RETICULUM FROM CASTOR BEAN AND  
THE CLONING OF A PLANT  
PHOSPHATASE

A Basis for Comprehensive Plant Organelle  
Proteomics Research

Daniel James Maltman

Thesis for Ph.D

University of Durham  
Department of Biological Sciences

2000

Supervisor Professor A R Slabas

The copyright of this thesis rests with the author. No quotation from it should be published in any form, including Electronic and the Internet, without the author's prior written consent. All information derived from this thesis must be acknowledged appropriately.



17 SEP 2001

Thesis

2000/  
MAL

## Abstract

The plant endoplasmic reticulum is the location of storage oil and membrane lipid assembly, and for fatty acid modifying reactions (desaturation, elongation, hydroxylation). It therefore represents a source of enzymes involved in these processes. Many of these defy traditional purification strategies. In this study, ER membranes have been isolated biochemically pure and in milligram quantities from the endosperm of developing and germinating castor bean. One-dimensional SDS-PAGE, used to routinely assess sample integrity, showed resolution limitations. Two dimensional gel electrophoresis was optimized regarding sample preparation and solubilization, and reproducible profiles confirmed its suitability as a sound basis for analysis of stage-specific ER components. In large format 2-D experiments, preparative loadings were reproducibly resolved.

MALDI TOF mass spectrometry was evaluated for high throughput peptide signature generation with individual ER components. Resolution problems were again highlighted with 1-D separations, although some functional assignments were made. Subsequently analysis of selected spots from a preparative 2-D gel of germinating ER was used to establish the limitations of the procedure. Database matching of a single component at very low levels of mass error tolerance also demonstrated the power and accuracy of the technology.

Membranes were subfractionated to simplify protein patterns. It is proposed that an organellar approach, including subfractionation, provides enrichment of specific subsets of cellular components. A putative plant phosphatidic acid phosphatase gene has been investigated following identification from the EST database.

The aim of this research is the identification of proteins involved in storage lipid synthesis in castor bean in reactions specific to the endoplasmic reticulum.

# List of Contents

|   | Page         |
|---|--------------|
| Title   | i            |
| Abstract  | ii           |
| List of Contents  | iii          |
| List of Figures   | viii         |
| List of Tables  | xi           |
| Declaration & Statement of Copyright  | xii          |
| Acknowledgements  | xiii         |
| Abbreviations   | xiv          |
| <br><b>CHAPTER 1: INTRODUCTION</b>  | <br><b>1</b> |
| <b>1.1 Biological membranes</b>   | <b>1</b>     |
| <b>1.2 Membrane components</b>  | <b>4</b>     |
| <i>Protein</i>  | 4            |
| <i>Lipid</i>  | 9            |
| <b>1.3 Membrane compositions</b>  | <b>12</b>    |
| <b>1.4 Endomembranes</b>  | <b>13</b>    |
| <b>1.5 Biochemistry of the endoplasmic reticulum</b>                          | <b>16</b>    |
| <i>ER as a site of protein synthesis</i>                                      | 16           |
| <i>Protein glycosylation at the ER membrane</i>                               | 20           |
| <i>Lipid metabolism</i>   | 22           |
| <b>1.6 Cellular protein sorting begins in the ER</b>                          | <b>31</b>    |
| <i>Retention</i>  | 35           |
| <i>Retrieval</i>  | 37           |
| <b>1.7 Endomembranes in seeds</b>   | <b>38</b>    |
| <b>1.8 Experimental perspective of the endoplasmic reticulum</b>              | <b>38</b>    |
| <i>Animal systems</i>   | 43           |
| <i>Yeast and fungi</i>  | 44           |
| <i>Plant systems</i>  | 45           |
| <b>1.9 Endoplasmic reticulum as a rich source of diverse proteins/enzymes</b> | <b>48</b>    |
| <b>1.10 Castor Bean (<i>Ricinus communis</i>)</b>                             | <b>51</b>    |
| <b>1.11 Two-dimensional protein electrophoresis</b>                           | <b>54</b>    |
| <b>1.12 Proteomics – screening expressed components of the genome</b>         | <b>56</b>    |

|                                  |  |    |
|----------------------------------|--|----|
| 1.13                             | Functional genomics  | 58 |
| 1.14                             | The present study  | 60 |
| CHAPTER 2: MATERIALS AND METHODS |  | 63 |
| 2.1                              | Materials  | 63 |
| 2.2                              | Methods  | 63 |
| 2.2.1                            | Protein methods  | 63 |
| 2.2.1.1                          | Bradford protein estimation  | 63 |
| 2.2.1.2                          | TCA-Lowry method   | 64 |
|                                  | <i>DOC-TCA precipitation</i>   | 64 |
|                                  | <i>The Lowry assay</i>   | 64 |
| 2.2.1.3                          | Coomassie Blue R-250 protein staining  | 65 |
| 2.2.1.4                          | Silver staining of <i>in gel</i> protein   | 66 |
| 2.2.1.5                          | Drying of SDS-PAGE Mini Gels   | 66 |
| 2.2.1.6                          | Chloroform/methanol protein precipitation  | 66 |
| 2.2.1.7                          | Trichloroacetic acid/acetone protein precipitation                               | 67 |
| 2.2.1.8                          | N-terminal protein sequencing  | 68 |
| 2.2.2                            | Subcellular fractionation methods  | 68 |
| 2.2.2.1                          | Preparation of biochemically pure endoplasmic reticulum<br>from castor endosperm | 68 |
| 2.2.2.2                          | Preparation of <i>E. coli</i> microsomes   | 70 |
| 2.2.2.3                          | Preparation of avocado microsomes  | 70 |
| 2.2.3                            | Membrane subfractionation procedures   | 70 |
| 2.2.3.1                          | Sodium chloride  | 70 |
| 2.2.3.2                          | Sucrose gradient isolation of a salt-washed ER membrane fraction                 | 71 |
| 2.2.3.3                          | Chloroform/methanol extraction of hydrophobic<br>proteins from ER membranes      | 71 |
| 2.2.4                            | Electrophoretic procedures   | 72 |
| 2.2.4.1                          | Mini SDS-polyacrylamide gel electrophoresis (PAGE)                               | 72 |
| 2.2.4.2                          | PhastGel analysis  | 73 |
| 2.2.4.3                          | Two-dimensional gel electrophoresis  | 73 |
| 2.2.4.3.1                        | Flatbed precast 2-D gels   | 73 |
|                                  | <i>Immobilized pH gradient rehydration</i>                                       | 74 |
|                                  | <i>Preparation and running of first dimension IEF</i>                            | 75 |
|                                  | <i>Preparation and running of second dimension SDS-PAGE</i>                      | 76 |
| 2.2.4.3.2                        | Large format 2-D gels  | 77 |
| 2.2.4.4                          | Electroblotting to PVDF membrane   | 78 |
| 2.2.5                            | Methods associated with MALDI mass spectrometric analysis                        | 79 |
| 2.2.5.1                          | Preparation of Coomassie stained proteins for tryptic digestion                  | 79 |
| 2.2.5.2                          | Tryptic digestion  | 80 |
| 2.2.5.3                          | Preparation of tryptic peptides for MALDI ms                                     | 80 |
|                                  | <i>Direct loading</i>  | 81 |
|                                  | <i>Loading following acetonitrile extraction</i>                                 | 81 |
| 2.2.5.4                          | Silver staining for subsequent ms analysis                                       | 81 |
| 2.2.5.5                          | Preparation of in gel silver-stained protein for MALDI ms                        | 82 |
|                                  | <i>Destaining</i>  | 82 |
|                                  | <i>Digestion</i>   | 83 |
| 2.2.5.6                          | Peptide clean-up by binding to C18 silica  | 84 |

|   |            |
|---|------------|
| 2.2.5.7 Matrix-assisted laser desorption ionization time-of-flight mass spectrometry  | 84         |
| 2.2.6 Molecular biology methods   | 84         |
| 2.2.6.1 Plasmid isolation   | 84         |
| 2.2.6.2 DNA sequencing  | 85         |
| 2.2.6.3 Polymerase chain reaction (PCR)   | 85         |
| 2.2.6.4 Restriction digestion of DNA  | 86         |
| 2.2.6.5 Gel purification of DNA   | 86         |
| 2.2.6.6 Ligation reactions  | 88         |
| 2.2.6.7 Bacterial transformation  | 89         |
| 2.2.6.8 Radiolabelled probe generation  | 90         |
| 2.2.6.9 Northern hybridization analysis   | 91         |
| 2.2.6.10 Southern hybridization analysis  | 92         |
| 2.2.6.11 Autoradiography  | 92         |
| 2.2.7 Biochemical assays  | 92         |
| 2.2.7.1 Phosphatidic acid phosphatase (PAP) assay   | 92         |
| 2.2.7.2 Para-nitrophenylphosphate ( <i>p</i> -NPP) assay  | 93         |
| <br>  |            |
| <b>CHAPTER 3: PREPARATIVE ISOLATION OF ENDOPLASMIC RETICULUM FROM CASTOR ENDOSPERM, CHARACTERIZATION AND SUBFRACTIONATION OF PROTEIN CONSTITUENTS</b> | <b>94</b>  |
| 3.1 Introduction  | 94         |
| 3.2 Isolation of ER from castor endosperm   | 98         |
| 3.2.1 The first step – sucrose density gradient ultracentrifugation   | 98         |
| 3.2.2 The second step - sucrose density gradient flotation ultracentrifugation  | 100        |
| 3.3 Analysis of ER membrane preparations by SDS-PAGE  | 104        |
| 3.4 N-terminal amino acid sequencing – identification of protein disulphide-isomerase and calreticulin  | 106        |
| 3.5 Special considerations regarding sequencing technology and ER protein complexity  | 109        |
| 3.6 Membrane subfractionation procedures  | 110        |
| 3.6.1 Studies on an <i>E.coli</i> microsome fraction  | 110        |
| 3.6.2 Sodium chloride extraction of castor ER membranes   | 113        |
| 3.7 Extraction of ER membranes with chloroform/methanol   | 118        |
| 3.8 Discussion  | 120        |
| <br>  |            |
| <b>CHAPTER 4: TWO-DIMENSIONAL ELECTROPHORETIC ANALYSIS – PARTIAL CHARACTERIZATION AND A COMPARISON OF GERMINATING AND DEVELOPING ER</b>               | <b>123</b> |
| 4.1 Introduction  | 123        |
| 4.2 Two-dimensional electrophoresis of castor endosperm ER  | 125        |
| 4.2.1 Sample preparation  | 125        |
| 4.2.2 2-D electrophoresis of germinating and developing ER  | 126        |



|   |  |         |
|---|--|---------|
| 4.2.3   | Assessing and optimizing reproducibility between 2-D profiles  | 129     |
| 4.2.4   | 2-D electrophoresis for direct comparison of germinating and developing ER                             | 131     |
| 4.3   | 2-D analysis of salt-generated subfractions of castor ER   | 137     |
| 4.4   | Correlating total and subfraction ER profiles  | 143     |
| 4.5   | Discussion   | 149     |
| <br><b>CHAPTER 5: APPLICATION OF MATRIX-ASSISTED LASER DESORPTION IONIZATION TIME OF FLIGHT MS TO PROTEINS OF CASTOR ER</b> |  | <br>152 |
| 5.1   | Introduction   | 152     |
| 5.2   | Generation of sample for proteomic investigation   | 155     |
| 5.3   | Preparative 1-D SDS-PAGE   | 156     |
| 5.4   | MALDI analysis of 1-D separated ER and database searching  | 158     |
| 5.4.1   | Comparison of experimental data with theoretical mass fingerprint profiles of known castor ER proteins | 175     |
| 5.4.2   | Database searching using edited MALDI data   | 176     |
| 5.5   | 2-D gels for proteomic analysis  | 179     |
| 5.5.1   | MALDI analysis of 2-D protein spots and database analysis  | 185     |
| 5.6   | Discussion   | 203     |
| <br><b>CHAPTER 6: ANALYSIS OF A PUTATIVE PLANT PHOSPHATIDIC ACID PHOSPHATASE: FROM EST TOWARDS FUNCTION</b>                 |  | <br>208 |
| 6.1   | Introduction   | 208     |
| 6.2   | Methods specific to this chapter   | 209     |
| 6.2.1   | Identification of an Arabidopsis EST encoding a putative PAP   | 209     |
| 6.2.2   | Ordering of EST 158J20T7   | 210     |
| 6.2.3   | Preparation of 158J20T7 and restriction digest analysis  | 210     |
| 6.2.4   | DNA sequencing of the 158J20T7 insert  | 210     |
| 6.3   | Sequence analysis of the 158J20T7 insert   | 213     |
| 6.4   | Database searches with the putative PAP amino acid sequence  | 213     |
| 6.5   | Other plant PAPs   | 219     |
| 6.6   | Southern hybridization analysis using ARAPAP ORF as probe  | 219     |
| 6.7   | Expression studies of the ARAPAP transcript and protein  | 221     |
| 6.7.1   | Northern analysis  | 221     |
| 6.7.2   | PCR subcloning of the ARAPAP ORF into a bacterial expression system                                    | 225     |
| 6.7.2.1   | PCR amplification  | 225     |
| 6.7.2.2   | Preparation of PCR insert and pET-24a(+) for ligation  | 228     |
| 6.7.2.3   | Ligation reactions   | 228     |

|  |            |
|--|------------|
| <b>6.7.2.2 Attempted overexpression in <i>E. coli</i></b>                    | <b>231</b> |
| <b>6.7.4 Biochemical assay analysis of ARAPAP-induced bacterial cultures</b> | <b>234</b> |
| <b>6.7.4.1 The PAP assay</b>   | <b>235</b> |
| <b>6.7.4.2 Non-specific phosphatase assay</b>                                | <b>238</b> |
| <b>6.8 Discussion</b>  | <b>240</b> |
| <b>General Discussion</b>  | <b>243</b> |
| <b>Appendices</b>  |            |
| <b>Appendix 1a Tryptic digest map – calreticulin</b>                         | <b>247</b> |
| <b>Appendix 1b Tryptic digest map – PDI</b>                                  | <b>248</b> |
| <b>Appendix 2a Tryptic peptides – human keratins</b>                         | <b>250</b> |
| <b>Appendix 2b Porcine trypsin autolysis peptides</b>                        | <b>251</b> |
| <b>Appendix 3 Expression vector pET-24a(+)</b>                               | <b>252</b> |
| <b>References</b>  | <b>253</b> |

## List of Figures

|   |            |
|---|------------|
| <b>1.1 Schematic diagram of a biological membrane showing integral and peripheral membrane proteins</b>                       | <b>2</b>   |
| <b>1.2 Structure of a typical GPI anchor</b>  | <b>6</b>   |
| <b>1.3 Integral membrane protein topologies</b>   | <b>8</b>   |
| <b>1.4 The major fatty acids and complex lipids found in the membranes of plant cells</b>                                     | <b>11</b>  |
| <b>1.5 The compartmentalization of lipid metabolism in oil storing plant organs</b>   | <b>23</b>  |
| <b>1.6 Phospholipid biosynthesis in castor bean endosperm</b>   | <b>28</b>  |
| <b>1.7 Vesicle shuttling is a process mediated by a variety of protein factors and requires GTP and ATP</b>                   | <b>32</b>  |
| <b>1.8 Model for maize embryo oil body synthesis and degradation during seed maturation and postgermination</b>               | <b>39</b>  |
| <b>3.1 Various states of 5-day germinated castor seed</b>   | <b>99</b>  |
| <b>3.2 Appearance of the first preparative sucrose density gradient following ultracentrifugation</b>                         | <b>101</b> |
| <b>3.3 Typical Lowry estimation of ER protein content using a BSA standard curve</b>  | <b>103</b> |
| <b>3.4 Progressive fractions of castor endosperm endoplasmic reticulum preparations</b>                                       | <b>105</b> |
| <b>3.5 Identification of protein disulphide-isomerase (PDI) and calreticulin from germinating ER</b>                          | <b>108</b> |
| <b>3.6 Sequential extraction of E.coli microsomes with 500 mM sodium chloride</b>   | <b>112</b> |
| <b>3.7 Micro-scale (50 <math>\mu</math>l) sequential extraction of germinating ER in 500 mM sodium chloride</b>               | <b>114</b> |
| <b>3.8 500mM sodium chloride sequential extract and twice-washed membrane subfractions from germinating and developing ER</b> | <b>116</b> |
| <b>3.9 Mini SDS-PAGE profiles of germinating and developing ER sodium chloride subfractions</b>                               | <b>117</b> |
| <b>3.10 Chloroform/methanol extraction of developing ER membranes</b>   | <b>119</b> |

|  |         |
|--|---------|
| 4.1 2-D separation of ER samples isolated from endosperm of germinating and developing castor seeds          | 127     |
| 4.2 Evaluating gel-to-gel reproducibility of 2-D protein profiles  | 130     |
| 4.3 Two-dimensional calibration grid   | 132     |
| 4.4 Germinating and developing ER 2-D protein profiles   | 133     |
| 4.5 Identification of spot differences between germinating and developing ER profiles                        | 135-136 |
| 4.6 Preliminary 2-D analysis of sodium chloride subfractions of germinating                                  | 138     |
| 4.7 Schematic representation of withdrawal of salt-washed membrane fractions from the resolving sucrose step | 140     |
| 4.8 Salt-extracted and sucrose resolved ER membranes from germinating and developing endosperm               | 142     |
| 4.9 Identification of proteins specific to salt-extracted germinating and developing ER membranes            | 144-145 |
| 4.10 Germinating and developing ER salt extract subfractions   | 146     |
| 4.11 Identification of proteins specific to germinating and developing ER salt extract subfractions          | 147-148 |
| 5.1 Preparative 1-D SDS-PAGE profiles of germinating and developing ER                                       | 157     |
| 5.2-5.16 MALDI TOF ms spectra from 1-D separated ER proteins   | 160-174 |
| 5.17-5.20 Large format 2-D gels for analytical and preparative separation of germinating and developing ER   | 180-183 |
| 5.21 Annotated large-scale 2-D gel profile of germinating ER   | 186     |
| 5.22-5.34 MALDI TOF ms spectra from 2-D separated proteins   | 187-199 |
| 5.35 Database report returned for four query peptides from germinating ER 2-D gel, spot 14                   | 202     |
| 6.1 Enzymatic excision of 158J20T7 from plasmid vector pZL 1   | 211     |
| 6.2 Oligonucleotide primers used for sequencing the 158J20T7 insert  | 212     |
| 6.3 Schematic representation of the 158J20T7 insert sequence   | 214     |
| 6.4 Three frame conceptual translation of the full 158J20T7 insert DNA                                       |         |

|   |     |
|---|-----|
| sequence  | 215 |
| 6.5 Kyte-Doolittle hydropathy plot of the <i>ARAPAP</i> primary sequence  | 216 |
| 6.6 Full alignment between the <i>Arabidopsis</i> ( <i>ARAPAP</i> ) and <i>Saccharomyces</i> amino acid sequences using the Clustal program                                     | 218 |
| 6.7 Clustal amino acid alignment between <i>ARAPAP</i> and a putative PAP from genomic clone AC006200   | 220 |
| 6.8 Southern blot autoradiogram showing hybridization of the <i>ARAPAP</i> ORF to <i>Arabidopsis</i> and <i>Brassica</i> genomic DNA  | 222 |
| 6.9 Northern blot autoradiogram showing expression patterns of an <i>ARAPAP</i> homolog in <i>Brassica napus</i>  | 224 |
| 6.10 Oligonucleotide primers designed for the PCR subcloning of the <i>ARAPAP</i> open reading frame into pET-24a PCR amplified <i>ARAPAP</i> ORF from primers ARAP1b and ARAP2 | 226 |
| 6.11 PCR amplified <i>ARAPAP</i> ORF from primers ARAP1b and ARAP2  | 227 |
| 6.12 Preparative agarose gel for ligation of <i>ARAPAP</i> ORF into pET-24a   | 229 |
| 6.13 Plasmid preparations from 6 positively selected recombinant pET clones   | 230 |
| 6.14 SDS-PAGE whole cell induction profiles from BL21 DE3 (pLysS) empty pET24a(+) and pET24a(+)/ <i>ARAPAP</i> recombinants   | 233 |
| 6.15 Autoradiogram showing the radiolabeled components of the lipid fraction of the phosphatidic acid synthesis   | 236 |
| 6.16 Phosphatidic acid phosphatase assay on avocado microsomes  | 237 |
| 6.17 Measurement of whole-cell non-specific phosphatase activity in <i>E.coli</i> cultures  | 239 |

## List of Tables

|   |            |
|---|------------|
| <b>1.1 Protein biochemical markers for endoplasmic reticulum</b>                                | <b>41</b>  |
| <b>1.2 List of proteins expected to be ER resident/associated</b>                               | <b>49</b>  |
| <b>2.1 Running conditions for Immobiline DryStrip<sup>®</sup> pH 3-10</b>                       | <b>75</b>  |
| <b>2.2 Running conditions for second dimension ExcelGel<sup>®</sup> SDS-PAGE</b>                | <b>77</b>  |
| <b>3.1 Identification of PDI and calreticulin from germinating ER</b>                           | <b>107</b> |
| <b>5.1 MALDI TOF ms data from SDS-PAGE separated castor ER proteins</b>                         | <b>177</b> |
| <b>5.2 Summary of peptide mass fingerprint data giving rise to significant database matches</b> | <b>178</b> |
| <b>5.3 MALDI TOF ms data from 2-D gel spots</b>   | <b>200</b> |
| <b>6.1 Comparison of yeast <i>DPP1</i> and <i>ARAPAP</i> products</b>                           | <b>217</b> |
| <b>6.2 Optimized reaction conditions for PCR amplification of the <i>ARAPAP</i> ORF</b>         | <b>228</b> |
| <b>6.3 Amino acid substitutions in <i>ARAPAP</i> ORF fragments PCR subcloned into pET-24a</b>   | <b>231</b> |

## **Declaration**

None of the material contained in this thesis has previously been submitted for a degree in this or any other university. I declare that, unless otherwise indicated, this thesis is entirely the result of my own work.

## **Statement of copyright**

The copyright of this thesis rests with the author. No quotation from it should be published without his prior written consent and information derived from it should be acknowledged.

Daniel J Maltman

September, 2000

## Acknowledgements

The author wishes to express his gratitude to Professor Toni Slabas for his committed support and guidance, and for his unfailing interest in the work.

He also acknowledges help gained from other members of the Lipid Molecular Biology Group. In particular Bill Simon offered generous technical expertise, and Adrian Brown advice on specific aspects of molecular biology. Thanks are due to John Gilroy for protein sequence information, and Paul Sidney for colour photography.

Elsewhere, Sean Coughlan kindly supervised an important training secondment at Pioneer Hi-Bred Int. in Iowa, and Mike Dunn at Harefield Hospital provided some valuable assistance on larger scale 2-D gel analysis.

Finally, acknowledgement is made of the post-graduate funding for the work provided by the BBSRC.



## Abbreviations

|                    |  |
|--------------------|--|
| ABRC               | Arabidopsis Biological Resource Center                         |
| ATP                | Adenosine triphosphate   |
| BLAST              | Basic Local Alignment Search Tool                              |
| bp                 | Base pair  |
| BSA                | Bovine serum albumin   |
| CAPS               | 3-[cyclohexamino]-1-propanesulphonic acid                      |
| CHAPS              | 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulphonic acid |
| cDNA               | Complementary DNA  |
| CoA                | Coenzyme A   |
| DAF                | Days after flowering   |
| DAG                | Diacylglycerol   |
| dbEST              | Database of expressed sequence tags                            |
| ddH <sub>2</sub> O | Double-distilled water   |
| DPM                | Disintegrations per minute                                     |
| DTT                | Dithiothreitol   |
| EDTA               | Ethylenediaminetetraacetic acid                                |
| ER                 | Endoplasmic reticulum  |
| EST                | Expressed sequence tag   |
| <i>g</i>           | Gravitational force  |
| IEF                | Isoelectric focusing   |
| IPG                | Immobilized pH gradient  |
| kDa                | KiloDalton   |
| MALDI              | Matrix-assisted laser desorption/ionization                    |

|               |                                       |
|---------------|---------------------------------------|
| MOWSE         | MOlecular Weight SEarch               |
| ms            | Mass spectrometry                     |
| OD            | Optical density                       |
| ORF           | Open reading frame                    |
| PA            | Phosphatidic acid                     |
| PAGE          | Polyacrylamide gel electrophoresis    |
| PAP           | Phosphatidic acid phosphatase         |
| PCR           | Polymerase chain reaction             |
| PDI           | Protein disulphide-isomerase          |
| Pi            | Inorganic phosphate                   |
| pmol          | Picomole                              |
| PMSF          | Phenylmethanesulphonylfluoride        |
| <i>p</i> -NPP | <i>p</i> -Nitrophenylphosphate        |
| PVDF          | Polyvinylidene fluoride               |
| rpm           | Revolutions per minute                |
| SDS           | Sodium dodecylsulphate                |
| TAG           | Triacylglycerol                       |
| TEMED         | N,N,N',N'-tetramethylethylene diamine |
| TOF           | Time of flight                        |
| Tris          | Tris(hydroxymethyl)aminoethane        |
| w/v           | Weight for volume                     |
| w/w           | Weight for weight                     |

# CHAPTER 1

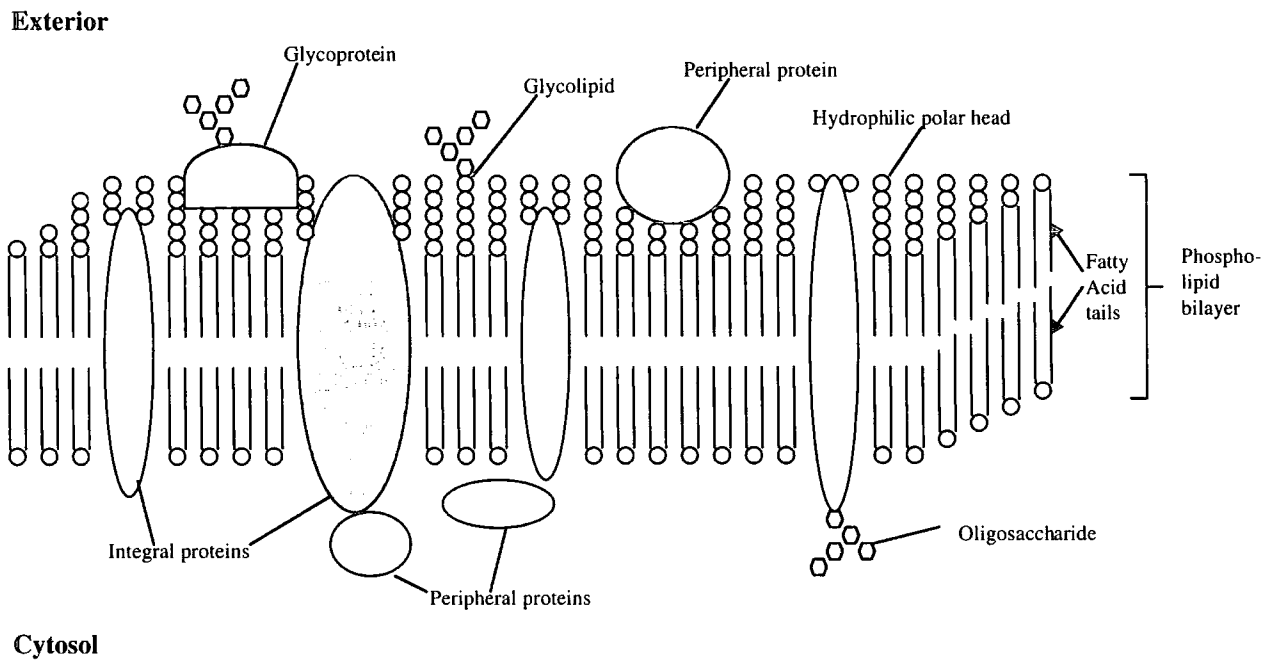
## Introduction

### 1.1 Biological membranes

Biological membranes are composed mainly of proteins and lipids. They define the outer perimeter of the cytosol as well as intracellular organelles. The major membrane lipid components are phospholipids which give rise to a bilayer structure. Bilayer formation results from the amphipathic nature of these lipids; a non polar portion comprising two fatty acyl 'tail' groups esterified to a glycerol backbone, which in turn has a polar 'head' group attached. The size of the hydrocarbon tails favours bilayer and not micelle formation in aqueous solution. In addition, bilayers have a tendency to be extensive, to close in on themselves, and to self seal. These features result from shielding of the hydrophobic tails from aqueous surroundings.

Proteins may be associated with the membrane in a number of ways, and generally it is the proteins within a membrane which mediate its specific functions. The fluid mosaic model (Singer and Nicolson, 1972) describes biological membranes as two-dimensional solutions of proteins and lipids in which the proteins are free to diffuse in the lateral plane (see figure 1.1). A membrane can be considered as a selective permeability barrier which controls the passage of certain molecules and solutes and thereby serves to maintain the internal composition of the cell or organelle which it defines. It can also be regarded as a lipid





**Figure 1.1: Schematic diagram of a biological membrane showing intergal and peripheral membrane proteins.** Phosphoglycerolipids provide the basic membrane bilayer structure. Integral proteins are held in thhe membrane by hydrophobic regions of amino acid sequence. Peripheral proteins interact with integral membrane proteins or membrane lipids.. Carbohydrate components can be attached to both lipid and protein.

Taken from Darnell, Lodish & Baltimore, 1990 in *Molecular Cell Biology* 2nd Ed.

Scientific American Books.

medium for the action of various associated proteins. Lipids can diffuse in the membrane to a limited extent. A general feature of membranes is molecular asymmetry, with the two faces of the structure having different compositions. Asymmetry depends on the manner in which different constituents are inserted and is maintained during membrane synthesis by extension of pre-existing membrane and also because proteins cannot rotate from one side of the membrane to the other. Membrane lipids are also restricted (but not absolutely) in their rotational motion within the bilayer.

Besides compartmentalization and action as permeability barriers, and providing a lipophilic environment for some proteins, membranes provide the site for a wide variety of biochemical/cellular processes and it is the specific proteins contained within them which mediate these functions. Many biochemical processes take place at membranes, including biosyntheses, energy conversion, and even control of the water content of the cell. Membranes also control communication processes between cells, signalling within cells, and regulate transport across and between themselves of simple metabolites such as sugars and amino acids (Reinhold and Kaplan, 1984), as well as complex macromolecules. The ability of internal membranes to break down and subsequently reassemble is required for the equal (or nearly) distribution of this material between daughter cells during cell division (Warren, 1993). Some of these functions will be discussed in more detail presently.

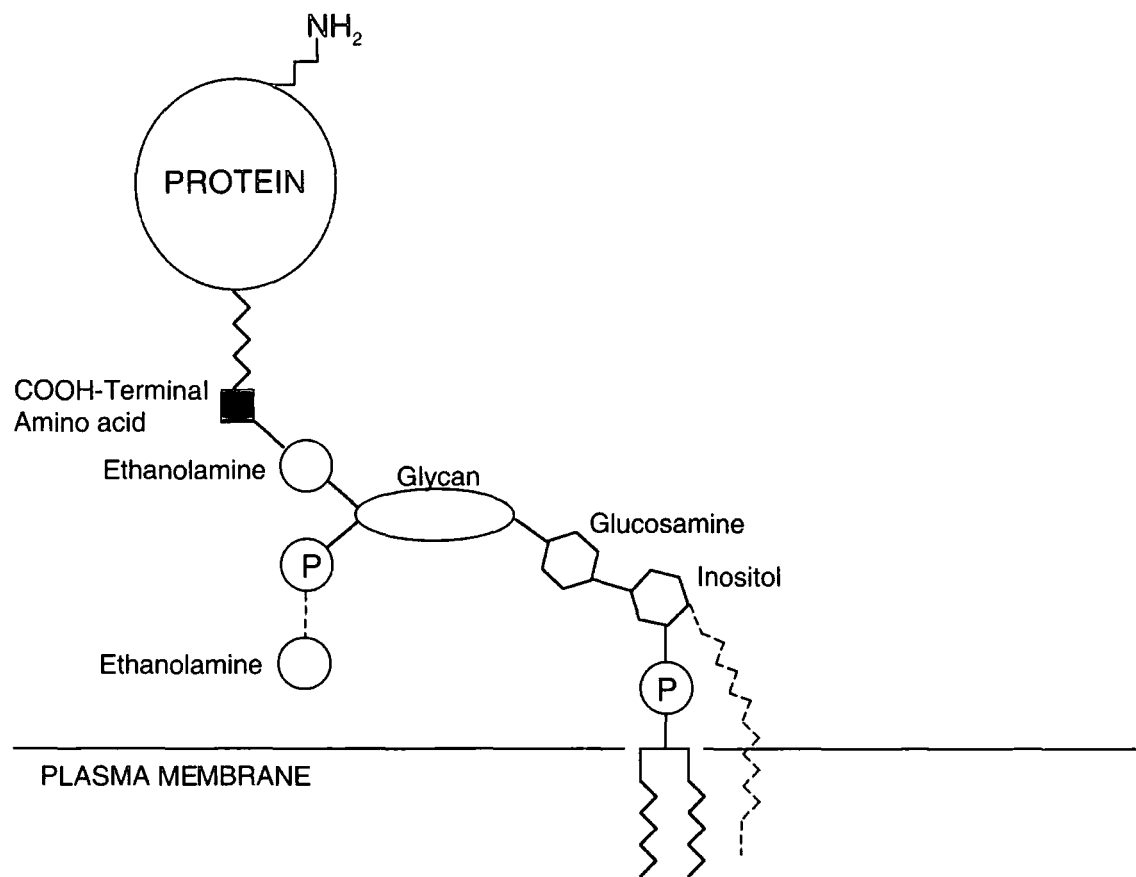
## 1.2 Membrane components

### *Protein*

The nature of its interaction with a membrane can be used to classify a protein as either integral or peripheral. Peripheral proteins are commonly held to membranes via interaction with long chain acyl groups contained in the bilayer (Schultz *et al*, 1988). These fatty acids may be attached directly to an amino acid in the protein or may form part of a more complex structure. Direct attachment to amino acid occurs in several different ways, including ester linkage at serine or threonine side chains, cysteine thioesters, unspecified esters and amide linkage (examples are to be found in Schultz *et al*, 1988). The fatty acid attachment increases overall hydrophobicity of a protein and many acylated proteins are associated with membranes. However, this membrane association may not depend on the fatty acid since some acylated proteins are soluble and secreted, for example human apolipoprotein B which contains palmitate and stearate (Huang *et al*, 1988). Enzymes involved in the attachment of acyl groups to proteins by oxyester bonds have not been well studied. On the other hand, the enzymology of thioesterification at cysteine residues is slightly better understood and activities have been associated with microsomal, particularly endoplasmic reticulum (Berger and Schmidt, 1985) and Golgi-rich, membrane subcellular fractions. The process of thioesterification can occur as an early or late stage of post translational modification. Proteins may similarly be anchored to membranes by farnesylation (eg. ras p21) or geranylgeranylation (eg. rab family of small G-proteins). Protein isoprenylation in plants has recently been reviewed (Crowell, 2000), and previous studies on tobacco suspension cultures have shown that

proteins modified by isoprenoid groups are mostly associated with intracellular membrane fractions (Randall *et al*, 1993). Acylation of proteins via amide linkage at the amino-terminal amino acid residue is interesting in that myristate, an otherwise rare fatty acid, is specifically required. A common example of a more complex attachment between a protein and lipid species is the glycerophosphatidylinositol (GPI) anchor (figure 1.2) which is a major means of anchoring proteins to the surface of eukaryotic cells where they function in cell-cell recognition, cell adhesion, transport, cell surface modulation, or as enzymes or receptors. The biosynthesis of GPI proteins involves two separate pathways; the production of GPI and the attachment of this moiety to the post translationally processed protein. Biosynthesis of the GPI unit occurs on the cytosolic face of the endoplasmic reticulum membrane whereupon it is translocated so that it is oriented to the lumen in association with the luminal face (Vidugiriene and Menon, 1994). A carboxy-terminal hydrophobic signal peptide on the nascent GPI protein is recognised by the putative transamidase which forms the attachment between the protein and the GPI anchor. GPI anchors are found in plants and the existence of multiple GPI-anchored proteins has recently been demonstrated in *Arabidopsis* (Sherrier *et al*, 1999). At least one of these is an arabinogalactan protein (AGP) which together with other GPI anchored proteins is secreted into the extracellular matrix *in vivo*.

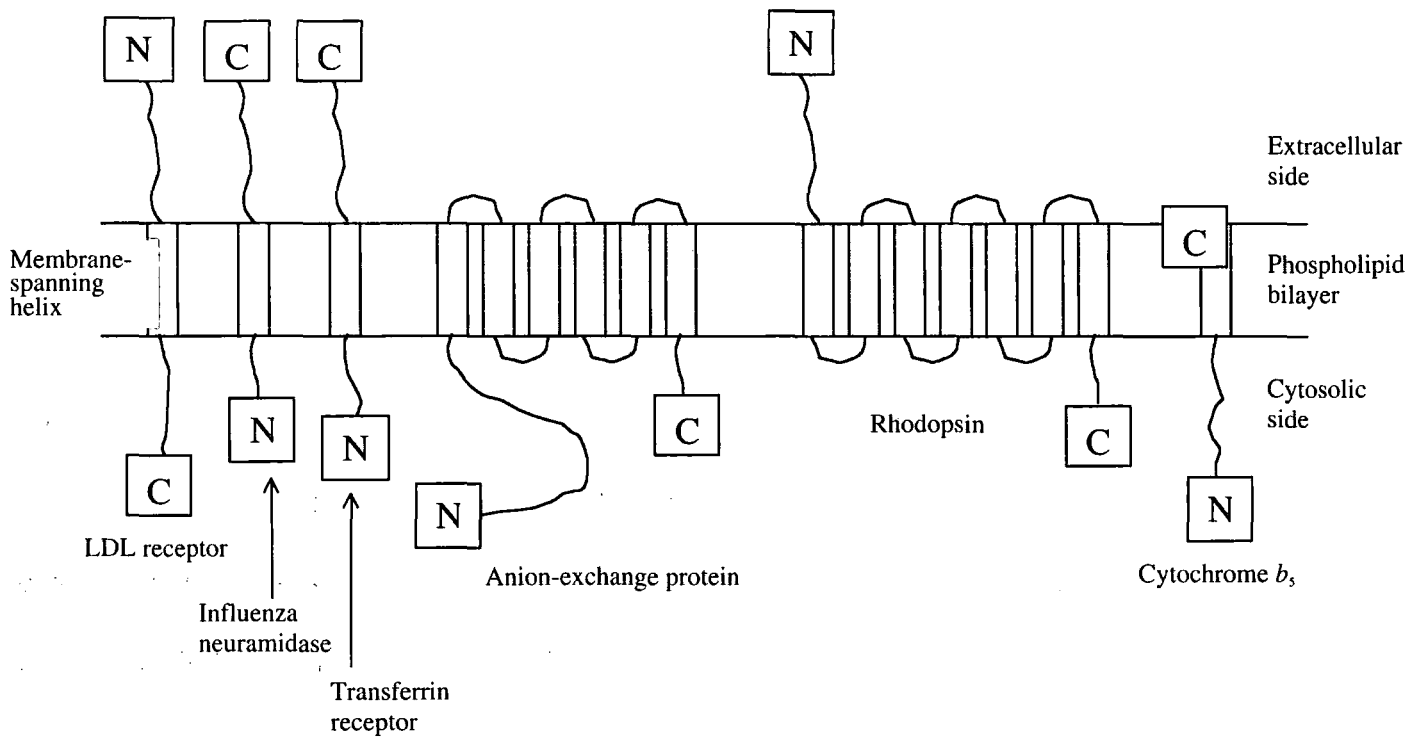
The sidedness of proteins with regard to each face of the membrane is crucially important to their correct functioning. Peripheral proteins are mostly found at the cell surface, and sidedness refers to whether a protein faces the outside or the inside of the



**Figure 1.2: Structure of a typical GPI anchor.** The anchor is a complex structure containing neutral and amino sugars (ethanolamine and phosphatidylinositol) linked to the protein by an amide bond between the C-terminal amino acid residue and the ethanolamine moiety. Redrawn from Udenfriend and Kodukula, 1995.



cell. Integral membrane proteins contain regions of amino acid sequence which actually insert into and across the lipid bilayer. As would be expected, these regions are generally hydrophobic by nature. Different integral membrane proteins vary widely in their topologies (see figure 1.3). They may traverse the membrane in which case their N-terminus might be on the cytosolic side (eg. influenza neuraminidase) or on the extracytosolic face (eg. VSV G-protein, glycophorin) of the membrane. Various multipass topologies have been identified in which the membrane is traversed several times by the polypeptide chain (eg. the seven transmembrane helices receptor superfamily which includes rhodopsin). However, it is important to note that only a small number of such proteins have actually been well characterized and most models are based on amino acid sequence homology to these. Multiple integral membrane proteins may be required to come together to form a functional assembly. This is commonly seen with membrane channels. A common feature of models for such proteins is membrane spanning  $\alpha$ -helices, each of at least 22 amino acid residues in length. The assumption that these helices are directly involved in the transport function of such a channel has been questioned and alternative proposals made (Lodish, 1988). It is clear that membrane proteins often require the lipid environment afforded by the membrane in order to function efficiently, or indeed to function at all. Examples of membrane requirement include bacterial glycerol-3-phosphate acyltransferase (Larson *et al*, 1980) and oligosaccharyltransferase (Chalifour and Spiro, 1988). It has been known for some years that integral membrane proteins restrict the mobilities of lipids which are in their immediate vicinity and that they modify lipid properties when introduced into model membrane systems. It is not



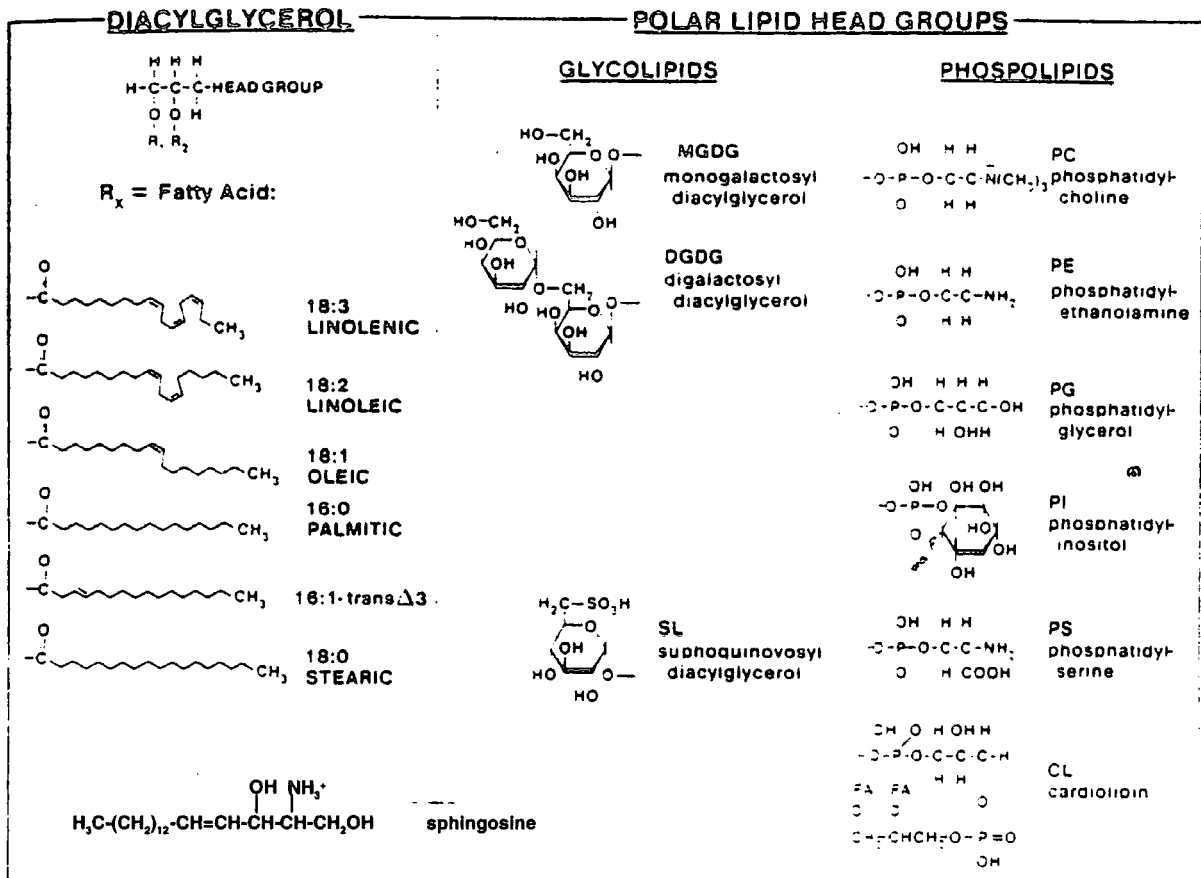
**Figure 1.3: Integral membrane protein topologies.** A single polypeptide may traverse the bilayer once or several times, and the N-terminus may be oriented on either side (ie Cytosolic or 'extracellular' face) of the membrane. After Wickner & Lodish, 1985.

surprising then that the opposite also applies, activities of certain proteins being sensitive to the lipid environment. The activity can alter due to the chemical nature of surrounding lipid or in response to environmental change (eg. temperature) which modifies the physical state of the lipid. The importance of lipids with respect to protein activities may come down to maintenance of active configuration, orientation of the protein at the edge of the membrane, or to simply the provision of a non-polar environment for optimal interaction with a certain substrate. Different membrane-associated enzymes appear to have varying degrees of specificity in their lipid requirement. Factors such as optimum chain lengths and level of unsaturation of acyl groups, and in some cases even the specific head group, may be important. For purified sarcoplasmic reticulum  $\text{Ca}^{2+}$ -ATPase (which accumulates calcium ions), the acyl chains of immediately surrounding lipid must be greater than 14 carbon atoms and unsaturated. Also, Phosphatidylcholine is more effective than phosphatidylethanolamine, although both are required for activity. An example of even greater specificity is mitochondrial  $\beta$ -hydroxybutyrate dehydrogenase which has an absolute requirement for phosphatidylcholine.

### *Lipid*

All animal cell membranes contain glycerophospholipids as the predominant lipid species but sphingolipids and cholesterol are also present in significant proportions, particularly in the plasma membrane. Sphingolipids are also synthesized in plants and serine palmitoyltransferase (which catalyses the first step in the synthesis of the sphingolipid chain) has been localized to the endoplasmic reticulum in squash (Lynch

and Fairfield, 1993). The differences in membrane lipid compositions between different organisms (animals, plants, yeast) and also between different intracellular organelles will be considered later. Glycerophospholipids (or phospholipids) can vary in the type of polar head group they possess or the fatty acyl groups it contains (see figure 1.4). They tend to have preferred acyl groups depending on their head group. The most common types are phosphatidyl-inositol (PI), -serine (PS), -ethanolamine (PE) and -choline (PC). Sphingolipids differ from glycerolipids in that the substituted backbone is sphingosine, giving rise to a hydrophobic portion called ceramide. While in animal cells this molecule potently inhibits protein kinase C, sphingosine in plants has been shown to strongly stimulate activity of the vacuolar pyrophosphatase in tonoplast-enriched microsomes and intact vacuoles (Bille *et al*, 1992). The headgroup may be phosphate containing as in phospholipids or may be a sugar species of varying chain length (glycosphingolipid). Carbohydrate is often attached to both lipids and proteins. An example of the former is the sulphated lipo-oligosaccharide involved in host specificity for nodulation of *R.meliloti* (Truchet *et al*, 1991). Glycosylation of proteins is discussed in detail later. The other main membrane lipid is sterol. Sterols are all based around a common structure. Those commonly found in animals, plants and yeast are all exclusive to these organisms. Other lipids found in membranes include carotenoids for example  $\beta$ -carotene which has antioxidant properties, and dolichols which are involved in core glycosylation.



**Figure 1.4: The major fatty acids and complex lipids found in the membranes of plant cells.** All are amphipathic, containing both a polar and nonpolar moiety. Sphingosine, which substitutes the glycerol backbone in sphingolipids, is shown. Adapted from Ohlogge & Browse, 1995.

### 1.3 Membrane compositions

Although, in general, biological membranes contain many of the same components, it is evident that differences exist not only between membranes of the cells of different organisms, but also between cells within an organism and indeed between the membranes of a single cell. One would expect significant protein compositional variety amongst cell/membrane types since specialized membrane functions are mediated by proteins. However, differences are also observed for membrane lipids. In erythrocyte membranes from different animals, PC and SM (Sphingomyelin) joint contributions are maintained relatively constant whilst their individual proportions can vary widely. PS, PI and DPG (diphosphatidylglycerol) also constitute a relatively constant collective proportion. Cholesterol is often equimolar with phospholipids in a wide range of animal plasma membranes which are also relatively rich in sphingolipids. Mitochondria, nuclear envelope (NE) and endoplasmic reticulum (ER) on the other hand contain little sphingolipid or cholesterol. Golgi, vesicles and secondary lysosomal membranes have lipid compositions intermediate between plasma membrane and major organelles. Mitochondria contain substantial amounts of DPG, which is strongly negatively charged, but which is not found in other animal membranes (this provides support for the idea that the mitochondrion evolved from symbiotic eubacteria, many of which contain DPG as a major membrane lipid). Difference between the classes of membrane lipids are seen in the fatty acyl groups they contain (which are unbranched, contain even numbers of carbon atoms and have cis- double bonds). The characteristic differences might result in significantly higher levels of unsaturation in the lipids of some cytoplasmic organelles than in the plasma

membrane lipids of the same cell, and the combined effect of lower unsaturation and higher cholesterol may be that the plasma membrane is less fluid than the membranes of organelles.

The organelles common to plant and animal cells appear to be of very similar lipid compositions (except sterols which are exclusive). Chloroplast membranes however contain mono- and di- galactosyldiacylglycerol (MGDG, DGDG), sulphated glycosyl diacylglycerol and phosphatidylglycerol (PG). As with mitochondria, these observations give an indication of evolutionary origin, in this case suggesting that chloroplasts and blue-green algae share a common ancestor. Other generally occurring phospholipids (PC, PE and PI) found in chloroplasts could be acquired from the cell's general biosynthesis.

The bulk of the yeast protoplast membrane lipid is ergosterol, and phospholipid accounts for only 15-20% of total lipid. Major individual phospholipids are PE and PC and charged phospholipids (PS, PI and PA) account for the rest (almost 40%).

#### **1.4 Endomembranes**

Just as the plasma membrane defines the outer confines of the cytosol, biological membranes also contain various intracellular compartments, or organelles. In the case of mitochondria and chloroplasts, a double membrane arrangement is seen, and indeed a separate membrane system (thylakoid stacks) exists within the chloroplast.

Other endomembrane-bound compartments include the endoplasmic reticulum, nuclear envelope, Golgi apparatus, lysosomes, secretory granules, endosomes and transport vesicles. In addition, plant cells contain a vacuole bound by what is termed the tonoplast, and also microbodies (glyoxysomes and peroxisomes: Beevers, 1979).

The ER is classically regarded as comprising three subdivisions; rough ER (rER), smooth ER (sER), and the nuclear envelope (NE). Rough ER is so called because of a studded appearance under the electron microscope which is due to the presence of actively translating ribosomes at its surface. As well as this role in protein synthesis, other functions associated with this organelle include core protein N-glycosylation, fatty acid modification, and the biosynthesis of sterols, phospholipids and triacylglycerols. In addition, the ER is the starting point of the secretory and sorting pathways responsible for the direction of different membrane, lumenal organellar and secretory proteins to their eventual destinations. In the developing seeds of plants, formation of protein and lipid storage bodies occurs from this organelle and plant protein and lipid bodies can be regarded as ER subdomains. Another plant-specific ER function is plasmodesmata-mediated cell-to-cell communication (Cantrill *et al*, 1999). Indeed, the tripartite structure of ER mentioned above appears to be an over simplification and the rough ER for example contains a transitional area of smooth membranes involved in the transport between the ER and Golgi. Although such structures are yet to be confirmed in plants, the use of live cell imaging in the study of endomembrane processes has revealed previously unrecognized phenomena. For example, KDEL-tagged green fluorescent proteins (GFPs) have been used to study



plant ER (Haseloff *et al*, 1997; Boevink *et al*, 1999; Ridge *et al*, 1999) and have identified novel ER membranes of unknown function, namely a fenestrated sheet network and fusiform ER bodies (Ridge *et al*, 1999). Other live cell imaging techniques include the use of dyes (Quader & Schnepf, 1986) and video-enhanced differential interference contrast (Allen & Brown, 1988). Other functional subdivisions of the ER have been identified and include domains involved in actin binding and anchoring to mitochondria, plasma membrane, and vacuole (Lichtscheidl *et al*, 1990), lipid recycling domains (Craig and Staehelin, 1988), and ER which extends between adjacent plant cells through plasmodesmata (Hepler, 1982; Lucas *et al*, 1993).

The Golgi apparatus appears as a stack of flattened membranes and is popularly subdivided into cis-, medial-, and trans-compartments, the cis- face being the closest to and in association with the ER (Pelham, 1995). The Golgi is the major protein sorting centre of the cell and is the site of alteration and elaboration of glycoproteins (fashioning of O-linked sugar units and modification of N-linked oligosaccharide). Live cell images of the Golgi apparatus have also demonstrated novel phenomena involving directional movement through the cytoplasm displaying characteristic stop and start patterns. This motility which is associated with tubular structures of the EF appears as a Golgi network which moves tethered to the ER (Boevink *et al*, 1999). The use of technology of this kind is therefore very powerful for observing the dynamics of intracellular membrane systems. Further novel processes are likely to be revealed through the use of random cDNA:GFP fusions, an approach which has

already been applied in *Arabidopsis* (Culter *et al*, 2000). Markers which simultaneously labeled Golgi and ER (Boevink *et al*, 1998) are capable of highlighting events which closely link the two organelles (for example 'stacks-on-stacks) which would not be seen by single-organelle imaging.

## 1.5 Biochemistry of the endoplasmic reticulum

### *ER as a site of protein synthesis*

The membrane of the endoplasmic reticulum is the site of the synthesis of proteins which, after translocation into the lumen, are either directed by a secretory/sorting pathway or retained in the ER (Palade, 1975). Proteins enter the ER by either a co-translational or post-translational insertion mechanism. Whether a nascent polypeptide emerging from an active ribosome remains in the cytosol for completion of translation, or enters the ER depends upon the presence or absence of a signal peptide. This is a sequence of amino acids (13-36 residues) which serves as an ER localization motif (von Heijne, 1985). Transit sequences can also direct proteins to the nucleus, mitochondria or chloroplasts. The ER signal specifies translocation into the lumen simultaneously with translation. Lysosomal proteins, secretory proteins and plasma membrane-bound proteins are the three major classes synthesized by ER-localized ribosomes. ER signal sequences generally possess a hydrophobic stretch (10-15 residues) preceded by at least one positively charged residue at the amino-terminal end. This amino terminal feature can be neutral or negatively charged in some plant signal peptides (Hesse *et al*, 1993). The signals also contain a cleavage

site at their carboxy-terminal end (preceded by several residues which are more polar than the hydrophobic portion) which is removed by a signal peptidase complex in the lumen (Shelness and Blobel, 1990). These signals are conserved between plants and animals as demonstrated by the correct recognition and processing in transgenic plants of human serum albumin and human immunoglobulin  $\kappa$  chain (Hein *et al*, 1991; Sijmons *et al*, 1990). Evidence suggests that bacterial signals are not equally functional in plants, only a fraction of secreted chitinase being glycosylated in tobacco (Lund *et al*, 1989). Furthermore substitution of a plant signal peptide led to a dramatic increase in glycosylation and secretion efficiency of this protein (Lund & Dunsmuir, 1992). Some proteins have an internal signal which unlike those just mentioned is not cleaved upon entry into the lumen. The generally accepted mechanism for ER localization (via co-translational translocation) of ribosomes involves recognition of the signal sequence by a signal recognition particle (SRP), which binds the exposed sequence in the cytosol and temporarily arrests further elongation. This complex diffuses to the ER membrane (Walter & Blobel, 1981) where SRP binds the SRP receptor or 'docking protein' (Meyer *et al*, 1982). Delivery of the ribosome to the translocase (Gilmore, 1993) involves interaction between ribosome and proteins in the membrane (Kalies *et al*, 1994; Wanker *et al*, 1995), such as the ribosome-binding proteins ribophorin I and II. This causes SRP to dissociate, thus allowing translation to continue. Translocation is an ATP-dependent process. However, the exact mechanisms of the translocation process are not fully understood. Sec61p is a key component of the translocation complex in yeast (Gorlich and Rapoport, 1993). The protein forms cylindrical oligomers within the ER membrane

that can be directionally visualized (Haneim *et al*, 1996). *Arabidopsis* homologues of the  $\beta$  and  $\gamma$  subunits have been described (Hartmann *et al*, 1994). Such findings present evidence for similarities in the translocation process between yeast, mammals and plants. Translocation of integral membrane proteins is more complex than for soluble proteins and involves additional stop-transfer signals on the polypeptide concerned. The additional signal-anchor sequence, like the signal sequence, interacts transiently with the ER translocase, but unlike the signal sequence is not cleaved and moves laterally out of the translocase to become a permanent membrane anchor. The interactions between the signal- and signal-anchor sequences with the translocase have been investigated using poly-leucine hydrophobic peptide regions of various lengths (Nilsson *et al*, 1994). The assays used were based on the fact that the active site of oligosaccharyl-transferase, an enzyme which is an integral part of the ER translocase, could serve as a fixed reference point from which the position of the two types of signal sequence in the translocase could be measured. This is because a 'minimal glycosylation distance' dictates that the acceptor asparagine residue will only be glycosylated when positioned no closer than 12-13 residues downstream of an internal signal-anchor sequence (Nilsson and Von Heijne, 1993). Post-translational translocation is SRP-independent and is mediated by the heptameric Sec/Kar2p complex (trimeric Sec61 complex plus four additional subunits) which is sufficient for insertion into reconstituted proteoliposomes (Panzner *et al*, 1995). Kar2p is the yeast homologue of BiP (see below).

Protein folding in the ER lumen is assisted by a variety of soluble and membrane-associated molecular chaperones. The best characterized chaperone is the binding protein BiP (or HSP70, GRP78). GRP78 was initially identified in mammalian cells and was referred to as BiP subsequent to the finding that it was identical to the immunoglobulin heavy chain binding protein (Munro & Pelham, 1986). In plants, detailed studies have shown that nascent bean phaseolin polypeptides transiently associate with BiP within the ER before they form export-competent trimers (Vitale *et al*, 1995).

Protein disulfide isomerase (PDI) is a homodimer which ensures the correct formation of intra-chain disulfide bonds in the ER. A chaperone-like activity is also served by PDI, and the two activities are performed by different domains (Noiva *et al*, 1993). There is evidence to indicate that PDI-catalysed disulfide bond formation is essential for the correct folding and assembly of maize gliadins into protein bodies (Galili *et al*, 1998). Mammalian studies have shown that the PDI monomer is also the  $\beta$  subunit of prolyl hydroxylase (Koivu *et al*, 1987), a function likely to also be served in plants (Denecke, 1996).

Calnexin and calreticulin are closely-related ER chaperones and have been shown in animals to assist in the folding of glycoproteins (Helenius *et al*, 1997). Calreticulin complexes with BiP, in an ATP-dependent manner, under stress conditions in tobacco. Such associations (Crofts *et al*, 1998; Li *et al*, 1998) have led to the theory of a chaperone matrix in the ER of plants (Hammond & Helenius, 1995). Several

other roles for calreticulin are proposed involving  $\text{Ca}^{2+}$ -binding and storage,  $\text{Ca}^{2+}$ -signaling, cell adhesion, and gene expression.

Another class of luminal chaperones, HSP90s, belongs to the 90 kDa heat shock protein family and appears to have similar function to BiP. Plant homologues have been identified (see Denecke, 1996).

### *Protein glycosylation at the ER membrane*

Once inside the lumen of the ER, many proteins are subjected (following cleavage of signal sequence where applicable) to modification by the acquisition of oligosaccharide units (Czichi and Lennarz, 1977). Sugar moieties can be attached to proteins in either so called N- or O-linked manner. O-linkage occurs at the side group oxygen of serine or threonine residues and is performed in the Golgi apparatus whereas N-linked oligosaccharides are joined to proteins at certain asparagine residues occurring in the sequence Asn-X-Ser/Thr, a process which is initiated in the rough ER. A common pentasaccharide core is seen in all N-linked oligosaccharides after trimming of an initial larger species. N-linked glycosylation is a highly conserved protein modification which occurs in all eukaryotes. Although closely linked to the processes of protein folding and intracellular transport, any dependence of the former on correct N-glycan attachment has yet to be demonstrated. The process of stepwise addition of monosaccharides and then transfer onto protein follows a pathway which is intimately linked to the ER membrane. Synthesis of an immature N-linked oligosaccharide ends with a molecule of composition three glucose (Glc), nine

mannose (Man), and two N-acetyl glucosamine (GlcNAc) residues. The growing oligosaccharide is held to the ER membrane by attachment to an activated lipid carrier, dolichol phosphate. Initially, dolichol phosphate is in the cytosolic face of the membrane and picks up activated sugar residues from the cytoplasm (2 UDP-GlcNAc and 5 GTP-Man). At this point the carrier translocates to the luminal face (Hanover and Lennarz, 1980), and subsequently the oligosaccharide receives new monosaccharides from membrane-bound dolichol phosphate derivatives. Upon completion, the 14 sugar unit is transferred onto the appropriate asparagine residue. This transfer of the precursor oligosaccharide (Glc x 3, Man x 9, GlcNAc x 2) is catalysed by the membrane-bound oligosaccharyltransferase (OST) enzyme. OST is part of the ER translocase assembly and is a heterooligomeric complex. The yeast OST was initially purified as a complex of six subunits:  $\alpha$ ; 60/62/64 kDa (glycoforms),  $\beta$ ; 48 kDa,  $\gamma$ ; 34 kDa,  $\delta$ ; 30 kDa,  $\epsilon$ ; 16 kDa, and  $\zeta$ ; 9 kDa (Kelleher and Gilmore, 1994). Six yeast OST subunit genes have been identified: WBP1 (Te Heesen *et al*, 1992), SWP1 (Te Heesen *et al*, 1993), OST1 (Silberstein *et al*, 1995a), OST2 (Silberstein *et al*, 1995b), and OST3 (Karaoglu *et al*, 1995). Studies on mammalian OST have revealed that both necessary and sufficient for activity is a complex of a 48/50 kD protein and two proteins of approximately 65 kDa related to ribophorins I and II (Kumar *et al*, 1995). The 50 kD subunit has been sequenced and found to be 98% identical to the canine and 93% identical to the avian (Kumar *et al*, 1994) homologs, and 25% identical to the yeast  $\beta$  subunit (Kumar *et al*, 1995).

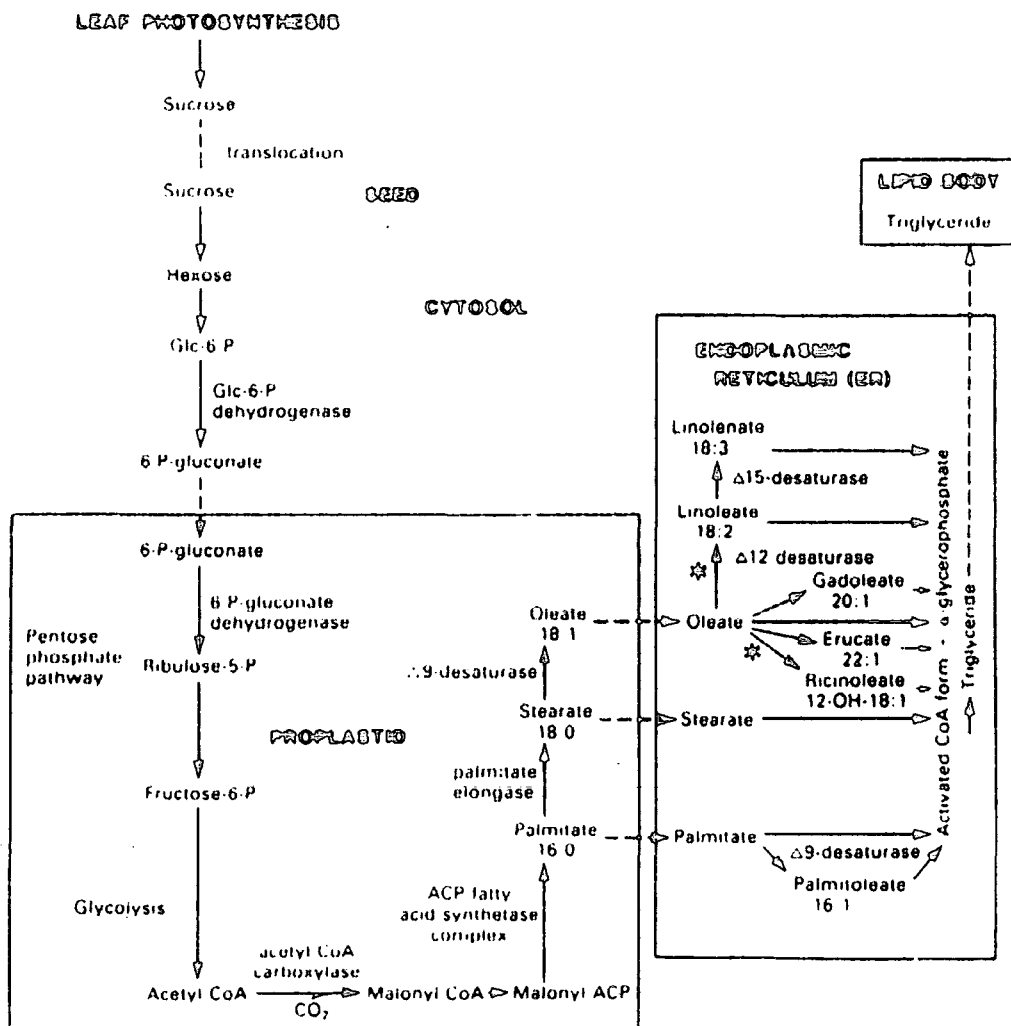
Three glucoses and one mannose are rapidly removed and such trimming is necessary for export from the ER for many glycoproteins. As stated previously, further alteration and elaboration of the carbohydrate portions occurs in the Golgi. Glycosyl units participate in cell-cell recognition and in targeting, for example mannose-6-phosphate targets lysosomal enzymes.

### *Lipid metabolism*

Another aspect of endomembrane-mediated metabolism is the alteration of fatty acids and the assembly of complex lipids. Some of these lipids are membrane components (sphingolipids, MGDG, DGDG, phospholipids) whilst others, triacylglycerols (TAGs) function as storage compounds and do not perform a structural role. Many of the steps of sterol biosynthesis also occur at the ER (reactions with hydrophobic pathway intermediates).

Uniquely in plants, fatty acid synthesis takes place in an organelle, the plastid. Fatty acids are either retained in the plastid or allowed to pass into the cytosol, depending on whether they are transferred from ACP (acyl carrier protein) onto glycerol-3-phosphate or are cleaved from ACP by an acyl-ACP thioesterase respectively. The acyltransferase which catalyses the former is soluble, whereas a second acyltransferase is bound on the inner membrane of this organelle and transfers acyl groups onto the sn-2 position of monoacylglycerol-3-phosphate. The two enzymes have different preferred acyl units and the result is a membrane system which is operating to selectively retain some fatty acids for use in its own composition. Figure





**Figure 1.5: Scheme for the compartmentalization of lipid metabolism in oil storing plant organs.** Reactions take place in the cytosol, the plastid, and the endoplasmic reticulum. Lipid bodies derive from the ER (see also figure 1.6). Asterisks indicate reactions which probably first require incorporation of oleoyl-CoA (18:1) into Phosphatidylcholine and then desaturation to linoleoyl (18:2)-phosphatidylcholine before release of the free fatty acid. Similarly, hydroxylation in ricinoleate formation may also occur on phosphatidylcholine. (Taken from Bewley & Black, 1985. In *Seeds: Physiology of Development and Germination*. Plenum Press, New York).

1.5 shows the scheme for the compartmentalization of lipid metabolism in the oil-storing organs of plants.

More than 75% of the fatty acids in plant tissues are unsaturated. Desaturation reactions take place after fatty acid synthesis. A soluble chloroplast desaturase introduces double bonds at the  $\Delta 9$  position. This enzyme is unique in that all other known fatty acid desaturases are membrane-bound. Membrane-bound desaturases are to be found in the ER, and also in the chloroplast in plants. Fatty acids attached to coenzyme A (those that leave the chloroplast in plants) are available for incorporation into membrane lipids or TAG at the ER and are also subject to desaturation, elongation and other modifications here. Desaturation has a profound effect on the properties of fatty acids and consequently on membranes. Many plants respond to prolonged exposure to cold temperatures by increasing the level of unsaturation in membrane lipids to maintain fluidity in their membranes (Murata, 1983). Glycerolipids in plants are synthesized from palmitoyl and oleoyl groups to start with and are subsequently desaturated to a high degree. Membrane-bound desaturases have proven difficult to characterize because of problems with purification. Alternative approaches including creation of mutants (Somerville and Browse, 1991) in *Arabidopsis*, the use of T-DNA tagging to clone FAD2 which desaturates oleoyl to linoleoyl groups on PC (Okuley *et al*, 1994), and map-based cloning of FAD3 - the gene encoding omega-3 linoleate desaturase (Arondel *et al*, 1992), represent significant advances in the understanding of desaturation. Once one desaturase is

cloned the possibility exists for the isolation of other desaturase genes on the basis of homology as is the case across molecular biology.

Diacylglycerol (DAG) is a common precursor to TAGs and membrane phospholipids. The four enzymes involved in the assembly of TAG in higher plants are also bound in the ER membrane. These include three acyltransferases and a phosphatase and produce TAG from glycerol-3-phosphate by what is known as the Kennedy pathway. Glycerol-3-phosphate acyltransferase (G-3-P-AT) transfers a fatty acyl group from CoA to the *sn*-1 position of glycerol-3-phosphate to produce lysophosphatidic acid (LPA). The plastid G-3-P-AT which is soluble has been purified and even cloned from a number of plants. Membrane-bound enzymes, as with desaturases, are more difficult to characterize. However the *E. coli* enzyme has been purified (Larson *et al*, 1980) and cloned (Lightner *et al*, 1980) and has a molecular mass of 83 kDa. The next enzyme in the sequence, LPA-AT, similarly transfers an acyl group to the *sn*-2 position in the formation of phosphatidic acid (PA), a key intermediate in TAG/phospholipid biosynthesis. Phosphatidic acid phosphatase (PAP) dephosphorylates PA to form diacylglycerol (DAG), after which DAG-AT catalyses the final acyl transfer reaction. At the DAG stage the alternative route is the synthesis of phosphatidylcholine by addition of phosphorylcholine from CDP-choline. It is not clear how control of entry of different fatty acids into either TAG or PC occurs or whether spatial separation of the two forms of synthesis is responsible. PAP clearly lies at a branch point in complex lipid synthesis and is therefore likely to be centrally involved in the control and flux of lipid into each of these groups. There is only

limited knowledge of the various properties, roles and regulation of PAP enzymes. In mammalian systems two isoforms have been identified. PAP-1 represents a cytosolic pool which translocates to the microsomes, in response to certain stimuli, to become actively involved in lipid metabolism. This isoform is characterized by sensitivity to *N*-ethylmaleimide (NEM) and a dependence on magnesium ions. PAP-2 is integral to the plasma membrane where roles in signal transduction have been implicated (Brindley and Waggoner, 1996). PAP-2 is NEM-sensitive and Mg<sup>2+</sup>-independent.

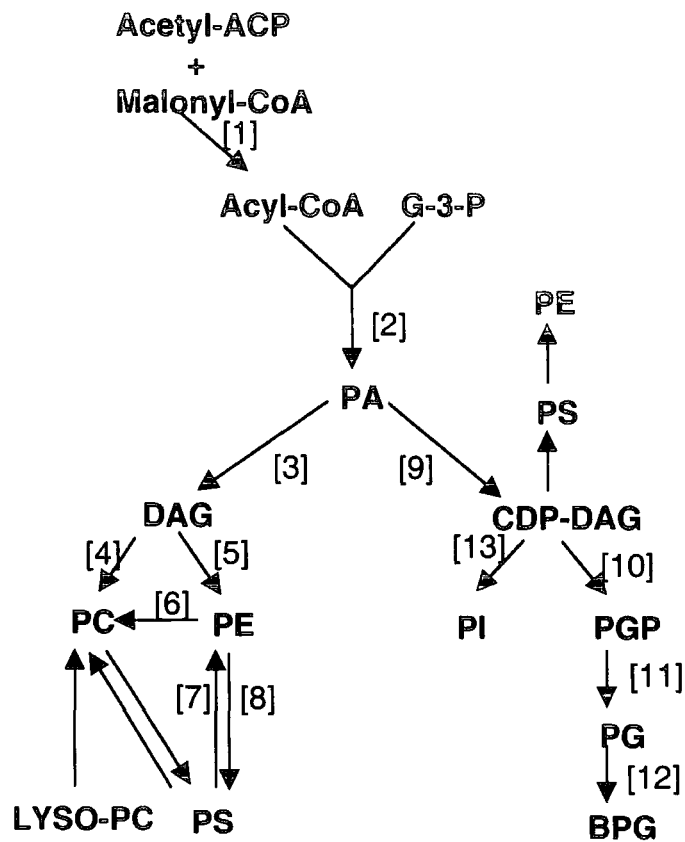
Yeast PAPs have been comprehensively studied in *Saccharomyces cerevisiae*. A Mg<sup>2+</sup>-dependent microsomal PAP was purified (Lin and Carman, 1989). Activity was also dependent on reducing agent and is likely to be involved in glycerolipid metabolism. The 91 kDa enzyme was actually the proteolytic product of a 104 kDa parent protein (Morlock *et al*, 1991) and its purification from a protease-deficient strain revealed it to be enzymatically indistinguishable from the 91 kDa species. An additional 45 kDa isoform from microsomes and mitochondria was identified (using anti-91 kDa antibodies) and was similar enzymatically to the 91 kDa activity (Morlock *et al*, 1991; Carman & Quinlan, 1992). A 75 kDa yeast PAP also exists, in the cytosol (Hosaka and Yamashita, 1984), and has similar properties to the 91 kDa component. Mammalian-like translocation to the ER has however been non-demonstrable for these two enzymes (Morlock *et al*, 1991). More recently, two genes have been isolated which lipid phosphatase activity in yeast. Both activities are independent of Mg<sup>2+</sup> ions and catalyse the phosphohydrolysis of PA. Between them the genes (called DPP1 and LPP1) encode almost all Mg<sup>2+</sup>-independent PAP and lyso-

PAP activities and all diacylglycerol pyrophosphate phosphatase activity in *S. cerevisiae*. Mutant strains revealed that the products of these genes were involved in regulating phospholipid metabolism and cellular levels of PI and PA. Roles in lipid signalling are also proposed (Toke *et al*, 1998a; Toke *et al*, 1998b).

PAP is comparatively poorly understood in plant systems and findings are confined to biochemical analyses. Activity has been detected in the cytosol, the microsomes and in the plastid.

Membrane-bound 1-acylglycerol-3-phosphate acyltransferases (LPA-ATs), mainly in the ER, exist in isofunctional forms displaying different fatty acid specificities and selectivities (Frentzen *et al*, 1983; Hares and Frentzen, 1991). The ER is the major site of phospholipid biosynthesis (Moore, 1982) and therefore the location, although often not exclusive, of many of the enzymes involved. Figure 1.6 outlines the pathway for phospholipid biosynthesis in castor bean endosperm. Although the degree of characterization of some of the enzymes is limited, activities for steps 2-14 (with the exception of reaction 12) have been found in the ER or microsomal fractions. The mitochondria are also the location of many of these enzymes.

Two pathways exist for the synthesis of membrane lipids in plant leaves: the prokaryotic pathway which takes place in the inner membrane of the chloroplast, and the eukaryotic pathway of the ER which begins with PA synthesis. The main differences between the two pathways are the acyltransferase substrates (acyl-ACP



**Figure 1.6: Phospholipid biosynthesis in castor bean endosperm.** 1. fatty acid synthesis; 2. acylation; 3. phosphatidic acid phosphatase (PAP); 4. CDP choline:DAG cholinephosphotransferase; 5. CDP ethanolamine:DAG ethanolaminephosphotransferase; 6. S-adenosylmethionine:phosphatidylethanolamine N-methyltransferase; 7. phosphatidylethanolamine:serine phosphatidyltransferase; 8. phosphatidylserine decarboxylase; 9. CTP:phosphatidic acid cytidyltransferase; 10. glycerophosphate:CDP DAG phosphatidyltransferase; 11. phosphatidylglycerophosphate phosphatase; 12. phosphatidylglycerol:CDP DAG phosphatidyltransferase; 13. CDP DAG:myo-inositol phosphatidyltransferase.

for the prokaryotic pathway and acyl-CoA for the eukaryotic pathway) and the fatty acid compositions of the end products owing to different specificities between the enzymes of the two compartments. Phosphatidylcholine synthesized by the eukaryotic pathway can be transferred to the plastid to contribute to the synthesis of lipids there. Movement of lipids between organelles has been thought to involve lipid-binding proteins (or lipid transfer proteins) although such proteins are now seen as more likely to be concerned with extracellular processes such as wax production. To summarize, phospholipid synthesis mainly takes place in three compartments: the plastid is the site of fatty acid synthesis, with subsequent modifications taking place notably in the ER; headgroups and the glycerol-3-phosphate backbone are mostly formed in the cytosol; the ER plays a major, but by no means exclusive role in the headgroup addition reactions.

The formation of uncommon fatty acids in the seed oils of many plants by elongation of 18:1-CoA is thought to be catalysed by an enzyme system associated with either the ER or the oil body membrane. Fatty acid hydroxylation is another modification which takes place at the ER. In castor bean, a  $\Delta$ -12 hydroxylase has been identified (Van de Loo *et al*, 1995). This enzyme closely resembles a  $\Delta$ -15 desaturase in amino acid sequence terms, although its principle activity is the hydroxylation of oleic acid (18:1) to form ricinoleic acid which is an important ingredient in certain industrial processes. While castor oil contains approximately 90% ricinoleate it is desirable to persuade other species, via transgenic technology, to synthesize this fatty acid. This is because toxic components of the castor seed pose processing problems. Although

the exact mechanisms of ricinoleate are not known, it has been demonstrated that the pathway proceeds from a 2-oleoyl-phosphatidylcholine substrate (Lin *et al*, 1998). It is probable therefore that castor contains acyltransferase enzymes which preferentially incorporate ricinoleate from PC into TAG. The identification of such enzymes presents obvious financial exploitation potential.

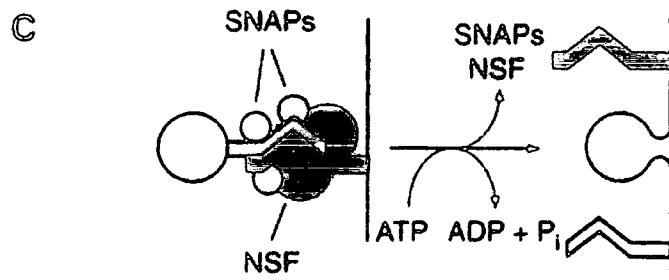
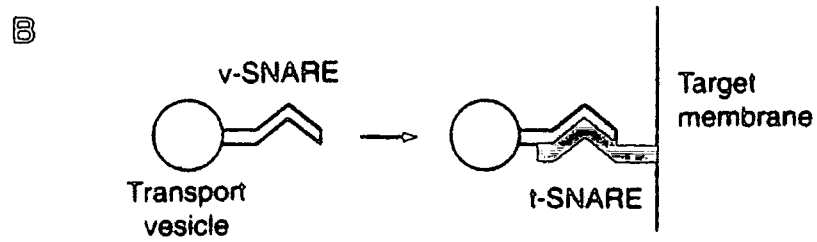
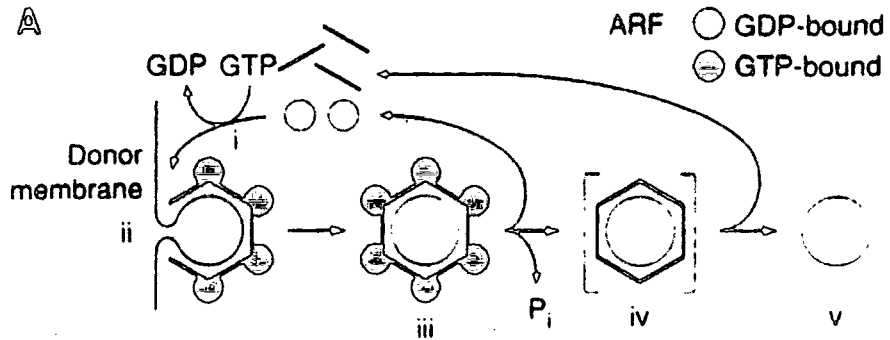
Increased knowledge of plant lipid metabolism can be applied to the development of improved plant oils and therefore wealth creating crops. Specifically, cloning of acyltransferases of different specificities and the creation of transgenic plants can be applied to the production of storage oils of desired fatty acid content at all three positions on the glycerol backbone. An example of such work is the production in oil seed rape of trierucic acid following genetic engineering to introduce a 2-acyltransferase from *Limnanthes douglasii* which specifically introduces erucic acid at the *sn*-2 position of triacylglycerols (Brough *et al*, 1996). No trierucin was detectable in non-transformed rape. Research of this nature aims to specifically reduce the levels of particular fatty acids (eg. palmitic acid in edible oils) or in producing an industrially useful TAG. Proteins involved in lipid metabolism are commonly used as biochemical markers for ER during its isolation. Examples include phosphorylcholine transferase and electron transfer components involved in desaturation such as NADH-cytochrome c reductase.



## 1.6 Cellular protein sorting begins in the ER

Endomembrane systems are responsible for the transport and sorting of proteins (and other molecules) throughout the cell (see Pfeffer and Rothman, 1987 & Rothman and Wieland, 1996 for reviews). Intracellular transfers between membrane-bound compartments of the secretory pathway are essential for cell growth and maintenance. Transport between membranes requires vesicle budding, diffusion, and docking and fusion processes, and is referred to as vesicle shuttling (see figure 1.7). These processes are mediated by the action of a variety of protein factors including vesicle coat proteins, factors which identify membranes called v- and t-SNAREs (vesicle- and target-specific respectively, SNARE standing for SNAP receptor), NSF (N-ethylmaleimide-sensitive fusion protein), and SNAP (soluble NSF attachment protein). Coat subunits from the cytosol assemble on the donor membrane to bring about vesicle formation (budding). Budding is best understood for the COP (coat protein) type coats I and II but two types of clathrin coats have also been identified. COP I and II coats are employed at Golgi and ER membranes respectively and have associated GTPases, ARF (ADP-ribosylation factor: Serefini *et al*, 1991) for COP I and SAR (secretion-associated and ras-superfamily-related) for COP II, which control the budding process. ARF is also required for assembly of clathrin coats. The protein was originally identified as a protein cofactor required in the ribosylation of the heterotrimeric G-protein Gs by cholera toxin. A member of the ARF family of GTP-binding proteins has been cloned from *Arabidopsis* (Regad *et al*, 1993) and has a calculated molecular mass of approximately 20.5 kD. COP I coats have been shown

**Figure 1.7: Vesicle shuttling is a process mediated by a variety of protein factors and requires GTP and ATP.** Assembly of coat components from the cytosol (A) results in the formation of a vesicle which pinches off from the donor membrane. Recruitment of the coat subunits requires that the associated GTP-binding protein (eg. ARF or Sar1p) exchanges GDP for GTP (i). Budding (ii) of COP I coats requires the presence of fatty acyl-CoA. Hydrolysis of bound GTP by the GTP-binding proteins stimulates their own release from the vesicle (iii), leaving a less stable coat (which immediately disassembles in the case of COP I vesicles and also subsequently does the same in COP II vesicles). The uncoated vesicle (v) is now capable of recognizing and binding to the target membrane by means of complementary docking proteins (B). Fusion of vesicle and target membranes is mediated by NSF and SNAP proteins, the latter binding to the SNARE docking complex (or SNAP receptor complex). NSF hydrolyses ATP to disrupt the complex and initiate fusion. (Taken from Rothman and Wieland, 1996).



to mediate vesiculation from ER as well as from the Golgi and carry different cargo to the COP II coated vesicles (Bednarek *et al*, 1995). The GTP-bound forms of the GTPases trigger coat assembly whereas hydrolysis to GDP stimulates coat release. Association of COP I subunits with the Golgi is prevented by the fungal metabolite Brefeldin A which interferes with the initial interaction of ARF with the membrane. The coat proteins of COP I vesicles are collectively referred to as coatomer (Waters *et al*, 1991) which consists of seven subunits:  $\alpha$  (160 kDa);  $\beta$  (107 kDa);  $\beta'$  (102 kDa);  $\gamma$  (97 kDa);  $\delta$  (57 kDa);  $\epsilon$  (34.5 kDa);  $\zeta$  (20 kDa). COP I coats bind dilysine retrieval motifs of the cytosolic domains of secretory pathway membrane proteins via the  $\gamma$ -subunit (Harter *et al*, 1996). Budding of COP II vesicles is satisfied by a group of proteins: Sar1p, Sec23p complex (a 400 kDa complex containing two polypeptides: Sec23p, an 85 kDa Sar1p-specific GTPase activating protein, and a 105 kDa Sec24p protein) and Sec13p complex (a 700 kDa complex composed of two proteins: 34 kDa Sec13p and a 105 kDa polypeptide) in the presence of GTP (Barlowe *et al*, 1994). If ATP is included, these vesicles can go on to fuse with the Golgi membrane. Sar1p is the coat associated GTP-binding protein (referred to as SAR above). Exchange of GDP for GTP on Sar1p is catalysed on the cytosolic side of the ER membrane by Sec12p, an integral membrane glycoprotein reported to reside in the ER and Golgi in yeast (Nakano *et al*, 1988). Homologues of Sar1p and Sec12p have been cloned from *Arabidopsis* (D'Enfert *et al*, 1992). Sec23p induces Sar1p to hydrolyse bound GTP after budding, resulting in the dissociation of Sar1p from the vesicle. If a non-hydrolysable alternative form of GTP is used in in vitro studies, the budded vesicle is unable to fuse with the Golgi (Barlowe *et al*, 1994) because failure of Sar1p and

subsequently COP II proteins to dissociate prevent access of the vesicle targeting (v-SNAREs) such as Sec22p and Bos1p (Lian and Ferro-Novick, 1993) to complementary proteins (t-SNAREs) on the Golgi such as Sed5p (Hardwick and Pelham, 1992). It would seem that the coat must link to the cargo protein, its receptor, and docking components in order for packaging to occur. It is possible that budding is coordinated with signalling pathways as indicated by the fact that coat proteins might associate with polyphosphoinositides or PA as well as with their respective GTPases. Alternatively this could be a mechanism of lipid selection by vesicles. Another constituent of COPII-coated ER-derived transport vesicles, Emp24p, has been characterized by Schimmoller and colleagues (Schimmoller *et al*, 1995). Absence of functional Emp24p in mutant yeast cells causes a delay in the transport of certain secretory proteins to the Golgi. It was proposed that the protein is involved in the sorting or concentration of vesicle cargo since it did not appear to perform a protein folding/assembly function. Periplasmic fusion, that is the pinching off of the vesicle from the donor membrane, requires fatty acyl-CoA for COP I vesicles (Osterman *et al*, 1993). Membrane identifiers mediate docking but it is not clear whether other proteins are involved in vesicle targeting. Cytosolic fusion components which are non-specific include SNAP and NSF. NSF hydrolyses ATP to disrupt the docking assembly, perhaps to irreversibly commit this assembly to the process of fusion by inducing conformational change.

Certain types of transport vesicles may remain docked to target membranes until fusion is triggered by a specific stimulus. For example in nerve endings the rise in

calcium ion concentration accompanying a nervous impulse triggers the fusion of neurotransmitter-containing synaptic vesicles to the presynaptic plasma membrane. VAMP (vesicle-associated membrane protein) is an integral membrane protein of the vesicle and serves as the v-SNARE (Trimble *et al*, 1988). The t-SNARE component consists of two subunits, syntaxin and SNAP-25 (Oyler *et al*, 1989). Similarly in muscle and fat cells, GLUT-4, a facilitative glucose transporter, is vesicle-shuttled from the trans-Golgi to the plasma membrane (from which it is normally excluded) in response to insulin and also in response to exercise in muscle (James and Piper, 1994). Another member of this family of glucose transporters, GLUT-7, is localized to the ER where it forms part of the glucose-6-phosphatase complex involved in gluconeogenesis.

Protein transport by vesicles is less well understood in plants than it is in animals and yeast. For example, a role for COP proteins in ER-Golgi transport is uncertain. However, homologues of other vesicle components of animals and yeast have been identified in *Arabidopsis* (eg. ARF; Sec12p) and other plant species.

### *Retention*

Some of the nascent polypeptides in the ER lumen stay in this organelle despite the efficient export of many other proteins. This phenomenon is termed retention. The mechanism of retention may well involve a signal sequence and such sequences are generally thought to reside in the membrane region in integral membrane proteins. One possible mechanism is that proteins may be selectively retained because they

have energetically favourable interactions with the lipid compositions of the ER (Bretscher and Munro, 1993). In this model, selection is unlikely to be absolute, resulting in the escape of some proteins which ideally would be retained. Retention might arise at regions of the ER (or Golgi) membranes which are unable to vesiculate, this immobility being caused by a matrix deposited on the membrane or the formation of patches on the membrane by retained proteins which are too large to enter vesicles. Retention by protein-protein interactions giving rise to aggregates or protein bodies is a popular notion. Indeed, plants routinely use the ER to store proteins in protein bodies by a series of orderly events of protein synthesis and assembly of properly folded proteins. Such processes seem to involve chaperones. Various putative signals for retention of stored plant proteins have been reported (Altschuler and Galili, 1994; Geli *et al*, 1994) although the differences in structures and lack of universality of these signals amongst retained proteins has cast doubt on whether signals actually exist for this purpose. Quality of interaction between proteins and chaperones may determine which are retained and which are not. Thus a general feature of a protein such as hydrophobicity could determine retention. Retention to some extent may also rely on a protein's dependence on chaperones, less dependent ones being ready for export sooner, never accumulating to a concentration required for aggregation. The luminal chaperones BiP (binding protein), calreticulin and PDI (protein disulphide-isomerase) are all likely to be involved in retention and protein body formation in plants. Calnexin is a novel ER chaperone in that it is membrane-bound (Bergeron *et al*, 1994). It is in fact a major membrane protein of the ER. Interestingly, in binding both soluble and membrane-bound proteins, it appears to play a holding role while

core glycosylation takes place. Functions other than simply that of a molecular chaperone have been proposed for calnexin, for example calcium-dependent retention of ER resident proteins and a role in the sensing or alteration of intralumenal calcium levels (Wada *et al*, 1991). Calnexin-like proteins have been identified in *Arabidopsis* and pea. The *Arabidopsis* cDNA encodes a protein with a predicted molecular mass of 60 kD (Huang *et al*, 1993).

### *Retrieval*

Residency of a protein in the ER or Golgi is not permanent and results not only from retention but also from mechanisms which operate to retrieve resident proteins which have escaped retention. Such transport is mediated by vesicles and is termed retrograde. A carboxy-terminal Lys-Asp-Glu-Leu (KDEL) signal is present in lumenal ER residents of animal cells. In yeast, a similar signal, HDEL is involved in ER retention / retrieval (Pelham *et al*, 1988), for which two receptors, ERD1 and ERD2 exist in the Golgi apparatus (Hardwick *et al*, 1990; Semenza *et al*, 1990). In plant ER lumen-resident proteins, KDEL (Shorosh *et al*, 1993) and HDEL (Vitale *et al*, 1993) can both be seen. For ER membrane-bound proteins, the best known retrieval signal is a tetrapeptide containing a pair of lysine residues (KKXX), also at the carboxy-terminus (Jackson *et al*, 1990). It is likely that the vesicles involved in retrieval are COPI-coated as indicated by binding of the coatomer to the KKXX signal (Cosson and Letourneur, 1994).



A protein directed to the ER which does not possess any sorting or retention / retrieval signals will move through the secretory system by default at a rate dependent on its prevailing ('bulk') concentration in the ER (Wieland *et al*, 1987).

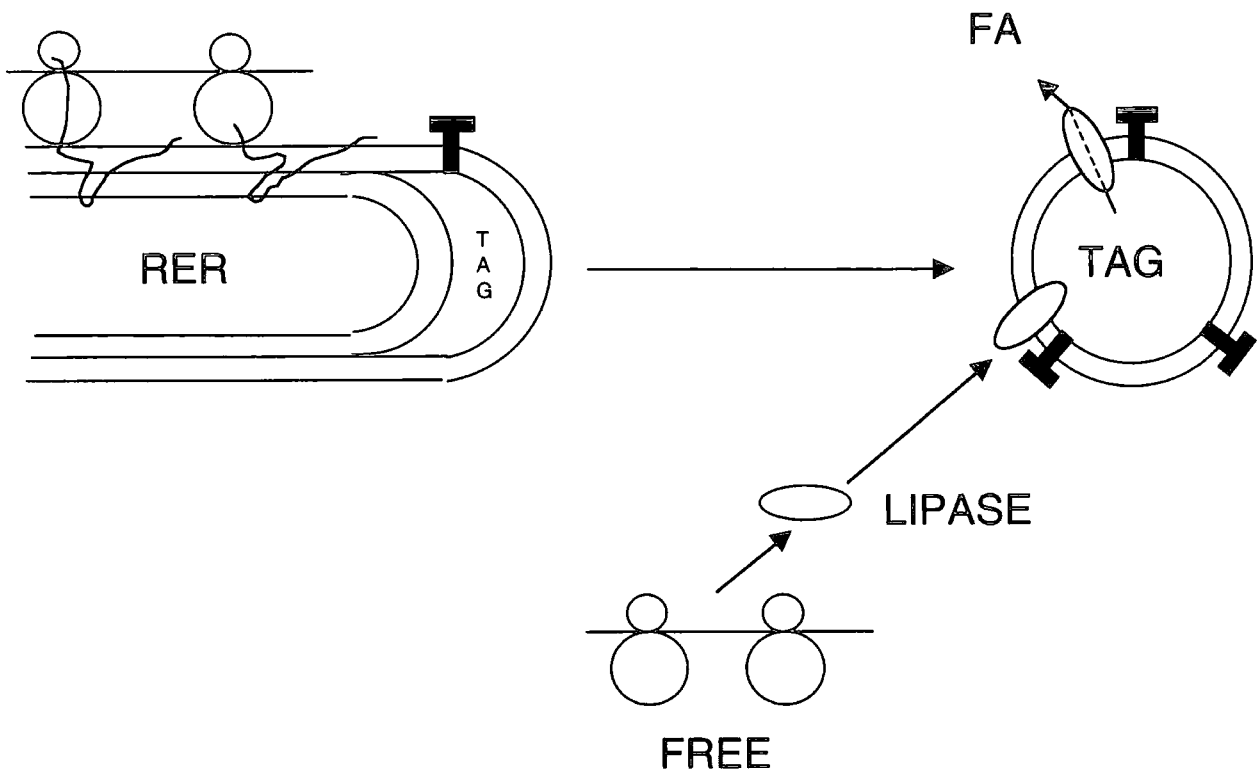
### **1.7 Endomembranes in seeds**

Together with protein bodies, discrete membrane-bound lipid bodies represent the food reserves for the germinating seed. Lipid bodies, or oil bodies (since their lipid contents are usually in the liquid state) are discrete subcellular particles and provide suitable surface area for the action of lipases during the mobilization of the stored TAG. The first extensive review on the subject was provided by Huang, 1992. There is general support for the idea that oil body formation occurs at the ER (see figure 1.8). The oil body consists of TAG surrounded by a phospholipid layer containing proteins called oleosins (Huang *et al*, 1992). Utilization of TAGs in lipid bodies is closely associated with the biogenesis of glyoxysomes and ER has therefore been implicated as the membrane donor for this microbody.

### **1.8 Experimental perspective of the endoplasmic reticulum**

In attempting to isolate a particular endomembrane for the first time there is a requirement for a specific subcellular marker. Such a marker is a characteristic component unique to the organelle concerned which can be assayed throughout isolation procedures. Markers are usually membrane-bound or luminal proteins/

**Figure 1.8: Model for maize embryo oil body synthesis and degradation during seed maturation and postgermination.** The structure comprises a core of neutral lipids surrounded by a phospholipid monolayer containing some proteins, and is formed by budding from the endoplasmic reticulum (ER). During the degradative processes of germination and postgerminative growth triacylglycerol (TAG) is hydrolysed to glycerol and fatty acids (FA) which are converted to carbohydrate for seedling growth. The catalytic lipase is absent in mature seed in most species but increases during these degradative stages. In maize the lipase becomes tightly associated with the oil body prior to hydrolysis. Castor bean is unusual in that an active lipase is already present in the mature, ungerminated seed. This enzyme has a pH optimum of 5.0 and is inactive at pH 7.0. After germination, castor oil bodies contain a lipase activity with converse pH characteristics. It is not known if the lipase in oilbodies is either distinct from or a modified form of the acid lipase. RER: rough ER; FREE: free polyribosomes. (Modified from Huang, 1992).



enzymes (see table 1.1). Alternatively, an exogenous compound which is somehow labelled (usually radiolabelled) may be used which will localize to the compartment for isolation. When choosing a marker, certain criteria should be considered: (i) The marker should be appropriate to the cell type chosen as the experimental material. Certain enzymes for example are not always present in all cell types of an organism/tissue. (ii) The marker should be evenly distributed in the organelle and should not be present to any significant degree in any other organelle/membrane. (iii) Stability of the prospective marker is important so as to allow for easy assaying after purification steps. (iv) Markers should remain with the fraction during isolation to avoid the occurrence of migration artifacts. This is not normally relevant with membrane-bound proteins but is a potential problem when using soluble markers. (v) Attention should be paid to the possibility of inhibition of enzyme markers either by components of buffers or natural factors in the cell homogenate. (vi) The marker should be assayed across the whole range of subcellular fractions obtained and should be capable of providing quantitative data.

Membrane-bound proteins are commonly chosen as biochemical markers for intracellular membrane fractions, for example phosphorylcholinetransferase which is an enzyme involved in ER specific complex lipid assembly (see table 1.1). In addition, electron microscopy can be applied to examine the appearance of the membranes. The effect of magnesium ions (which prevent the dissociation of ribosomes from rER) on the appearance of the membranes and hence their mobility in sucrose gradients can be used as further evidence for successful isolation.

| Fraction              | Marker   |
|-----------------------|--|
| Endoplasmic reticulum | Glucose-6-phosphatase, Phosphoryl-<br>cholinetransferase, cyt b5, cyt P450,<br>Antimycin-insensitive NADH-cyt <i>c</i> -<br>reductase. |

**Table 1.1: Protein biochemical markers for endoplasmic reticulum.**

It is worth noting that the ER is by no means homogenous throughout the cell either morphologically or biochemically. ER exists as subdivisions (rER, sER and NE) and even subdomains (eg. transitional ER, or tER, which is a smooth region of rER observed to be involved in vesicle transport between the ER and Golgi). Biochemical heterogeneity has been reported for various metabolic processes including glycosylation (Conder and Lord, 1983) and lipid synthesis (Lacey and Hills, 1996). The term heterogeneity has also been used in reference to the fact that microsomal fractions from cell homogenates contain membranes derived from ER, Golgi, plasma membrane, mitochondria and other compartments (Beaufay *et al*, 1974), especially when fractionating secretory cells in which different membrane systems have particularly close relationships. Indeed, even with the development of more sophisticated isolation procedures which result in highly purified ER devoid of contaminating endomembranes, the problem of contamination by non-ER derived proteins which adhere to the membrane upon cell disruption remains a possibility.

Such contaminants will falsely appear to be ER-associated and therefore be difficult to expose. Electron microscopy/gold techniques are useful in this respect in that they allow visualization of the membrane location (or otherwise) of specific proteins in the intact cell situation. Similarly, immunofluorescence and immunocytochemistry (see Xu et al, 1996) can be used for observing the location of particular proteins in vivo and to identify in vitro artifacts.

Historically, most work to isolate ER has involved cells from various animal sources, most commonly rat liver. In comparison, plant ER has received relatively little attention mainly because of difficulty in obtaining preparative amounts. Seed material has proved the most useful source of ER from plants and the most widely used has been castor bean endosperm which, being largely liquid is convenient for this type of work. ER from yeast (Swida *et al*, 1982) and fungi (Borgeson and Bowman, 1983) has also been isolated and characterized. Of particular interest with regard to plant ER is the identification and characterization of membrane-bound enzymes of complex lipid metabolism responsible for the assembly of storage triacylglycerol. Detailed knowledge of these proteins is important in that it provides the potential to manipulate lipid metabolism by genetic engineering for the purpose of obtaining valuable oils of desired fatty acid composition. Purification of such proteins has proved difficult, and most of those cloned have been identified with genetic approaches utilizing mutants (Arondel *et al*, 1992; Okuley *et al*, 1994).

Some of the strategies and markers which have been used in isolating ER from various sources, but particularly from plants, are outlined below.

### *Animal Systems*

Microsomes were first studied biochemically in 1938 (Claude, 1938), and later Palade and Siekevitz described the fraction as containing three major components (rough vesicles, smooth vesicles and ribosomes) which were concluded to derive from the ER during the homogenization process (Palade and Siekevitz, 1956). Microsomes were subfractionated using differential centrifugation in homogenous medium, centrifugation through density gradients (Glaumann and Dallner, 1970), through sucrose layers (Chauvea *et al*, 1962) and also by density gradient isopycnic equilibration (Rothschild, 1963). Rat liver has been the most common starting material in the study of ER from animal sources. In a series of papers by Beaufay and colleagues (Beaufay *et al*, 1974b; Amar-Costesec *et al*, 1974; Beaufay *et al*, 1974a), the results of a detailed analysis of the microsomal fraction from rat liver obtained by density gradient centrifugation were reported. By the early 1970s much work had been conducted on microsomal fractions and subfractions. However, apart from the role of associated ribosomes in protein synthesis, little had been learnt from the structural and biochemical relationships in the various subcomponents of microsomal fractions. Beaufay and colleagues took an analytical rather than a preparative approach to addressing the subject of enzymic distribution throughout microsomal fractions from rat liver. This work made use of density gradient centrifugation and assays for a range of enzymes. The particulate subfractions obtained were assigned

nuclear, large granules, and microsomes. Livers from freshly killed rats were cut into pieces and homogenised in a tissue grinder containing buffered sucrose. Homogenates were centrifuged and pellets homogenized in buffer; nuclei were then pelleted also. Supernatant was used for sedimentation of the large granular fraction and this pellet was resuspended and the microsomal fraction separated from the postmitochondrial supernatant by further centrifugation. The various fractions were assayed for enzyme activities and observed under an electron microscope. These experiments led to the eventual grouping of microsomes into four constituent groups based upon the activities they contained. Two of these were concluded to be ER-derived and contained NADH- and NADPH-cytochrome *c* reductase, aminopyrine demethylase, cytochromes b5 and P450, glucose-6-phosphatase, nucleoside diphosphatase, esterase,  $\beta$ -glucuronidase and glucuronyltransferase, and also RNA. These assays exposed a biochemical heterogeneity of ER which was largely attributed to contaminating endomembranes.

#### *Yeast and fungi*

Endoplasmic reticulum from yeast and fungi has been isolated and investigated. Swida and colleagues (Swida *et al*, 1982) described work on *Saccharomyces cerevisiae*. The aim of this study was to produce highly enriched rER, clearly separated from sER for the investigation of mRNA recruitment by ribosomes at the ER. A French Pressure cell was used to break yeast cells suspended in a buffer. Supernatant obtained after centrifugation for 5 minutes at 3 000 x g was centrifuged three times at 16 000 x g for 10 minutes to remove nuclei and mitochondria. The



final supernatant was subjected to centrifugation for 1 hour at 90 000 x g. The total microsome pellet was resuspended in 1.33 M STKM buffer (containing sucrose, Tris/HCl, KCl, magnesium chloride) and layered between 2 M and 0.2 M STKM buffer, to be centrifuged for 16 hours at 131 000 x g. sER was collected from the top of the 1.33 M zone and rER from the top of the 2 M zone. sER was further purified in a similar repeated fashion and sucrose was removed from both rER and sER fractions. Characterization of these subfractions was on the basis of peptidylpuromycin synthesis and in vitro peptide synthesis capability.

Isolation and characterization of *Neurospora crassa* has also been described (Borgeson and Bowman, 1983). ER was identified by the marker enzyme phosphatidylcholine glyceride transferase. Cell suspensions were mechanically broken with glass beads and the resulting homogenate centrifuged at 1 000 x g for 10 minutes to remove debris and nuclei. Supernatant was subjected to differential centrifugation, pellets being obtained at 14 000, 40 000, and 100 000 x g. Alternatively, supernatant was subjected to sucrose density gradient (14-42%) centrifugation for 2 hours at 95 000 x g. Fractions were analysed by enzyme assays (phosphatidylcholine glyceride transferase, cyt *c* reductase, glucose-6-phosphatase) and protein concentration determination or by electron microscopy.

### *Plant systems*

As noted previously, ER from plant cells is less easily characterized. ER membranes were isolated from castor bean endosperm by Lord *et al* in a study into lecithin

formation which concluded that the exclusive site of this metabolism is the ER (Lord *et al*, 1973). Endosperm halves from 4-day old seedlings were manually homogenized with a razor blade on ice in a grinding medium. After filtration through nylon cloth, the homogenate was centrifuged at low speed (270 x g / 10 min.) to remove cellular debris. Subsequently, cellular components were fractionated by sucrose density gradient centrifugation. The membrane fraction identified as ER on the basis of containing NADH- and NADPH-cytochrome *c* reductase, NADH diaphorase, and cytochromes b5 and P450 had a density of 1.12 g/cm<sup>3</sup> on sucrose gradients. Also, treatment with magnesium ions had the effect of increasing this value to 1.16 g/cm<sup>3</sup> by preventing the dissociation of ribosomes from these membranes, a phenomenon which could be observed in the electron microscope. Later, ER membranes isolated in this way from germinating castor bean endosperm were subjected to a second fractionation on a flotation gradient (Conder and Lord, 1983). Isolated ER fractions were mixed with sucrose (to give a solution of approx. 50% w/w) and overlaid with successively less concentrated sucrose solutions (40-20% w/w). Centrifugation was performed for 17 hours at 24 000 rpm and 2<sup>o</sup>C. Enzyme assays for ER were cholinephosphotransferase and NADH-cytochrome *c* reductase and indicated the presence of these proteins in two separate membrane bands. In addition, both subfractions were shown to possess glycoprotein fucosyltransferase activity, which is confined to the ER (rather than Golgi) in nondividing endosperm. However, heterogeneity was shown between the ER subfractions with respect to distribution of glycosyltransferases which catalyze the transfer of sugar from nucleoside diphosphate derivatives to lipid acceptor and the

assembly of monosaccharide lipids into oligosaccharide lipids. The heterogeneity was proposed to result from the existence of distinct regions of the ER, specialized for particular functions.

Fucosyltransferase activity has also been analysed in microsomal fractions from cucumber cotyledons (Sturm and Kindl, 1983) which are biochemically comparable to castor bean endosperm. In this study, filtered homogenates from 3-day old seedlings were subjected to an initial centrifugation (5000 x g / 15 min.) and the microsome-containing supernatant was then centrifuged on a sucrose density gradient for 2 hours at 27 000 x g. The microsomal membranes were further purified by flotation in a sucrose gradient for 18 hours 28 000 rpm. The marker for ER was NADH-cytochrome *c* reductase and the conclusions drawn were that GlcNAc transferase is located primarily in the ER whereas fucosyltransferase is located in the Golgi. This study also used sodium carbonate to separate luminal from membrane constituents (see Fujiki *et al*, 1982).

Recently, germinating and developing castor bean endosperm has been exploited for a molecular characterization of plant ER (Coughlan *et al*, 1996). This work included an effective strategy for obtaining preparative amounts of ER, devoid of contaminating endomembranes. Typically, endosperm halves from 25 seedlings were manually chopped with a razor blade for 10 minutes in a homogenization medium on ice. Filtering of homogenate through nylon preceded a 5 minute period of centrifugation 1000 x g and 2<sup>o</sup>C. Supernatant was centrifuged on a sucrose gradient (20-60%) for 2

hours at 25 000 rpm and 2 C. Fractions were assayed for markers , those for ER being antimycin-insensitive NADH-cytochrome *c* reductase and CDP: 1,2-DAG choline phosphotransferase. Refractometry revealed that the partially purified ER fractions lay at the 20/30% m/v sucrose interface (1.12g/cm<sup>3</sup>). These fractions were pooled and mixed with sucrose (to give a solution of approx. 50% m/v) and overlaid with successively decreasing concentrations of sucrose solutions (40-20% m/v) for flotation at 25 000 rpm and 2<sup>0</sup>C for 24 hours. Sodium carbonate was used to separate the purified ER into lumenal and integral membrane fractions. This work showed that large amounts of ER can be purified to apparent homogeneity (by enzyme marker and SDS-PAGE criteria) in 2 days, and it therefore represents a substantial step forward in the study of the ER of plant cells.

### **1.9 Endoplasmic reticulum as a rich source of diverse proteins / enzymes**

As the location of a wide range of cellular processes, the endoplasmic reticulum is a source of many different proteins. A significant proportion of these proteins will not have been characterized, particularly from plants which, until recently have not provided substantial quantities of pure ER (Coughlan *et al*, 1996). Membrane-bound proteins present the greatest difficulty with regard to purification.

Table 1.2 is a list of proteins which, from a consideration of functions of the ER, would be expected to be found there either membrane-associated or lumenal, and as

---

### Protein synthesis

SRP receptor: ( $\alpha$  chain; interacts with a  $\beta$  subunit which may be yeast Yk1154p)  
SSR $\alpha$  and SSR $\beta$  (signal sequence receptor subunits)  
translocation machinery: Sec61p/62p/63p(member of DNAJ protein family)/66p/72p/Kar2p complex.  
Sec61p/Sss1p/Sbh1p complex  
chaperones  
ribosome binding proteins (ribophorins): p34  
signal sequence peptidase: Sec11p

### \*Chaperones

calreticulin  
calnexin  
GRP 94 (yeast: glucose regulatory protein)  
PDI (protein disulphide-isomerase)  
hsp70: Kar2p (BiP homologue)  
BiP  
DNAJ protein  
peptidylprolyl *cis-trans* isomerase

### Glycosylation

dolichol kinase: Sec59p  
glycosyltransferases:  $\beta$ -mannosyltransferase  
dolichol-P-glucose synthetase  
UDP-GlcNAc-1-P-transferase (GPT)  
glucosyltransferase  
dolichol-P-mannosyltransferase  
glycosidases:  $\alpha$ -mannosidase  
glucosidase  
oligosaccharyltransferase: mammalian subunits: ribophorins I and II, and OST48  
yeast subunits (encoded by the 'OST' genes):  $\alpha$  (heterogeneously glycosylated protein with 3  
glycoforms - 60/62/64 kD),  $\beta$  (45 kD Wbp1),  $\gamma$  (34 kD),  $\delta$  (30 kD Swp1),  $\epsilon$  (16 kD), and  $\zeta$  (9 kD)  
ost4 (~3.6 kD): required for oligosaccharyltransferase activity  
calnexin

### Retention

SSR $\alpha$  and calnexin  
chaperones  
HDEL recognition proteins

### Lipid body formation

oleosins

### \*\*Vesicle mediated transport

coat proteins

COP I coats: coatomer ( $\alpha, \beta, \beta', \gamma, \delta, \epsilon, \zeta$  - subunits)  
ARF (ADP ribosylation factor; GTPase activity)  
Sed4p (GDP/GTP exchange ?)  
COP II coats: Sec13p/Sec31p complex  
Sec23p/Sec24p complex  
SAR (Sar1p: secretion associated and RAS superfamily-related; GTPase)  
Sec12p (GDP/GTP exchange)  
cargo receptors  
vSNAREs: Sec22p, Bos1p, Bet1p  
tSNAREs: Ufe1p  
ARF receptor

### Retrieval

KDEL/HDEL receptors  
KKXX receptor

### Lipid metabolism

phospholipids:

CDP-cholinephosphotransferase  
CDP-ethanolaminophosphotransferase  
SAM: phosphatidylethanolamine N-methyltransferase  
phosphatidylethanolamine L-serine phosphatidyltransferase  
phosphatidylserine decarboxylase  
CTP phosphatidate cytidyltransferase  
CDP-DAG serine O-phosphatidyltransferase  
CDP-DAG inositol phosphatidyltransferase  
phosphatidylinositol: inositol phosphatidyltransferase  
glycerophosphate CDP-DAG phosphatidyltransferase  
phosphatidylglycerophosphate phosphohydrolase (PGP phosphatase)  
phosphatidate cytidyltransferase  
PLD (phospholipase D)

triacylglycerols:

glycerol-3-P-acyltransferase (1-AT)  
LPA-AT (2-AT)  
DAG-AT (3-AT)  
phosphatidate phosphatase

desaturases: stearoyl ( $\Delta 9$ ) desaturase  
oleoyl ( $\Delta 12$ ) desaturase  
omega-3 desaturase  
omega-6 desaturase

enzymes of sterol synthesis: HMG-CoA reductase  
SAM: sterol-C-methylase  
desaturases

fatty acid modifying enzymes:  $\Delta 12$ -hydroxylase  
elongases

#### GPI anchoring

enzymes of GPI anchor synthesis  
N-terminal signal peptidase  
C-terminal processing enzyme  
putative transamidase: gaalp

#### Electron transfer

cytochrome b5  
cytochrome P450  
NADH-cytochrome *c* reductase  
NADPH-cytochrome *c* reductase

#### Others

phosphorylases (phosphorylate membrane proteins) eg. casein kinase II-like protein  
glucose-6-phosphatase  
glucose transporter: GLUT-7  
glucose-6-phosphate transporter  
phosphate transporters  
protein fatty acyltransferase  
protein kinases  
ubiquitin-conjugating enzyme  
calcium-ATPase

---

\* Chaperones are likely to be involved in retention mechanisms for the purposes of quality control, protein body formation, and the holding back of resident ER proteins from the secretory pathway.

\*\* Coat and associated proteins are generally cytosolic but will be transiently associated with the ER prior to vesicle budding.  
NB Not all of the proteins listed are exclusive to ER.

---

**Table 1.2: List of proteins expected to be ER resident/associated.**

such represents a subcellular proteome, some components of which will be encountered during this study.

### 1.10 Castor bean (*Ricinus communis*)

The castor bean plant is a dicotyledonous angiosperm. As previously illustrated, it is commonly chosen as the experimental material for biochemical studies. The seed is oil-rich, and has a largely liquid endosperm which greatly facilitates fractionation studies aimed at isolating intracellular membrane systems such as ER, which is obtained pure and in preparative amounts from this tissue (Coughlan *et al*, 1996). This section includes a brief background to the development, maturation and germination of dicot seeds, followed by a detailed morphological description relevant to seed selection for experimental use.

Fertilization in angiosperms is unique in that two male nuclei participate. One is immotile and, after release from the pollen tube, fuses with the female egg giving rise to a diploid zygote. The other fuses with two polar nuclei to produce a triploid nucleus. The first division in endosperm (derived from the triple fusion nucleus) development results in an axial (distal or apical) cell and a basal cell. The latter gives rise to the suspensor in dicot seeds and may contribute to the embryo. The axial cell contributes mostly to the embryo. The developing endosperm draws nutrients from adjacent tissues and also lays down new products, resulting in a growing embryo which is closely surrounded by readily available food reserves required during maturation and subsequent germination/growth stages. Endosperm is retained as a permanent storage tissue in this type of plant. Imbibition (water uptake) marks the start of germination, during which the embryo is active in numerous metabolic processes including protein hydration, subcellular structural changes, respiration,

macromolecular synthesis and cell elongation. All this leads to growth. Strictly, germination does not include emergence of the seedling and so, seedling processes which include fuel reserve mobilization should be classed as postgermination events. Quiescence describes the state of seeds in which no germinating processes are taking place and represents a state in which the seed can survive for long periods (years) and from which germination can occur if provided with the correct conditions (eg. hydration, temperature, oxygen). Alternatively, germination processes do not always lead to radicle emergence, a condition termed dormancy. Seeds dispersed from mother plants already blocked in germination show primary dormancy. Environmental change can induce secondary dormancy by blocking entry into germination.

The use of plant seeds as experimental material for biochemical studies requires strict attention to developmental uniformity in the selected seeds. It is necessary to be able to accurately describe the stage of development of a particular tissue to allow reliable comparisons of results to be made between different experimenters. For castor, the subject of the present study, a morphological timetable has been compiled (Greenwood and Bewley, 1981) to facilitate selection of seeds which are at very close stages of development. Ten morphological stages were established by the subdivision of the three major physiological periods of development characterized by changes in size and fresh and dry weights of the seed and seed parts. The subdivisions are more complicated and are based on more than one morphological trait. These stages are as



follows. I - proembryo; II - globular; III - heart; IV-VIII - cotyledon; IX & X - maturation stages.

Public access nucleotide and protein databases provide valuable tools in gene cloning/protein identification. Effective and regular searching is necessary due to the quantity of information contained and the rapid rates of expansion. A variety of data bases can be accessed via the internet. Of particular interest is that of expressed sequence tags (dbEST) at the National Center for Biotechnology Information which results from the accumulation of partial gene sequences from anonymous cDNA clones derived from a diverse range of organisms. Large scale sequencing data of this nature was first deposited by Adams and colleagues (Adams *et al*, 1992) who were working with human brain genes. Common plant entries include *Arabidopsis thaliana*, *Brassica napus*, *Oryza sativa* and *Ricinus communis*. Results of a large-scale *Arabidopsis* EST project are presented by Newman *et al* (Newman *et al*, 1994) in a report which provides a description of methods for accessing such information. Similarly, several hundred ESTs from castor bean seeds have been selectively processed as potentially ER-associated, assigned function based on homology with known sequences present in the data bases from both plant and non-plant sources (Van de Loo *et al*, 1996). Work of this nature was also carried out using rice endosperm as the experimental material (Liu *et al*, 1995).

### 1.11 Two-dimensional protein electrophoresis

Two-dimensional electrophoresis (2-D electrophoresis) is the most powerful and widely used tool for the separation and analysis of proteins from complex biological sources. The technology was first introduced by O'Farrel (O'Farrel, 1975). Conventionally, proteins are separated according to their isoelectric point (pI) by isoelectric focusing polyacrylamide gel electrophoresis (IEF-PAGE) in the first dimension. Subsequently the second dimension separates on the basis of molecular weight using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). A single high quality 2-D experiment can be expected to result in the resolution of up to several thousand protein spots depending upon biological complexity and detection sensitivity. Furthermore this resolution may be observed as an almost uniform pattern across most of the surface of the gel. Such a distribution is achievable due to the fact that each dimension separates according to two different physical protein properties. Other 2-D electrophoretic methods do not exploit independent parameters in this way and their limitation is evident in the tendency for a diagonal spot distribution across the final gel (for example see Hartinger *et al*, 1996).

Once separated, the proteins in a 2-D gel can be detected/visualized in a number of ways according to the nature of the experiment and/or the level of sensitivity required. The usual means of detection involve Coomassie blue or silver staining protocols which directly colour protein. Rudimentary visual analysis normally allows only limited conclusions to be drawn regarding the nature of the sample but can indicate

potential improvements (for example in sample preparation) for future experiments. Commonly 2-D gels are used to compare two samples which are expected to differ, for a defined reason, in their protein composition – an application that relies upon an adequate level of experimental reproducibility. More detailed inspection yields approximate isoelectric point and molecular weight data for individual spots. This is facilitated through the inclusion of standard marker proteins, either in the sample solution, or exogenously alongside. With satisfactory gel-to-gel reproducibility, markers may be neglected altogether.

The resolving power of 2-D electrophoresis, which enables very detailed sample comparisons to be made, has long been appreciated and applied to a variety of studies. Until relatively recent technological advances however, the extent to which it can provide meaningful qualitative data has been severely limited. Molecular characterisation from 2-D gels involving various forms of amino acid sequence generation and mass spectrometric analysis is now enabling the true potential of the procedure to be realised. 2-D electrophoresis as the core technology for proteomics research, and which ultimately aims towards functional assignment, provides an extremely powerful analytical tool to complement, and in some ways surpass, genomics. Proteomics is discussed separately in the following section.

Despite the undeniable justification of the use of 2-DE as the widespread precursor to proteomic identification, it is apparent that the technique is subject to a number of limitations. Although 2-DE can easily separate several hundred proteins in a single

experiment it is important to realize that certain subsets of proteins are unlikely to be represented, at least not without special consideration. In particular, strongly hydrophobic, highly basic, and very high molecular weight proteins are notoriously difficult to capture in 2-D profiles. It is noteworthy that the dynamic range of the individual protein quantities in any given sample will, for complex mixtures, render many of the lower abundant components invisible, regardless of the staining procedure employed. This is an issue which can be addressed by the incorporation of some form of sample prefractionation. Furthermore, in spite of the high resolving power there can be no guarantee that a single spot does not represent the position of more than one protein out of a population of potentially thousands. Apart from the technological limitations outlined above, it is always worth remembering that high quality, reproducible 2-DE is difficult and time-consuming to perform.

The successful application of 2-D electrophoresis whether as an analytical tool in its own right or as a preliminary step in a proteomics program depends upon careful consideration of certain fundamental principles and these are discussed in chapter 4.

### **1.12 Proteomics – screening expressed components of the genome**

Proteomics is the large scale screening of the proteins of a complex mix such as those of a certain cell type, tissue, biological fluid, or, pertinent to this work, subcellular compartments. The word proteome describes all of the PROTeins expressed from the genOME in the particular system under investigation. As noted previously, 2-

dimensional electrophoresis forms the basic separation method of choice for proteomic characterization. Downstream of 2-D electrophoresis, proteomics includes molecular protein analysis involving either amino acid sequencing (N-terminal microsequencing or by electrospray tandem mass spectrometry), or peptide mass fingerprinting by MALDI (matrix-assisted laser desorption/ionization) time-of-flight (TOF) mass spectrometry (ms). Such data generation is aimed ultimately at the assignment of function to individual protein spots. The application of mass spectrometric tools to proteomics has enormously increased its potential as smaller quantities of protein become amenable to analysis. Ultimately proteomics relies upon the field of genomics, that is the availability of searchable cognate or homologous DNA data for rapid data extension or functional insight. The rapid growth of DNA databases thus greatly facilitates such identification. While many species are already fully sequenced, or are nearing completion, at the genomic level (for example *Arabidopsis*), expressed sequence tag (EST) sequencing also represents a rapid and relatively inexpensive means of generating proteomics compatible DNA data. At present, Genbank lists only 830 nucleotide entries (EST and full cDNA clone sequences) for castor (van de Loo *et al*, 1996). However, *Arabidopsis* and castor share a high level of homology both at the amino acid and gene nucleotide levels. This is potentially of use during the proteomic characterization of castor in light of its relatively limited DNA representation in the databases.

As noted, paucity of species-specific genome data will restrict the success rate of proteomics approaches, particularly those generating peptide mass fingerprint data

alone. The high sensitivity of mass spectrometers commonly used leads to increased contamination problems where the quantity of protein for analysis is very low and data from trypsin autolysis and airborne keratins merge with sample peaks. Other limitations to mass spectrometric analyses include the efficiency with which specific peptide ions 'fly' in the instrument and also the degree of sequence coverage provided by resultant spectra.

### **1.13 Functional genomics**

As has been discussed proteomics and genomics are closely linked and it is becoming increasingly evident that one cannot stand entirely alone from the other if the fullest possible understanding of how proteins are expressed is to be gained. Identification of proteins through proteomic strategies ultimately relies upon the presence of corresponding genomic data in a searchable database. On the other hand, proteomics based data provides information which the study of the genome and levels of mRNA cannot. Nonetheless functional genomics encompasses a range of strategies which can give considerable insight into the functions of gene sequences and also into the patterns of their expression.

EST sequencing can provide preliminary information on expression patterns of more abundant transcripts representing as it does a rapid means of obtaining an inventory of expressed genes. However, rare transcripts (or those induced under specific conditions) may not be represented in an EST database and only complete genomic

sequences can guarantee the inclusion of all gene data. Genome sequence is of limited value in terms of what it can reveal about patterns of expression and the position of non-coding (intron) sequences.

Technologies have been developed to directly screen mRNA transcript levels from specific biological tissues thus providing insights into when and where a given gene is expressed. Tentative implications for protein expression levels can also be summarised. Serial analysis of gene expression (SAGE) technology is based on cDNA 3' sequence tags (9-11bp) concatenated and cloned and then sequenced. Relative abundance of a specific tag can indicate expression pattern for the cognate gene as recognised in the database through homology. Levels of mRNA levels for large numbers of genes can be performed in parallel using high-density DNA arrays. Open reading frames are gridded onto a support such as nylon or a glass slide and probed with single stranded cDNA prepared from total mRNA from the cell/tissue under study. Broadly, such nucleic acid-based hybridization approaches can provide a systematic description of mRNA levels in different tissues. Global expression studies can be used to classify genes based on their spatial and kinetic expression patterns; new regulatory elements may be identified by comparing the regulatory sequences of the same class. Finally, the kinetics of changes in gene expression, combined with expression profiles of mutants for known regulatory genes, should allow the study of expression networks.

Genetics approaches involving the study of mutants have been extremely useful for the study of the genetic and molecular bases for any trait in plant biology. Insertional

mutagenesis involves the creation of loss-of-function mutants either by transposons (DNA elements able to insert at random within chromosomes) or the T-DNA of *Agrobacterium tumefaciens*. Because the sequence of the inserted element is known, the gene in which it is inserted can be recovered using various cloning or PCR-based strategies. Insertional mutagenesis can be engineered with reporter cassettes (enhancer or gene traps) that will report on the expression of the chromosomal gene at the site of insertion. Mutagenized populations can be screened for lines expressing the reporter gene in specific cell types or in specific environmental conditions. Genes with interesting expression patterns and their promoters can be isolated from such lines.

A direct approach to investigating function from gene sequence data is to specifically create a gene 'knockout' plant. Such mutants may be created directly by homologous recombination or antisense technology. Alternatively, mutagenized in the gene of interest can be obtained from a population of T-DNA or transposon mutants, using PCR.

#### **1.14 The Present Study**

The first objectives of this research are to prepare pure endoplasmic reticulum from germinating and developing castor endosperm and to investigate protein complexity using electrophoresis techniques. Following primary analysis it is intended that a differential analysis of the two states is performed and including membrane



subfractionation experiments. Such investigations were expected to require the high resolving power of 2-D electrophoresis, a technique which would need optimizing for the specific samples under study.

The incorporation of membrane subfractionation methodology was intended to allow the simplification of gel profiles by enriching for specific subsets of proteins. Enrichment is important where a relatively small number of highly abundant, possibly 'housekeeping', components account for a significant proportion of total protein. Without subfractionation gel loadings determined by estimation of total ER protein content may render some of the more 'interesting' lower abundance proteins undetectable.

N-terminal Edman sequencing is evaluated as a means of identifying proteins from gel profiles. Subsequent establishment of MALDI TOF ms as a means of characterizing ER proteins from 2-D gels was expected to generate novel peptide mass fingerprint data for selected components. Efficiency of functional inference would rely on quality of data and data editing as well as familiarity with the database search programs available. Ultimately, however the results would depend upon the presence of corresponding castor data or suitably similar clones from other plant species.

Although much of the work is consequently presented as an evaluation and optimization of enabling technologies, it was expected that novel data would be

generated. With database considerations in mind it is worth note that for castor bean at least, some of this data may require genome input to 'catch up' before functional assignments are adequately addressed. Nonetheless the ultimate aim in this laboratory is the identification and cloning of new genes encoding complex lipid biosynthetic enzymes. This work represents a basis from which this goal worked towards by a proteomics approach.

## CHAPTER 2

### Materials and methods

#### 2.1 Materials

All chemicals used throughout this study, unless otherwise stated in the relevant sections, were obtained from either Sigma Chemical Company Ltd., Fancy Road, Poole, Dorset BH17 7NH or BDH Chemicals Ltd., Merck Ltd., Merck House, Poole, Dorset BH15 1TD.

Dry castor seeds (*Ricinus communis* L var. Hale) were kindly provided by Dr Sean J Coughlan (formerly Pioneer Hi-Bred Int.) and were originally obtained from Bothwell Seed Company (Plainview, Texas). Developing castor fruit were harvested at Pioneer Hi-Bred from plants derived from the same seed.

#### 2.2 Methods

##### 2.2.1 Protein Methods

##### 2.2.1.1 Bradford protein estimation

The reagent used was the Protein Assay Dye Concentrate (Bio-Rad Laboratories Ltd., BioRad, Marylands Avenue, Hemel Hempstead, Hertfordshire HP2 TTD). Estimation is made by a dye-binding assay based on the differential change of the absorbance maximum for an acidic solution of Coomassie Brilliant Blue G-250 from 465nm to 595nm when binding to protein occurs (Bradford 1976). The assay was performed according to the manufacturer's recommendations for the micro-assay procedure. Standard curves were constructed using a series of BSA standard solutions in the range 1-20µg.

### 2.2.1.2 TCA Lowry method

Sample total protein concentration was estimated using an adaptation of the method of Lowry (Lowry *et al*, 1951) as described by Peterson (Peterson, 1979). The basis for the assay is the reduction of molybdic-tungstic mixed acid chromagen (Folin and Ciocalteu reagent) by tryptophan and tyrosine aromatic side chains, and the reaction of copper chelates with polar groups and peptide chains. The reduction of the chromagen results in an increase in absorbance at 750nm which is maximal at 20-30 minutes after initiation. The method incorporates a deoxycholate-trichloroacetic acid (DOC-TCA) precipitation step to remove lipid and other potentially interfering compounds.

#### *DOC-TCA precipitation*

Samples were prepared in a total volume of 1ml to which was added 100 $\mu$ l 0.15% (w/v) deoxycholic acid. The mixture was vortexed, left to stand at room temperature for 10 minutes, and mixed by further vortexing with 100 $\mu$ l 72% (w/v) trichloroacetic acid. Samples were spun for 10 minutes in a microcentrifuge and the supernatant immediately decanted.

#### *The Lowry assay*

Pellets were resuspended in 1ml solution C, vortexed, and left to stand at room temperature for 5-10 minutes. 100 $\mu$ l solution D (freshly prepared) was added followed by vortexing and a 20-30 minute incubation period, at room temperature. Absorbance was measured at 750nm. Standard curves were constructed according to a standard BSA series in the range 1-20 $\mu$ g.

Solution A: 0.5% (w/v) CuSO<sub>4</sub>·5H<sub>2</sub>O in 1% (w/v) sodium tartrate

Solution B: 2% Na<sub>2</sub>CO<sub>3</sub> in 0.1M NaOH

Solution C: Solution A: Solution B, 1:50 ratio (freshly prepared)

Solution D: ddH<sub>2</sub>O: Folin-Ciocateu phenol reagent, 1:1 ratio (freshly prepared)

### 2.2.1.3 Coomassie Blue R-250 Protein Staining

Gels were sequentially treated with three different Coomassie solutions designated I, II, and III. Each incubation involved 2 minutes' heating (full power) in a microwave oven followed by a 10 minute stand. The final incubation was followed by prolonged soaking in a destaining solution. The sequential staining method results in enhanced staining of protein accompanied by the gradual background destaining of the gel matrix.

|                     |  |
|---------------------|--|
| Coomassie stain I   | 10% glacial acetic acid, 25% propan-2-ol, 2%<br>Coomassie<br>stock |
| Coomassie stain II  | 10% glacial acetic acid, 10% propan-2-ol, 0.25%<br>Coomassie stock |
| Coomassie stain III | 10% glacial acetic acid, 0.25% Coomassie stock                     |
| Coomassie stock     | 1.25% (w/v) Coomassie brilliant blue R250                          |
| Destain solution    | 10% glacial acetic acid, 10% glycerol                              |

### 2.2.1.4 Silver-Staining of *In Gel* Proteins

Following electrophoresis gels were placed in fixing solution (40% ethanol, 10% glacial acetic acid) for 30 minutes. Fixed gels were transferred to a sensitizing

solution (30% ethanol, 0.13 % w/v glutaraldehyde, 6.8% w/v sodium acetate.3H<sub>2</sub>O, 0.2% w/v sodium thiosulphate.5H<sub>2</sub>O) and incubated for 30 minutes. Washing (3 times in fresh ddH<sub>2</sub>O, 5 minutes each) was followed by a 40 minute incubation in 0.1% w/v silver nitrate, 0.008% w/v formaldehyde. After brief rinsing in ddH<sub>2</sub>O, protein bands/spots were developed in 2.5% w/v anhydrous sodium carbonate, 0.004% w/v formaldehyde for up to 15 minutes. The developing reaction was stopped by transfer into 1.46% w/v EDTA.Na<sub>2</sub>.2H<sub>2</sub>O for 10 minutes. Gels were washed (2 x 10 minutes in ddH<sub>2</sub>O), before drying. SDS-PAGE mini gels were dried directly (see section below). Pharmacia Excel gels were first soaked in a preserving solution of 10% glycerol and overlaid with a sheet of cellophane soaked in the same solution. Drying took place at room temperature.

#### **2.2.1.5 Drying of SDS-PAGE Mini Gels**

Gels were held between water soaked sheets of cellophane using a drying frame (Hoefer Scientific Instruments Easy Breeze). Drying took place in the Hoefer Easy Breeze Gel Dryer.

#### **2.2.1.6 Chloroform/Methanol Protein Precipitation**

This procedure (Wessel and Flugge, 1984) is effective at desalting and delipidation as well as concentration of protein samples, and was performed in microcentrifuge tubes. Samples were made up in 100µl with ddH<sub>2</sub>O to which 400µl methanol was added. This was vortexed and spun briefly, then mixed with 100µl chloroform by further vortexing, and briefly spun again. 300µl ddH<sub>2</sub>O was added, and the suspension vortexed vigorously before phase resolution by centrifugation (full speed, microcentrifuge) for 1 minute. The upper (aqueous) phase was withdrawn and

discarded, leaving behind protein precipitate (visible at the phase interface) on top of the lower phase, to which was added 300µl methanol. After vortexing, protein was pelleted by centrifugation for 5 minutes. Supernatants were discarded and pellets air-dried for an appropriate period (typically 10-15 minutes). Care was taken to avoid pellets becoming tough and unmanageable due to over-drying.

#### **2.2.1.7 Trichloroacetic acid/Acetone Protein Precipitation**

This method is effective at removing various non-protein contaminants. In removing much of the lipid intrinsic to the preparations it optimizes solubilization of hydrophobic protein domains by the sample buffer detergent. In addition the procedure is effective in removing salt which can interfere with electrophoresis. Samples were precipitated by addition of 50% w/v TCA in acetone (stored at room temperature) to give a final concentration of 10% w/v in a solution which included 20mM DTT (Melford Laboratories Ltd.). Incubation at -20°C for 1 hour was followed by centrifugation (full speed, microcentrifuge) for 15 minutes. Supernatants were discarded and pellets washed with an equal volume ice-cold acetone. Samples were re-spun for 5 minutes and acetone removed. Pellets were allowed to air dry for 10-15 minutes. Care was taken to avoid fibrous pellets, caused by overdrying, which are difficult to resuspend.

#### **2.2.1.8 N-terminal protein sequencing**

Protein sequencing was performed in a gas phase protein sequencer (Applied Biosystems, 850 Lincoln Center Drive, Foster City, California 94404, USA). The service was carried out by the Protein Sequencing Facility, University of Durham.

## **2.2.2 Subcellular Fractionation methods**

### **2.2.2.1 Preparation of biochemically pure endoplasmic reticulum from castor endosperm**

#### *Germinating seeds*

Between 75 and 100 mature castor seeds (*Ricinus communis* L. var. Hale) were surface sterilized in 10% by vol. hypochlorite (industrial) and soaked overnight in running tap water. Soaked seeds were sown in moist vermiculite and germinated in the dark for between 3 and 5 days at 30°C. Seedlings were removed from vermiculite and rinsed before discarding testa and roots. Endosperm were dissected and placed in ice-cold homogenization buffer: 500mM sucrose, 10mM KCl, 1mM EDTA, 1mM MgCl<sub>2</sub>, 2mM DTT (Melford Laboratories Ltd., Chelsworth, Ipswich, Suffolk IP7 7LE), 150mM Tricine/KOH pH 7.5, few crystals PMSF. Care was taken to exclude any cotyledonous material. Homogenization involved manual chopping with two single-edged razor blades in a large glass petri dish on ice, for 15 minutes. Filtration of the crude homogenate through a 100µm nylon mesh was followed by centrifugation in a Sorvall® RC-5B centrifuge, HB-4 rotor (Sorvall Products, P.P., Newtown, Connecticut, USA.) at 2000 rpm and 4°C for 15 minutes. The surface fat pad was carefully removed with a glass rod. Decanted supernatant was layered onto a discontinuous sucrose density gradient step. Steps consisted of 12 ml 20% (w/w) sucrose on top of 7 ml 30% (w/w) sucrose solutions in Beckman Ultra-Clear ultracentrifuge tubes (Beckman Coulter Inc., Harbour Boulevard, Fullerton, California, USA.). These solutions were used ice-cold and contained 1mM EDTA and PMSF. Centrifugation was at 25,000 rpm on a Beckman SW 28 rotor for 2 hours at 2°C in a Beckman Optima™ L-8 ultracentrifuge. Tubes were mounted vertically in a laboratory clamp under a lamp, and membranes were clearly visible at the 20/30%



sucrose interface (~1.12 density). Membranes were removed after piercing tube walls with a hypodermic syringe. These fractions were pooled, mixed with an equal volume of ice-cold 60% (w/w) sucrose and subjected to flotation through a second discontinuous gradient (12 ml 40% (w/w), 12 ml 30% (w/w) and 5 ml 20% (w/w) sucrose solutions, all ice-cold and containing 1mM EDTA and PMSF). Centrifugation was performed at 25,000 rpm, for 22 hrs, 2°C. Membranes were again visible at the 20/30% interface and removed as described above. Pooled fractions were mixed with an equal volume of ice-cold distilled water and the membranes pelleted by centrifugation at 55,000 rpm in a Beckman Optima™ TLX ultracentrifuge, Ti 80 rotor (45 min., 2°C). Pellets were resuspended in a protective 10% (v/v) glycerol solution prior to snap-freezing in liquid nitrogen and storage at -80°C.

#### *Developing seeds*

Inflorescences were harvested from castor plants at 25 days after flowering. Individual seed pods were dissected and endosperm removed from the seeds inside and placed into ice-cold homogenization buffer, care being taken to avoid spongy aleurone tissue and the developing leaf material.

#### **2.2.2.2 Preparation of *E. coli* microsomes**

French cell disrupted *E. coli* cells were subjected to a 40 000 x g centrifugation. The supernatant was centrifuged at 100 000 x g and the resultant pellet resuspended, with the aid of a glass homogenizer, in 20 mM Tris-HCl, pH8 containing 1 mM β-mercapto-ethanol.

### **2.2.2.3 Preparation of Avocado Microsomes**

A microsomal fraction was prepared for use in PAP (phosphatidic acid phosphatase) assays according to the method described by Pearce (PhD thesis, 1998). Avocados (Haas variety) were purchased from Sainsbury's Supermarket, Durham. Medium sized fruit were ripened at 30°C, until the skin appeared black, and incubated for 1 hour at 4°C prior to dissection. Microsomes were prepared from the ripened mesocarp (green/yellow in colour).

### **2.2.3 Membrane Subfractionation Procedures**

#### **2.2.3.1 Sodium chloride**

A variety of salt-based membrane extractions were investigated and applied to the bacterial and plant membrane preparations. All solutions were maintained ice cold and followed the same incubation conditions as for Triton washes. The principal wash involved sequential extraction with 500 mM NaCl, containing 10 mM Tris-HCl, pH8 at both 50 000 and 100 000 x g.

#### **2.2.3.2 Sucrose Gradient Isolation of a salt-washed ER membrane fraction**

Equal concentrations of total ER suspensions of germinating and developing ER were incubated with 500mM NaCl on ice as described previously. Following incubation the salt/ER solution was layered onto a discontinuous sucrose density step (10% w/v resolving layer on top of a 60% w/w sucrose barrier). Ultracentrifugation was performed on a Beckman SW28 rotor (25,000rpm, 2 hours, 4°C). Subsequently, membranes resting on the sucrose barrier were clearly visible through the wall of the tube. Membranes were removed through the tube wall with a hypodermic syringe. This procedure utilised only a single salt for minimal sample handling/loss.

### **2.2.3.3 Chloroform/methanol extraction of hydrophobic proteins from ER membranes**

A chloroform/methanol mixture was used to solubilise moderately hydrophobic membrane proteins as described by Sieneurin-Berny and co-workers (Sieneurin-Berny *et al*, 1999). ER membrane preparations (50µg protein) were mixed with 9 volumes of a chloroform/methanol (2:1 v/v) mixture and incubated on ice for 1 hour. Samples were spun at 12 000 x g for 20 minutes, at 4°C, giving rise to a hydrophobic supernatant fraction and an insoluble white pellet. Supernatant proteins were TCA/acetone precipitated and reconstituted in SDS-PAGE loading buffer. The insoluble pellets were also resuspended in an equal volume of SDS buffer for analysis by SDS-PAGE.

## **2.2.4 Electrophoretic Procedures**

### **2.2.4.1 Mini SDS-Polyacrylamide Gel Electrophoresis (PAGE)**

Samples were analysed using Bio-Rad's Mini Protean II vertical gel apparatus and resolving gels of 10% concentration and 0.75 mm thickness, with 5% stacking gels. The buffer system used was based on that of Laemmli (Laemmli, 1970). Resolving gel solutions were prepared to give 10% (w/v) acrylamide (BioRad acrylamide:bis acrylamide 37.5:1), 0.375M Tris-HCl (pH 8.8), 0.1% (w/v) SDS, 0.01% (w/v) ammonium persulphate (BioRad), and 20µl TEMED (BioRad) in a total volume of 10ml. Stacking gels were similarly prepared adjusting acrylamide concentration accordingly and with 0.125M Tris-HCl (pH6.8) as the buffer. These were poured onto the polymerized resolving gel, and Teflon combs inserted to create the desired

sample wells. Prior to protein loading, wells were freed of any remaining acrylamide by rinsing.

Samples were loaded in up to 20µl final volume, containing 5 x SDS sample buffer (10% w/v SDS, 5% w/v DTT (Melford Laboratories Ltd.), 0.05% w/v bromophenol blue, 0.312M Tris-HCl pH 6.8, 50% v/v glycerol diluted to 1 x final concentration.

Running buffer was made up of 25mM Tris-HCl, pH 8.8 , 0.1% (w/v) SDS, and 192mM glycine, kept as a 10 x stock at room temperature. Electrophoresis was conducted at 100V through the stacking gel and 200V through the resolving phase, until the dye front had reached the bottom of the gel.

Gels were carefully detached from the apparatus and proteins visualized either by Coomassie Blue R-250 or silver-staining.

Sample preparation for SDS-PAGE ranged from direct loading of ER suspensions to prior precipitation using either a chloroform/methanol or a TCA/acetone procedure. Exact sample preparation procedures are included in the appropriate results sections.

#### **2.2.4.2 PhastGel<sup>®</sup> Analyses**

During salt subfractionation procedures the PhastGel<sup>®</sup> system (Amersham Pharmacia Biotech) was chosen for resolution of ‘salty’ subfractions due to its fast running time and thinness of gels which appeared to provide increased tolerance to salt. Gradient (8-25%) PhastGels were run on the PhastSystem<sup>®</sup> according to the manufacturer’s

recommended running conditions. Gels were either Coomassie or silver stained according to the protocols outlined in sections 2.2.1.3 and 2.2.1.4, respectively

### **2.2.4.3 Two-Dimensional Gel Electrophoresis**

#### **2.2.4.3.1 Flatbed Precast 2-D gels**

Two-dimensional electrophoresis was performed using Pharmacia's MultiPhor<sup>®</sup> II Electrophoresis System. First dimension isoelectric focusing was through immobilised pH gradients (Pharmacia Immobiline DryStrip<sup>®</sup> pH 3-10, 11cm). Second dimension SDS-PAGE used 8-18% gradient gels (Pharmacia ExceGel<sup>®</sup> SDS).

#### *Immobilised pH gradient rehydration*

Immobiline gel strips were rehydrated in a glass 'U-frame' (Pharmacia) by a minimum of six hours in reswelling solution: 48% w/v urea, 0.5% v/v Triton X-100 or 0.1-1% w/v octylglucoside, 0.15% w/v DTT (Melford Laboratories Ltd.), 2mM acetic acid, 0.5mg orange G. Reswelling solution was routinely prepared in 50ml volumes (sufficient for approximately three applications) and stored at -20°C. The U-frame was treated with dichloro-dimethyl silane to prevent the gels from adhering to its surface on reswelling. The opposing glass plate was moistened with ddH<sub>2</sub>O to hold the plastic backing of the gel strips in place during assembly and subsequent filling of the cassette, and following removal of the protective foil from each strip, was placed against the U-frame. The assembly was reinforced with clips and reswelling solution was pumped into the cassette from a plastic syringe, along rubber tubing connected to the plug-hole of the U-frame.

### *Preparation and running of first dimension IEF*

The Multiphor unit was set up with a Grant thermostatic circulating water bath (Grant Instruments (Cambridge) Ltd., 29 Station Road, Shepreth, Roysten, Hertfordshire SG6 6P2). Water at 15°C was pumped through the ceramic cooling plate. 1ml silicone oil was pipetted onto the cooling plate, and onto this was placed the first dimension electrode tray. Into the tray was poured 10ml silicone oil, onto which the gel strip aligning sheet was layered (care being taken to eliminate any air bubbles between the tray and aligner).

Rehydrated IEF strips were removed from the reswelling cassette and drained of residual reswelling solution by gentle pressing of tissue paper along the gel surface. The time between removal of strips and immersion under silicone oil was kept to a minimum to avoid formation of urea crystals on the gel, which could impede subsequent electrophoresis steps. Strips were placed into the gel aligner and each end covered in a moistened electrode wick (Pharmacia) which had been cut to the width of the containing electrode tray. Electrode adapters were inserted along the wicks and sample cups positioned at the origins (anodic ends) of the gels. Once the sample cups had been brought gently into contact with the gel surfaces, silicone oil was poured into the tray so that gel surfaces were completely submerged and sample cups filled.

Samples were loaded in a total final volume of 50µl (see sample preparation for 2-dimensional electrophoresis) directly into sample cups, and isoelectric focusing performed as a three-step program according to the manufacturer's recommendation:

| Phase | V   | mA | W | Time (h) | Vh  |
|-------|-----|----|---|----------|-----|
| 1     | 300 | 1  | 5 | 3        | 900 |

|       |      |   |   |    |       |
|-------|------|---|---|----|-------|
| 2     | 300  | 1 | 5 | 5  | 5750  |
| 3     | 2000 | 1 | 5 | 8  | 16000 |
| Total |      |   |   | 16 | 22650 |

**Figure 2.1: Running conditions for Immobiline DryStrip<sup>®</sup> pH 3-10**

On completion of electrofocusing the first dimension gel strips were either stored temporarily at -20°C or immediately equilibrated for SDS-PAGE second dimensions.

Equilibration of gel strips for SDS-PAGE involved sequential soaking with gentle agitation for 10 minutes in each of two solutions. The first contained 36% (w/v) urea, 50mM Tris-HCl pH6.8, 30% (v/v) glycerol, 1% (w/v) SDS and 0.25% (w/v) DTT (Melford Laboratories Ltd.). The second equilibration solution contained urea, Tris, glycerol, and SDS in the same concentrations and 4.5% (w/v) iodoacetamide, but no DTT. Addition of bromophenol blue to the second solution enables visualization of a dye front during SDS-PAGE.

#### *Preparation and running of second dimension SDS-PAGE*

Equilibrated first dimension gel strips were drained of fluid on blotting paper before positioning on the second dimension (horizontal) gel. SDS-PAGE was performed on the cooling plate at 15°C. The contact between the plastic backing of the SDS gel and cooling plate was insulated with 1ml silicone oil (air bubbles were avoided). Protective foil was removed from gels approximately 20 minutes prior to electrophoresis to allow surface moisture to evaporate. During the second equilibration step anodic and cathodic buffer strips (ExcelGel<sup>®</sup> SDS, Pharmacia) were laid onto their respective ends of the SDS gel and any air bubbles removed by gentle pressure. Handling of buffer strips required that gloves were moistened (to avoid

tearing) and it was essential not to allow water droplets to fall onto the SDS gel surface since this would result in serious distortion of the final 2-D pattern. Drained IPG strips were laid gel side down, with the use of forceps, onto the SDS gel at a maximum of 1mm from and alongside the cathodic buffer strip. Gel strips were stroked to expel air bubbles and thereby ensure total contact with the SDS gel. On the SDS gels used, which were 18cm wide, two IEF strips could be arranged side by side and resolved simultaneously. Absorbant sample application pieces were placed at each end of each IEF strip to prevent water released from the back of the gel strips during electrophoresis from travelling into the path of the second dimension edges. The electrode holder/electrofocusing electrode combination was positioned onto the buffer strips and electrophoresis performed as three manual programs as recommended:

| Step | V   | mA | W  | Time (min)        |
|------|-----|----|----|-------------------|
| 1    | 600 | 20 | 30 | 30 <sup>(1)</sup> |
| 2    | 600 | 50 | 30 | 5 <sup>(2)</sup>  |
| 3    | 600 | 50 | 30 | 70 <sup>(3)</sup> |

**Figure 2.2: Running conditions for second dimension ExcelGel<sup>®</sup> SDS-PAGE**

<sup>(1)</sup>After step 1 (dye front = 1-2mm from IEF gel) IEF gels and application pieces removed

<sup>(2)</sup>After additional 2mm migration of the dye front (ie. approx. 5 min.) cathodic buffer strip moved to previous position of IEF gels

<sup>(3)</sup>Completion of electrophoresis; dye front migrated into anodic buffer strip

Upon completion of the second dimension, SDS gels were either Coomassie or silver-stained. In cases where Coomassie was found to be insufficiently sensitive, gels were subsequently silver stained.



#### **2.2.4.3.2 Large Format 2-D Gels**

12% large format (1mm x 20cm x 20cm) SDS gels were cast in a Hoefer-Dalt gel casting block (Pharmacia). Isoelectric focusing was carried out using 18cm Immobiline Drustrips, pH3-10. Samples were included in the reswelling solution and applied to the IEF strips in a reswelling tray (Pharmacia). Focusing was performed on the MultiPhor apparatus at 20°C. Running conditions were as follows: 0.05mA/strip, 5W, 3000V. As runs proceeded current dropped and was periodically adjusted back to 0.05mA/strip until 3000V was reached. Focusing was completed overnight for up to 60kV hours. Strips were equilibrated for the second dimension. A stock equilibration solution of 6M urea, 30% glycerol, 2% SDS in running buffer was divided into two 10ml aliquots. To the first was added 100mg DTT (Melford Laboratories Ltd.) and bromophenol blue. To the second was added 480mg iodoacetamide and bromophenol blue. Strips were equilibrated for 15 minutes in each solution and rinsed with ddH<sub>2</sub>O before being placed onto the SDS gel. Strips were sealed in position by 1% low melting point agarose in running buffer (containing bromophenol blue). SDS-PAGE was performed at 10°C in a Hoefer-Dalt electrophoresis tank (Pharmacia). Gels were run overnight at 50V.

#### **2.2.4.4 Electroblotting of protein to PVDF membrane**

Semi-dry blotting was performed using Pharmacia's MultiPhor II apparatus. Six pieces of blotting paper (Whatman International Ltd., St. Leonard's Road, Maidstone, Kent ME16 0LS) were cut to gel size, briefly soaked in HPLC grade methanol, and placed onto the anodic plate of the blotting apparatus. Gel size membrane (ProBlott™, from Applied Biosystems) was soaked for 5 minutes in electroblotting buffer 1 x stock buffer in 10% methanol and placed onto the blotting paper (stock buffer:

100mM CAPS, pH 11, stored at 4°C). Care was taken to eliminate air bubbles. Similarly the gel was placed onto the membrane, and on top of the gel was placed a further three pieces of soaked blotting paper. The cathodic plate was positioned and electroblotting performed at 56mA for 45 minutes

## **2.2.5 Methods associated with MALDI mass spectrometric analysis**

### **2.2.5.1 Preparation of Coomassie-stained proteins for SDS-polacrylamide in-gel tryptic digestion**

Selected protein bands were excised from one-dimensional Coomassie-stained SDS-polyacrylamide gels for digestion with trypsin prior to MALDI MS peptide analyses. Gels (12%) were prepared and run with non-specialized reagents/apparatus in normal laboratory space. At the end of the run gels were stained with Coomassie I only, and placed in destain solution (see section 2.2.1.3). Destained gels were transferred to a dedicated proteomics laboratory for all tryptic and mass spectrometry preparative procedures. Bands were cut from the surrounding gel and subsequently manipulated in a laminar flow hood. Excisions were performed using clean separate razor blades to produce strips of approximate dimensions 1mm x 5mm. These were placed into separate 0.5ml microcentrifuge tubes (pre-rinsed with ddH<sub>2</sub>O) and incubated with 100µl Coomassie/protein destain solution (40% ethanol, 50mM NH<sub>4</sub>HCO<sub>3</sub>). Samples were periodically vortexed and destain solution was routinely replaced until gel bands retained no visible Coomassie colouration (for intensely stained, predominant bands, destaining could take several days). Once completely destained, strips were washed by incubation in 100µl 25mM NH<sub>4</sub>HCO<sub>3</sub> for 10 minutes at room temperature. Washing buffer was withdrawn and discarded, and gel strips were removed onto separate sterile razor blades for dissection into several smaller pieces using sterile

hypodermic needle tips. Gel pieces were returned to their respective tubes and washed three times with 100µl aliquots of acetonitrile (brief centrifugation after each addition). During removal of acetonitrile, care was taken to avoid disturbance of the gel piece pellets. These were white and fluffy in appearance, and fine gel-loading pipette tips were used (as was the case for other manipulations from destaining through to final treatments). After removal of the final acetonitrile solution, sample tubes were placed in a speed vacuum centrifuge (Gyrovap GT, V.A.Howe & Co. Ltd., UK) with separate clean syringe-pierced lids. Samples were dried for 30 minutes, followed by break-up of the resultant crystalline pellets using separate syringe needles via the lid holes. Debris was collected by brief centrifugation and incubated on ice for 30 minutes.

#### **2.2.5.2 Tryptic digestion**

An ice-cold trypsin solution was prepared as follows: 5µl stock trypsin, consisting of 20µg porcine trypsin (from Promega Ltd., Delta House, Enterprise Road, Chilworth Research Centre, Southampton SO16 7NS) in 100µl 1mM HCl, was added to 95µl 25mM NH<sub>4</sub>HCO<sub>3</sub>. 25µl of this mix was added to each dry and chilled sample. Reswelling of gel pieces took place in the trypsin solution over a 30 minute incubation, on ice (during which trypsin remains inactive). Excess trypsin solution (approximately 20µl) was replaced with 25mM NH<sub>4</sub>HCO<sub>3</sub>. Digestions were allowed to proceed at 37°C in an overnight incubation.

#### **2.2.5.3 Preparation of tryptic peptides for MALDI ms**

Samples were introduced to the mass spectrometer, either as direct tryptic supernatants or following acetonitrile/trifluoroacetic acid extraction, or both.

Calibration was made from a calibration mix applied as close external standards spotted as close as possible on the grid to several samples.

#### *Direct loading*

A 5 $\mu$ l aliquot was removed from the tryptic digest solution. For direct analysis, a 0.5 $\mu$ l sample of this was mixed with an equal volume of matrix solution (1% w/v  $\alpha$ -cyano-4-hydroxy cinnamic acid in 50% CH<sub>3</sub>CN, 0.3% TFA). Mixing of sample and matrix was achieved by gentle pipetting, care being taken not to touch the surface of the plate. Sample was allowed to dry in air (a few minutes) before entry into the mass spectrometer.

#### *Loading following acetonitrile extraction*

The remainder of the tryptic digest solution (approximately 20 $\mu$ l) was withdrawn to a fresh tube (pre-rinsed with ddH<sub>2</sub>O) and retained. Serial extraction of the residual polyacrylamide gel was performed with 25 $\mu$ l aliquots of CH<sub>3</sub>CN/5% TFA (1:1). Each extraction involved a 10 minute incubation at room temperature followed by brief centrifugation. Supernatants were pooled with the original 20 $\mu$ l tryptic digest supernatant. The final pool (approximately 95 $\mu$ l) was concentrated in a speed vacuum centrifuge for 30 minutes to a volume of around 5 $\mu$ l. This solution was processed as described above for the direct tryptic supernatant analysis.

#### **2.2.5.4 Silver staining for subsequent mass spectrometric analysis**

Silver staining was performed according to the PlusOne<sup>®</sup> Staining Kit Protein protocol (Pharmacia). Specific important modifications to conventional silver staining procedures (see section 2.2.1.4) are necessary if detection is to be followed by mass

spectrometric experimentation (these are omission of glutaraldehyde from the sensitizing solution, and of formaldehyde from the silver solution, as well as increased formaldehyde in the developing solution and a more concentrated silver nitrate solution – see below).

Gels were stained using 250 ml volumes of silver staining solutions. Protein was fixed by incubation twice for 15 minutes in 10% acetic acid, 40% ethanol. Sensitization (30 minutes) in 30% ethanol, 0.2% (w/v) sodium thiosulphate, 6.8% (w/v) sodium acetate was followed by three 5 minute washes in dH<sub>2</sub>O. Silver reaction proceeded for 20 minutes in 0.25% (w/v) silver nitrate, followed by development in 2.5% (w/v) sodium carbonate including (100µl formaldehyde) for 4 minutes. Gels were transferred to stopping solution (1.46% w/v EDTA) for 10 minutes and then washed for 5 minutes each in three changes of dH<sub>2</sub>O.

#### **2.2.5.5 Preparation of in gel silver stained protein for MALDI MS**

Proteins were excised from gels and chopped as for Coomassie-stained bands. Such samples were placed into 0.5ml microcentrifuge tubes for destaining and processing by a method compatible with the PlusOne<sup>®</sup> stain and other silver stains procedures which omit glutaraldehyde.

##### *Destaining*

A destain working solution was prepared by mixing equal volumes of two stock solutions; 300mM potassium ferricyanide and 100mM sodium thiosulfate. Sufficient of this was added to each sample to cover the gel pieces (30-50µl). Samples were

vortexed occasionally and once all colouration had disappeared, destain was removed and the gel pieces washed in several changes of dH<sub>2</sub>O.

Excess fluid was removed from the gel and 400µl 200mM NH<sub>4</sub>HCO<sub>3</sub> added followed by incubation at room temperature for 20 minutes. This solution was removed and the gel extracted twice for 20 minutes each with two 75µl aliquots of 100mM acetate in 50% acetonitrile. Extracts were discarded and the gel pieces were dried in a vacuum centrifuge.

### *Digestion*

A working enzyme solution was prepared from stock trypsin solution (0.1µg/µl in dH<sub>2</sub>O) and digestion buffer (200mM NH<sub>4</sub>HCO<sub>3</sub>). The working solution consisted of a dilution in stock trypsin in digestion buffer to give 0.02µg/µl trypsin. 10µl working enzyme solution was added per sample gel pieces and left to reswell. Gel was further rehydrated by the addition of 10µl digestion buffer. Digestion took place at room temperature, overnight. As with coomassie-treated gels, tryptic peptides were analysed directly (2µl aliquots were removed from the incubation for this purpose) and following solvent extraction from the polyacrylamide. Extraction was performed by the addition of 150µl 60% acetonitrile, 1% formic acid to the digest solution/gel and sonication in a water bath for 20 minutes. Supernatant was removed to a fresh tube and concentrated to a volume of approximately 5µl in a vacuum centrifuge. Samples were prepared with matrix solution for application to a target grid and mass spectrometric analysis as described in section...

#### **2.2.5.6 Peptide clean-up by sample binding to C18 silica**

Where clean distinctive spectra were not obtained from the samples prepared as described above, peptide samples were further cleaned using ZipTip™ from Millipore (UK) Ltd. (The Boulevard, Blackmore Lane, Watford WD1 8YW). Prewetting, equilibration, sample binding, and sample elution were performed according to the recommendations of the manufacturer.

#### **2.2.5.7 Matrix-assisted laser desorption ionization mass spectrometry**

Time-of-flight (TOF) MALDI analyses were performed in a PerSeptive Biosystems Voyager-DE™ STR BioSpectrometry™ Workstation (Applied Biosystems). Running methods and parameters were altered according to the nature of individual samples. Altered parameters included laser intensity, guide wire voltage and resolution, and are listed on individual spectra pages in chapter 5.

### **2.2.6 Molecular Biology Methods**

#### **2.2.6.1 Plasmid isolation**

Double stranded plasmid DNA was prepared from bacterial cultures using the Hybaid Recovery™ Plasmid and Cosmid Prep Kit (Hybaid-AGS). The procedure is based on a silica gel and is predicted to deliver high yields of supercoiled plasmid which is suitable for downstream enzymatic manipulations such as restriction enzyme digestion and PCR.

Cells were grown overnight at 37°C as 5ml cultures in LB media containing appropriate antibiotic selection. Bacteria were harvested from 1.5ml portions of this culture by spinning in a microcentrifuge for 30 seconds. Supernatant was completely discarded and the cell pellet resuspended, by pipetting, in 50µl Pre-Lysis Buffer. To

this was added 100µl Alkaline Lysis Solution and mixing was continued until the solution was clear and viscous. Following addition of 75µl Neutralizing Solution the mixture was briefly vortexed and spun for 2 minutes. Supernatant was removed with a pipette to a spin filter, care being taken to avoid uptake of white precipitate. Binding Buffer was shaken thoroughly and 250µl was added to the spin filter where it was pipette-mixed with the cell lysate prior to centrifugation of the filter (in a capture tube) for 1 minute. The filter/DNA was washed by the addition of 350µl Wash Solution followed by a 1 minute spin. The collection tube was emptied and the filter re-spun to dry the pellet. The filter was transferred to a clean capture tube into which DNA was eluted (by addition of 50µl sterile ddH<sub>2</sub>O to the matrix/DNA and brief vortexing followed by a 30 second spin).

#### **2.2.6.2 DNA sequencing**

DNA sequencing was performed by the DNA Sequencing Laboratory, University of Durham, on an Applied Biosystems 373A DNA Stretch Sequencer using double stranded DNA templates and dye-dideoxy terminator chemistries.

#### **2.2.6.3 Polymerase chain reaction (PCR)**

PCR reactions were conducted in a Robocycler<sup>®</sup> thermal cycler (Stratagene Ltd., Cambridge Innovation Centre, Cambridge Science Park, Milton Road, Cambridge CB4 4GF). Specific running conditions are provided in chapter 6. Reaction mixtures (50µl per reaction) consisted of 0.5µl Vent<sub>r</sub> DNA Polymerase (New England Biolabs UK Ltd., 67 Knowl Piece, Wilbury Way, Hitchin, Hertfordshire SG4 0TY), 1x reaction buffer (New England Biolabs), 2mM MgSO<sub>4</sub> (New England Biolabs), 25mM each dNTP (Pharmacia), 10pmol primer ARAP1b, 10pmol primer ARAP2,



approximately 1ng template (plasmid 158J20T7). Oligonucleotide primers for DNA sequencing and PCR were ordered from MWG-Biotech UK Ltd., Waterside House, Peartree Bridge, Milton Keynes MK6 3BY.

#### **2.2.6.4 Restriction digestion of DNA**

Restriction endonucleases were obtained from Boehringer Mannheim UK (Diagnostics & Biochemicals) Ltd., Bell Lane, Lewes, East Sussex BN7 1LG) and were used in conjunction with accompanying buffers as recommended by the manufacturer.

#### **2.2.6.5 Gel purification of DNA**

As a precursor to ligation reactions, PCR fragment inserts and recipient plasmids were purified by extraction from agarose following gel electrophoresis.

In the PCR subcloning of the *Arabidopsis* EST 158J20T7 putative PAP open reading frame into the PCR cloning vector pGEM<sup>®</sup>-T Easy (Promega Ltd.), both DNA species were purified in this manner prior to ligation using the QIAquick Gel Extraction Kit (Qiagen Ltd., Unit 1, Tillingbourne Court, Dorking Business Park, Dorking, Surrey RH4 1HJ). The procedure is suitable for providing purified DNA of 70bp to 10kb for use in manipulations such as ligations. The DNA fragment was excised from the agarose gel on a UV transilluminator using a clean razor blade. The gel slice was cut as close to the DNA as possible, weighed and placed in a microcentrifuge tube. To this was added 3 volumes of Buffer QG to 1 volume of gel (ie. 300µl per 100mg). The tube was incubated at 50°C for 10 minutes, with occasional vortexing, to completely dissolve the gel slice. One gel volume (100µl per original 100mg) of

isopropanol was added to the sample and mixed, and this was applied to a QIAquick spin column contained in a 2ml collection tube. The column was centrifuged for 1 minute to bind DNA, flow-through was discarded, and the column was replaced in the collection tube. Washing was achieved by addition of 0.75ml of Buffer PE to the column, with centrifugation for 1 minute. Flow-through was discarded and the column was re-centrifuged for 1 minute.

At this point the column was transferred to a clean microcentrifuge tube. DNA was eluted by addition of 50µl ddH<sub>2</sub>O onto the centre of the column matrix with centrifugation for 1 minute.

Preparation of DNA for ligation of the *Arabidopsis* EST open reading frame into the expression vector pET-24(a) (Novagen CN Biosciences UK, Boulevard Industrial Park, Padge Road, Beeston, Nottingham NG9 JR) was from agarose using the GFX™ PCR DNA and Gel Band Purification Kit (Pharmacia). As with the Qiagen procedure, DNA purified by this method can be used in a variety of applications including restriction digestion and subcloning.

As previously, a DNA/gel slice was excised and weighed, and placed in a clean microcentrifuge tube. One volume of Capture Buffer was added to 1 volume of gel (100µl per 100mg). Vigorous vortexing was followed by incubation at 60°C until the agarose had become completely dissolved (5-15 minutes). The solution was briefly centrifuged to collect sample at the bottom of the tube and applied to a GFX column placed in a collection tube. Incubation at room temperature for 1 minute was followed by centrifugation for 30 seconds. Flow through was discarded from the collection tube and the column replaced. Washing was achieved by addition of 500µl

Wash Buffer to the column followed by centrifugation for 30 seconds. The column was transferred to a clean microcentrifuge tube and 50µl sterile ddH<sub>2</sub>O was added directly to the top of the column matrix. Following incubation at room temperature for 1 minute, elution was achieved through centrifugation for 1 minute.

Small aliquots of gel-purified DNA samples were routinely analysed by agarose gel electrophoresis as a means of assessing recovery.

#### 2.2.6.6 Ligation reactions

Ligations were catalysed by T4 DNA ligase (Promega Ltd.) typically with a vector to insert molar ratio of 1:3, as calculated using the following equation:

$$\frac{\text{ng of vector} \times \text{size (kb) insert}}{\text{size (kb) vector}} \times \text{insert: vector ratio} = \text{ng insert}$$

Appropriate amounts of the vector and insert were mixed with the following components in a final total volume of 10µl:

|                          |     |
|--------------------------|-----|
| T4 DNA Ligase 10x Buffer | 1µl |
| T4 DNA Ligase (3U/µl)    | 1µl |

Control reactions were performed in which vector but no insert DNA was included. Ligation reactions were incubated at 16°C, overnight, and the total reaction volume used directly for transformation into the appropriate bacterial strain.

#### **2.2.6.7 Bacterial transformation**

Competent bacterial (*E.coli*) cells used were XL1-Blue, BL21 (DE3) and BL21-CodonPlus™(DE3)-RIL (Stratagene Ltd.). These were stored at -80°C as 100µl aliquots. For each transformation, plasmid DNA solution was made up to 10µl containing approximately 50ng (either as a pre-existing intact plasmid or as a completed ligation reaction mixture). This was added to a competent cell aliquot and the mixture then incubated on ice, for 30 minutes. The cells were heat shocked at 42°C for 45 seconds and placed back onto ice for 2 minutes. The solution was transferred to a sterile 15ml Falcon tube and mixed with 0.9ml LB media. This culture was incubated at 37°C for 1 hour after which cells were plated onto solid LB agar media containing the appropriate antibiotic selection. Cells were grown as (i) a 1/10 plate (100µl culture directly spread onto the plate) and (ii) a 9/10 plate (cells in the remaining 0.9ml culture were pelleted and, after decantation of supernatant, resuspended in residual media for spreading). Background false positive levels were assessed by reference to a control transformation which consisted of cells alone with no plasmid addition. Verification of positive transformants was by growth of culture from colony and plasmid isolation by agarose gel electrophoresis, to allow comparison with non-recombinant plasmid DNA or following restriction digestion to demonstrate fragment excision.

#### **2.2.6.8 Radiolabeled probe generation**

Radiolabeled probe was generated by random priming from a 870 bp gel-purified fragment (corresponding to the open reading frame of a putative *Arabidopsis* PAP) using the *rediprime*™ Random Primer Labelling kit (Amersham Life Science Ltd., Amersham Place, Little Chalfont, Amersham, Buckinghamshire HP7 9NA).

Denatured DNA template (50µg) was added (in 10mM Tris-HCl pH 8, 1mM EDTA) to the *rediprime*<sup>™</sup> reaction mix to give a final volume of 45µl. Addition of 5µl [<sup>32</sup>P]dCTP (Amersham International Plc., Little Chalfont, Amersham Buckinghamshire HP7 9NA) was followed by incubation at 37°C for 10 minutes. The reaction was stopped by adding 5µl of 0.2M EDTA. The reaction mixture was passed through a micro-spin chromatography column (Bio-Rad) according to instructions. Sample was applied directly to the centre of the column followed by centrifugation for 4 minutes at 1000 x g. The purified sample was eluted in 50µl Tris buffer. Radioactive incorporation was approximately 10<sup>8</sup> cpm/µg.

#### **2.2.6.9 Northern hybridization analysis**

RNA (poly-A<sup>+</sup>) from *Brassica napus* leaf, root and embryo was prepared for electrophoresis. Samples (2µg) of each were separated on a 1% agarose gel (containing 1 x MOPS buffer and 17.5% formaldehyde) at 100V for 3 hours. The gel was soaked in several changes of sterile DEPC-treated H<sub>2</sub>O before transfer onto double thickness Whatman paper soaked in 20 x SSC with ends in contact with a 20 x SSC reservoir. A gel-size piece of Hybond<sup>™</sup>-N nylon membrane (Amersham Life Science Ltd.) was wetted with H<sub>2</sub>O and 20 x SSC and placed onto the gel. The membrane was overlaid with 3 pieces of soaked, and 3 pieces of dry Whatman paper (gel-size). On top of the assembly was placed a cushion of 3 nappy layers and a glass bottle as weight. The gel was left to blot overnight. Probe was generated as described above. Prehybridization was performed in a total volume of 30ml containing the following: 5x SSPE, 5x Denhardt's solution, 0.5% (w/v) SDS, 25% deionized formamide. To this was added 1mg denatured herring sperm DNA and incubation was carried out at 42°C for 3 hours. Hybridization was performed in a

12.5ml volume comprising fresh prehybridization solution with probe and 10% dextran sulphate included. Incubation was at 42°C, overnight. Blots were washed first in 4x SSC (copious) at 42°C for several minutes. Subsequently washing took place twice in 2x SSC for 15 minutes, at 60°C. Conditions of low stringency were employed following failure to detect target signals at higher stringencies. Blots were wrapped in polythene and exposed to X-ray film in an autoradiography cassette for 1 week.

#### **2.2.6.10 Southern hybridization analysis**

Duplicate genomic DNA samples from *Arabidopsis* and *Brassica napus* (2µg and 10 µg respectively) were digested with different restriction endonucleases precipitated with 70% ethanol, 150 mM NaCl prior to loading into separate wells of a 0.5% agarose gel.

After electrophoresis the gel was soaked in denaturing solution (30 minutes) followed by neutralizing solution (2 x 15 minutes). Hybridizations were performed using Hybond™-N nylon membrane as described for Northern blotting and radiolabeled probe was generated by random priming from the *ARAPAP* ORF. Prehybridization was performed in a 20ml volume of the following: 5x SSPE, 5x Denhardt's solution, 0.1% (w/v) SDS, 4mg denatured herring sperm DNA. Incubation was at 65°C for 3hrs. For hybridization, probe was directly added to the prehybridization mixture followed by incubation at 65°C, overnight. Blots were washed twice for 15 minutes at 60°C in 2x SSC. Membrane was exposed to X-ray film for 3 weeks.

#### **2.2.6.11 Autoradiography**

Exposed X-ray film (RX-100) (Fuju Photo Film UK Ltd) was developed in a Compact X4 automatic x-ray film processor (X-ograph Imaging Systems, Malmesbury, Wiltshire SN16 9JS).

## 2.2.7 Biochemical assays

### 2.2.7.1 Phosphatidic acid phosphatase (PAP) assay

The assay for PAP activity was performed and radiolabeled phosphatidic acid (PA) substrate synthesized as described by Pearce (PhD thesis, 1997). DAG kinase (*sn*-1,2-diacylglycerol kinase, *E.coli*) was purchase from Calbiochem CN Biosciences UK, Boulevard Industrial Estate, Padge Road, Beeston, Nottingham NG9 2JR.

The DAG kinase synthesis of [<sup>32</sup>P]PA was performed in a 100 µl reaction volume in a pyrex glass reaction tube. The reaction consisted of 51mM octylglucoside, 1mM cardiolipin, 60mM imidazole-HCl pH6.6, 50mM NaCl, 12,5mM MgCl<sub>2</sub>, 1mM EGTA, 0.3mM DTPA, 2mM DTT, 2mM dioleoylglycerol and 50µM [ $\gamma$ -<sup>32</sup>P]ATP (30Ci/mmol, 10mCi/ml, 150µCi total) and the reaction was initiated with 0.05 unit of DAG kinase. After an incubation of 1 hour at 25°C the reaction was terminated with 60µl 10% perchloric acid. To extract the [<sup>32</sup>P]PA from the [ $\gamma$ -<sup>32</sup>P]ATP 2.5 ml ChCl<sub>3</sub>:CH<sub>3</sub>OH (1:1) was added and the mixture vortexed for 5 minutes. Two phases were created by the addition of 1 ml H<sub>3</sub>PO<sub>4</sub> in 1M KCl. The phases were separated and the ChCl<sub>3</sub> phase was extracted with fresh aqueous phase (1 ml H<sub>3</sub>PO<sub>4</sub> in 1M KCl) while the aqueous phase was extracted with 2.5ml fresh ChCl<sub>3</sub>:CH<sub>3</sub>OH (1:1). The two chloroform phases were pooled and stored at -20°C until use.

PAP assays were performed in a total volume of 100 $\mu$ l and were made up of 0.5mM [<sup>32</sup>P]PA (1000-2000 dpm/nmol), 50mM Tris-maleate pH 6.0, 5mM Triton X-100, 2mM MgCl<sub>2</sub>, and protein sample. Reactions proceeded at 25°C.

#### 2.2.7.2 Para-nitrophenyl phosphate (*p*-NPP) assay

Non-specific phosphatase activity was monitored using the Sigma 104™ Phosphatase Substrate adapted from the method described by Pearce (PhD thesis, 1997). The assay monitors the change in absorbance at 405nm when *para*-nitrophenylphosphate (*p*-NPP) is phosphohydrolysed to produce *para*-nitrophenol, a yellow coloured compound. Each reaction consisted of 50mM Tris-maleate pH 6.0, 2mM MgCl<sub>2</sub>, protein sample, and 2mM *p*-NPP in 100 $\mu$ l. Reactions were incubated at 30°C and were terminated by the addition of 0.5 ml 0.1M HCl in CH<sub>3</sub>OH. Water was added to a final volume of 1 ml and absorbance was measured at 405nm.



## CHAPTER 3

# Preparative Isolation of Endoplasmic Reticulum from Castor Endosperm, Characterization and Subfractionation of Protein Constituents

### 3.1 Introduction

In the storage tissues of higher plants the endoplasmic reticulum is a particularly important organelle. Studies on this fraction are specifically important to this laboratory since it will allow the identification of proteins specific to both germination and maturation, and is a potential rich source of proteins involved in triglyceride/oil body formation. These events take place during maturation of the endosperm and seed embryo, in preparation for subsequent energy-requiring germination processes.

Castor has long been recognised as particularly amenable to subcellular fractionation because it has a liquid endosperm which allows for relatively gentle homogenization which allows for membrane systems and organelles to remain intact during isolation. However, bulk preparative isolation, required for detailed protein characterization, has not been attempted. Despite undoubted challenges likely to confront such scale-up, it was expected that a combination of optimizing individual isolations, during a period of high throughput, would provide the most effective approach for efficient generation of required quantities of ER.

Rapid advances have been made in the proteomics area allowing for the identification of proteins in complex samples. These have centred around 2-dimensional electrophoresis with mass spectrometry techniques of peptide mass fingerprinting and protein sequencing. Key to any study however is the purity of the membrane fraction being analysed. Previous ER-based investigations have made use of a microsomal membrane fraction which despite being enriched in ER are clearly contaminated with other membranes such as mitochondria. Such 'microsomal' fractions are therefore unsuitable for this work and their interpretation could be held to be subjective. The approach we would prefer to use therefore resides around the generation of a pure ER fraction.

Lord and colleagues in 1973 isolated castor ER membranes employing a single sucrose density gradient (Lord *et al*, 1973) and found that inclusion of magnesium ions in the sucrose prevented dissociation of ribosomes from the membranes and consequently increased isopycnic density from 1.12 to 1.16 g/cm<sup>3</sup>. In 1983 Conder and Lord reported on the judicious application of a second gradient through which membranes were floated and thus further purified (Conder & Lord, 1983). Finally this procedure was refined by extension of the flotation step from 17 to 22 hours, resulting in high purity and fully floated smooth ER containing milligram quantities of protein (Coughlan *et al*, 1996). Identity of the fraction as ER was confirmed by assaying for the ER marker enzymes antimycin-insensitive NADH:cytochrome-*c* reductase and CDPcholine:1,2-diacylglycerol choline phosphotransferase. Purity was measured by the absence of contaminating endomembrane markers. This method of isolation was adopted here, and during the work

described I was privileged to gain expertise in the techniques first hand during a period in the research group of Dr Sean Coughlan (then of Pioneer Hi-Bred Int.)

The success of any individual ER preparation, in terms of quality (purity, membrane integrity) and yield (total protein), was found to be critically affected by several factors at various points in the procedure. Before isolation could be undertaken reliable sources of 5 day-germinated and 25 days-after-flowering seed material were required.

Dry seed was be routinely germinated for ER preparation in this laboratory. Supply of developing endosperm was more limiting and the preparative phase of the work made use of fully grown plants kindly made available by D Coughlan.

Similarly, during the 5 day incubation period, some seedlings proceeded beyond germination and endosperm appeared depleted and soft. There is thus a heterogeneity in the stage of development following sowing and careful note needs to be made of the morphological appearance to obtain suitable experimental material. Ideally endosperm, to be used in isolations, were free of the testa and were more typically bean shaped. The seed had also begun to split revealing cotyledons which were yellow to orange in colour and sufficiently differentiated to separate from endosperm with ease. Variation amongst developing seeds was observed even at the level of single inflorescences, and selection was made at 25 days after flowering (DAF) to include endosperm which was sufficiently mature to be distinguished from surrounding inner integument tissue but had not yet reached the dryer stages of maturation. This corresponds to the mid-to-late cotyledon stages (V-VII) of castor seed development (Greenwood & Bewley, 1981) At this stage endosperm was very moist and easily sliced through with a razor blade.

Given the availability of sufficient starting material, thoroughness of homogenization was believed to be the single most important processing factor in determining final yield. Increasing the homogenization time increased the final ER protein yield from as little as 0.3g to 1mg. The aim of homogenization was to achieve maximal cell breakage in as gentle a way as possible. This was considered to have been reached after 20 minutes, and resulted in a finely-diced paste.

During the work performed for this chapter it was also considered advantageous to generate distinct ER subfractions in an attempt to simplify gel profiles and at the same time enrich for separate protein classes. Specific to castor ER membranes, two selective washes have been reported, namely sodium carbonate (Coughlan *et al*, 1996) and Triton X-100 (Coughlan *et al*, 1997). Carbonate is effective in removing peripheral as well as luminal components from enclosed membrane structures (Fujiki *et al*, 1982). Its application to castor ER revealed a capability to extract approximately 75% of soluble protein (as represented by the major 55kDa band) from the membrane. Triton X-100 was similarly effective, in conjunction with 100 mM NaCl, in generating two ER subfractions, designated reticuloplasmin and membrane on the basis of extractability and non-extractability respectively, in a 0.1% Triton X-100 (v/v) treatment. Developing ER appeared to be more resistant to this extraction. There was residual contamination of membranes by predominant reticuloplasmins after a single wash, postulated to reflect a temporal higher molar abundance of these chaperones. In this laboratory sodium chloride alone has been shown to release phosphatidic acid phosphatase from plant microsomes

despite a close metabolic proximity to membrane-bound enzymes of complex lipid synthesis (Pearce and Slabas, 1998).

Since samples were prepared both at Pioneer and in Durham this allowed for comparison between preparations and also critical assessment of purity using known 'homogenous' samples generated by Dr Sean Coughlan. Individual characterization of two abundant components was used to verify the authenticity of these preparations as endoplasmic reticulum. Furthermore, primary analyses have been extended to include ER subfractions generated by treatment specifically with sodium chloride and by chloroform/methanol solubilization.

### **3.2 Isolation of endoplasmic reticulum from castor endosperm**

The importance of seed selection and homogenization has already been discussed and figure 3.1 shows the various states of germinating/ germinated seeds encountered after 5 days' incubation. Once suitable endosperm (see figure 3.1b) had been homogenized fractionation could proceed.

#### **3.2.1 The first step – sucrose density gradient ultracentrifugation**

The first separation procedure following tissue homogenization involved centrifugation of ER membranes onto a 20/30% (w/w) sucrose interface. The first clear indication of the quality and yield of a preparation came from observation of the membranes at this interface. A strong and discrete band at this position was indicative of both efficient homogenization and a high quality sucrose step. It also indicated that there was likely to

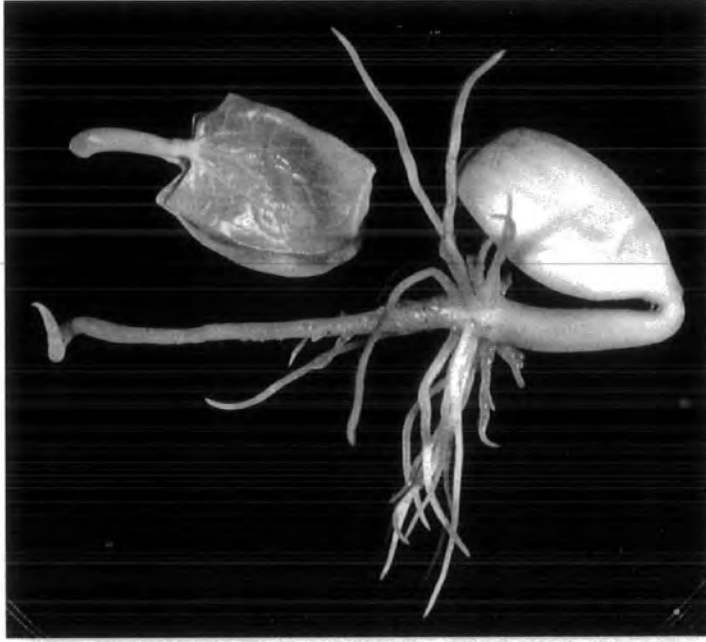
**Figure 3.1: Various states of 5-day germinated castor seed.** Dry seeds (*Ricinus communis* L. var. Hale) were surface-sterilized and soaked overnight in running tap water before planting in moist vermiculite and dark incubation at 30°C. (A) Early germination; rounded endosperm (~1.5cm long axis), short primary root with little secondary root growth. Endosperm remains in tight association with testa. (B) Mid germination selected for ER preparation. Endosperm has become bean-shaped (~2.2cm long axis) and is free of testa. Cotyledons are yellow and easily separated from the dissected bean. Primary root length is ~5cm with substantial secondary root growth. (C) Late germination. Depleted endosperm is soft and dull in appearance (~2.7cm long axis). Cotyledons are deep orange. Extensive secondary root growth.

**A**



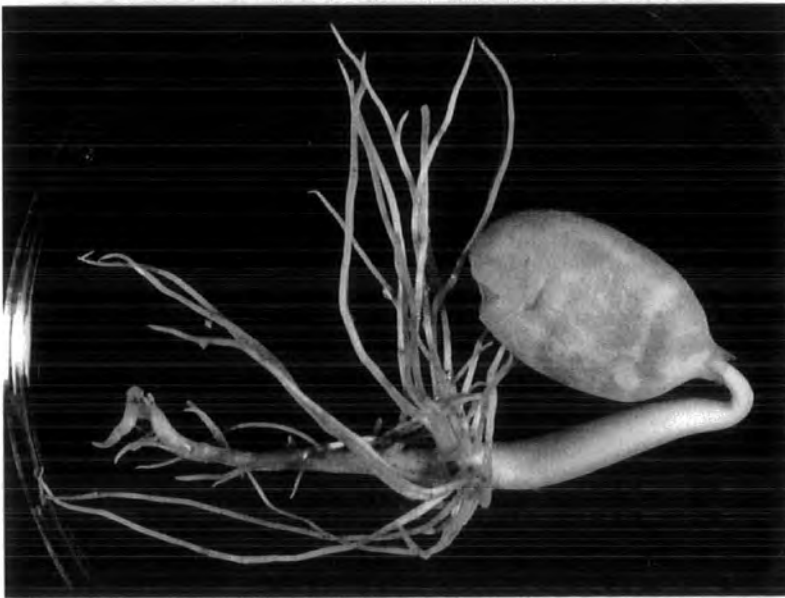
1 cm

**B**



1 cm

**C**



1 cm

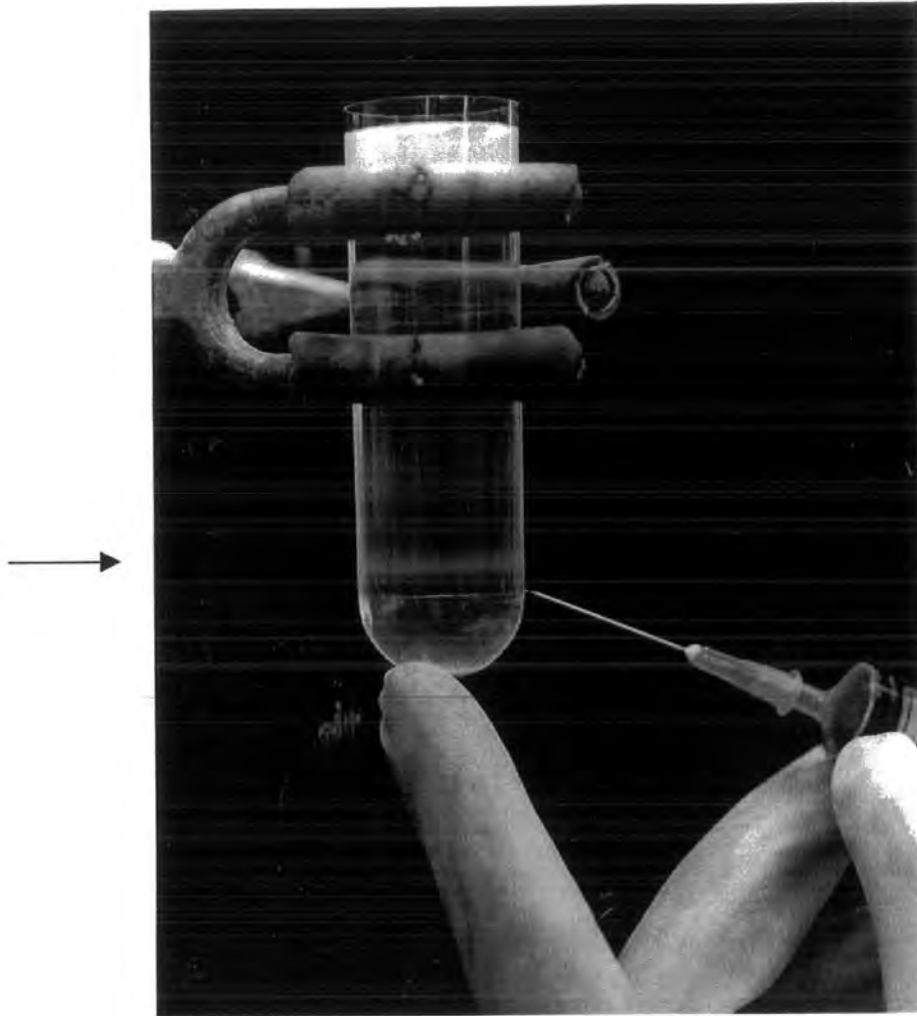
be a high yield of ER. Figure 3.2 shows the banded fraction from a germinating ER preparation. Although robust to an extent, great care was taken in generating discontinuous sucrose gradients since integrity of the 20/30% interface ensured optimal focusing of ER into a narrow region and therefore allowed collection of a fraction of the highest purity possible at this stage in a convenient manner. Of particular importance are problems in gradient overloading (eg. following particularly efficient germination and homogenization). In this situation the membrane band appeared exceptionally strong, and broader than usual, potentially compromising fraction purity. Aggregation occasionally associated with such high membrane concentrations also had the effect of impeding flotation during the second ultracentrifugation step, thereby effecting the final yield. Aggregates which did float were however pure as revealed by SDS-PAGE. These preparations behaved as normal upon standard pelleting and thorough resuspension.

Withdrawal and dispensing of membranes through hypodermic needles was performed very slowly to avoid disruption to organellar structure. The SDS profiles of the first step ER fraction are shown in representative gels as described in section 3.3 (figure 3.4a and b, lane 4).

### **3.2.2 The second step – sucrose density gradient flotation ultracentrifugation**

The partially pure ER fraction from the first sucrose step was divided between 4 narrower tubes for flotation. Construction of discontinuous flotation gradients was consequently a





**Figure 3.2: Appearance of the first preparative sucrose gradient following ultracentrifugation.** Crude cell supernatant was layered onto the gradient and spun for 2 hours at 25 000 rpm (Beckman SW28 rotor). ER membranes were visible as a discrete band at the position of the 20/30% (w/w) sucrose interface (density = 1.12g/cm<sup>3</sup>). This photograph illustrates the effective implementation of an intact sucrose interface. The fraction was withdrawn into a hypodermic syringe through the tube wall.



delicate procedure and, the integrity of the 20/30% sucrose junction being particularly important for obtaining a pure fraction.

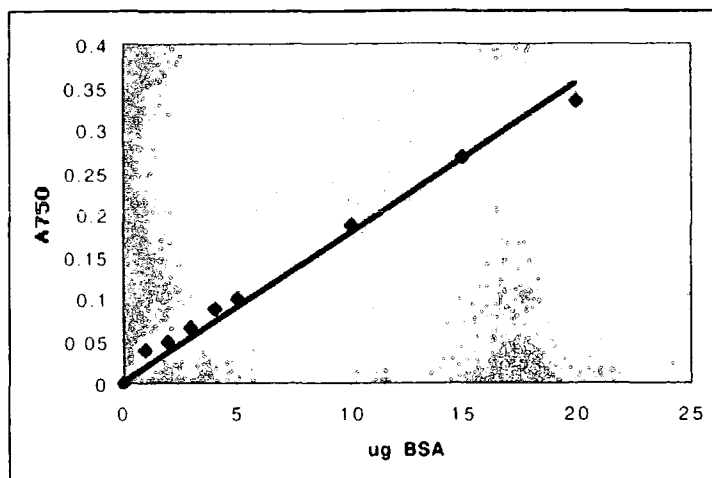
Occasionally during developing ER isolation a second membrane band was visible at the 30/40% sucrose interface. Rather than discard this component, comparison of the two fractions was made by SDS-PAGE which revealed them to be identical (figure 3.4b, lane 6a). This suggested that the 30/40% fraction was residual ER resulting from a combination of sample ER-richness and incomplete flotation. The occurrence was rare however, and it was estimated that 30/40% membranes did not exceed 25% of total ER in the gradient. It is worth noting that this fraction was restricted to developing ER preparations and is therefore likely to reflect the relative high abundance of ER in the tissue at this stage in development.

Membrane pellets were light brown and green in appearance, and slightly gelatinous. Resuspension was routinely performed in 0.5 – 1.0 ml of 10% glycerol per preparation and was achieved by gently pipetting in increasing volumes of this solution until an aggregation-free and slightly turbid suspension was obtained. Where problems were encountered with aggregation glycerol solutions were buffered with 10mM Tris (pH 7.8) and 1mM EDTA.

The protein yields for representative germinating and developing ER samples are presented in figure 3.3, and were estimated using the TCA/Lowry assay method.

**Figure 3.3: Typical Lowry estimation of ER protein content using a BSA standard curve.** Samples were assayed in duplicate alongside a series of BSA concentration standards ranging from 0-20 $\mu$ g. (A) BSA standard curve with example germinating ER estimation according to the linear (zero-intercepting) trendline equation  $y = 0.0176 x$ . (B) Catalogue of representative germinating and developing ER preparations. Developing ER protein yields were typically higher than in germinating samples.

**A**



|                                | A750  | Mean   | $\mu\text{g}/\mu\text{l}$ |
|--------------------------------|-------|--------|---------------------------|
| Germ. ER<br>(5 $\mu\text{l}$ ) | 0.095 | 0.0935 | 1.06                      |
|                                | 0.092 |        |                           |

**B**

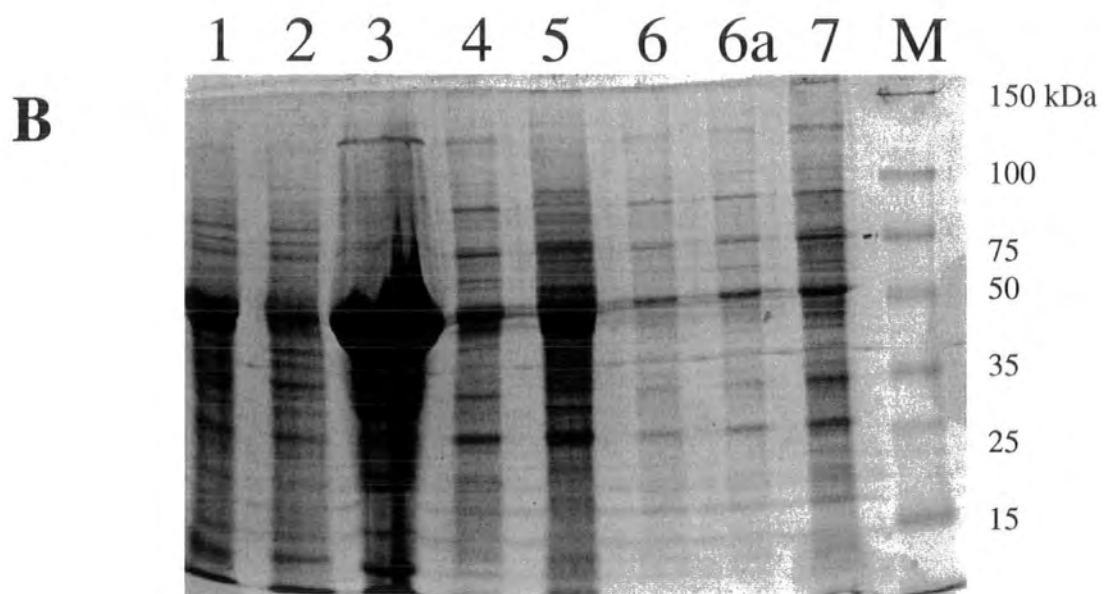
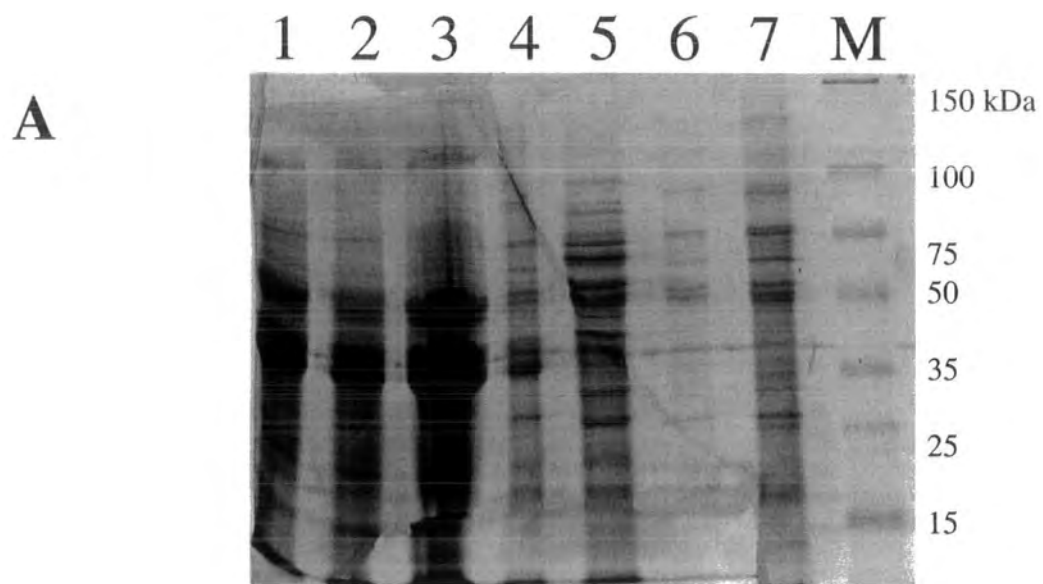
| Preparation | Type        | Date    | Yield (mg)        |
|-------------|-------------|---------|-------------------|
| 1           | germinating | 8.1.97  | 1.0               |
| 2           | germinating | 27.2.97 | 0.9               |
| 3           | germinating | 6.5.97  | 1.0               |
| 4           | germinating | 5.8.99  | 1.0               |
|             |             |         | average = 0.975mg |
| 1           | developing  | 22.1.97 | 1.5               |
| 2           | developing  | 14.5.97 | 1.7               |
| 3           | developing  | 29.5.97 | 1.8               |
| 4           | developing  | 9.6.97  | 1.8               |
|             |             |         | average = 1.7mg   |

### 3.3 Analysis of ER membrane preparations by SDS-PAGE

Endoplasmic reticulum preparations were subjected to SDS-PAGE as a routine means of visually assessing fraction quality and providing a further indication of protein concentration. These analyses provided a convenient assessment of inter-preparation reproducibility, and demonstrated that the ER isolation procedure was generating reproducible protein profiles. Assessment of the quality of ER preparation attainable was greatly assisted by means of working in the laboratory of Dr Coughlan which allowed for SDS profiles of preparations to be compared with gel profiles of authentic ER which had been assayed for biological activity. Gel profiles were identical. On this basis, biochemical marker assays were considered unnecessary. The assays of Coughlan *et al* (1996) had found that the membranes isolated from the first sucrose step "were heavily contaminated by soluble proteins" and that these "were removed by flotation centrifugation". Following flotation "all of the ER enzymatic marker activity (antimycin-A-insensitive cytochrome-c reductase) had been concentrated at the 20/30% sucrose interface. The soluble contaminating protein remained at the bottom of the gradient in the 50% sucrose step". Figure 3.4 shows full preparation profiles for both germinating and developing endosperm. The complexity of sample clearly decreased upon purification. The massively predominant bands present in the first three 'crude' fractions (lanes 1-3) are likely to represent various seed storage proteins precursors.

Comparison of the 20/30% ER fraction from the first and second sucrose gradients reveals clear difference and confirms the importance of the flotation gradient in removing extra soluble cytosolic contaminants (figure 3.4, lanes 4 and 6 respectively). In this

**Figure 3.4: Progressive fractions of castor endosperm endoplasmic reticulum preparations.** (A) Germinating preparation; (B) Developing preparation. Lane 1: Crude filtrate; Lane 2: Crude supernatant; Lane 3: Crude pellet; Lane 4: First sucrose gradient 20/30% (w/w) interface fraction; Lane 5: First sucrose gradient pellet; Lane 6: Flotation gradient 20/30% (w/w) sucrose interface fraction; Lane 6a: Flotation gradient 30/40% (w/w) sucrose interface fraction; Lane 7: Final ER suspension (20 $\mu$ g protein); M: Molecular weight markers (Novagen). Gel is Coomassie-stained.



regard, the profiles also confirm that the floated 20/30% fraction and the final ER pellets are identical in protein composition and that the pelleting step serves merely to concentrate the samples and allow transfer to a freezing-protective glycerol solution.

As noted previously, the developing 30/40% floated fraction appears identical to 20/30% ER (figure 3.4b, lanes 6 and 6a) by SDS-PAGE criteria.

It was clear from the SDS-PAGE profiles that significant protein difference existed between germinating and developing ER. Up to 40% of total protein was estimated to be accounted for by a few major components although the majority of differences between these samples was believed to be confined to less abundant components.

Several dominant common bands were putatively identified in germinating ER profiles based on previous publications (Coughlan *et al* 1996 and 1997). The following section provides verification of the identity of two of the most dominant protein components.

### **3.4 N-terminal amino acid sequencing – identification of protein disulphide-isomerase and calreticulin**

Total germinating ER protein was resolved using SDS-PAGE and electroblotted to PVDF membrane for N-terminal protein sequence analysis. A triplet of three major bands was selected for N-terminal sequencing. These bands, although not particularly well resolved following electroblotting, were chosen specifically because they were known to include protein disulphide-isomerase and calreticulin. As abundant proteins it was therefore



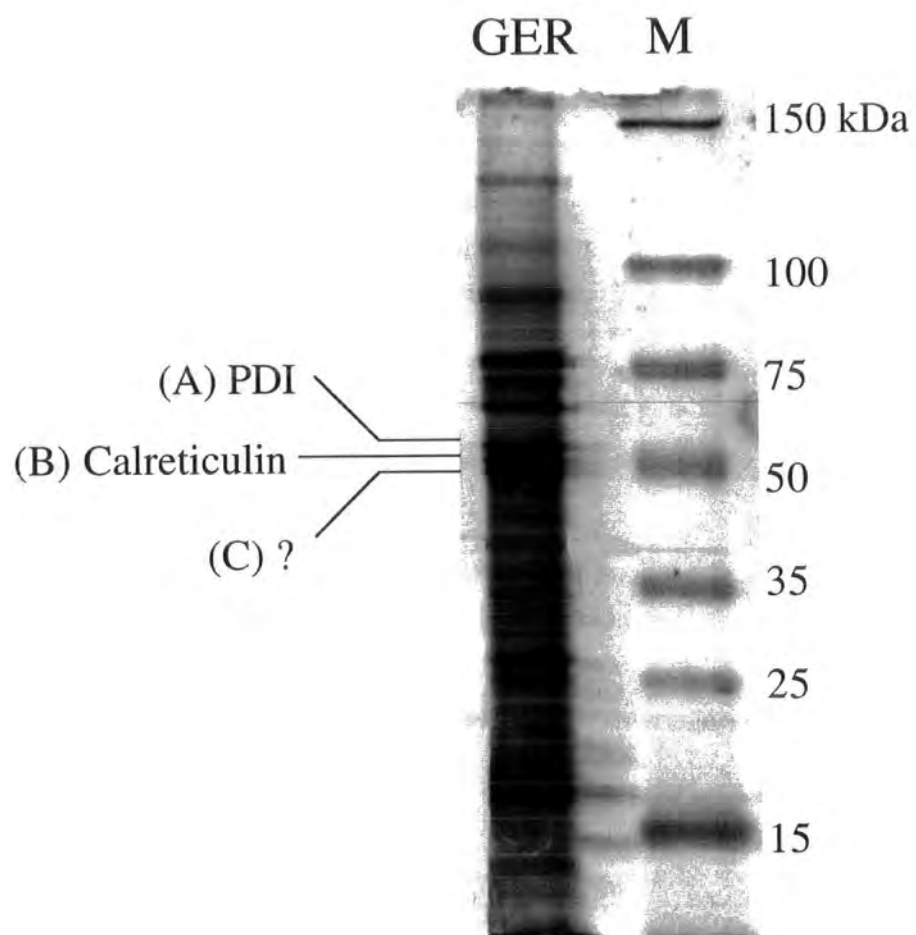
likely that positive identifications could be made and so the exercise would provide an indication of the suitability of N-terminal sequencing for a differential characterization. Assignment of these proteins would also support the contention that the sample was pure ER. A total of approximately 100µg protein was loaded across five lanes of the SDS mini gel. Bands excised were of approximate molecular weights 55, 50 and 45 kDa and designated A, B and C respectively (see figure 3.5). Although these components were clearly resolved by SDS-PAGE, following blotting some merging was observed and consequently bands A and B were both found to contain a mixture of the same two proteins. The relative intensity of individual amino acid peaks was used to separate the data into the correct respective band position (table 3.1) No sequence data was obtained from band C, which was assumed to be the result of N-terminal blockage.

| Protein | Molecular weight (kDa) | N-terminal amino acid Sequence Data | Sequencing yield (pmoles) |
|---------|------------------------|-------------------------------------|---------------------------|
| Band A  | 55                     | AESEEEQSSVL                         | 6                         |
| Band B  | 50                     | VFFEEQFEDG                          | 6                         |
| Band C  | 45                     | -                                   | -                         |

**Table 3.1: Identification of PDI and calreticulin from germinating ER.** N-terminal sequence data obtained from two predominant ER proteins. Band C yielded no data.

The data was used to query the public access protein database Swissprot in an attempt to assign functional identity to proteins A and B. Search results confirmed the 55kDa component as PDI for which an exact match was returned (Coughlan *et al*, 1996). The search results for band B returned a calreticulin-like sequence derived from spinach, which matched with nine out of the ten amino acids of the castor query (Menegazzi *et al*, 1993). This represents the first N-terminal data obtained for calreticulin from castor.

**Figure 3.5: Identification of protein disulphide-isomerase (PDI) and calreticulin from germinating ER.** Approximately 100 µg total germinating ER protein was resolved by SDS-PAGE (MiniGel, 10%) and semi-dry electroblotted to PVDF membrane (ProBlott). Excised Coomassie-stained bands were designated A, B and C and N-terminally sequenced in an Applied Biosystems gas phase sequencer. Band A yielded sequence from protein disulphide-isomerase: AESEEEQSSVL; Band B yielded calreticulin sequence: VFFEEQFEDG; Band C yielded no data and protein therein was assumed to be blocked.



Subsequent to this finding the cDNA cloning of castor calreticulin using an immunoscreening approach was reported (Coughlan *et al*, 1997). The data reported here represents the N-terminus of the mature (processed) protein since it matches exactly with the published sequence from residue 22 which follows a putative 18 amino acid signal peptide. The site of protein cleavage has thus been positively identified by direct sequencing of the mature protein.

### **3.5 Special considerations regarding sequencing technology and ER protein complexity**

It was estimated that approximately 100 pmol (5 $\mu$ g) of calreticulin was loaded onto the blotted gel and that in this particular experiment only 40% transfer was achieved through electrophoretic transfer onto PVDF. The calreticulin data was detected at a level of approximately 6pmoles which lies close to the sensitivity limit of the sequencer. In all, these figures represent protein loss of up to 90% from that originally loaded on the gel. Therefore, when blotting to PVDF for sequencing, experience indicated that the majority of bands in the sample were present in quantities too low to be amenable to characterization by this method. Furthermore, increased loading to overcome this problem runs the risk of band merging upon blotting, leading to contaminated data which could be difficult to decipher. There was consequently a clear requirement for the development of more sensitive technology (such as MALDI TOF mass spectrometry). An alternative means of addressing problems associated with sample resolution is to subfractionate protein constituents of total ER, in order to simplify gel profiles.

### **3.6 Membrane subfractionation procedures**

Since the intention was to subfractionate ER proteins using differential extraction with various reagents it was decided to develop the technique using readily available membranes from *E.coli* rather than waste castor material which was precious.

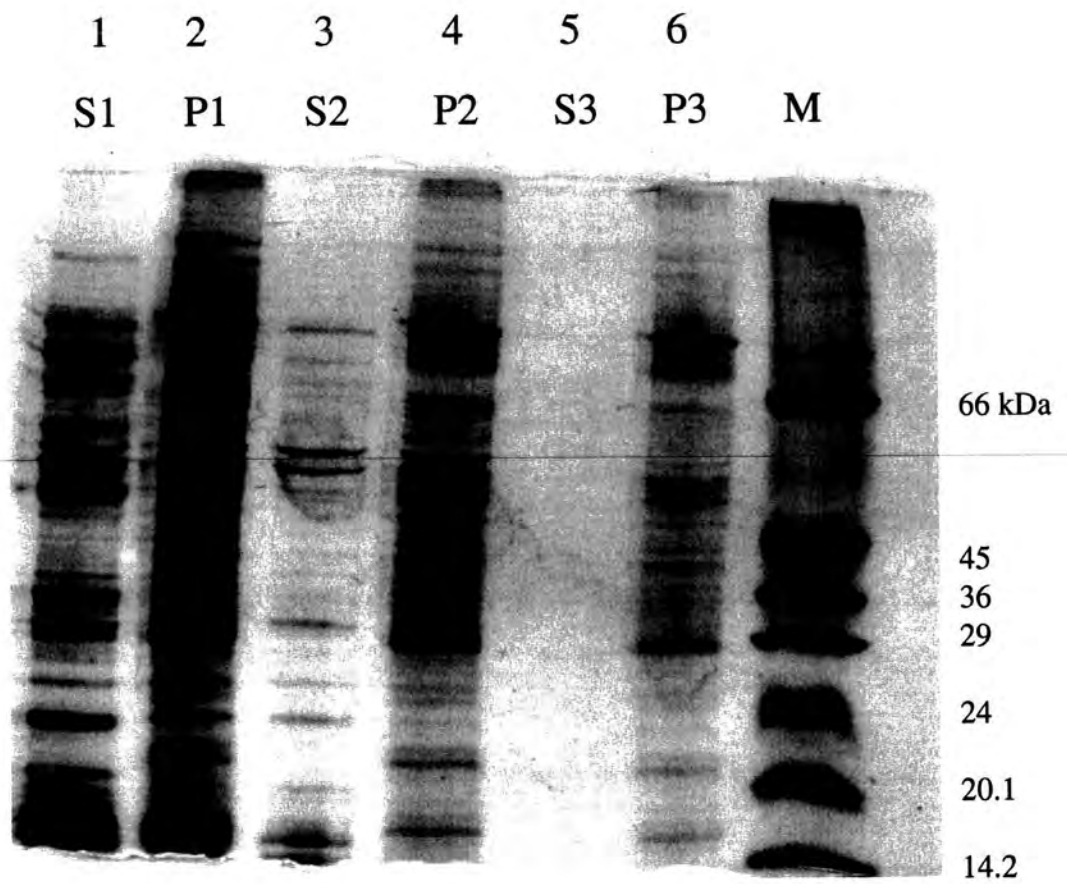
#### **3.6.1 Studies on an *E. coli* microsome fraction**

Subfractionation methods were developed on a small (100 $\mu$ l) scale to conserve material. Handling of membrane fractions at such a scale presents its own quantitative difficulties and practice in micromanipulation was obtained using an *E. coli* microsome fraction. *E.coli* microsomes were treated separately with 500mM sodium chloride, 100mM sodium hydroxide, 100mM sodium carbonate (pH 12), and 5mM EDTA (pH 7.5). All extractions involved a 30 minute incubation on ice, except the EDTA treatment which required a one hour incubation at room temperature, and all samples were subsequently collected by ultracentrifugation. Although all reagents tested were effective in removing a significant proportion of total protein (results not presented), sodium chloride was selected for further evaluation. This choice was made based on the fact that this reagent is likely to operate by a non-disruptive mechanism. The importance of this is illustrated by a subsequent requirement for sucrose gradient resolution of subfractions (chapter 4). Also, salt washing has been shown to release phosphatidic acid phosphatase from avocado microsomes (Pearce & Slabas, 1998).

Separate duplicate microsome samples were incubated with a range of sodium chloride concentrations. Extracted membranes were pelleted by ultracentrifugation (50 000 x g) and analysed, following resuspension, together with the corresponding supernatant (extract), following chloroform/methanol precipitation and resuspension, using SDS-PAGE. The experiment revealed all of the concentrations of extractant used to be effective in protein removal. It was expected that gradual increases in salt levels would have the overall effect of releasing more protein by disrupting increasingly tight ionic membrane-protein associations. This was evident with respect to several components which became specifically more strongly represented in the higher salt extracts. However, any general quantitative difference was not obvious. Salt extracts were from hereon performed using a 500mM concentration of NaCl since levels higher than this appeared not to result in any additional extraction. Also, subsequent requirement for salt removal made higher salt washes undesirable.

The possibility of incomplete extraction by a single wash was the subject of an extended procedure incorporating a sequential treatment to microsomal membranes. Once-extracted and resuspended membranes were subjected to a second, and then a third extraction in 500mM sodium chloride. The four fractions generated (three supernatants and three times-extracted membranes) are presented in figure 3.6. The results demonstrate that a single extraction removes most of the extractable protein (lane 1) and that the second treatment results in the complete release of the remaining material. The third extraction is therefore obsolete by this visualization criteria (compare lanes 3 and 5).

**Figure 3.6: Sequential extraction of *E. coli* microsomes with 500 mM sodium chloride.** Microsomes (20 mg) were incubated in the presence of 500 mM NaCl and subsequently ultracentrifuged 3 times. S: Supernatant (extract subfraction); P: Pellet (salt washed membrane subfraction). Lanes 1 and 2 represent samples illustrated in figure 3.6 and contain 1/20 and 1/50 of total subfraction respectively. Subsequent lanes represent further extraction/subfractionation of the membrane fraction in 500 mM NaCl. All extractable protein has been released by the end of the second wash. M: Molecular weight markers (Sigma). Loading of supernatants was increased 2.5-fold relative to the membrane fraction to aid visualization. Gel is Coomassie-stained.



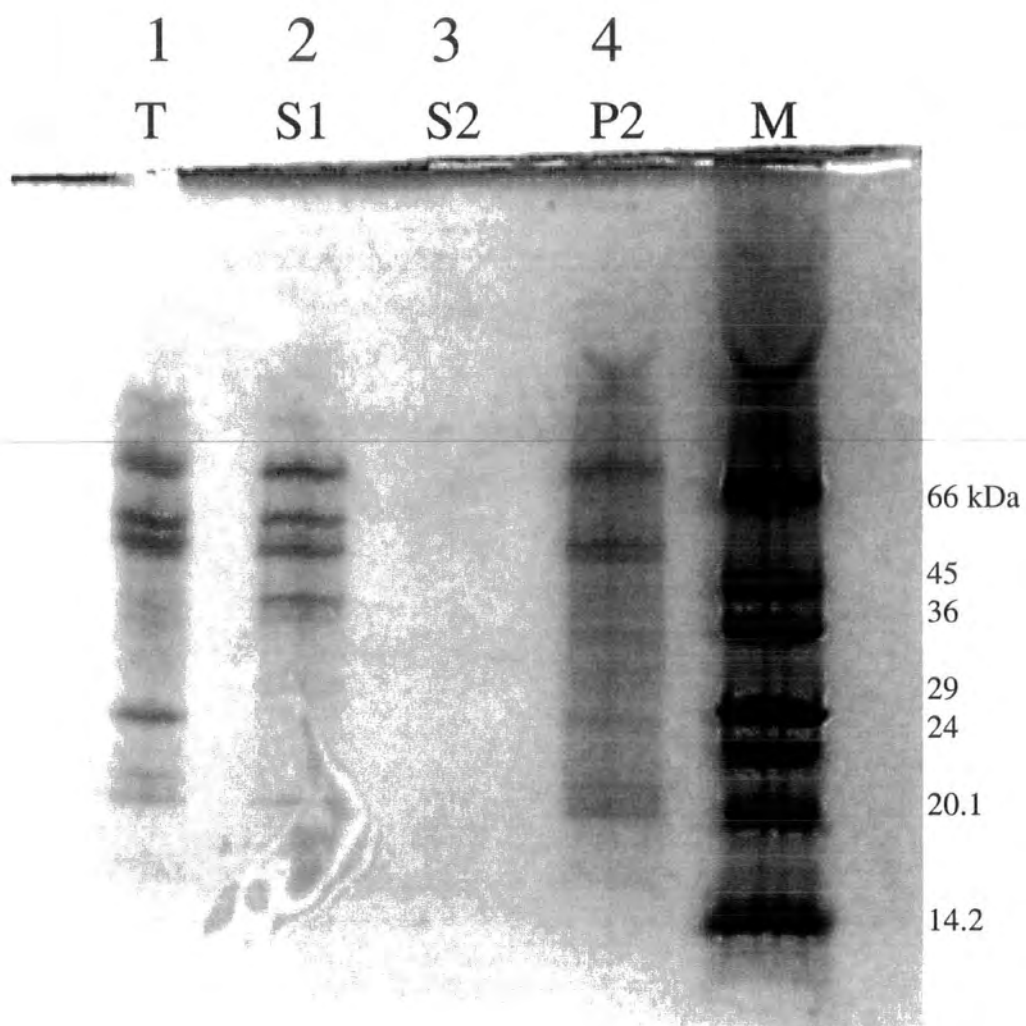


To further aid extraction efficiency in subsequent extractions of castor ER, the speed of ultracentrifugation was increased to 100 000 x g to ensure complete pelleting of the membranes. These experiments demonstrate the efficient implementation of a reagent-based membrane subfractionation technique at a micro-scale level, suitable for transfer to the experimental ER material.

### **3.6.2 Sodium chloride extraction of castor ER membranes**

Following preliminary trials in small-scale subfractionation and manipulation of membranes using *E. coli* microsomes a NaCl wash was developed for the subfractionation of castor ER. The method was first applied to germinating samples in order to conserve the more limited developing material. Figure 3.7 shows a sequential 500mM extraction and reveals an extremely effective first wash, no protein being visible in the second supernatant (lane 3). Several major components (notably at ~70, 60, 50, and 40 kDa) were stripped by the first extraction including a band unseen in the total ER, presumably now visible through enrichment (lane 2). Two predominant bands were visible in the membrane pellet fraction (lane 4), and although it is possible that these are residual contaminating major salt extractable components, they may equally appear due to their enrichment in the membrane fraction. On a first assessment it was apparent that the salt wash procedure was capable of generating two different fractions from germinating whole ER. Success in resolving and detecting the relatively low protein in a high salt system was mainly attributed to particular advantages (speed/thin gel surface) of the electrophoresis system employed (Phast System, Pharmacia). Differences between

**Figure 3.7: Micro-scale (50  $\mu$ l) sequential extraction of germinating ER in 500 mM sodium chloride.** PhastGel (gradient 8-25%, Pharmacia) showing a single membrane wash to be effective in releasing most of the salt-extractable protein visible at the Coomassie level of detection. ER membrane fraction ( $\sim$ 50 $\mu$ g) was incubated, on ice, in 500 mM NaCl and subfractionated by ultracentrifugation (100 000 x g). T: Total ER; S1: First supernatant; P1: Once-extracted membranes; S2: Second supernatant; P2: Twice-extracted membranes; M: Molecular weight markers (Sigma). Subfraction loadings are proportionally representative of the total sample (1/25).

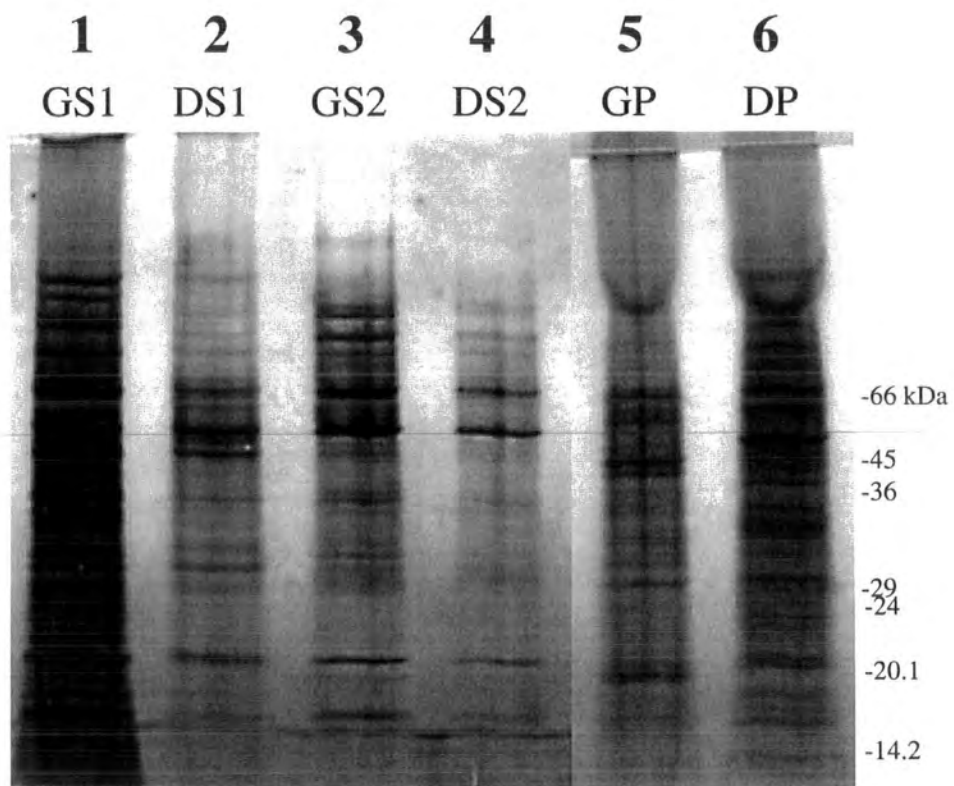


total ER profiles from the Phast system and conventional mini SDS-PAGE are due to gel concentration differences (8-25% gradient and 12% homogenous, respectively).

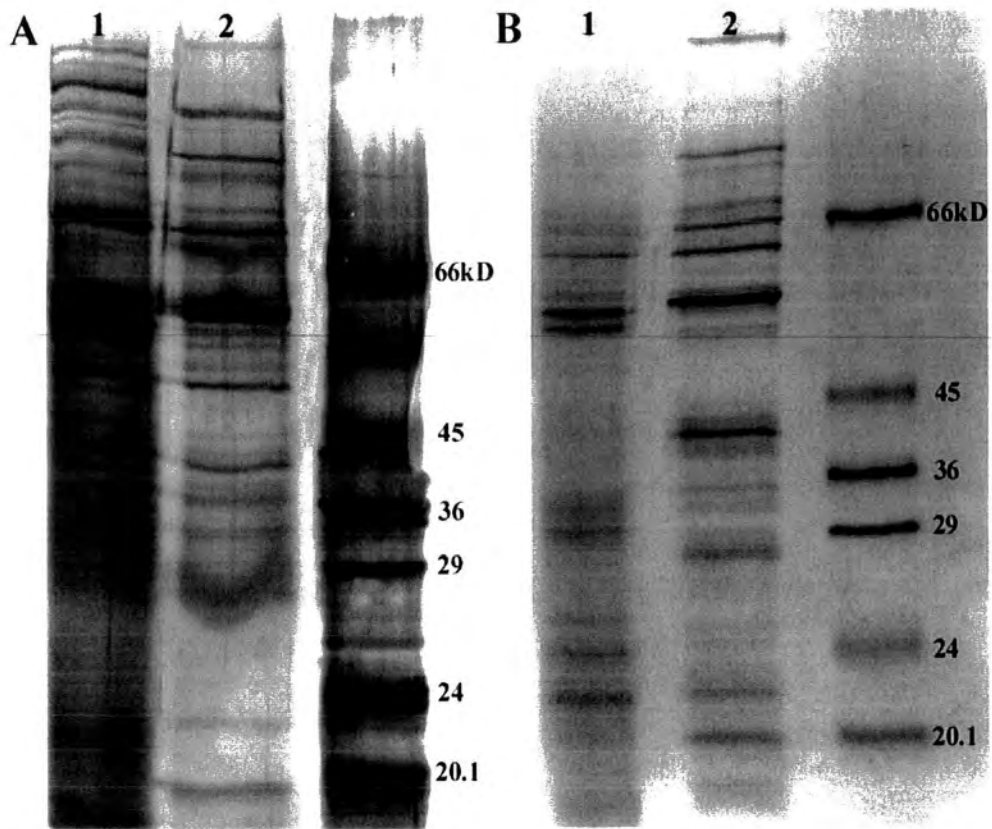
Developing ER was subsequently extracted and analysed alongside corresponding germinating salt subfractions (figure 3.8). Results indicated that developing ER membranes do not have so many proteins extracted with NaCl as seen with germinating material (lanes 3 and 4 versus lanes 1 and 2). This observation correlates with those previously reported for Triton X-100 (Coughlan *et al*, 1997). The high salt does however decrease the complexity of the sample.

At the salt subfraction level it seemed that sample complexity had been insufficiently reduced to allow high level resolution by 1-D SDS\_PAGE. As will be seen in the following chapter attempts to run these samples on 2-D gels presented aggregation problems with the extracted membrane fraction. This prompted the development of a novel method for resolving salt wash subfractions avoiding the need for ultracentrifugal pelleting of membranes. In the 1-D context of this chapter the SDS-PAGE profiles for extracted membranes generated by this method are shown together with conventionally prepared salt supernatants (figure 3.9). As before, profiles indicate that a substantial proportion of luminal protein remains associated with the membrane as evidenced by the apparent presence of the predominant bands corresponding to protein disulphide isomerase and/or calreticulin in the extracts. These results suggest that the organelle retains much of its structural integrity throughout subfractionation, and that the treatment mostly removes external peripherally associated components as was initially proposed.

**Figure 3.8: 500 mM Sodium Chloride sequential extract and twice-washed membrane subfractions from germinating and developing ER.** PhastGel (gradient 8-25%, Pharmacia) presenting evidence for the removal of residual extractable components by a second membrane wash. Developing ER appears the more resistant to extraction by this reagent. GS: Germinating ER salt supernatants (lanes 1 and 3); DS: Developing ER salt supernatants (lanes 2 and 4); GP: Germinating ER salt-washed membrane pellet (lane 5); DP: Developing ER salt-washed membrane pellet (lane 6). Position of molecular weight markers shown. Pellet loadings represent 1/100 of the subfraction generated. Supernatants represent 1/25 of subfraction. Gel is silver-stained.



**Figure 3.9: Mini SDS-PAGE profiles of germinating and developing ER sodium chloride subfractions.** (A) Conventionally isolated 500 mM NaCl extracts as an supernatant following ultracentrifugal pelleting of membranes. Lane 1: Germinating extract; Lane 2: Developing extract. (B) Salt-washed membrane fraction resolved onto sucrose barrier. Total ER (~1mg) was incubated in 500 mM NaCl and loaded onto a sucrose step: membranes were subfractionated by ultracentrifugation through a resolving 10% (w/v) sucrose solution, onto a 60% (w/w) sucrose barrier. The fraction (~2ml) was removed from the top of the barrier using a hypodermic syringe. Aliquots (10 $\mu$ l) were mixed with SDS buffer and loaded direct onto the gel. Lane 1: Germinating membranes; Lane 2: Developing membranes. Molecular weight markers (Sigma) are indicated. Gel is silver-stained.



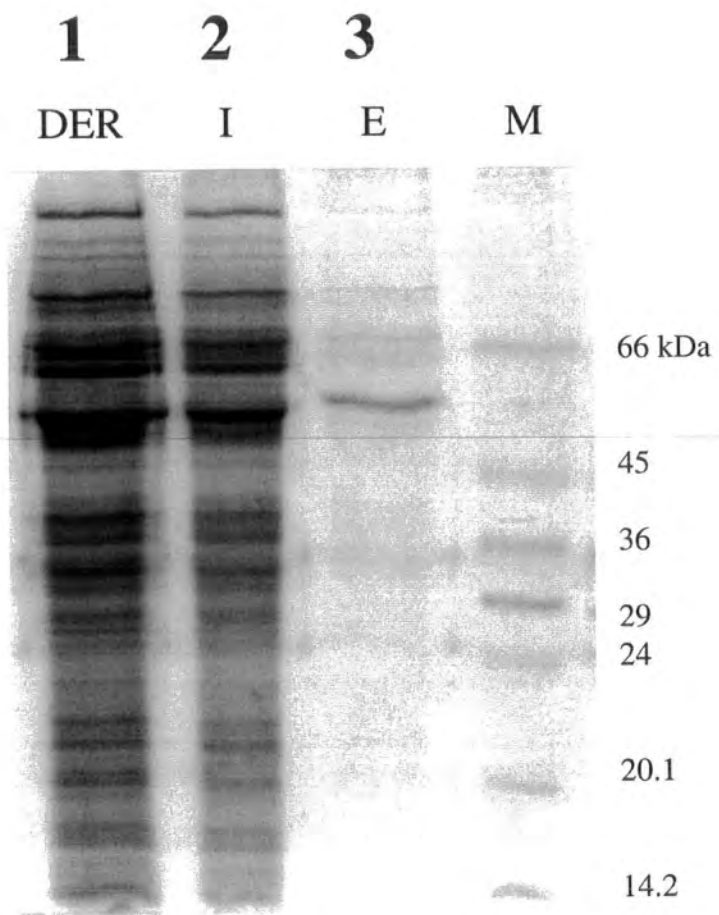


Although involving a more disruptive mechanism, sodium carbonate extraction was similarly observed to be incompletely efficient in removing the predominant 55 kDa band from castor ER (Coughlan *et al*, 1996). Attempts to analyse sodium carbonate ER subfractions were complicated by high viscosity in the extract fraction.

### **3.7 Extraction of ER membranes with chloroform/methanol**

In an attempt to specifically isolate hydrophobic proteins associated in particular with the developing ER membrane, the extraction procedure of Siegneurin-Berny and colleagues was adopted (Siegneurin-Berny *et al*, 1999). Total ER membranes were subfractionated on the basis of solubility/non-solubility in a 2:1 (v/v) chloroform/methanol mixture. Availability of sample, together with time limitations, meant that a more extensive investigation into different chloroform/methanol ratios and the analysis of increased quantities of starting material was not possible. Results of germinating ER analyses are omitted because loss of insoluble constituents during the fractionation prevented a quantitative comparison with the extract. SDS-PAGE analysis of the developing ER subfractions revealed a similarity with observations reported by Siegneurin-Berny *et al* (1999) on chloroplast envelope. In both the chloroplast and ER system only a small proportion (~1-5%) of overall protein is solubilised by this reagent, and this is largely represented by a single predominant band at ~60 kDa (figure 3.10). The soluble components of germinating ER was of a comparable quantity but produced a different SDS-PAGE profile, largely represented by 4 protein bands at ~55, 47.5, 40, and 35 kDa. There is thus a difference in the hydrophobic proteins in ER of these two developmental

**Figure 3.10: Chloroform/methanol extraction of developing ER membranes.** Total developing ER (~50 µg protein) was incubated, on ice, in 9 volumes 2:1 (v/v) chloroform/methanol. The mixture was centrifuged (12 000 x g) and the soluble supernatant separated from the insoluble pellet. Lane 1: Total developing ER (10 µg); lane 2: insoluble fraction; lane 3: Soluble extract fraction; M: Molecular weight markers (Sigma). Gel is Coomassie-stained.



stages. These experiments provide a potentially valuable means of differentially subfractioning ER from germinating and developing ER to provide a sample of hydrophobic membrane-associated proteins for proteomic characterization. The simplicity of the extract, together with the potential for sample scale-up, presents an opportunity for characterization from 1-D gels, thereby avoiding the need for 2-DE electrophoresis (which is likely to provide solubility problems with such an enriched source of hydrophobic proteins). The dominant 60 kDa developing ER hydrophobic protein was subjected to MALDI mass spectrometry analysis although no meaningful spectrum was derived.

### **3.8 Discussion**

Endoplasmic reticulum was efficiently and routinely isolated in milligram quantities from castor endosperm during an intense preparative period.

The first limiting factor affecting yield is homogenization and it is recommended that seeds are chopped to a pine paste for 20 minutes or longer. Developing ER fractions were generally higher than germinating ER in protein content, reflecting perhaps the metabolic demands placed on this organelle during seed maturation. An occasional developing-specific 30/40% sucrose fraction appeared identical to the 20/30% fraction and was concluded to be residual incompletely floated ER indicative of particularly rich preparations. Another more problematic feature of preparations containing a high content of ER is the potential for overloading of sucrose gradients. It was found that limiting the

amount of sample loaded per gradient resulted in more efficient fractionation in each of the sucrose gradient steps, leading to higher final yield.

SDS-PAGE profiles illustrated the gradual purification of ER from contaminating protein, most strikingly evident in the removal of massively predominant cellular proteins. ER fractions were found by SDS-PAGE visualization to be identical to those of Coughlan *et al* (1996) and as such were considered biochemically pure. Biochemical marker assays were therefore not performed. Considerable protein difference between developing and germinating ER was also evident.

The presence of two major known luminal proteins in the ER preparations, PDI and calreticulin, was confirmed by N-terminal amino acid sequencing, following electroblotting from a 1-D gel onto PVDF membrane. The calreticulin sequence revealed the N-terminus of the mature protein following signal peptide removal. However, considerations of the resolution attainable by 1-D SDS-PAGE and electroblotting, combined with sensitivity limitations of protein sequencing by this method led to the conclusion that a 2-D capability would need to be established. The resolving power of this technology would aid a detailed comparative study of the samples and provide the basis for protein characterization by mass spectrometric (eg. MALDI TOF) techniques.

ER was subfractionated by differential extraction in an attempt to simplify gel profiles while enriching for certain protein subsets (membrane proteins). Sodium chloride was chosen as an effective extraction. That major luminal components appeared to remain

associated with the membrane fraction is believed to be indicative organellar membrane intactness, a state which is necessary for resolution of the subfractions in the sucrose barrier system. Chloroform/methanol treatment generated a simple extract fraction, consisting of hydrophobic proteins, which is amenable to mass spectrometric characterization direct from 1-D gels. Scale up of this fraction may permit N-terminal sequencing of hydrophobic components. For a detailed characterization of such proteins, the investigation of a range of chloroform/methanol ratios would be recommended, together with the aforementioned sample scale up.

The aim of subsequent work was to focus on establishing reproducible 2-D electrophoresis techniques specifically suitable for the separation of castor ER proteins. The availability of this technology is likely to make a major contribution to evaluating sample complexity.

## CHAPTER 4

# Two-Dimensional Electrophoretic Analysis – Partial Characterization and a Comparison of Germinating and Developing ER

### 4.1 Introduction

In the previous chapter the preparation of ER from developing and germinating castor bean has been investigated. Separation of proteins using 1-D SDS-PAGE facilitated visual analysis and N-terminal sequencing of two proteins. This technique however was deemed inadequate for a detailed comparison of the two states, and it was concluded that a 2-dimensional approach was required. The resolving power of this technique is beyond that of any other biochemical tool for analysing complex protein mixtures, and when combined with the appropriate detection method is capable of analysing thousands of proteins in a single experiment. Furthermore, with mass spectrometry techniques (laser desorption ionization or electrospray) being routinely used for protein characterization, 2-D electrophoresis is now the core preparative technology for proteomics research. In plants, the separation of *Arabidopsis* plasma membrane fractions has recently been used to investigate the recovery of hydrophobic proteins on 2-D profiles using novel detergent systems (Santoni *et al*, 1999). This work found that solubilization in a buffer containing C8Ø as detergent, following subfractionation by treatment with alkaline carbonate solution, was effective in detecting integral membrane proteins. The carbonate wash which incompletely removed peripheral components from the plasma membranes was concluded to offer

an effective advantage over detergent subfractionation (Tritons X-100 and X-114), preserving integral proteins in a resolvable state.

SDS-PAGE produced significantly different profiles for the two samples under study. However differences occurred mainly amongst the less abundant proteins which were poorly resolved by this system. Highly reproducible silver stained 2-dimensional profiles were expected to allow direct comparison of the fractions at low nanogram levels. Reproducibility of the process is influenced by a number of factors including method of sample preparation, sample loading, and choice of detection stain.

For optimal results from 2-dimensional electrophoresis, appropriate sample preparation is a vital fundamental consideration. However, regardless of effective sample preparation, isoelectric focussing commonly presents post application solubility problems and it is important to avoid complications in gel separation due to poorly solubilized samples. Preparation of samples for IEF ideally results in a fully solubilized protein mix devoid of interfering contaminants and particulate material, precipitation within the gel of a particular protein will compromise resolution in the gel.

The membranous nature of the ER fractions posed potential problems since membrane proteins are generally regarded as being difficult to solubilize and hence resolve by 2-D gel analysis. Indeed, a previous study on a plasma membrane-enriched fraction from barley (Hurkman & Tanaka, 1986) included four different sample preparations. The present study made use of a chloroform/methanol precipitation which was subsequently replaced by a TCA/acetone procedure.



Although the number of possible preparation procedures investigated was limited by availability of sample, the method finally adopted as routine was regarded as particularly appropriate to the sample. Detergents tested were Triton X-100 and later octyl  $\beta$ -D-glucopyranoside.

Experiments generating 2-D profiles for direct comparison require an adequate level of reproducibility. This was achieved by the use of immobilized pH gradients (IPGs) for the first dimension isoelectric focusing step. Since their introduction (Gorg *et al*, 1985) IPG strips have revolutionised 2-D electrophoresis and allow reliable gel to gel, and even inter laboratory sample comparison. The use of immobilised pH gradients together with pre-cast gradient SDS gels for the analysis of castor ER is described here.

## **4.2 Two-dimensional electrophoresis of castor endosperm ER**

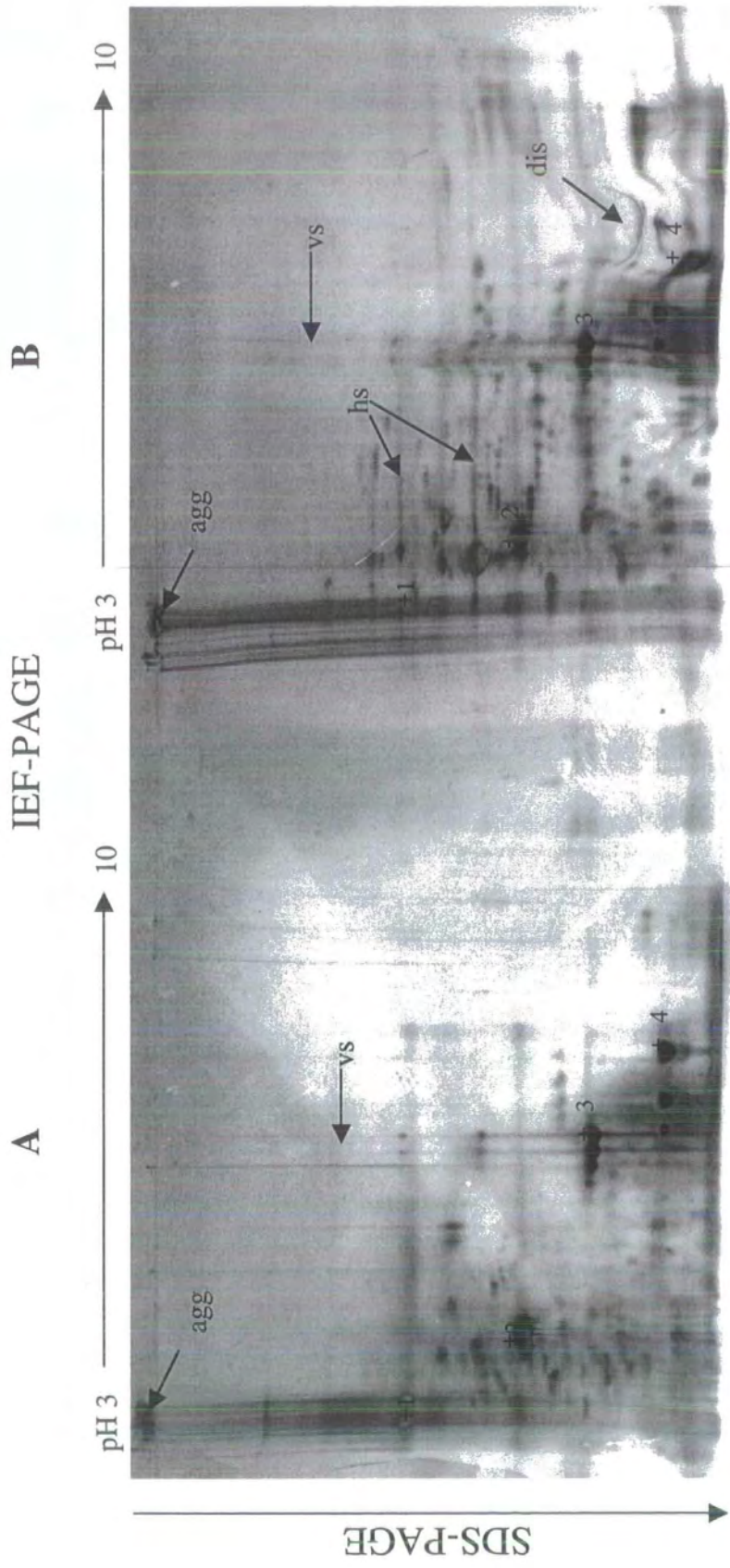
### **4.2.1 Sample preparation**

Initial experiments involved chloroform/methanol extraction for sample protein precipitation followed by resuspension in IEF lysis/sample loading buffers. During phase separation, protein was seen as a white precipitate at the phase interface. After re-extraction of the chloroform phase with methanol, the protein pellet appeared darker in colour and was of a fibrous nature. Resuspension required simple agitation and pipetting action, which was repeated periodically during the hour long 37°C incubation.

#### 4.2.2 2-D electrophoresis of germinating and developing ER

Approximately one-twentieth of a typical ER sample from each developmental stage was prepared via chloroform/methanol extraction for 2-D electrophoresis. Initially all experiments used this method of sample preparation for uptake into Triton X-100/urea IEF buffers. Samples were not centrifuged prior to loading as gels were required to act as diagnostic of sample solubility problems, including any aggregation existing prior to IEF initiation.. Standard marker proteins (Sigma 2-D markers) were also included in the sample solution. Figure 4.1 shows the profiles resulting from a typical separation. While indicating the suitability of the chosen pH range (3-10) for resolving the components of these particular fractions, the results highlighted a variety of artifacts. Appreciation and identification of the various artifacts common to this set of techniques is important when aiming for high quality reproducible 2-D profiles and is the reason for their inclusion in this section. It was clear from these images that the substantial aggregation (Agg) at the IEF origin (evident as granular spots and a prominent streak down the SDS-PAGE axis) significantly reduced the amount of resolving protein. Some aggregation at this position is not uncommon when sample cups are used for application. This can be due in part to the high voltage which rapidly drives the sample proteins into an immediately low pH and concentrated environment, and may be minimised by adjustment of sample volume and initial voltage. Aggregation arising from a dramatic drop in environmental pH would be expected to affect only a subset of the total proteins present but close examination of the aggregation streak reveals that a wide range of apparently readily soluble components are affected (corresponding bands within the streak). Therefore the main cause of aggregation in this instance was believed to be the degree of solubility of sample prior to application to, and initiation of, IEF. Providing a

Figure 4.1: 2-D separation of ER samples isolated from endosperm of germinating and developing castor seeds. Suspensions of ER (~50 $\mu$ g) were prepared for IEF by chloroform/methanol precipitation. Pellets were reconstituted in IEF buffers (including a 1 hour incubation at 37°C) before loading, via sample application cups, onto the acidic (anodic) end of Immobiline pH 3-10 IPGs (Pharmacia). Proteins were separated in the second dimension on SDS ExcelGel gradient gels, 8-18% (Pharmacia). (A): Germinating ER; (B): Developing ER. Profiles display several common 2-D artifactual features (see text): agg, aggregation; hs, horizontal streaking; vs, vertical streaking; dis, distortion. Internal 2-D standard markers (Sigma) were included in the sample; 1: Amyloglucosidase (89, 70 kDa; pI 3.8); 2: Ovalbumin (45 kDa; pI 5.1); 3: Carbonic anhydrase (29 kDa; pI 7.0); 4: Myoglobin (17 kDa; pI 7.6). Gel is silver-stained. In subsequent diagrams IEF and SDS-PAGE orientations are as shown here.



sufficient quantity of protein was loaded, the aggregation, although undesirable, did not itself prevent a rudimentary visual analysis. The findings reinforced the importance of sample preparation, an issue addressed in more detail in later sections.

The vertical streaking (vs) observed does not interfere with the detection of individual protein spots, although it does suggest either insufficient equilibration between the first and second dimension or the occurrence of electroendosmotic processes. These problems were kept to a minimum by the inclusion of high urea concentrations and glycerol in the equilibration solutions and by the use of absorbent application pieces at each end of the IPG strip during electrophoretic migration into the SDS gel, respectively.

The slight horizontal streaking (hs), more pronounced in the acidic side of the profile, is suggestive of a number of possible causes including ionic or other non-protein impurities in the sample, under/over focusing. Non-protein contaminants are best reduced by an effective sample preparation procedure such as TCA/acetone precipitation.

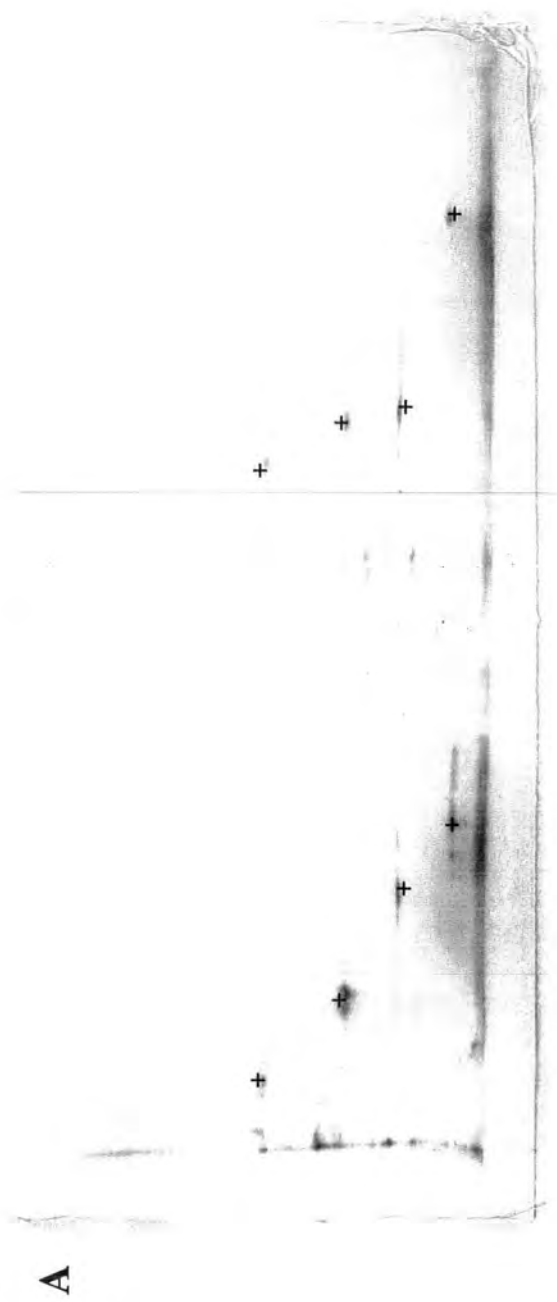
The distortion present in Figure 4.1b (dis) is the result of either an air bubble between the plastic support of the SDS gel and the cooling plate (which causes inefficient heat transfer and consequent uneven protein migration) or water contamination of the gel surface during positioning of buffer strips. Such distortions were subsequently eliminated as handling skills improved.

### 4.2.3 Assessing and optimising reproducibility between 2-D profiles

Reproducibility between two-dimensional profiles is essential if a meaningful sample-to-sample comparison is to be made. The system employed was conducive to achieving this since both dimensions were precast and the pH gradients of IEF gels were immobilised in the gel matrix. Gel-to-gel parameters such as thickness and overall dimensions were constant, and running buffer systems were also precast in polyacrylamide strips. Running conditions were identical for every experiment as programmed into the MultiPhor power pack, temperature was maintained at 15°C. The only likelihood of significant error was in the actual handling of apparatus and reagents, or in the preparation of samples and stock solutions/buffers.

Reproducibility was investigated with standard markers manufactured specifically for 2-D electrophoresis (Sigma). Assessment was made by the superimposibility of standard 2-D profiles. Increasing familiarity with the technique together with standardization of reagent preparation led to a level of reproducibility suitable for application to the direct comparison of ER samples. Figure 4.2a demonstrates uneven focussing which leads to non-reproducible 2-D profiles. As technique improved, superimposable standard protein distributions were obtained (figure 4.2b). It is worth noting that since two IEF dimensions were routinely resolved side by side on the same SDS gel, reproducibility was more readily achieved allowing meaningful comparison between the final profiles of a single experiment. Reproducibility between different experiments demanded particular care with the handling of gels, samples and reagents, and with the setting up of apparatus. From the marker profiles illustrated in figure 4.2b, a standard calibration grid was

Figure 4.2: Evaluating gel-to-gel reproducibility of 2-D protein profiles. Reproducibility of the 2-D process/experimental technique was assessed through the separation of standard marker proteins (as figure 4.1). Samples were reconstituted according to the manufacturer's instructions directly loaded with IEF buffers onto IEF gels. (A): Non-reproducible profiles. Proteins had not migrated to the same positions along the first dimensions (see text); (B): Reproducible profiles as determined by complete superimposibility. Gels are Coomassie-stained.



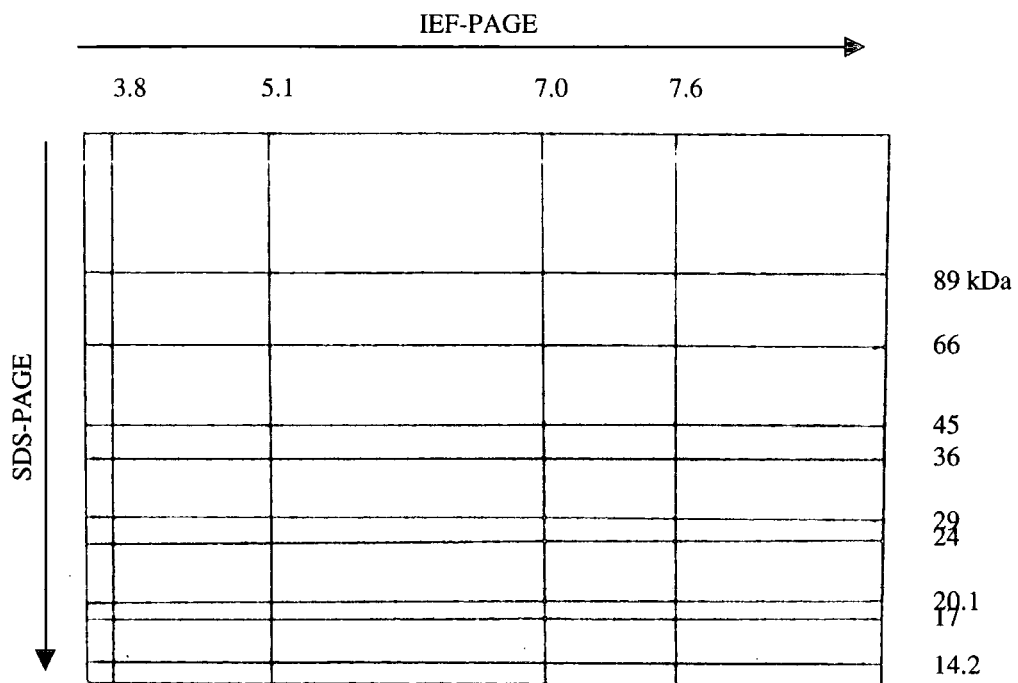


constructed to facilitate subsequent partial characterization (approximate isoelectric point and molecular weight assignment) of sample protein spots (see Figure 4.3). This grid was later extended to include additional molecular weight axes provided by standard SDS-PAGE markers (Sigma).

#### **4.2.4 2-D electrophoresis for direct comparison of germinating and developing ER**

Following achievement of routine reproducible 2-D profiles with standards, the technique was applied to a differential analysis of ER from different stages of development. As for previous experiments (section 4.2.2) samples were prepared using the chloroform/methanol procedure and uptake into the respective buffers for lysis and loading. Internal standard 2-D markers were again included and profiles are shown in figure 4.4. Standard markers confirmed reproducibility in resolution. Increased attention to rigorous sample solubilization prior to loading led to a marked decrease in the level of aggregation seen in figure 4.1. Limited horizontal streaking remained evident but did not interfere with the identification of individual protein spots. No vertical streaking from sample spots was observed although this was clearly visible from the standard carbonic anhydrase marker. Consequently, standard markers were considered to be over-loaded in the profiles and later experiments excluded their use, relying instead upon a calibration grid if approximation of pI and molecular weight was desired.

In the absence of dedicated image analysis computer software, gel profile comparisons were facilitated by analysis of enlarged digital photographs printed onto



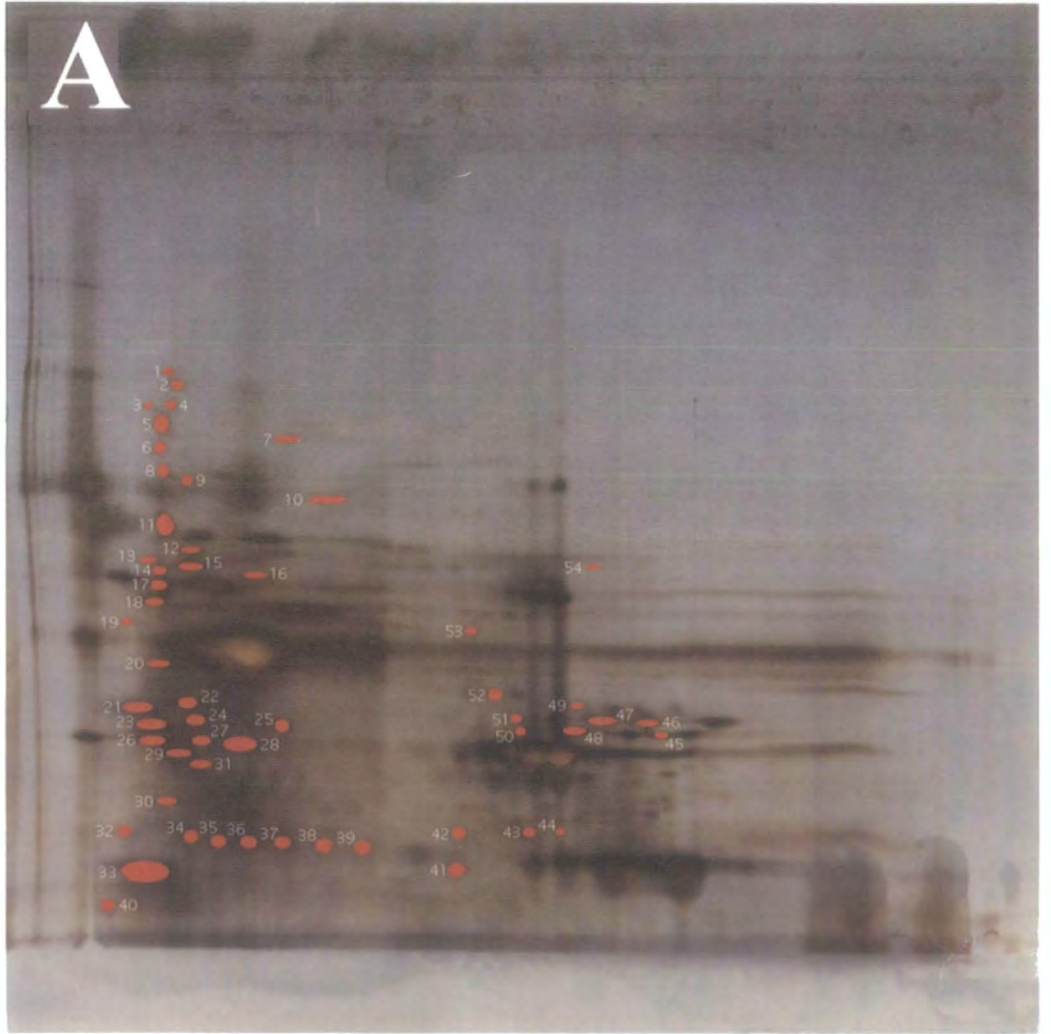
**Figure 4.3: Two-dimensional calibration grid.** Axes were drawn according to the positions of standard protein markers (Sigma). IEF-PAGE: Immobiline DryStrip pH 3-10, 11 cm (Pharmacia); SDS-PAGE: ExcelGel SDS Gradient 8-18% (Pharmacia).

Figure 4.4: Germinating and developing ER 2-D protein profiles. Samples (~50µg each) were prepared as for figure 4.1 for a direct proteomic characterization to investigate the differences in protein profiles of ER from two different stages of development. The profiles represent effective elimination of some of the previously observed artifacts, although horizontal streaking and apparent carbamylation trails remain fairly prominent. (A): Germinating ER; (B): Developing ER. Internal markers 1-4 as in figure 4.1. Gel is Silver-stained.

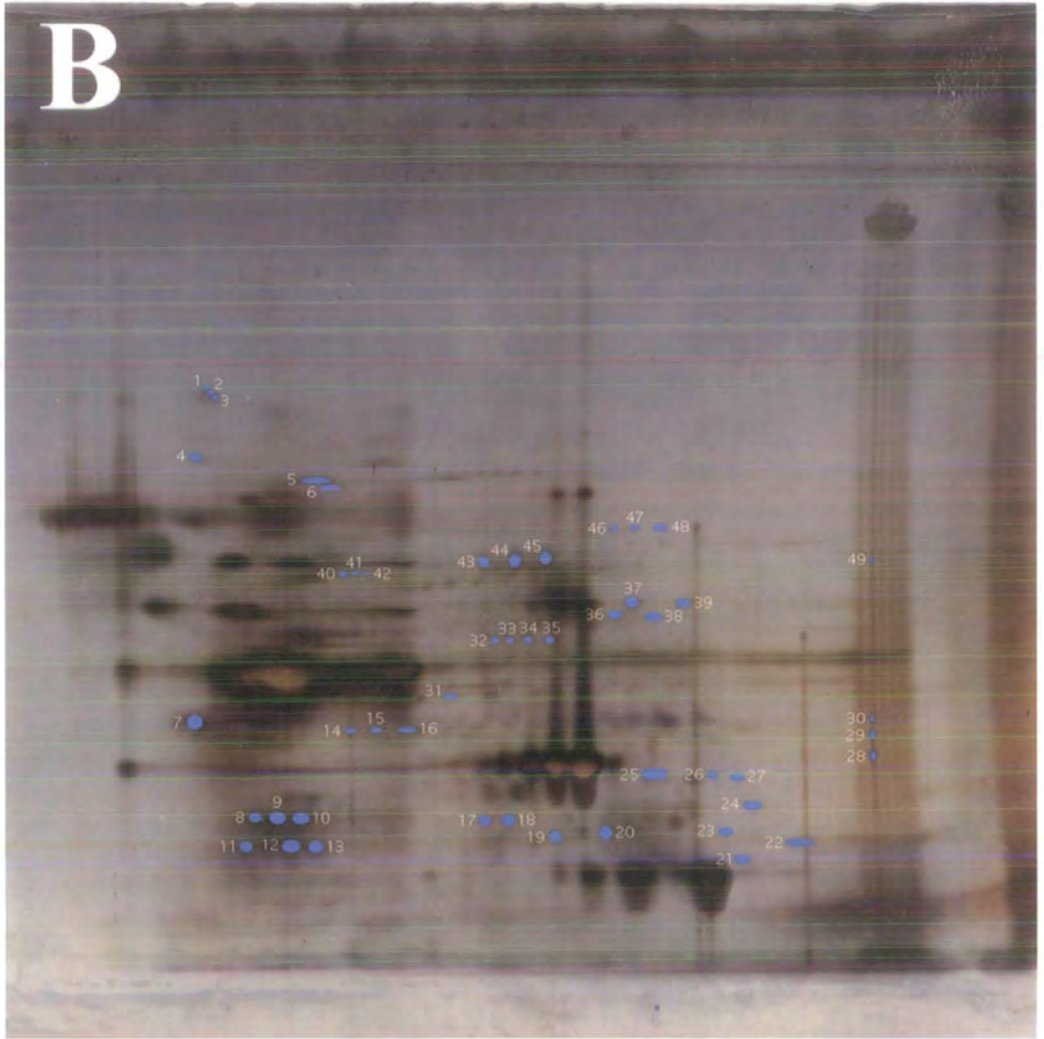


transparent sheets. These were superimposed for 'overlay' analysis on a light box and the components apparently specific to each stage in development highlighted. In all, 54 spots were identified as being stage-specific to germinating ER, and 49 were identified as developing unique (see fig 4.5). Each of these proteins could be assigned approximate isoelectric point and molecular weight values estimated according to their positions relative to standard markers. These spots represent potential targets for further characterization relevant to the present study. Proteins lying outside the outer axes of the region enclosed by standard markers were not assigned precise values. It is possible that the specific spots highlighted in this section do not in fact represent all of the stage-specific components detectable at this level, and that some masking from dominant internal marker proteins and persistent artifacts (eg. horizontal streaking) had occurred. A quantitative characterization was not made because of the non-quantitative nature of silver stained images. In terms of spots common to each stage of development, analyses from the germinating and developing gels were in agreement with figures of 67 and 68 respectively.

The fact that several groups of proteins lie at the same molecular weight and close isoelectric point positions infers the existence of charge isoforms of a single protein. Although such clusters could represent naturally occurring isoelectric point isoforms, the widespread nature of such 'trails' suggest charge modification through carbamylation, which is a common problem associated with urea derivatives in sample/rehydration solutions. The modification results from the reaction of cyanate with the  $\alpha$ -amino group of N-termini and the  $\epsilon$ -amino groups of lysine. Cyanate is derived from urea which in water exists in equilibrium with ammonium cyanate and increases with temperature and pH. All standard precautions to minimise urea



**Figure 4.5: Identification of spot differences between germinating and developing ER profiles.** Profiles presented in figure 4.4 were directly compared using differentially coloured and enlarged transparent digital photographs (see appendix). (A) Germinating-specific proteins. 54 spots were included in this category; (B): Developing-specific proteins included 49 distinguishable components. Gel is silver-stained. Table 1 lists approximate molecular weight and isoelectric point values assigned to these proteins.





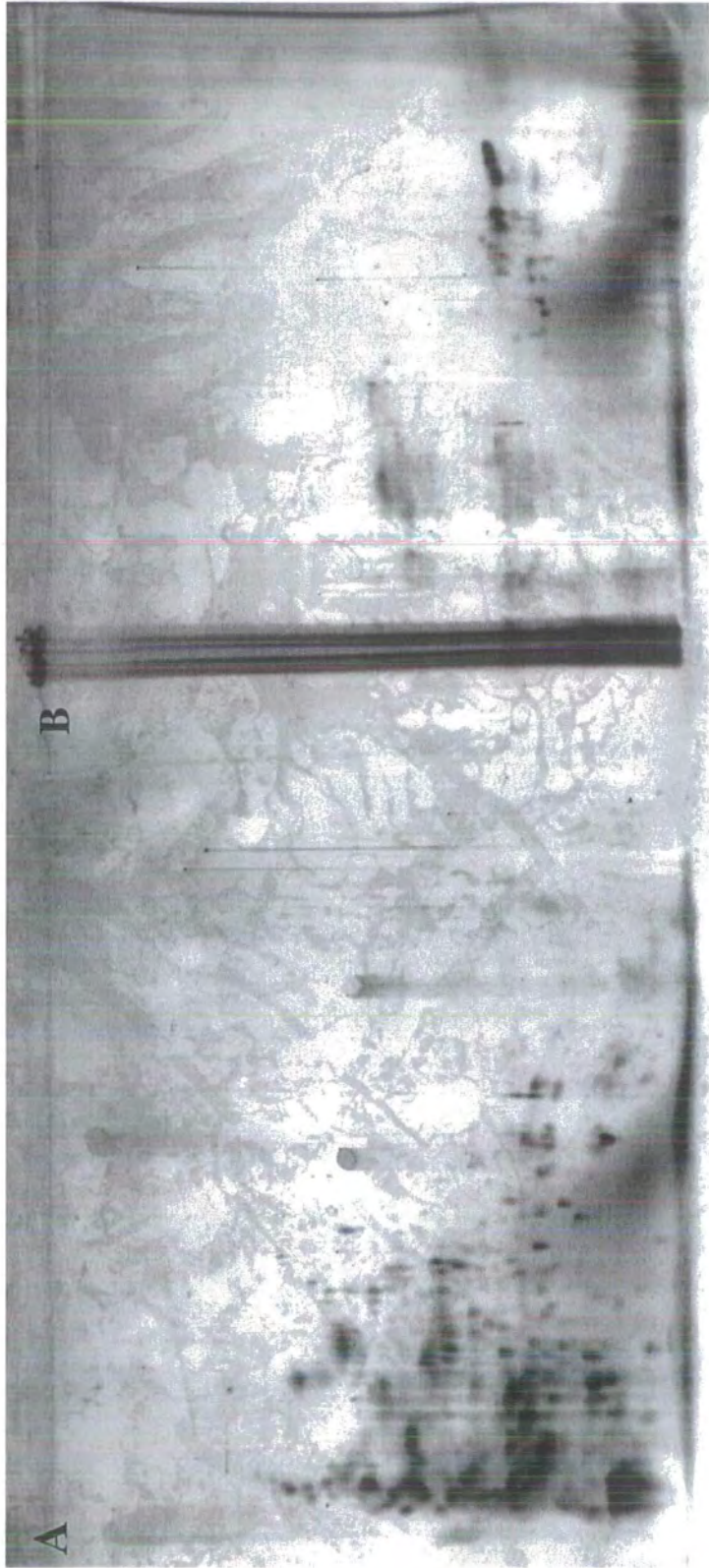
breakdown were taken (ie. urea-containing solutions were never heated above 37°C and were stored at -20°C, ultra pure grade urea was used, and carrier ampholytes included in first dimension buffers act as cyanate scavengers). The fact that carbamylation artifacts were observed despite these precautions led to all 2-D solutions being prepared fresh, as required, or at least being stored for only limited periods (ie. up to 2 months) at -20°C.

### **4.3 Two-dimensional analysis of salt-generated subfractions of castor ER**

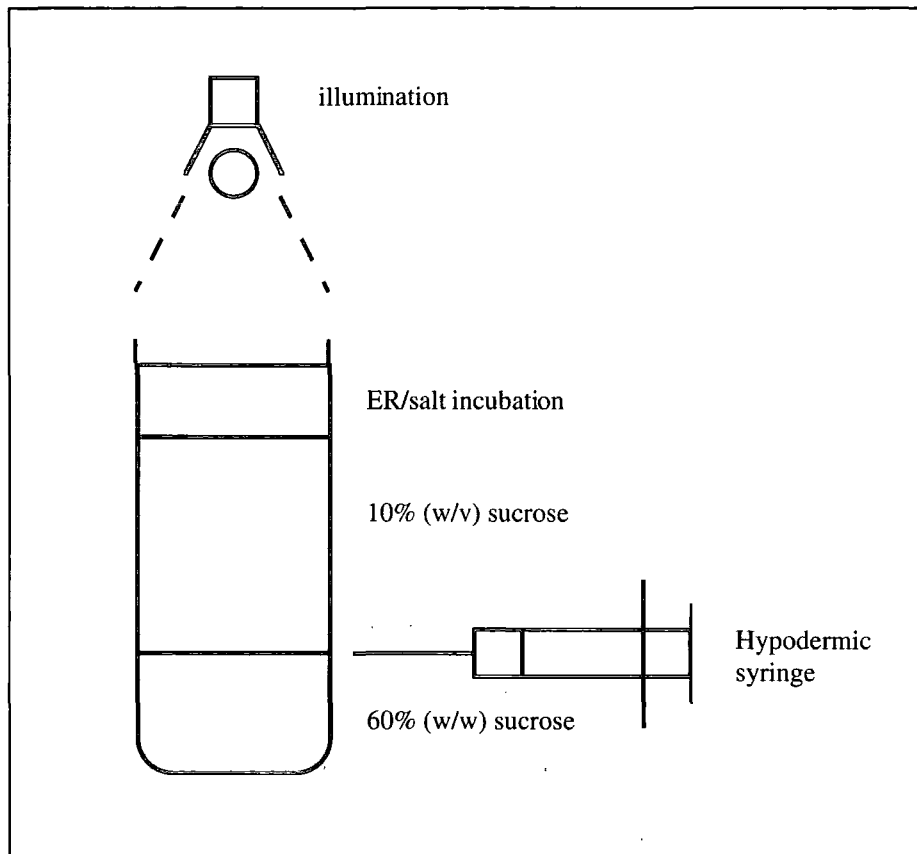
Two-dimensional electrophoresis was applied to the separation of proteins from ER subfractions proteins, as introduced in the previous chapter. The majority of subfractionation work focussed on the use of sodium chloride to release associated proteins from the ER membrane, presumably by the disruption of ionic interactions. The first attempts to resolve these subfractions in two dimensions involved repetition of methodology so far presented in this chapter, the main difference being the non-inclusion of internal marker proteins. As for much of the exploratory technology included in this thesis, 2-D resolution was at first attempted for germinating samples to avoid unnecessary consumption of developing material.

Sodium chloride (500 mM) extract and twice-extracted membranes from germinating ER were analysed on the Immobiline/ExcelGel system following salt removal by chloroform/methanol precipitation (figure 4.6), removal of higher than trace levels (eg. < 50 mM) of salt is essential if disturbance to IEF is to be avoided. A representative total ER profile was previously presented (figure 4.4a). The profiles showed that the salt extract fraction resolved relatively clearly with only limited

Figure 4.6: Preliminary 2-D analysis of sodium chloride subfractions of germinating ER. Germinating ER membranes were subjected to a sequential two-step extraction with 500 mM NaCl. Samples were subfractionated by ultracentrifugation. Equal volumes (proportionally representative of the total starting material) of each subfraction were prepared via chloroform/methanol precipitation for IEF. (A): Single salt extract; (B): Twice-extracted membrane fraction. Corresponding total ER profiles are shown in figure 4.4. Gel is silver-stained.



interference to the profile from vertical streaking. Conversely, a particularly severe level of aggregation in the membrane fraction IEF gel has resulted in few proteins being successfully separated along this axis. Aggregation to such an extent as to affect most of the proteins in a sample had not previously been encountered and was attributed to the nature of the sample in which hydrophobic components had been enriched. Although incompatibility between hydrophobic proteins and IEF is well appreciated a major problem with the sample in question was found to be aggregation during preparation for, and prior to, focussing. Steps were therefore taken to alleviate this tendency in order that an effectively solubilized sample could be presented. The realization that the main problem in effective sample preparation was the efficient reconstitution of the ultracentrifugal pellet prompted particular attention to this area. In addition, the chloroform/methanol precipitation step was generating a fibrous pellet which proved difficult to resuspend in IEF buffers. Two main areas therefore warranted development or alteration. Although subfractionation was regarded to require ultracentrifugal force, actual pelleting of the membrane fractions at high speed was ultimately avoided by the implementation of a sucrose resolution barrier system (figure 4.7). The ER/salt incubation was layered onto a discontinuous sucrose step consisting of a resolving 10% (w/v) solution on top of a 60% (w/w) barrier. This method could not be performed on the accustomed micro-scale and so milligram quantities of ER were applied. The sucrose system was contained in tubes as used for the first sucrose gradient of ER isolation (section 3.2). Fortunately, following ultracentrifugation (25 000 rpm, 2 hours) withdrawal of membranes from the tube was greatly facilitated by their visibility atop the sucrose barrier solution. Salt extract fraction was sacrificed by this method since effective recovery from the resolving phase of the step was not practical. Subsequent analysis of the novel membrane



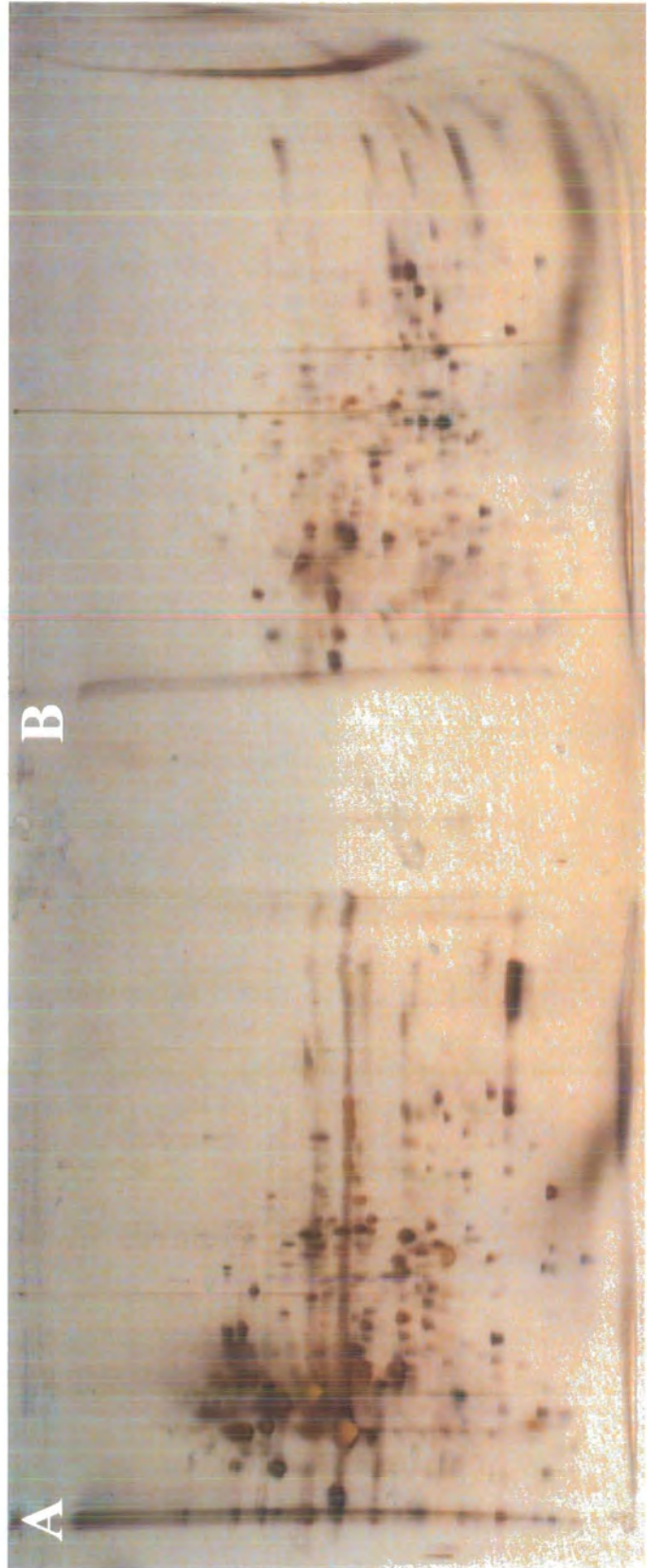
**Figure 4.7: Schematic representation of withdrawal of salt-washed membrane fractions from the resolving sucrose step.** The sample incubation was resolved through the 10% sucrose solution. The membrane fraction came to rest at the 60% sucrose barrier where removal was effected through hypodermic needle.

fraction was firstly performed by direct loading, with SDS sample buffer, onto SDS-PAGE (figure 3.10).

For 2-DE separation a sample of this membrane fraction was prepared by precipitation in a mixture of TCA and acetone (10% w/v and 20% v/v respectively). Uptake in IEF sample buffers was ensured by thorough homogenization using a 100 $\mu$ l glass homogenizer. Using this approach, the sample was prepared for IEF without an ultracentrifugal pelleting step which would therefore hopefully minimize aggregation. TCA/acetone precipitation generated a pellet more amenable to subsequent buffer uptake and effective breakdown of particulates by homogenization. The resulting 2-D profile was comparatively free of aggregation (figure 4.8). As expected, the more extraction-resistant developing membrane fraction contained more protein and displays slightly more aggregation at the IEF origin, than the corresponding germinating fraction. By combined application of a sucrose barrier and TCA/acetone preparation, a sample had been generated from the membrane subfraction which was amenable to 2-D electrophoretic analysis. A further possible factor contributing to the decrease in aggregation is the fact that the membranes applied to the sucrose system were only once-extracted as compared with that shown in figure 4.6. This would make the sample slightly less enriched in hydrophobicity, but since a single extraction has been shown to release the major part of extractable material, success was mostly attributed to the improved methodology of imposing a sucrose barrier and preventing high-speed membrane pelleting.

As has been shown, the salt extract fraction was not problematic in obtaining a representative 2-D profile. Figure 4.9 shows the 2-D distribution of proteins from

Figure 4.8: Salt-extracted and sucrose resolved ER membranes from germinating and developing endosperm. Salt-treated total ER membranes were layered onto a 10% (w/v) sucrose solution and resolved by ultracentrifugation (25 000 rpm, Beckman SW28 rotor). The subfractionated (salt-extracted) membranes were withdrawn from above a 60% (w/w) sucrose barrier. The fraction was prepared by TCA/acetone precipitation for uptake in IEF buffers, aided by mechanical homogenization. (A): Developing membrane subfraction; (B): Germinating membrane subfraction. Marker proteins were not included. Gel is silver-stained.





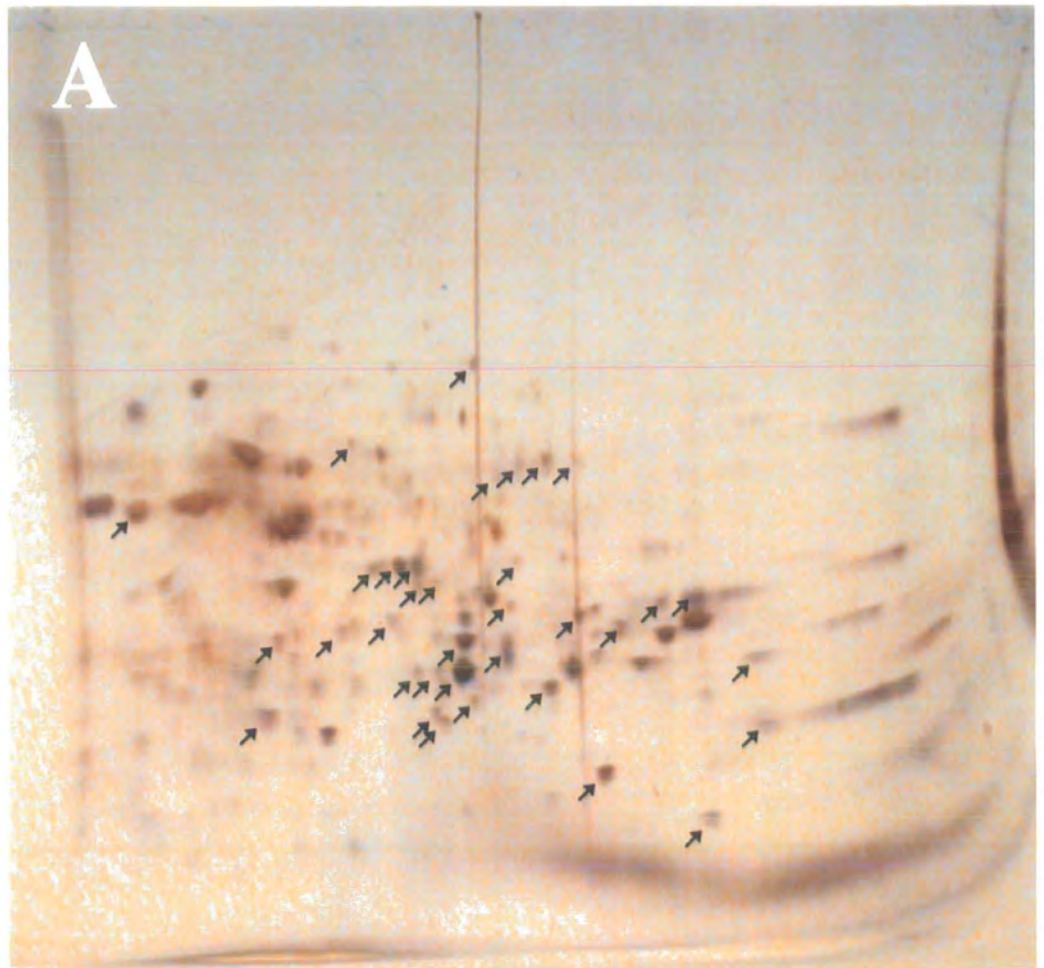
germinating and developing ER salt extracts, achieved following TCA/acetone sample preparation. A dramatic difference in the level of protein released from each sample type is immediately evident, consistent with the supposition that developing membranes display increased resistance to extraction in high salt. Comparison of the germinating extract profile (figure 4.9a) with that of figure 4.6a suggests that the TCA/acetone procedure is the preferable sample preparation method for obtaining better resolved patterns, with less streaking artifacts. The two images also underline the spatial reproducibility routinely obtainable.

The availability of well-resolved and reproducible 2-D profiles made possible the direct comparison of germinating and developing ER protein, this time at the subfraction level. Transparencies of enlarged gel images were again used to facilitate the identification of stage-specific spots, by superimposition on a light box. Figures 4.10 and 4.11 indicate these components. From the membrane fractions, 35 spots (from a total of 110) were specific to the germinating and 45 (from a total of 125) to the developing sample (figure 4.10). From the salt extracts, 62 from 160 were germinating specific, while 10 from only 38 were specific to the developing (figure 4.11).

#### **4.4 Correlating total and subfraction ER profiles**

Examination of total ER and subfraction 2-D profiles together revealed significantly improved clarity and detection of spots in the latter, while maintaining reproducibility. This had obvious visible consequence with regard to the numbers of stage-specific spots identified and is particularly evident in the germinating samples

**Figure 4.9: Identification of proteins specific to salt-extracted germinating and developing ER membranes.** The ER subfraction 2-D profiles presented in figure 4.8 were directly compared. Stage-specific spots were identified by the use of enlarged transparent gel images and are indicated by arrows. (A): Germinating-specific components (35 spots); (B): Developing-specific components (45 spots). 40 spots were common to each developmental stage.



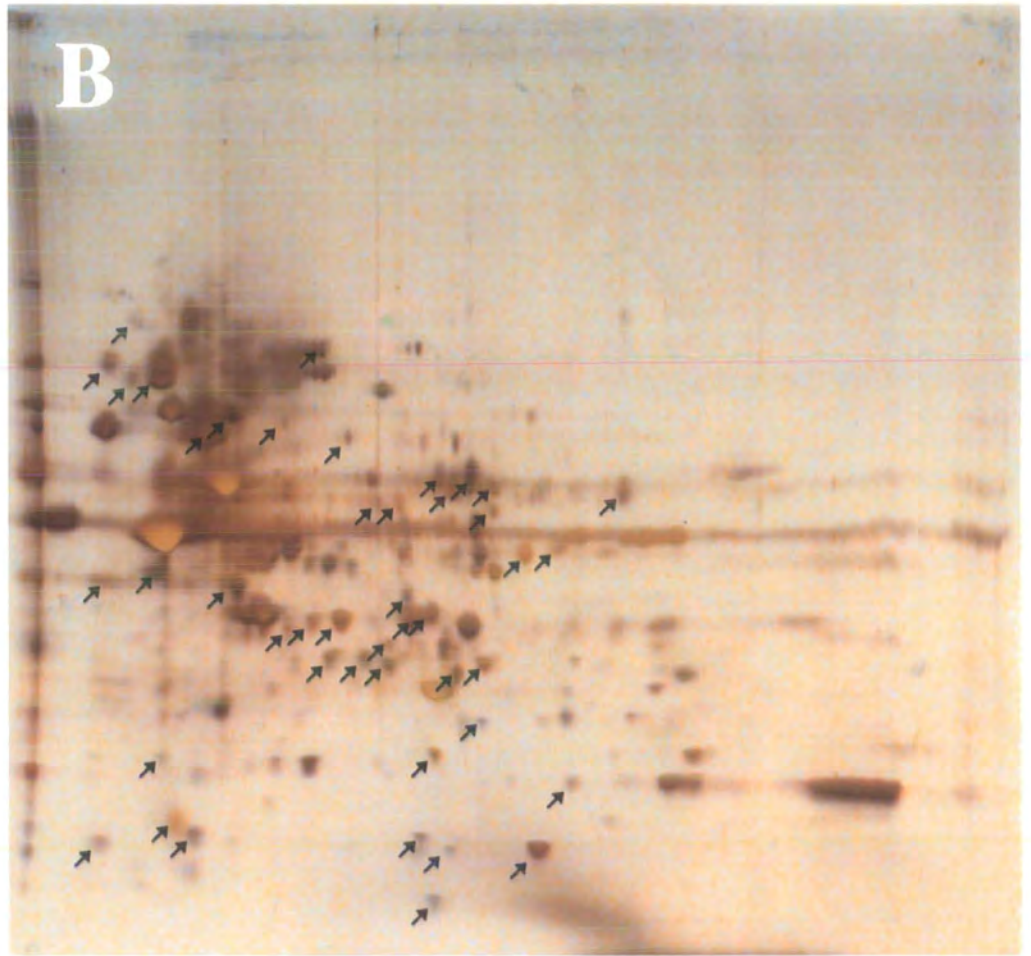
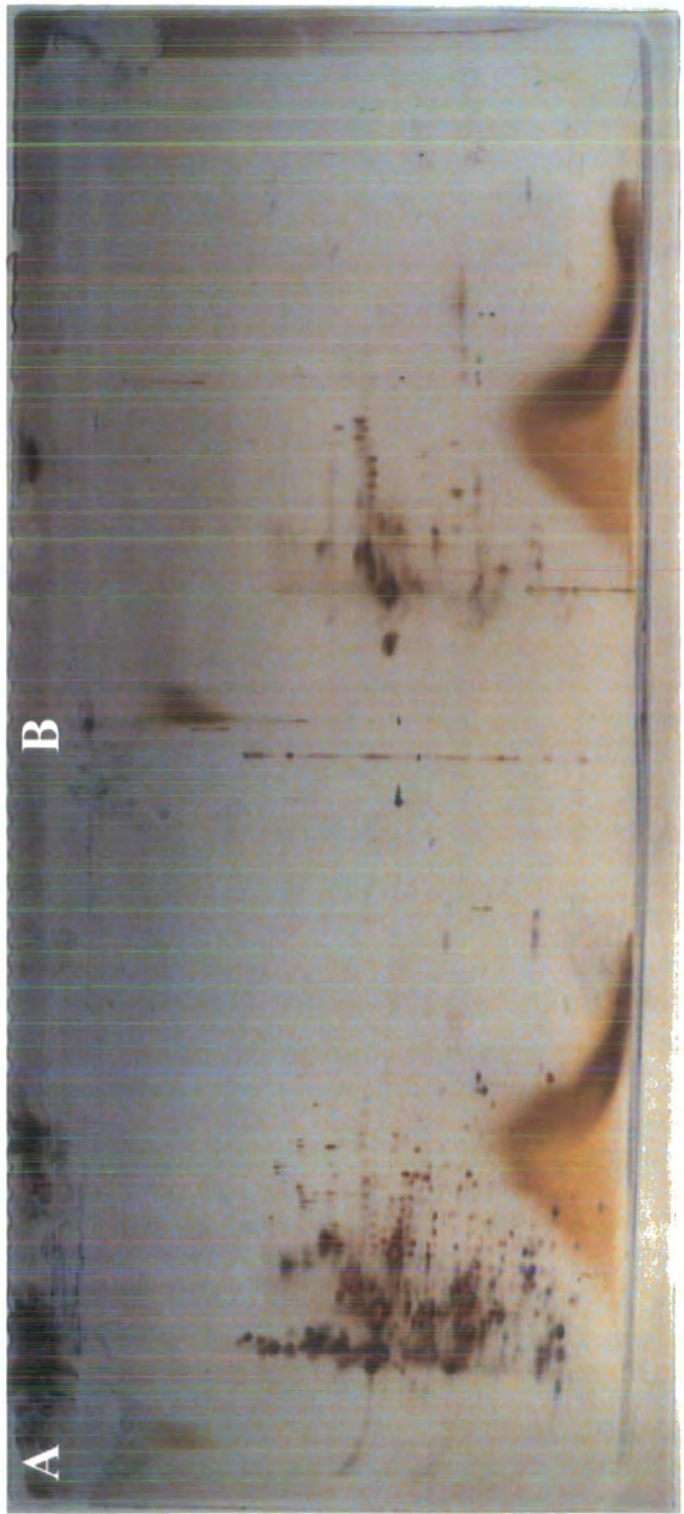
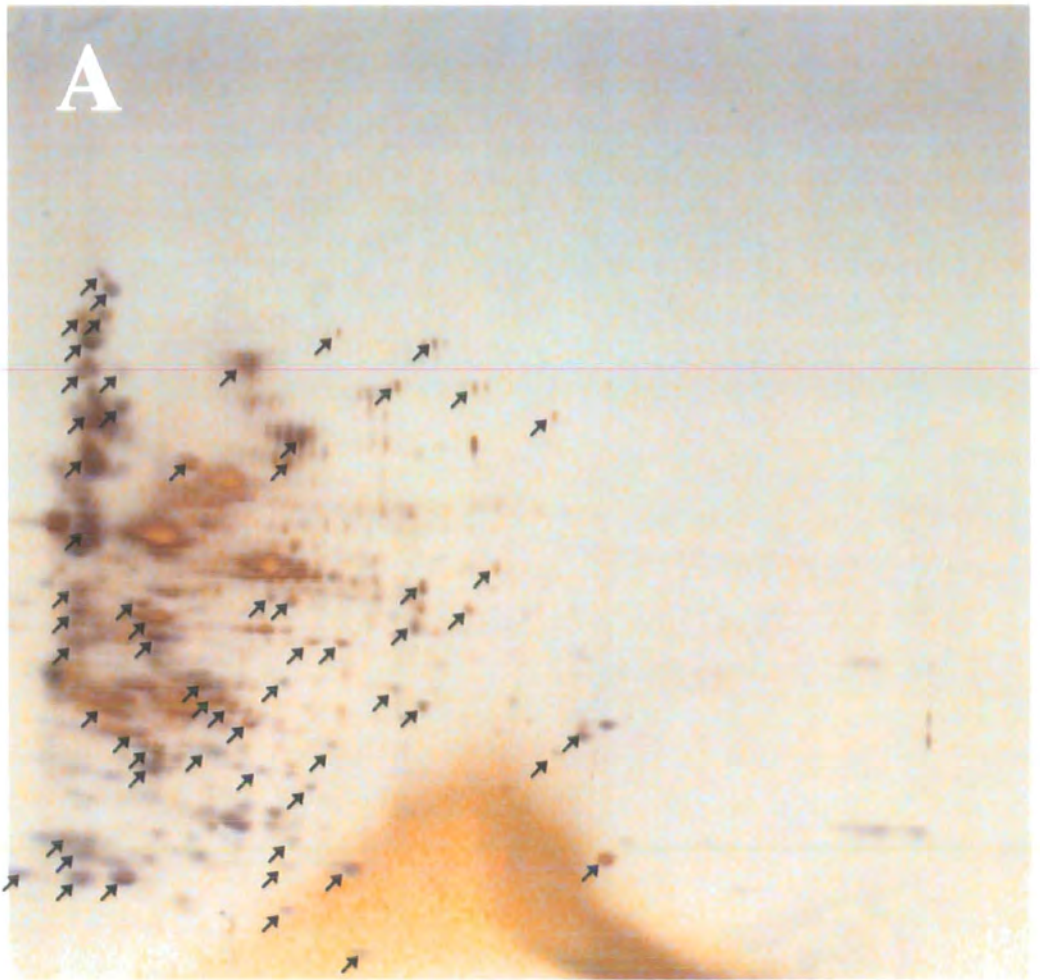


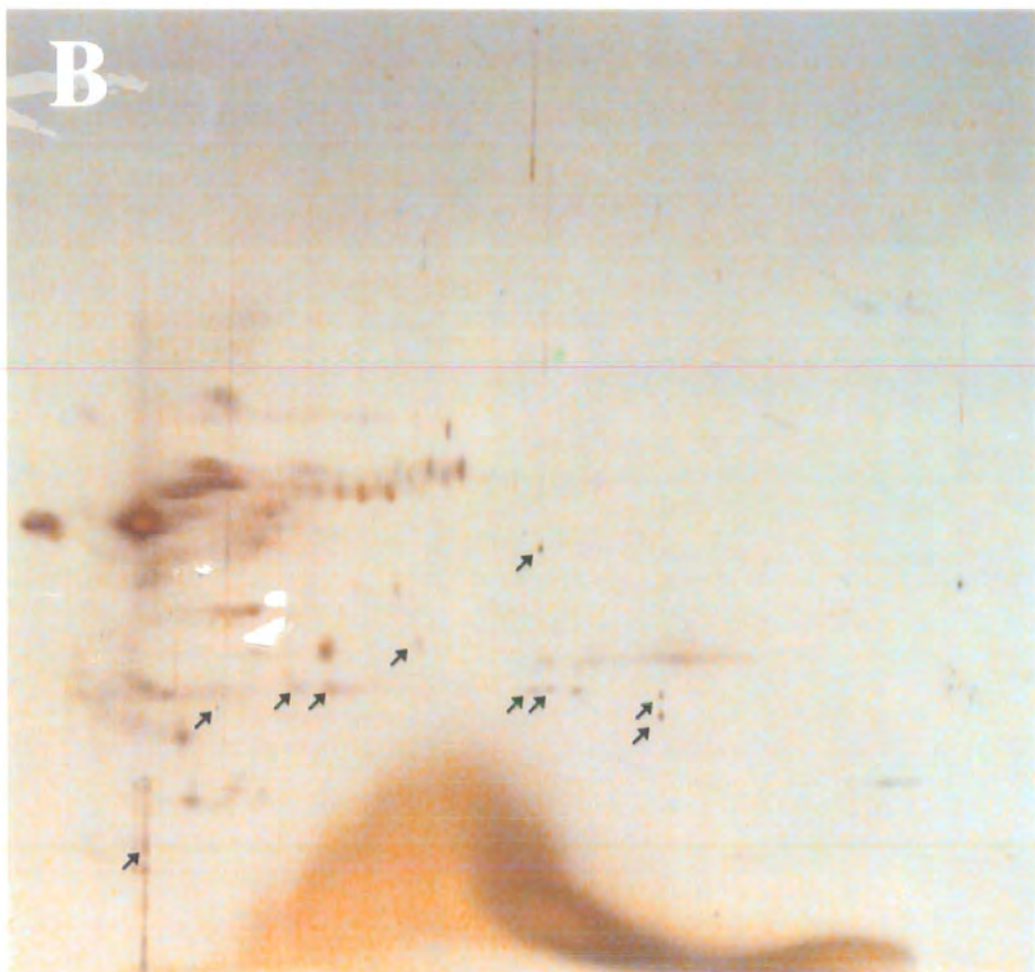
Figure 4.10: Germinating and developing ER salt extract subfractions. Salt-treated ER was subfractionated by ultracentrifugal pelleting of membranes. The supernatant (extract) fraction was TCA/acetone precipitated and solubilized in IEF buffers. (A): Germinating extract; (B): Developing extract. Marker proteins were not included. Gel is silver-stained.



**Figure 4.11: Identification of proteins specific to germinating and developing ER salt extract subfractions.** The profiles presented in figure 4.9 were directly compared using enlarged transparent images, and stage-specific components are indicated by arrows. (A): Germinating-specific components (62 spots); (B): Developing-specific components (10 spots). See text for discussion of numerical discrepancies between proteins identified as stage-specific in total ER versus subfraction profiles.







where a discrepancy of 43 spots can be attributed to a combination of two factors. First, the additional spots detected at the subfraction level, particularly in the mid pH range, give rise to an increased tally. The reason for the greater clarity of profiles at this level is believed to be largely due to improved sample preparation. In addition, and particularly with the germinating extract fraction, the removal of membrane components (ie. lipid, hydrophobic proteins) has likely contributed to resolving success. Second, the incomplete nature of extraction of a single salt wash results in the appearance of some components in each subfraction profile and consequent overestimation of stage-specific components. This problem is less pronounced with developing estimates for which there is an observed discrepancy of only 6 spots.

#### **4.5 Discussion**

Two dimensional electrophoresis was implemented to examine the extent of protein profile differences between maturing and germinating ER from castor endosperm. Initial experiments utilising a chloroform/methanol procedure for sample preparation immediately highlighted that this difference was significant. It was thus established that the system used, in conjunction with detection by silver visualization, was adequate for the resolution and subsequent analysis of sample proteins.

The early experiments however highlighted various 2-D artifacts which were diagnostic of specific experimental imperfections, and therefore provided valuable troubleshooting experience.

Subsequent experiments focussed on gaining familiarity with the techniques involved, with the aim of working towards a high level of reproducibility. Such attempts made use of standard markers and were concerned with obtaining directly comparable and superimposable protein distributions. Once a satisfactory level of experimental reproducibility had been reached, and a calibration grid constructed, attention was turned to direct intercomparison of 25 days-after-flowering and 5-day germinated ER using the same format. Following steps to resolve identified pitfalls, both samples produced well-resolved 2-D profiles despite the recurrence of minor artifacts (mainly associate with over-represented standard markers). A differential analysis was consequently made with the aid of enlarged transparent digital photographs and the internal markers. This exercise resulted in the partial characterization (pI and molecular weight) and tabulation of all spots identified as stage-specific/strongly enhanced. This category included 54 germinating ER and 49 developing ER protein components. Approximately 67 proteins were found to be common to each stage of development.

The 2-D methodology was adapted for application to salt-generated subfractions of total ER from both developmental stages. Aggregation patterns observed with the membrane subfraction were overcome by the development of a resolving sucrose barrier system in conjunction with improved sample preparation involving TCA/acetone precipitation. TCA/acetone precipitation is recommended as a general preparative method for 2-D electrophoresis of castor ER samples.

The value of 2-dimensional electrophoresis as the basis for a proteomic strategy to investigate differentially expressed components of the biological system under study

was confirmed. The following chapter presents the results of larger scale 2-DE technology as an analytical and preparative tool for protein characterization by MALDI TOF mass spectrometry.

## CHAPTER 5

### Application of Matrix-Assisted Laser Desorption Ionisation Time of Flight Mass Spectrometry to Proteins of Castor ER

#### 5.1 Introduction

Organelle isolation in combination with differential extraction procedures is expected to be increasingly applied as a powerful approach to proteomics research.

The preceding chapter established a reproducible 2-D strategy for comparing samples of biochemically pure ER from germinating and developing castor endosperm. This included an examination of salt-generated ER subfractions and a key challenge was the circumvention of aggregation problems associated with membrane subfractions. Analytical 2-D electrophoresis was developed to overcome resolving limitations of the 1-D approach. The differences between germinating and developing material can be analysed using mass spectrometric protein characterization by MALDI TOF.

Although mostly applied to medical research, and therefore mostly using various animal and microbiological material, proteomics is becoming increasingly applied to the study of plants. Particular applications in plant biology are wide ranging and include analysis of genetic variation, investigation of phylogenetic distances, mutant characterization, and monitoring the effects of transgenic manipulation. This is aided by genomic and EST sequence databases becoming available from plant sources and for this reason attention in plants has mainly focused on species with extensive available databases (*Arabidopsis*,

maize, rice). For example, proteins of green and etiolated rice shoots were separated by 2-DE for selective sequencing as part of a study intended to complement the rice Genome Research Project (Komatsu *et al*, 1999). *Arabidopsis* plasma membrane has also been the subject of large-scale differential 2-DE-based proteomic investigation (Santoni *et al*, 1999).

The application of spectrometry is rapidly increasing in the published literature. A particularly impressive illustration of the high throughput potential for identifying proteins using this technology is a 2-D approach used to identify yeast proteins (Shevchenko *et al*, 1996). In a situation in which the strain used has a fully sequenced genome, MALDI TOF analysis alone was sufficient to identify up to 90% of the proteins investigated. By a combination of 2-D electrophoresis plus MALDI and, where necessary nanoelectrospray tandem mass spectrometry, a total of 150 protein spots were unambiguously identified in yeast. Of these 32 were novel. The work of Shevchenko *et al* demonstrates the power and reality of high throughput mass spectrometry in coupling genomics data to the proteome following 2-D electrophoresis. Criteria for a high degree of success are (1) a biological sample which can be consistently generated and reproducibly resolved by 2-DE, (2) the availability of a substantial DNA database for the organism under study, (3) accurate mass characterization, and (4) experience in data editing and database searching. These factors can combine to make possible the routine identification of hundreds of proteins within a weekly timescale. In this way, modern proteomics far exceeds the power of more traditional characterization methods both in terms of throughput and sensitivity. The enabling technologies are however subject to

certain limitations. With regard to the present study the principle difficulties lie with the electrophoretic separation of a membranous cellular fraction, and the lack of a significant DNA or protein database.

The importance of a substantial database representation for the species concerned cannot be overemphasized when interpreting MALDI mass ion data. GenBank contains 827 nucleotide and 174 protein (12 of which are structural) entries for *Ricinus communis*. Of the 827 nucleotide *Ricinus* entries, 747 are EST clones. This modest number of EST entries was submitted to GenBank following an attempt to target clones derived from microsomes, involved in lipid assembly, in the developing seed (Van de Loo et.al. 1995). Putative functions of the proteins was made following homology-based database searches.

The aim of the work described in this chapter was to exploit the unique *Ricinus* ER fraction available for the generation of novel peptide mass fingerprint data following electrophoretic separation of the component proteins. The results of these analyses are presented to illustrate the current limitations and difficulties encountered with this system. This work provides the opportunity for a fully integrated genomic/proteomic initiative into the identification of important enzymes from this organelle. The chapter includes 2-D analyses and the generous help of Dr. Colin Wheeler and Dr. Michael J. Dunn of Harefield Hospital. All mass spectrometry, data editing, and database analysis was performed in this laboratory.

It was decided in the first instance to analyse proteins from Coomassie-stained 1-D profiles. Protein bands visible with Coomassie should yield peptides well within the limits of detection of the mass spectrometer, with minimal interference from contaminants. This exercise was intended to also confirm the identities of some of the castor ER proteins which had been previously characterized and were therefore present in the public access databases. PDI and calreticulin were targeted and results generated from these samples were used for a detailed evaluation of the requirements for data analysis and editing.

## **5.2 Generation of sample for investigation**

The same ER fractions were used throughout the work described in this chapter since inevitable slight variation between different preparations may only become apparent once high resolution and high sensitivity techniques are employed (2-DE, silver staining). This was especially important to allow meaningful comparison between the different gel loadings and staining methods which were to be examined.

The 2-D gels presented in this chapter were run at Harefield Hospital by Drs. Wheeler and Dunn, these being larger scale (18 cm Immobiline 3-10, vertical 12% SDS-PAGE; 18 cm x 21 cm x 1.5 mm). Analytical large-scale 2-D gels were stained by a conventional silver protocol (50 $\mu$ g loading). Preparative (200  $\mu$ g) gels were stained with a dedicated silver stain, eliminating the use of glutaraldehyde (PlusOne, Pharmacia), a method which is less sensitive but ms-compatible.



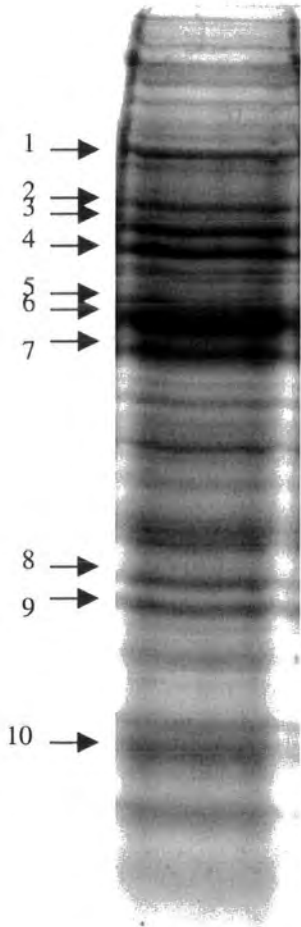
### 5.3 Preparative 1-D SDS-PAGE

Samples of germinating and developing ER (20 $\mu$ g total protein) were run side-by-side on a 12% acrylamide MiniGel, they were stained with Coomassie and selected bands were subjected to MALDI TOF analysis following processing. Ten bands designated G1-10 and D1-10 were excised, from germinating and developing profiles respectively, then subjected to tryptic digestion. Most of these candidates appeared common to both developmental stages, although several were stage-specific for example developing D1, D4, D9 and D10, and G10 (figure 5.1). Only bands which were relatively intense (i.e. > 200 ng) were selected. Clearly resolved intensive bands were chosen to minimise contaminating peaks in the resultant spectra.

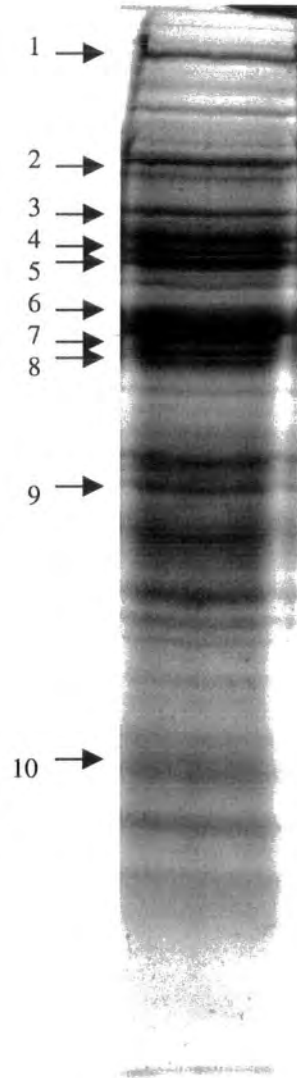
There were two main objectives when choosing bands for this preliminary analysis. The first was confirmation of several of the major known castor ER proteins, thereby permitting annotation of a limited 1-D map through peptide mass fingerprint characterization (such designations provide a potential means of future preparation authenticity checks). The second aim was generation of data from non-characterized proteins, some of which were stage-specific.

**Figure 5.1: Preparative 1-D SDS-PAGE profiles of germinating and developing ER.** 20  $\mu\text{g}$  of each sample was loaded, side by side, on a 12% acrylamide gel. 10 bands were selected from each profile for tryptic digestion and MALDI TOF ms. A blank gel slice was also processed as control. The gel was Coomassie-stained.

**Germinating  
ER**



**Developing  
ER**



#### 5.4 MALDI analysis of 1-D separated ER proteins and database searching

Trypic digestion mixtures were analysed both directly as aliquots of the digest solution, or following acetonitrile extraction of peptides from gel pieces and speed vacuum concentration. Comparison of spectra generated from each of these methods revealed little or no data similarity, and none of the directly loaded samples could be identified from the databases. The main reason for this difference is believed to be higher relative levels of contamination in the more dilute direct loadings, many of the peaks being common to several different samples. In other words, the ratio of contaminant peaks to real data is much higher in the dilute sample than in the more concentrated acetonitrile extracts. The acetonitrile extracts were deemed to provide the more reliable data set, and it was these data which were more thoroughly processed for database searches. MALDI calibration mixtures were positioned as close external standards.

In order to illustrate the quality of raw data, the manner in which results were edited prior to database searching is outlined here. Firstly, all mass values lower than 800 Da were removed. Below this limit, it was common to observe increased background 'noise' levels arising from ions fragmenting from the matrix. As might be expected such noise generally increased with laser intensity. Secondly, sample data sets were compared to each other to reveal common recurring peaks which could be background subtracted. Similarly, comparison with a control sample (blank gel slice) was made to identify non-sample peaks (e.g. keratin-derived and/or tryptic autodigest products). The decision on whether similar potentially contaminant peaks were in fact the same was made according to the machine's running ppm mass accuracy, as indicated by standard calibration data.

However, allowance was made for the fact that standards were external to samples and this has an effect of reducing mass accuracy of assignments. Although the accuracy of the machine was routinely within a value of 10 dpm during calibration it was found during database searches that error values could be as high as 150 ppm. Following editing procedures, almost all of the data obtained from direct loading of tryptic supernatants was disregarded, leaving few, if any, mass ions with which to interrogate public access databases.

The exercise demonstrated the need to perform acetonitrile extraction if meaningful results were to be obtained, and emphasised the importance of close examination of all generated data. All data presented in this chapter was obtained following acetonitrile extraction and concentration, and where indicated, following clean up through micro reverse phase columns.

Of the 20 protein samples prepared, 14 yielded data. The results are shown in table 5.1 and the respective spectra presented in figures. 5.2 to 5.16. Edited (non-sample) values are indicated in table 5.1.

It was expected that this set of data would include mass ions from known proteins represented in the databases (e.g. NCBI nr, SwissProt, OWL). All data generated was entered into the MASCOT search programme for database comparison. In addition, the experimental results from bands likely to be those of known proteins were compared with the results of theoretical tryptic digests of those sequences. This information was

**Figures 5.2-5.16: MALDI-TOF ms spectra from 1-D separated ER proteins.** Tryptic digest samples were analysed following acetonitrile extraction and concentration. Samples were mixed 1:1 with  $\alpha$ -cyano 4-hydroxy cinnamic acid (matrix) and loaded onto a sample plate. Measurements were made in reflector mode with an accelerating voltage of 20 000. No guide wire voltage was applied. The low mass gate was set at 600 and laser intensity ranged from 2000-2150. The mass spectrometer was run in reflector mode, mass values are monoisotopic.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\Dan\Dan11.ms  
This File #1 : C:\VOYAGER\DATA\BILL\DAN\ DAN11.MS  
Comment:

Method: BSLOWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.000 %  
Delay: 200 ON  
Sample: 34

Laser: 2000  
Scans Averaged: 71  
Pressure: 3.36e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 10/28/99 4:17 PM

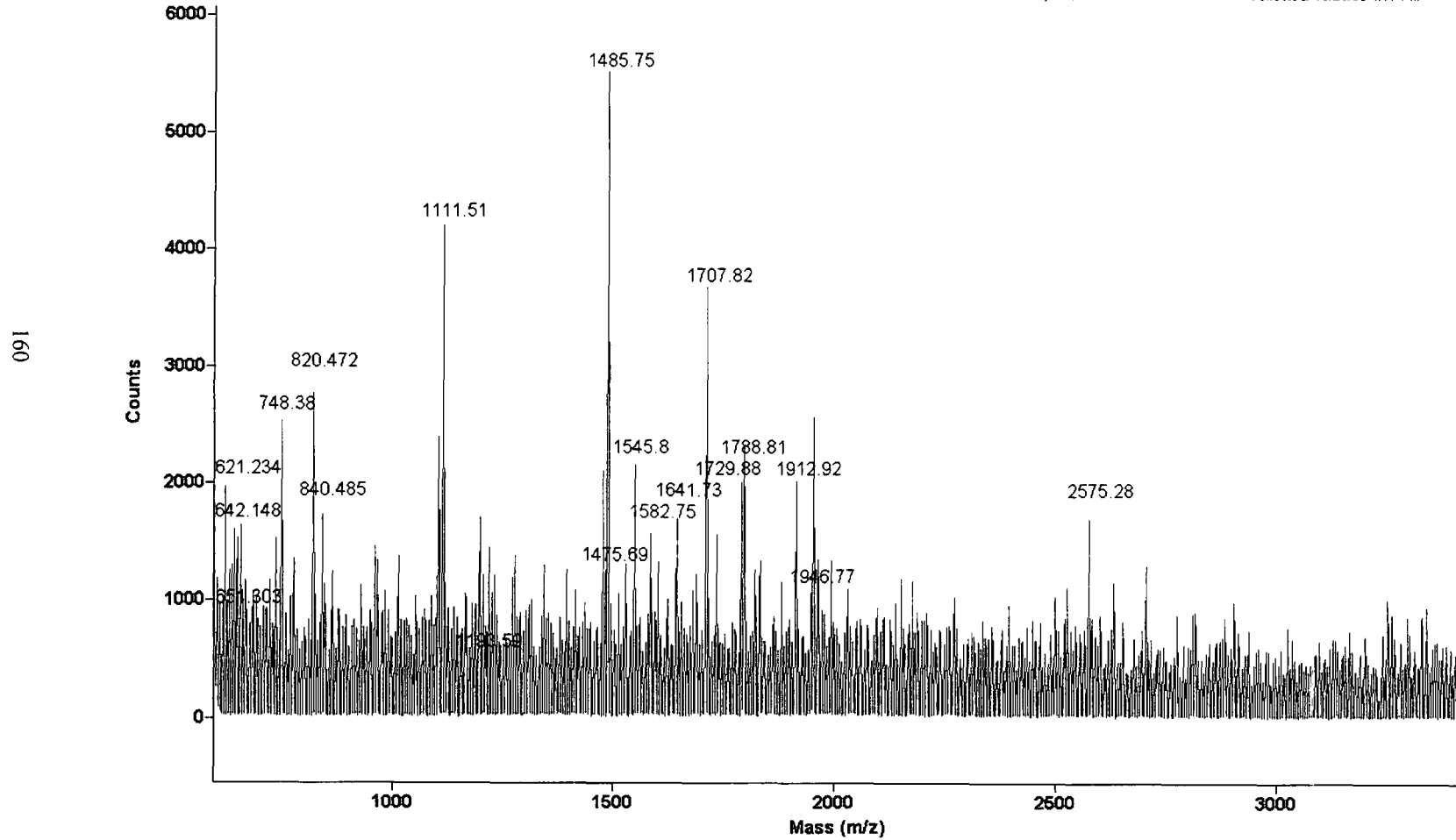


Figure 5.2: MALDI ms spectrum from germinating ER 1-D gel, sample G1.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\ndan\Dan12.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\NDAN\NDAN12.MS  
Comment:

Method: BSLWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.000 %  
Delay: 200 ON  
Sample: 35

Laser: 2150  
Scans Averaged: 147  
Pressure: 2.75e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 10/28/99 4:24 PM

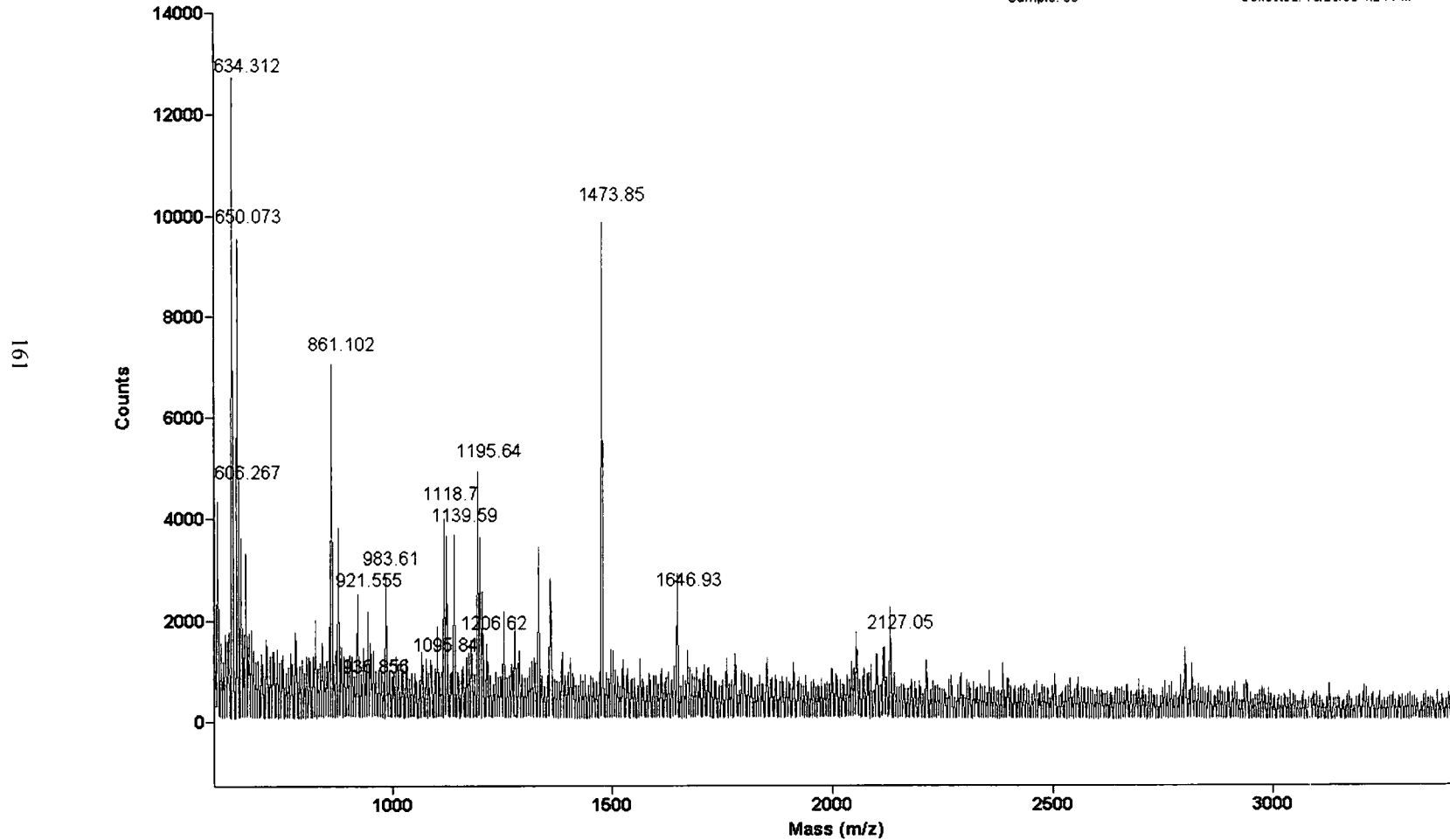


Figure 5.3: MALDI ms spectrum from germinating ER 1-D gel, sample G2.



# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\dan\Dan13.ms  
This File # 1: C:\VOYAGER\DATA\BILL\DAN\13.MS  
Comment:

Method: BSLWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.000 %  
Delay: 200 ON  
Sample: 36

Laser: 2150  
Scans Averaged: 66  
Pressure: 2.66e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 10/28/99 4:25 PM

162

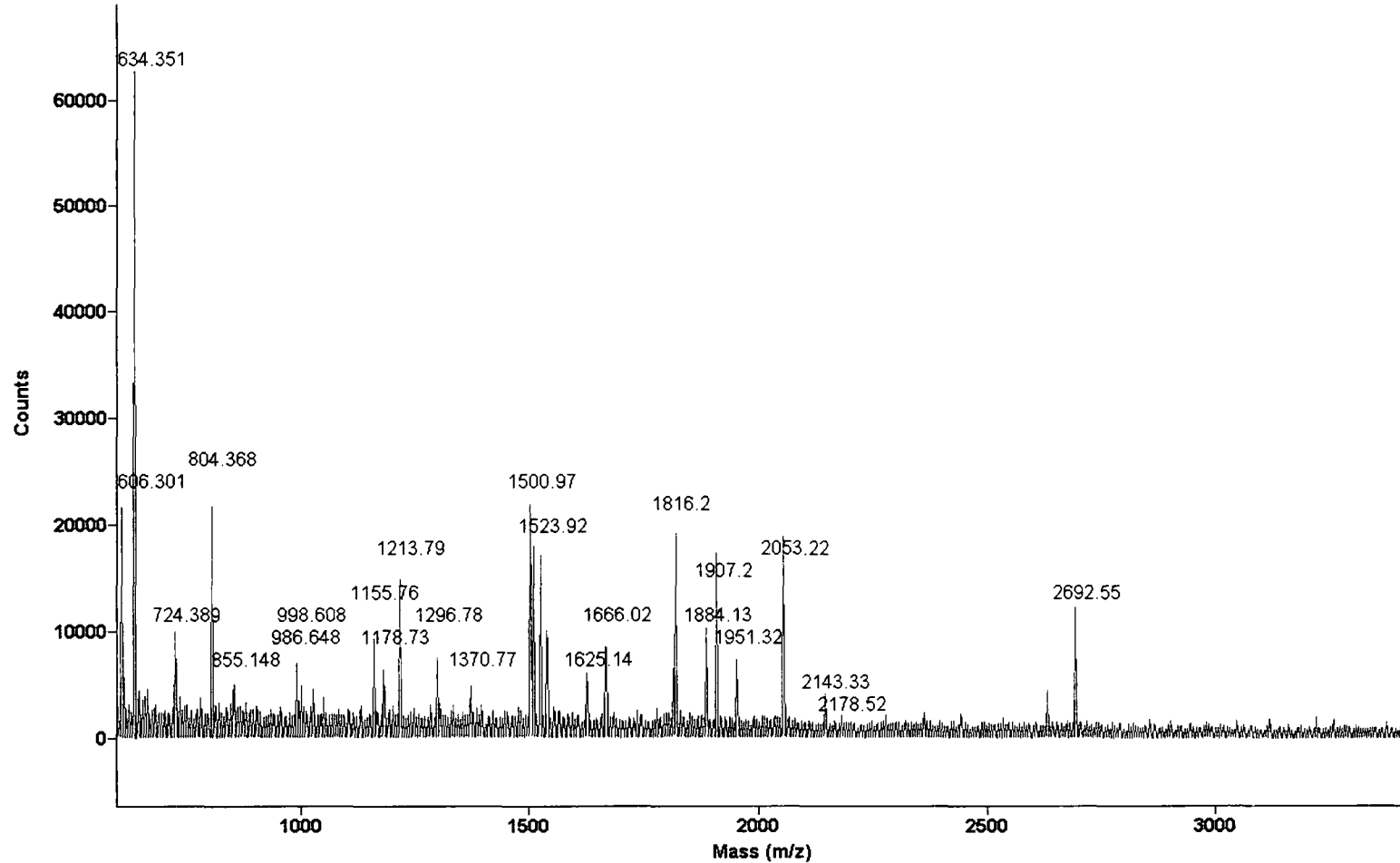


Figure 5.4: MALDI ms spectrum from germination ER 1-D gel, sample G3.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\dan\Dan25.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\DAN\ DAN25.MS  
Comment:

Method: BSL0WRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.000 %  
Delay: 200 ON  
Sample: 34

Laser: 2150  
Scans Averaged: 59  
Pressure: 2.47e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 10/29/99 3:27 PM

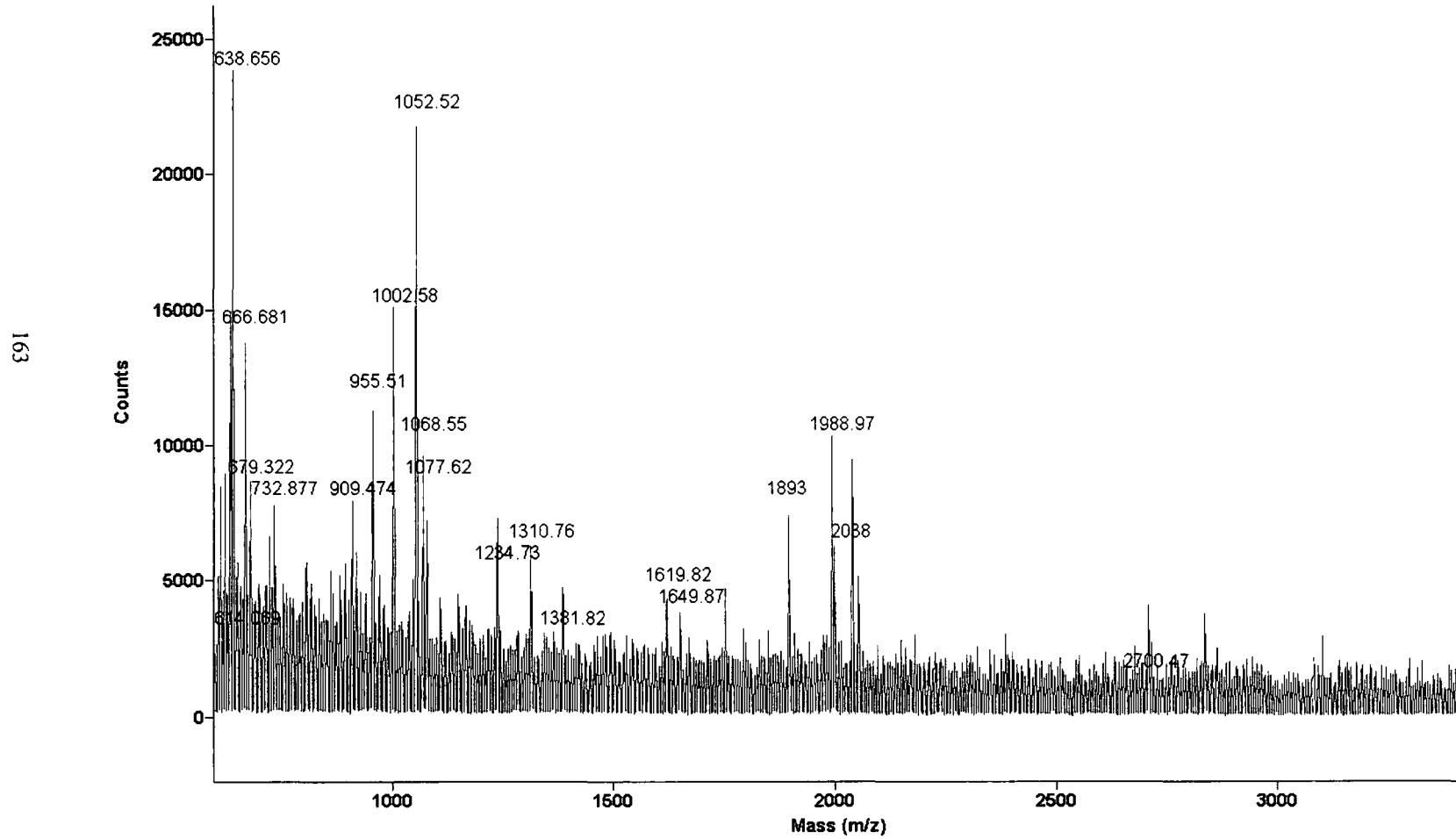


Figure 5.5: MALDI ms spectrum from germinating ER 1-D gel, sample G5.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\dan\Dan26.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\ DAN\ DAN26.MS  
Comment:

Method: BSLWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.000 %  
Delay: 200 ON  
Sample: 35

Laser: 2150  
Scans Averaged: 96  
Pressure: 2.46e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 10/29/99 3:28 PM

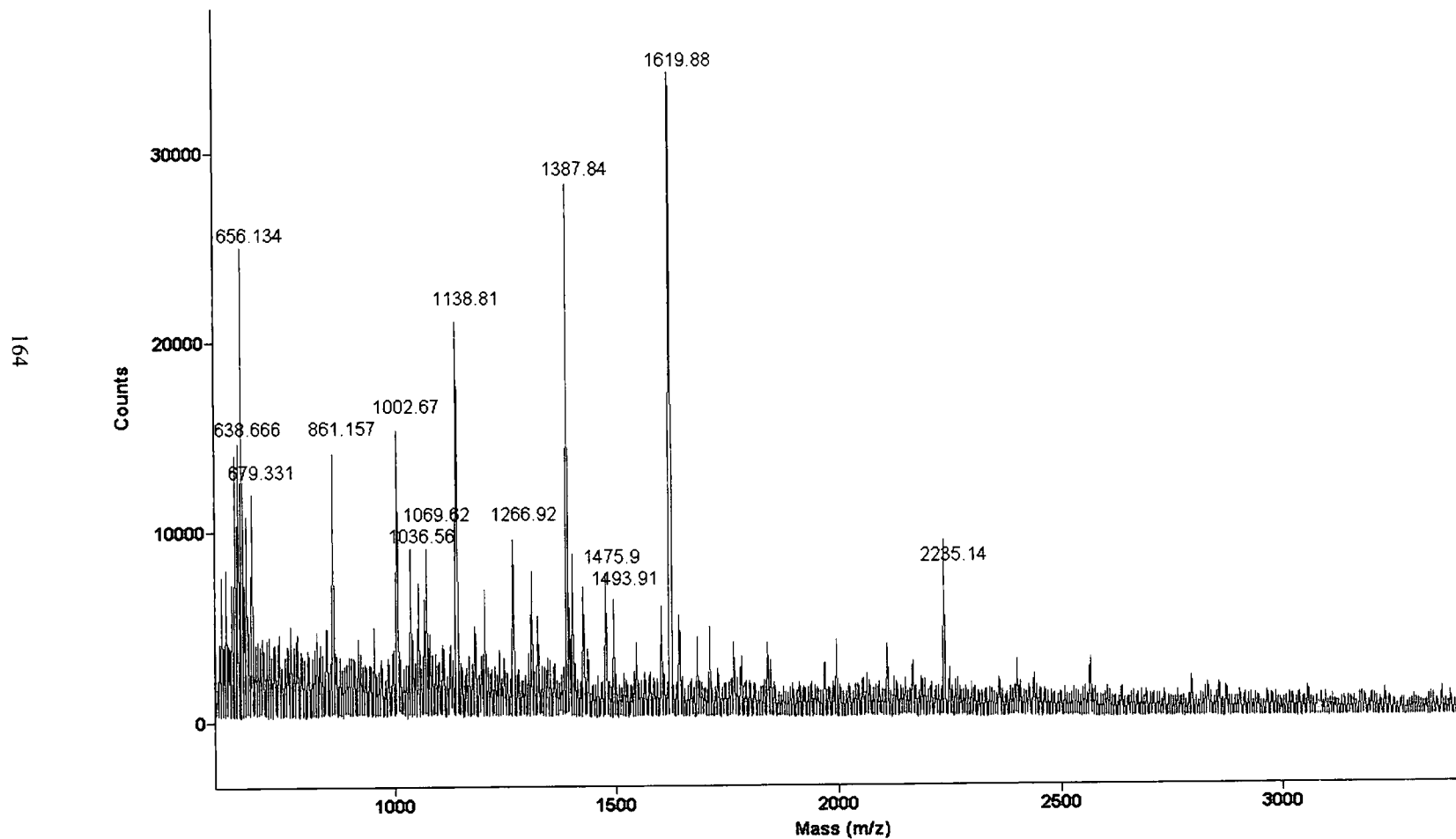


Figure 5.6: MALDI ms spectrum from germinating ER 1-D gel, sample G6.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\dan\Dan29.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\DAN\ DAN29.MS  
Comment:

Method: BSLOWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.000 %  
Delay: 200 ON  
Sample: 38

Laser: 2150  
Scans Averaged: 125  
Pressure: 2.21e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 10/29/99 3:38 PM

165

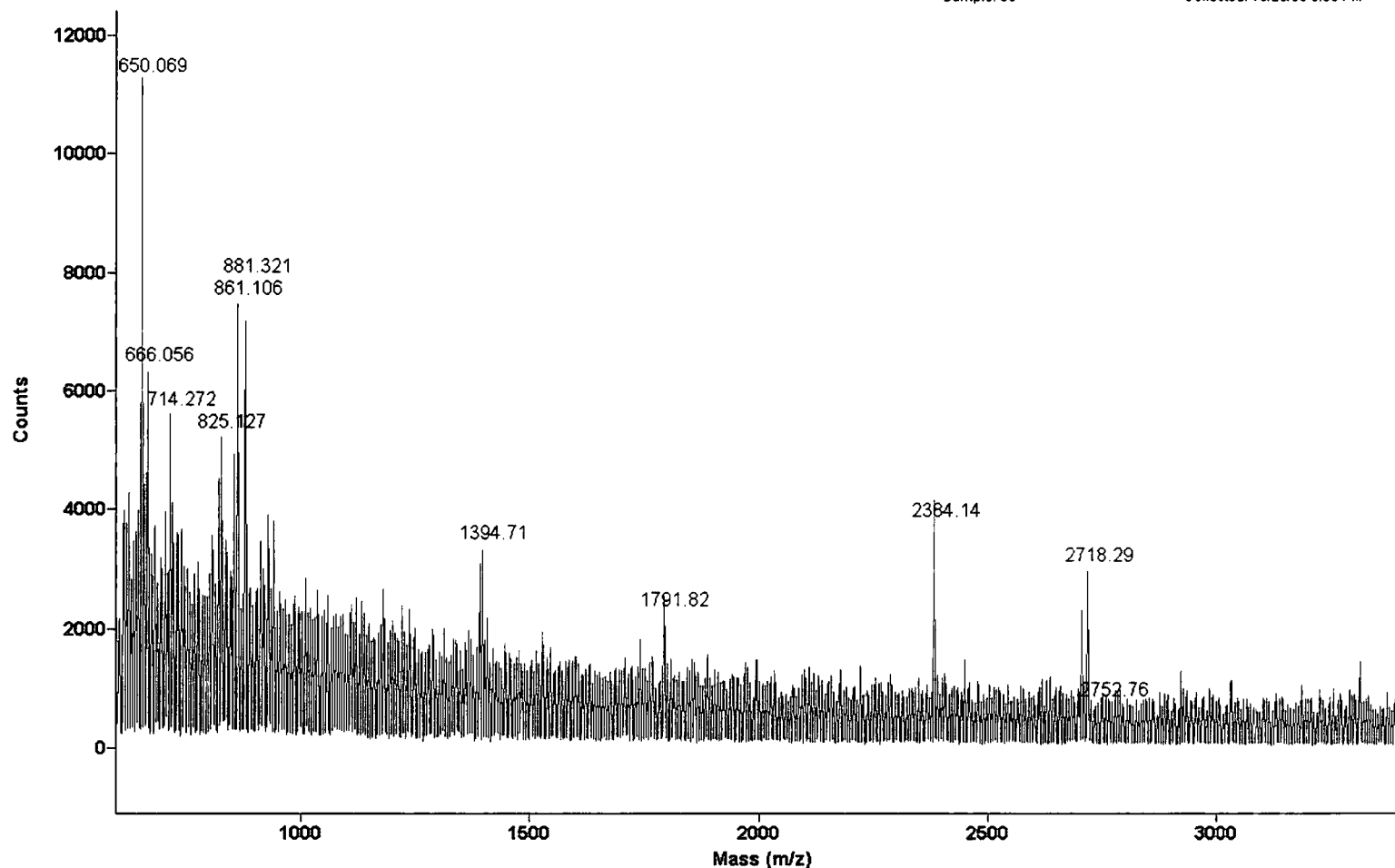


Figure 5.7: MALDI ms spectrum from germinating ER 1-D gel, sample G9.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\dan\Dan14.ms  
This File #1: C:\VOYAGER\DATA\BILL\DAN\ DAN14.MS  
Comment:

Method: BSLWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.000 %  
Delay: 200 ON  
Sample: 37

Laser: 2150  
Scans Averaged: 68  
Pressure: 2.61e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 10/28/99 4:26 PM

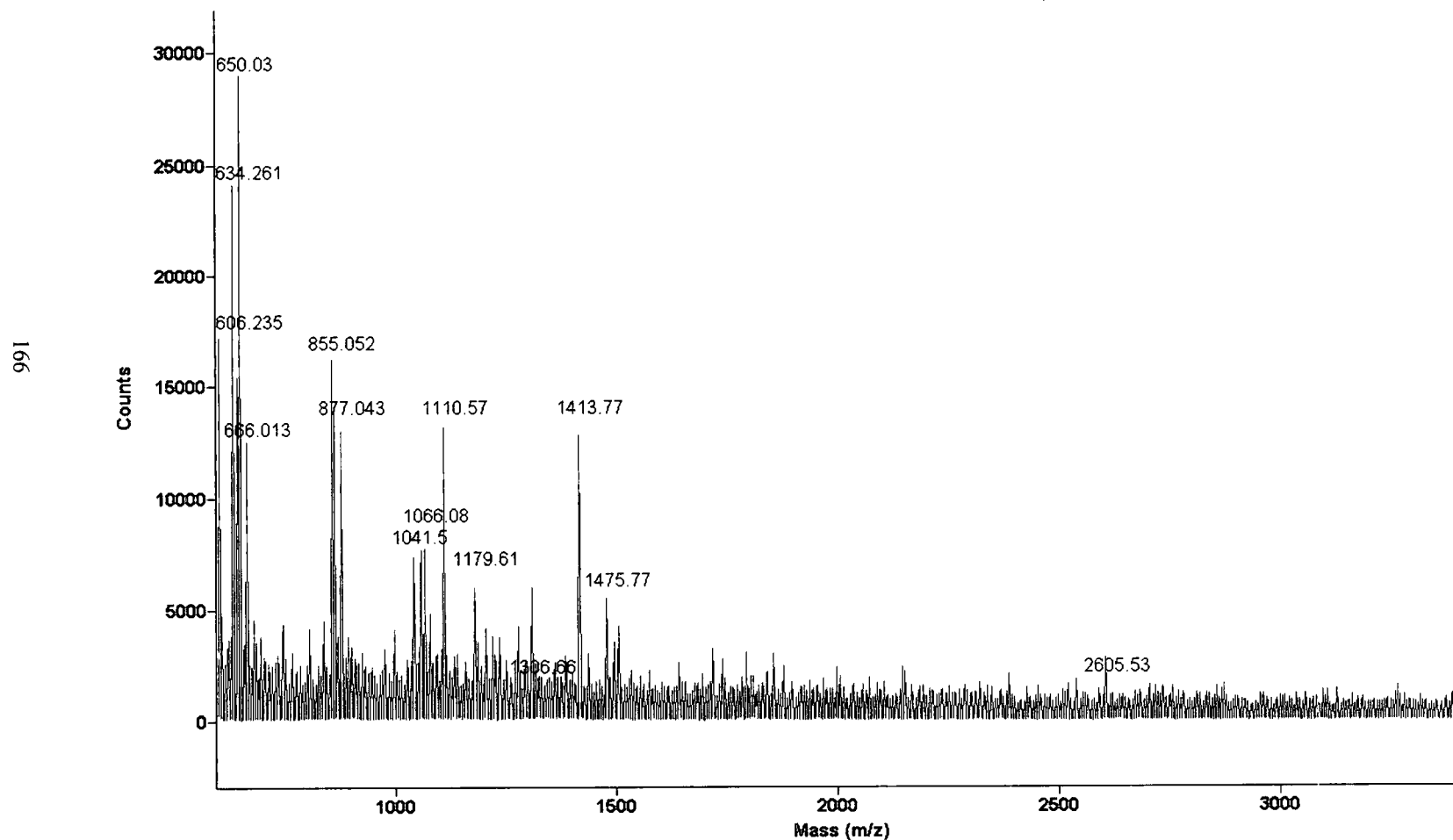


Figure 5.8: MALDI ms spectrum from germinating ER 1-D gel, sample G10.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\dan\Dan16.ms  
 This File # 1 : C:\VOYAGER\DATA\BILL\DAN\16.MS  
 Comment:

Method: BSLOWRDE  
 Mode: Reflector  
 Accelerating Voltage: 20000  
 Grid Voltage: 74.500 %  
 Guide Wire Voltage: 0.000 %  
 Delay: 200 ON  
 Sample: 43

Laser: 2150  
 Scans Averaged: 64  
 Pressure: 2.47e-07  
 Low Mass Gate: 600.0  
 Timed Ion Selector: 382.2 OFF  
 Negative Ions: OFF  
 Collected: 10/28/99 4:31 PM

167

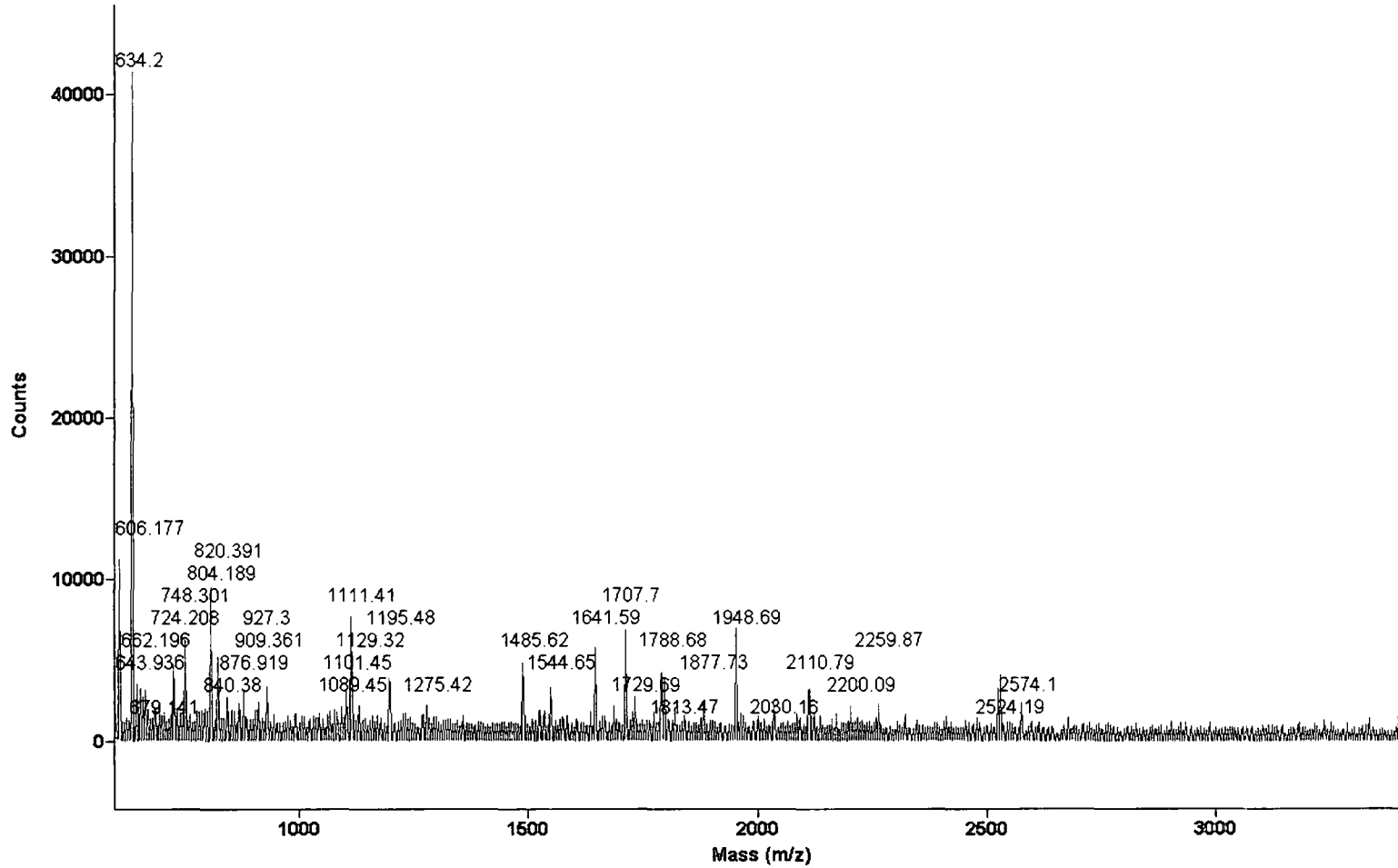


Figure 5.9: MALDI ms spectrum from developing ER 1-D gel, sample D2.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\dan\Dan32.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\DAN\ DAN32.MS  
Comment:

Method: BLOWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.000 %  
Delay: 200 ON  
Sample: 43

Laser: 2150  
Scans Averaged: 96  
Pressure: 2.17e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 10/29/99 3:46 PM

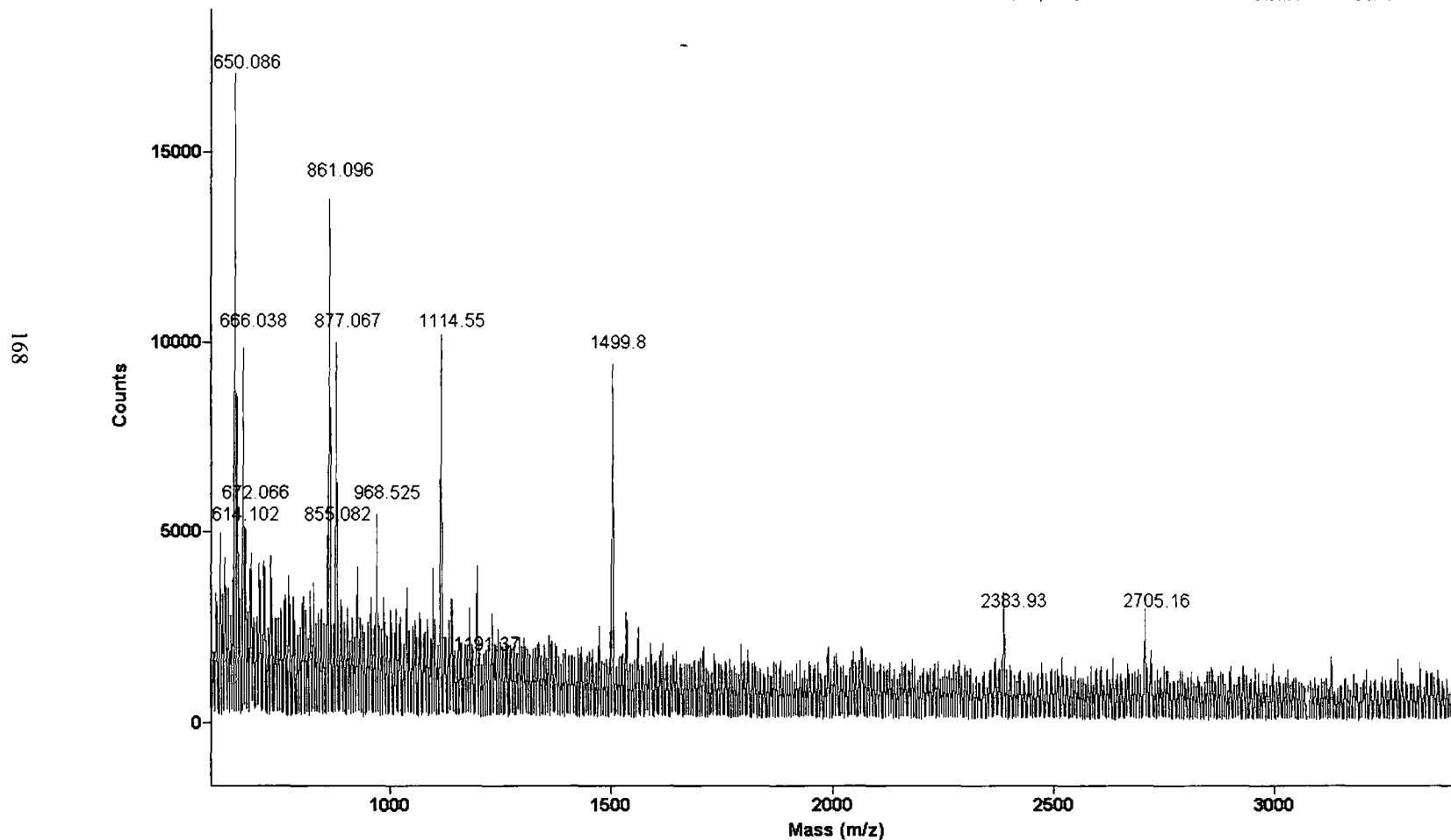


Figure 5.10: MALDI ms spectrum from developing ER 1-D gel, sample D3.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\dan\Dan33.ms  
This File # 1: C:\VOYAGER\DATA\BILL\DAN\ DAN33.MS  
Comment:

Method: BSLWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.000 %  
Delay: 200 ON  
Sample: 44

Laser: 2150  
Scans Averaged: 111  
Pressure: 2.17e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 10/29/99 3:48 PM

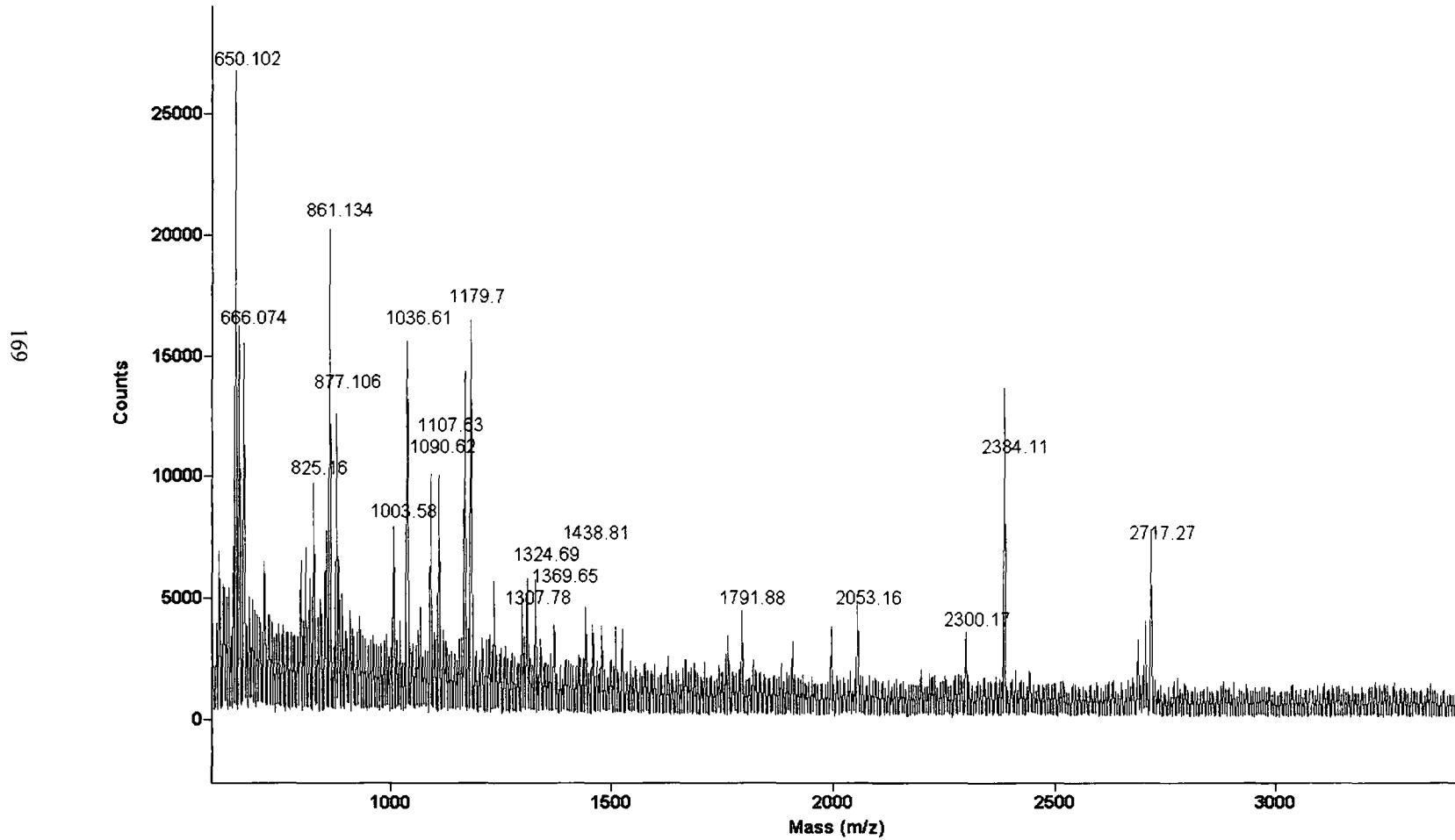


Figure 5.11: MALDI ms spectrum from developing ER 1-D gel, sample D4.



# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\dan\Dan35.ms  
This File #1: C:\VOYAGER\DATA\BILL\DAN\ DAN35.MS  
Comment:

Method: BSLOWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.000 %  
Delay: 200 ON  
Sample: 45

Laser: 2200  
Scans Averaged: 142  
Pressure: 2.12e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 10/29/99 3:51 PM

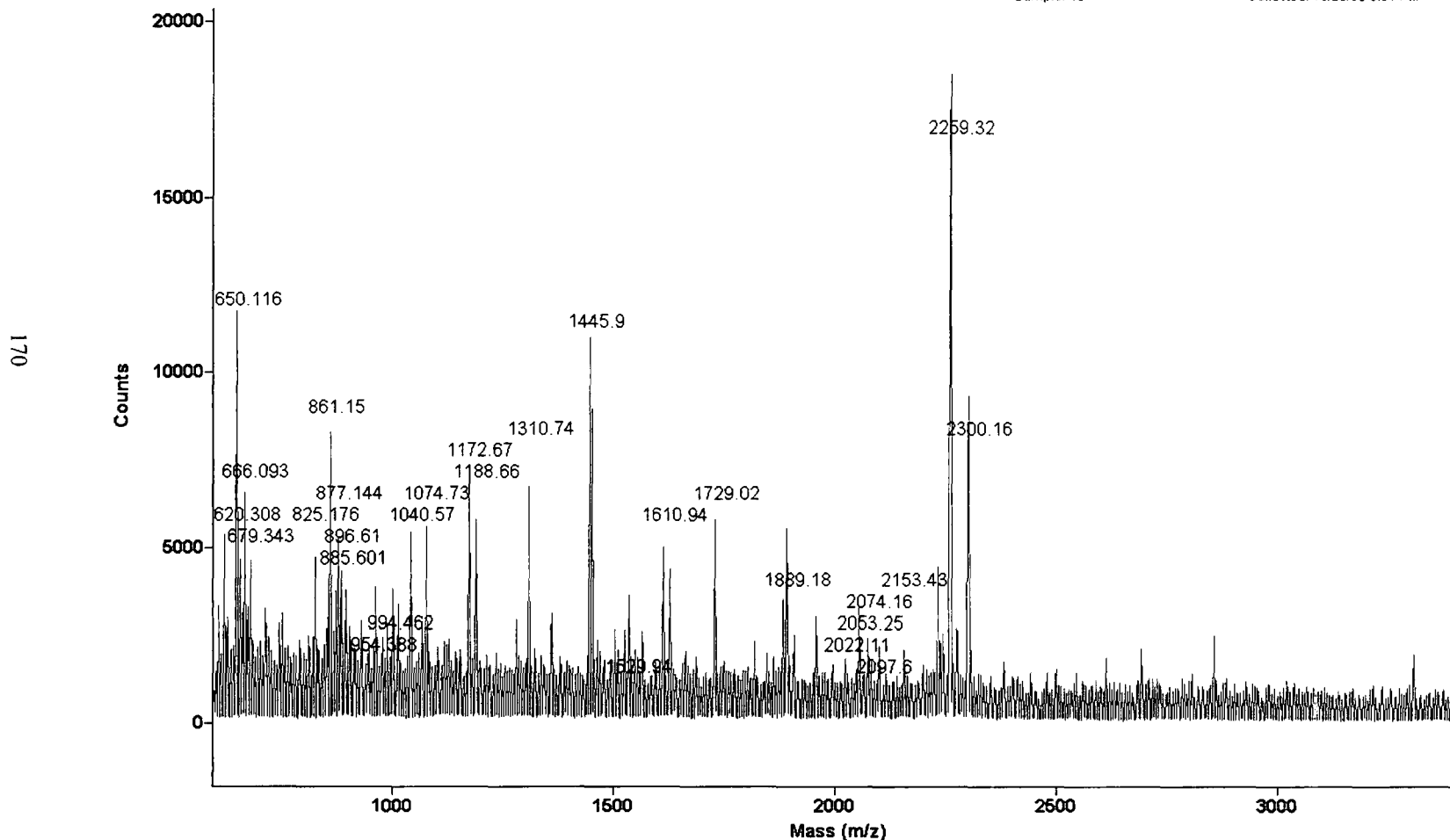


Figure 5.12: MALDI ms spectrum from developing ER 1-D gel, sample D5.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\dan\Dan36.ms  
This File #1 : C:\VOYAGER\DATA\BILL\DAN\DAN36.MS  
Comment:

Method: BSLOWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.000 %  
Delay: 200 ON  
Sample: 46

Laser: 2200  
Scans Averaged: 83  
Pressure: 2.10e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 10/29/99 3:52 PM

171

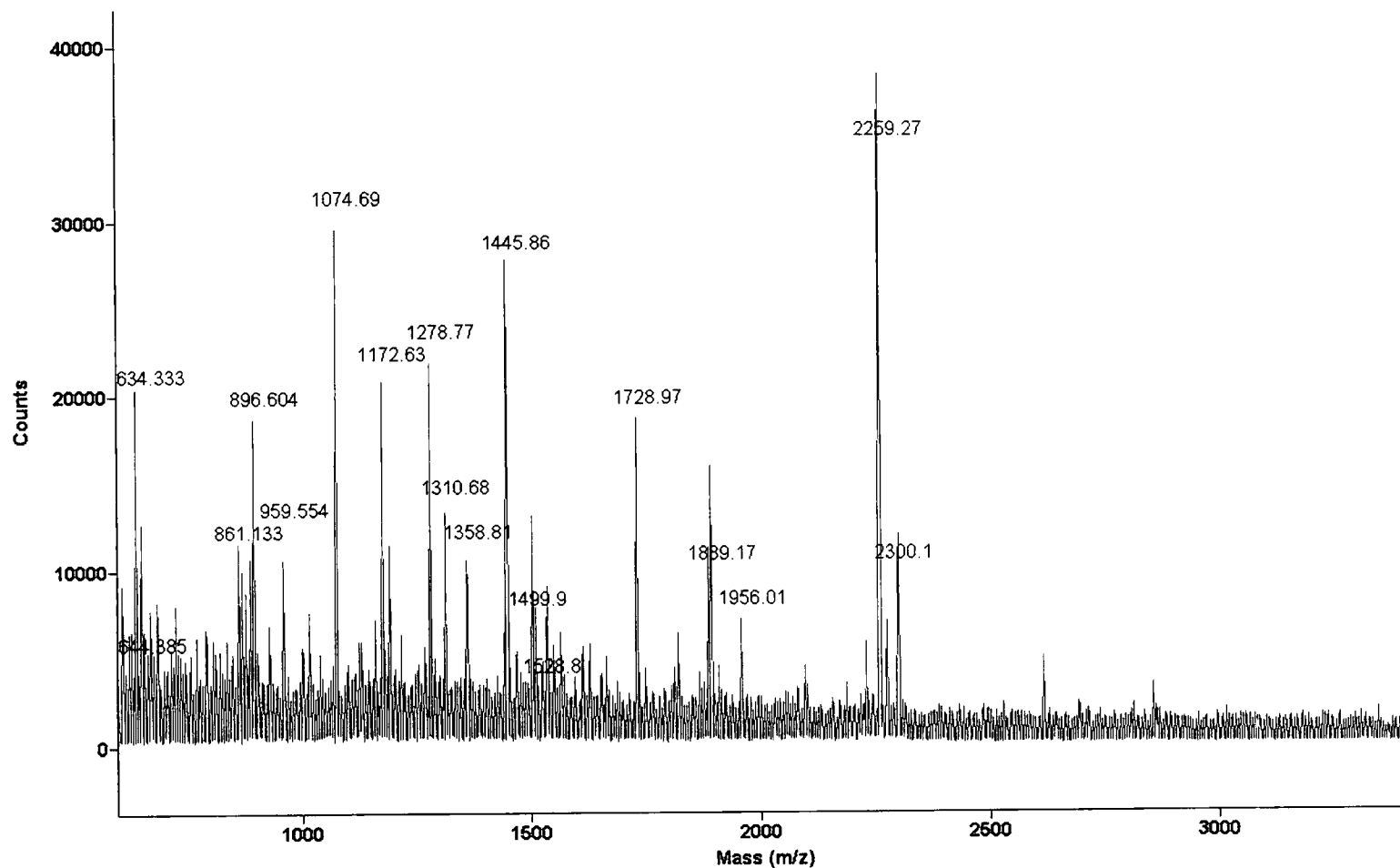


Figure 5.13: MALDI ms spectrum from developing ER 1-D gel, sample D6.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\dan\Dan17.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\DAN\17.MS  
Comment:

Method: BSLOWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.000 %  
Delay: 200 ON  
Sample: 44

Laser: 2150  
Scans Averaged: 98  
Pressure: 2.44e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 10/28/99 4:32 PM

172

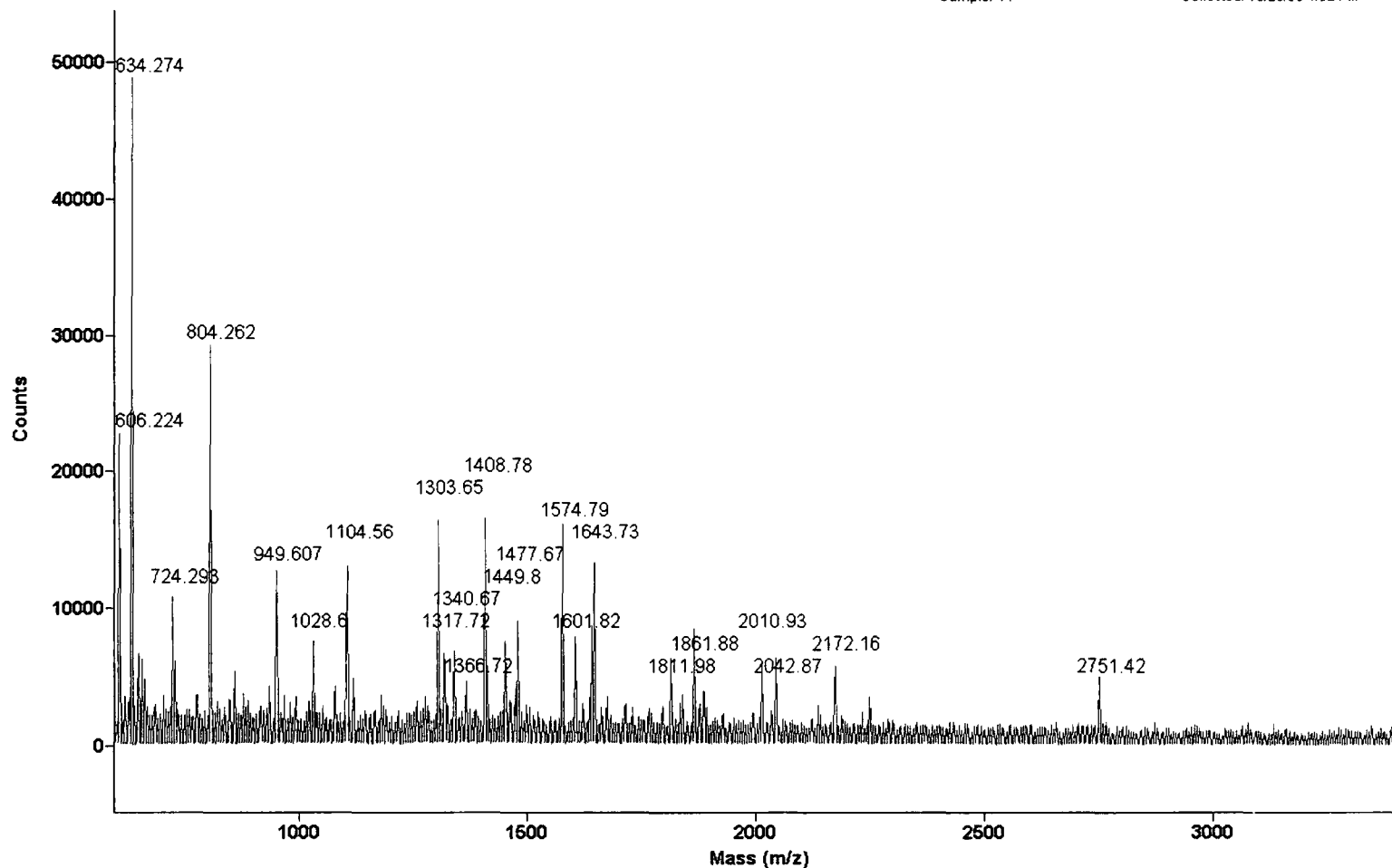


Figure 5.14: MALDI ms spectrum from developing ER 1-D gel, sample D7.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\dan\Dan38.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\DAM\ DAN38.MS  
Comment:

Method: BSLWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.000 %  
Delay: 200 ON  
Sample: 48

Laser: 2200  
Scans Averaged: 179  
Pressure: 2.08e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 10/29/99 3:56 PM

173

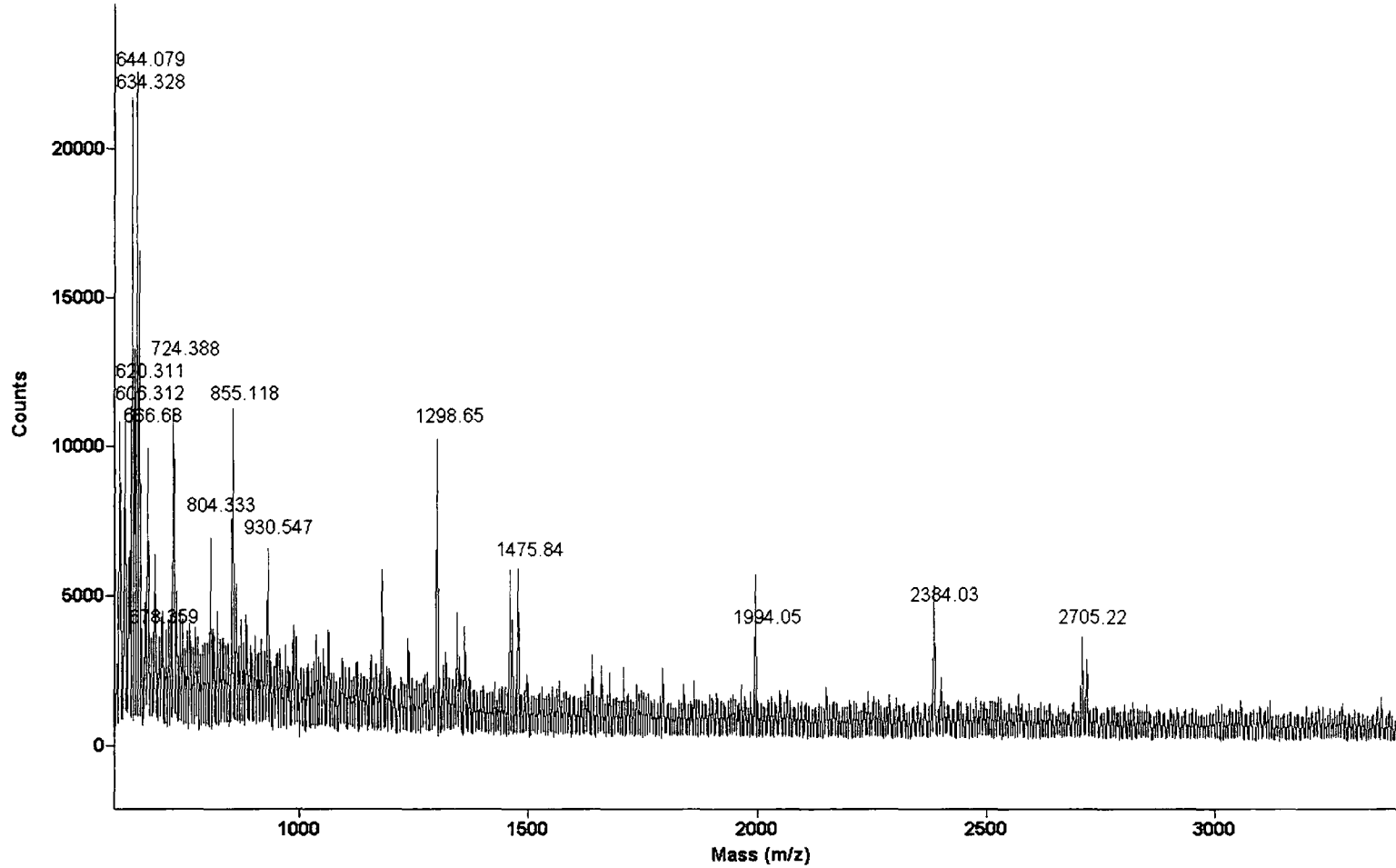


Figure 5.15: MALDI ms spectrum from developing ER 1-D gel, sample D9.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\dan\Dan22.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\DAN\ DAN22.MS  
Comment:

Method: BSLWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.000 %  
Delay: 200 ON  
Sample: 33

Laser: 2150  
Scans Averaged: 60  
Pressure: 2.35e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 10/28/99 4:37 PM

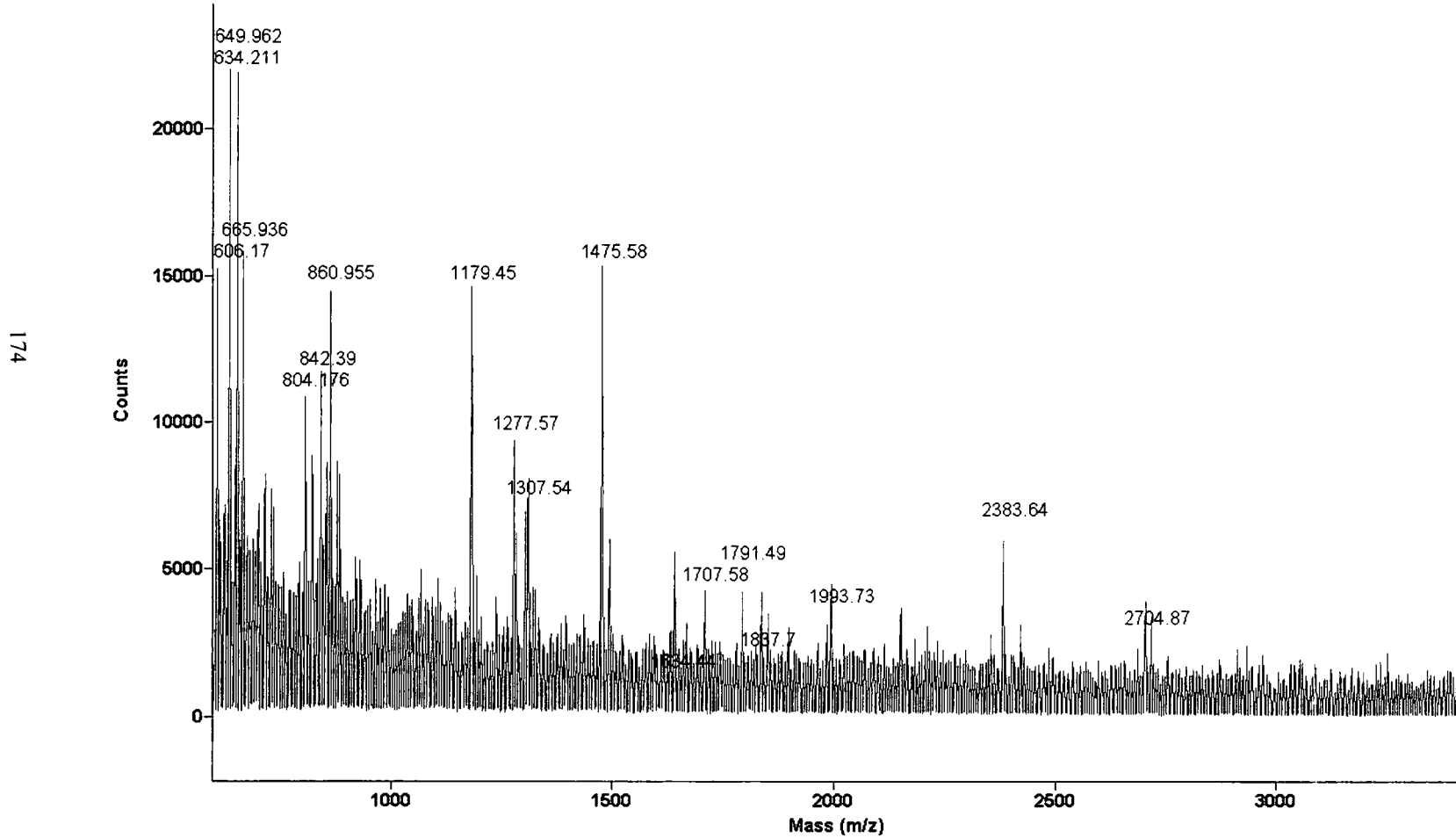


Figure 5.16: MALDI ms spectrum from 1-D SDS-PAGE, blank sample.

generated using the MS-Digest programme at the Protein Prospector website (<http://prospector.ucsf.edu>). Proteins selected for this exercise were protein disulphide-isomerase, calreticulin, non-specific lipid transfer protein, and the 2S albumin precursor (see appendix).

#### **5.4.1 Comparison of experimental data with theoretical mass fingerprint profiles of known castor ER proteins**

By checking the data from major bands in the 50-60 kDa region against theoretical tryptic digests of castor PDI and calreticulin, it was possible to identify several likely peptides from each in the G5 sample within a 100 ppm error range (PDI: 1077.62, 1381.82, 2038; calreticulin: 1052.52, 1068.55, 1234.73, 1649.87, 1893). Three of the peptides generated from sample D7 matched with calreticulin peptides, all within 50 ppm of the calculated value. Matching peptides were not found for the non-specific lipid transfer protein or 2S albumin precursor.

Inclusion of database-represented proteins in such an exercise allows confident editing of data giving an indication of the proportion of peptide peaks which are sample-derived. Such an exercise also provides valuable indications of machine accuracy with regard to measured sample data.

#### 5.4.2 Database searching using edited MALDI data

Most of the proteins selected from figure 5.1 for MALDI had not previously been characterized from castor. Therefore, much of the data generated in this exercise was likely to be without any corresponding sequence information in the databases, a fact which would complicate protein identification. Nonetheless, all edited data was entered into the MASCOT search program ([www.matrixscience.com](http://www.matrixscience.com)) which enabled peptide mass searching of the OWL and NCBI non-redundant protein databases. Results are summarized in table 5.2. The ability to 'refine' searches by editing out non-matching peptides provides a means of deciphering mixed protein data, in the case of a band containing more than one polypeptide. The edited values may be used in a separate search.

In support of the results of the previous exercise, the G5 data returned castor calreticulin with 5 significant peptide matches at a mass error tolerance of 75 ppm. Four of the remaining peptides (1002.58, 1077.62, 1381.82, 2038) matched with theoretical PDI digest masses (see appendix 1). This confirmation that sample G5 contained both PDI and calreticulin re-emphasizes the resolving limitations of SDS-PAGE. All other germinating ER protein data remained unassigned after database interrogation although an *Arabidopsis* unknown protein matched 5 of the peptides from the G1 sample with a 100 ppm tolerance setting.

Developing sample D4 data produced a significant match of 6 peptides at 150 ppm tolerance with a tobacco protein GRP78-5 (a 78 kDa ER-resident 'glucose regulatory

**Table 5.1: MALDI TOF ms data from SDS-PAGE separated castor ER proteins.**

Coomassie-stained protein bands were selected from a 12% 1-D minigel for MALDI analysis following tryptic digestion (see figure 5.1). G: germinating ER samples; D: developing ER samples. Sample numbers correspond to those in figure 5.1. Data was edited to remove mass values believed to derive from cross contamination and tryptic autolysis. Values represent  $[M+H^+]$  ions. Data from the blank sample are shown below:

**Blank:** 804.176; 842.39; 860.955; 1179.45; 1277.57; 1307.54; 1475.58; 1634.44;  
1707.58; 1791.49; 1837.7; 1993.73; 2383.64; 2704.87 (figure 5.17).



| <b>G1<br/>(fig 5.3)</b> | <b>G2<br/>(fig 5.4)</b> | <b>G3<br/>(fig 5.5)</b> | <b>G5<br/>(fig 5.6)</b> | <b>G6<br/>(fig 5.7)</b> | <b>G9<br/>(fig 5.8)</b> | <b>G10<br/>(fig 5.9)</b> |
|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|--------------------------|
| 1111.51                 | 921.555                 | 986.648                 | 909.474                 | 1002.67                 | 825.127                 | 1041.5                   |
| 1193.59                 | 936.856                 | 998.608                 | 955.51                  | 1036.56                 | 881.321                 | 1066.08                  |
| 1475.69                 | 983.61                  | 1155.76                 | 1002.58                 | 1069.62                 | 1394.71                 | 1110.57                  |
| 1545.8                  | 1095.84                 | 1178.73                 | 1052.52                 | 1138.81                 | 1781.82                 | 1306.66                  |
| 1582.75                 | 1118.7                  | 1213.79                 | 1068.55                 | 1266.92                 | 2384.14                 | 1413.77                  |
| 1641.73                 | 1139.59                 | 1296.78                 | 1077.62                 | 1387.84                 | 2718.29                 | 2605.53                  |
| 1707.82                 | 1195.64                 | 1370.77                 | 1234.73                 | 1493.91                 | 2752.76                 |                          |
| 1729.88                 | 1206.62                 | 1500.97                 | 1310.76                 | 1619.88                 |                         |                          |
| 1788.81                 | 1473.8                  | 1523.92                 | 1381.82                 | 2235.14                 |                         |                          |
| 1912.92                 | 1646.93                 | 1625.14                 | 1619.82                 |                         |                         |                          |
| 1946.77                 | 2127.05                 | 1666.02                 | 1649.87                 |                         |                         |                          |
| 2575.28                 |                         | 1816.2                  | 1893                    |                         |                         |                          |
|                         |                         | 1884.13                 | 1988.97                 |                         |                         |                          |
|                         |                         | 1907.2                  | 2038                    |                         |                         |                          |
|                         |                         | 1951.32                 |                         |                         |                         |                          |
|                         |                         | 2053.22                 |                         |                         |                         |                          |
|                         |                         | 2143.33                 |                         |                         |                         |                          |
|                         |                         | 2178.52                 |                         |                         |                         |                          |
|                         |                         | 2692.55                 |                         |                         |                         |                          |

| <b>D2<br/>(fig 5.10)</b> | <b>D3<br/>(fig 5.11)</b> | <b>D4<br/>(fig 5.12)</b> | <b>D5<br/>(fig 5.13)</b> | <b>D6<br/>(fig 5.14)</b> | <b>D7<br/>(fig 5.15)</b> | <b>D9<br/>(fig 5.16)</b> |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 909.361                  | 968.525                  | 1003.58                  | 885.601                  | 896.604                  | 949.607                  | 930.547                  |
| 927.3                    | 1114.55                  | 1036.61                  | 896.61                   | 959.554                  | 1104.56                  | 1298.65                  |
| 1089.45                  | 1191.37                  | 1090.62                  | 954.388                  | 1074.69                  | 1303.65                  | 1994.05                  |
| 1101.45                  | 1499.8                   | 1107.63                  | 994.462                  | 1278.77                  | 1408.78                  |                          |
| 1111.41                  |                          | 1307.78                  | 1040.57                  | 1310.68                  | 1574.79                  |                          |
| 1129.32                  |                          | 1324.69                  | 1074.73                  | 1358.81                  | 1643.73                  |                          |
| 1195.48                  |                          | 1369.65                  | 1188.66                  | 1445.86                  | 1861.88                  |                          |
| 1275.42                  |                          | 1438.81                  | 1310.74                  | 1499.9                   | 1885.98                  |                          |
| 1485.62                  |                          | 2053.16                  | 1445.9                   | 1528.8                   | 2010.93                  |                          |
| 1544.65                  |                          | 2300.17                  | 1529.94                  | 1728.97                  | 2042.87                  |                          |
| 1641.59                  |                          | 2717.27                  | 1610.94                  | 1889.17                  | 2132.99                  |                          |
| 1707.7                   |                          |                          | 1729.02                  | 1956.01                  | 2172.16                  |                          |
| 1729.69                  |                          |                          | 1889.18                  | 2300.1                   | 2246.08                  |                          |
| 1788.68                  |                          |                          | 2022.11                  |                          | 2751.42                  |                          |
| 1813.47                  |                          |                          | 2053.25                  |                          |                          |                          |
| 1877.73                  |                          |                          | 2074.16                  |                          |                          |                          |
| 1948.69                  |                          |                          | 2097.6                   |                          |                          |                          |
| 2030.16                  |                          |                          | 2153.43                  |                          |                          |                          |
| 2110.79                  |                          |                          | 2259.32                  |                          |                          |                          |
| 2200.09                  |                          |                          | 2300.16                  |                          |                          |                          |
| 2259.87                  |                          |                          |                          |                          |                          |                          |
| 2524.19                  |                          |                          |                          |                          |                          |                          |
| 2574.1                   |                          |                          |                          |                          |                          |                          |

**Table 5.2: Summary of peptide mass fingerprint data which returned significant database matches.** Data from Coomassie-stained germinating ER protein bands were compared against conceptual digests of all OWL and NCBI nr protein entries. Searches were made using the MASCOT program (<http://www.matrixscience.com>).

Settings/restrictions as follows: Viridiplantae (green plants); fixed modifications = oxidation (methionine); variable modifications = propionamide (cysteine); mass values = monoisotopic; protein mass = unrestricted; peptide charge state = 1+; maximum missed cleavages = 1; enzyme = trypsin; mass tolerance limits (ppm) and probability based MOWSE scores are shown.

| Sample | Protein database | Mass Tolerance (ppm) | Matching entry/probability based MOWSE score [observed/significant] | Matching peptides | Monoisotpic mass |
|--------|------------------|----------------------|---|-------------------|------------------|
| G1     | NCBIInr          | 100                  | <i>Arabidopsis</i><br>Unknown protein/<br>[104/61]                  | KSTDEIK           | 820.43           |
|        |                  |                      |   | MHLPTQFQPNT       | 1485.72          |
|        |                  |                      |   | SGNQSKEPQPHIQPK   | 1545.81          |
|        |                  |                      |   | QSLKTGAANNIETNLR  | 1729.91          |
|        |                  |                      |   | SRSNQSKEPQPHIQPK  | 1788.94          |
| G5     | OWL              | 75                   | <i>Ricinus</i><br>calreticulin/<br>[86/58]                          | FEDGWENR          | 1052.44          |
|        |                  |                      |   | GIQTSEDYR         | 1068.49          |
|        |                  |                      |   | YVGIELWQVK        | 1234.68          |
|        |                  |                      |   | FYAISAEFPFESNK    | 1649.78          |
|        |                  |                      |   | VHAILNYNDTNHLIKK  | 1893.03          |
| D4     | OWL              | 100                  | Tobacco<br>78 kDa GRP<br>homologue 5<br>precursor/<br>[77/58]       | LSQEEIER          | 1003.50          |
|        |                  |                      |   | KAMEDAGLEK        | 1107.53          |
|        |                  |                      |   | EAEFAEEDKK        | 1324.58          |
|        |                  |                      |   | VEIESLFDGVDFSEPLT | 2053.01          |
|        |                  |                      |   | DILLLDVAPLTLGIETV | 2384.33          |
| D5     | OWL              | 150                  | <i>Ricinus</i><br>Agglutinin<br>precursor/<br>[59/58]               | LEQLGGLR          | 885.51           |
|        |                  |                      |   | HEIPVLPNR         | 1074.60          |
|        |                  |                      |   | DNCLTTDANIK       | 1278.59          |
|        |                  |                      |   | FQYIEGEMRTR       | 1445.68          |
|        |                  |                      |   | SAPDPSVITLNSWGR   | 1728.85          |
|        |                  |                      |   | NDGTILNLYNGLVLDVR | 1889.01          |
|        |                  |                      |   | LSTAIQESNQGFASPI  | 2259.17          |
| D6     | NCBIInr          | 100                  | Ricin precursor/<br>[62/59]   | HEIPVLPNR         | 1074.60          |
|        |                  |                      |   | FQYIEGEMRTR       | 1445.68          |
|        |                  |                      |   | SAPDPSVITLNSWGR   | 1728.85          |
|        |                  |                      |   | NDGTILNLYNGLVLDVR | 1889.01          |
|        |                  |                      |   | LSTAIQESNQGFASPI  | 2259.17          |
| D7     | OWL              | 150                  | <i>Arabidopsis</i><br>Hypothetical<br>protein/<br>[85/58]           | NEIVESSPVSDK      | 1303.63          |
|        |                  |                      |   | QNPEADFKETSNILGAK | 1861.92          |
|        |                  |                      |   | WMELGEEEKQVYNSK   | 1885.86          |
|        |                  |                      |   | DCAETKRPSTPYILWCK | 2010.97          |
|        |                  |                      |   | SVLEEHPGINNSTVTAH | 2246.17          |
|        | NCBIInr          | 100                  | <i>Arabidopsis</i><br>98b like protein/<br>[69/59]                  | NEIVESSPVSDK      | 1303.63          |
|        |                  |                      |   | EIYLQEMEGYKR      | 1574.74          |
|        |                  |                      |   | QNPEADFKETSNILGAK | 1861.92          |
|        |                  |                      |   | WMELGEEEKQVYNSK   | 1885.86          |

protein'). Such glucose-regulated proteins are known to reside in ER. The position of this band in figure 5.1, at approximately 75 kDa, is in agreement with such an assignment.

Sample D5 produced a significant match of 6 peptides with *Ricinus communis* agglutinin (RCA) precursor. Similarly, RCA was returned with some ricin precursor fragments for the D6 sample. Both agglutinin and ricin are seed storage lectins and their passage through the ER in precursor form is well established (Lord, 1985). Temporally their presence in the ER is highest during development when germination energy reserves are being prepared.

Another unassigned *Arabidopsis* protein was suggested for sample D7. Unfortunately, the time consuming searches of dbEST provided no additional sequence information or insight into protein identity.

### **5.5 2-D gels for proteomic analysis**

Clearly there are limitations in using 1-D SDS-PAGE for protein resolution as a precursor to MALDI TOF have been presented. Large scale 2-DE, suitable for combination with mass spectrometric analysis, was performed on germinating and developing ER samples at analytical (50 µg) and preparative (200 µg) levels (figures 5.17-5.20). ER proteins resolved well in this system. Samples for analytical loadings were prepared via TCA/acetone precipitation prior to solubilization in the IPG

**Figures 5.17-5.20: Large format 2-D gels for analytical and preparative separation of germinating and developing ER.** Protein was applied to IEF as part of the rehydration solution. First dimension: Immobiline pH 3-10 non linear, 18 cm (Pharmacia). Second dimension: 12% homogeneous SDS-PAGE. Gels were run by Drs Wheeler and Dunn, Harefield Hospital. Molecular weight markers: Pharmacia. Analytical gels (figures 5.18 and 5.20) are silverstained according to standard conditions. Preparative gels (figures 5.19 and 5.21) are stained by the PlusOne silver method (Pharmacia). Molecular weight standards: Amersham Rainbow Markers.

180

12% SDS-PAGE (20 cm)

220 kDa

97.4

66

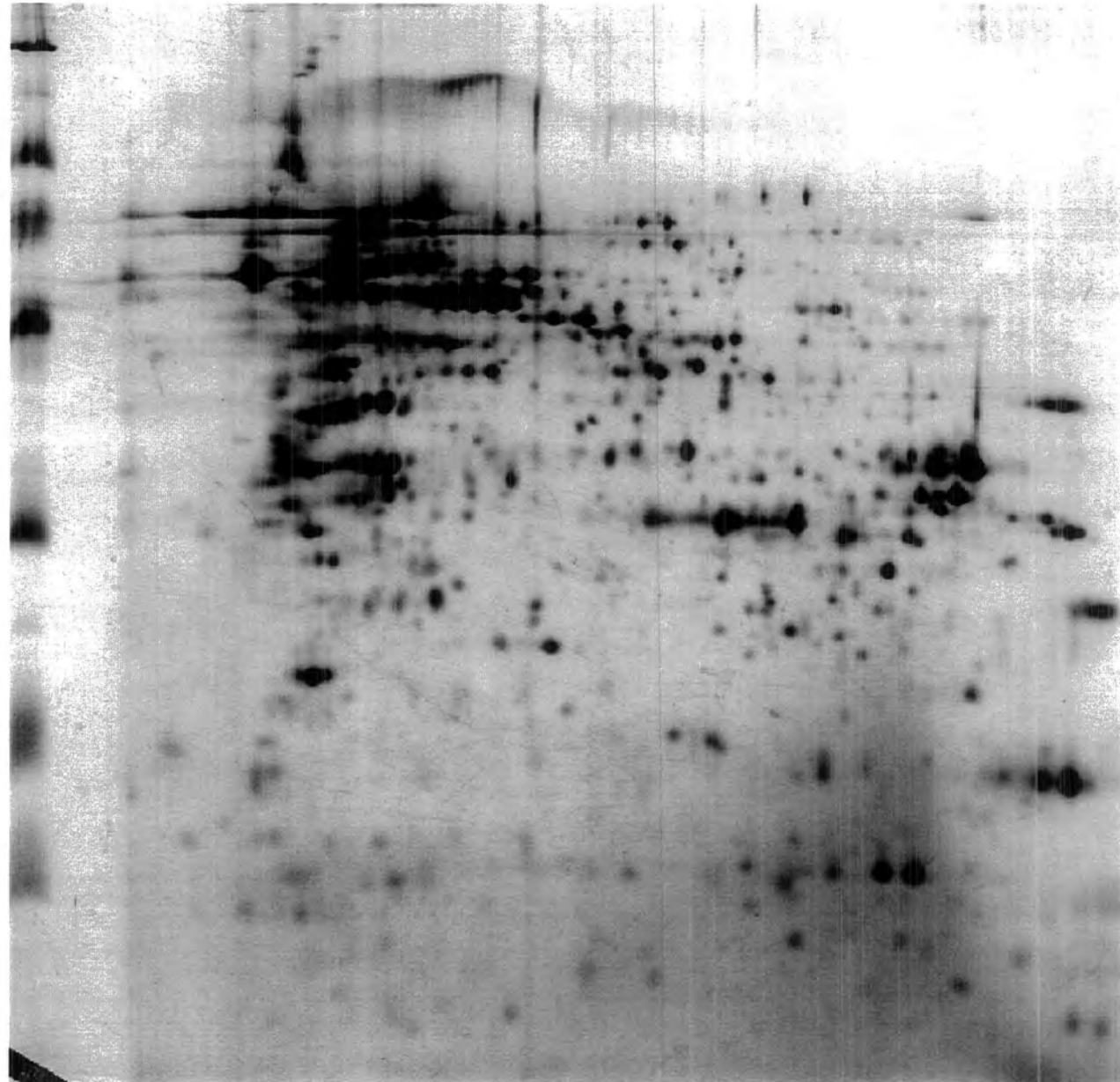
46

30

21.5

14.3

IPG 3-10 NL (18 cm)



**Figure 5.17:**  
**Germinating ER**  
50 $\mu$ g  
Standard silver stain

IPG 3-10 NL (18 cm)

181

12% SDS-PAGE (20 cm)

220 kDa

97.4

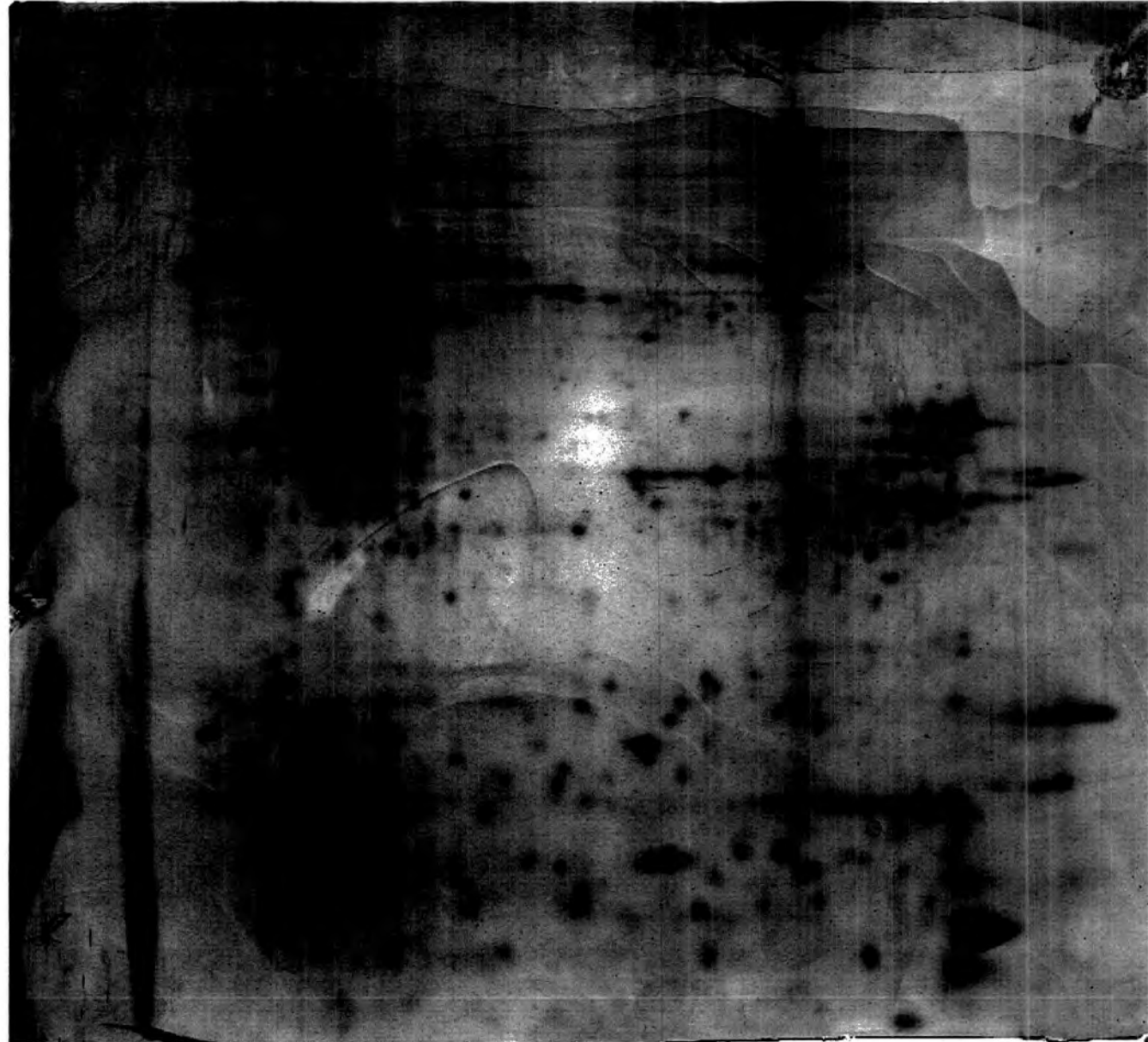
66

46

30

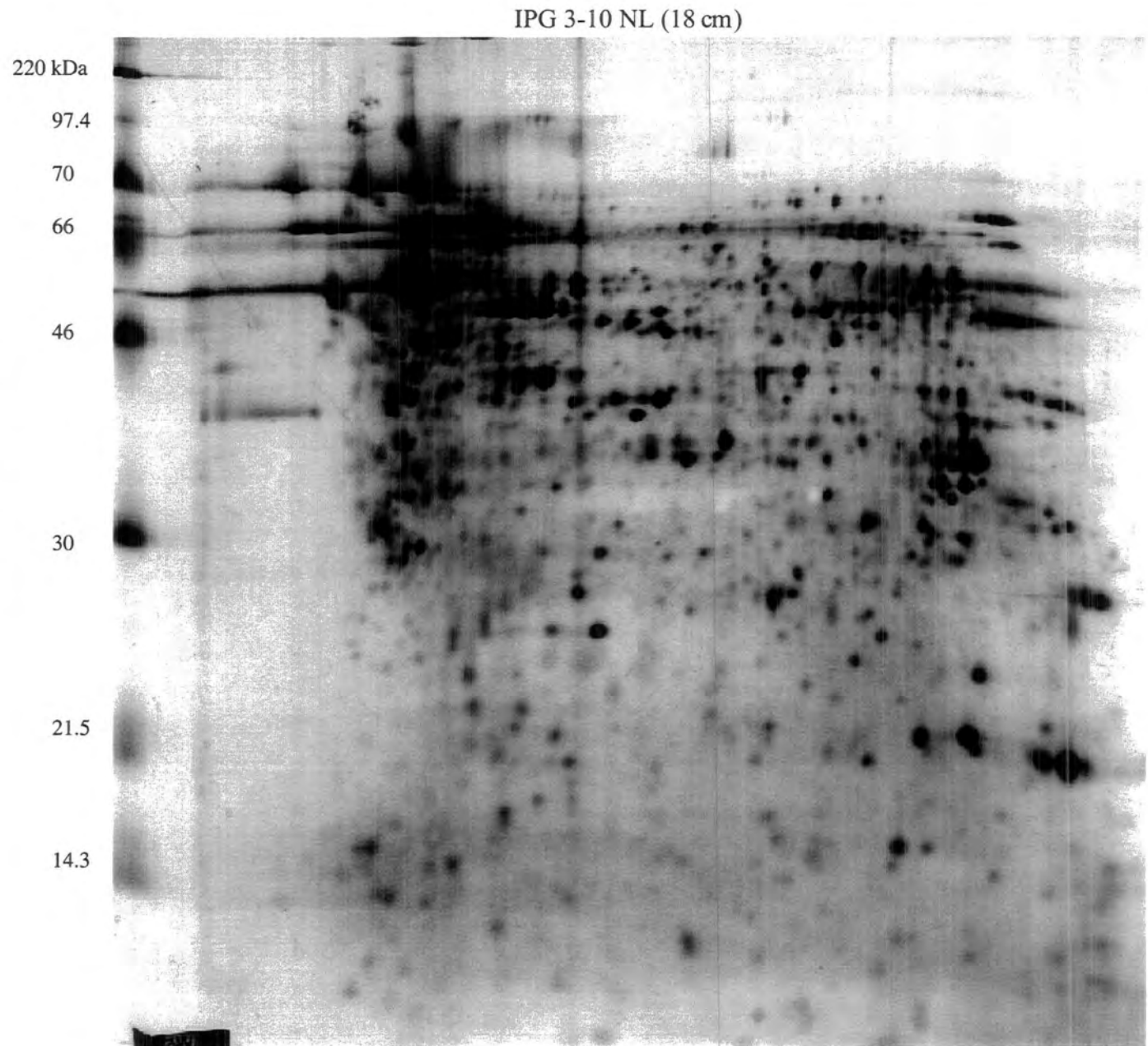
21.5

14.3



**Figure 5.18:**  
**Germinating ER**  
200 $\mu$ g  
PlusOne silver

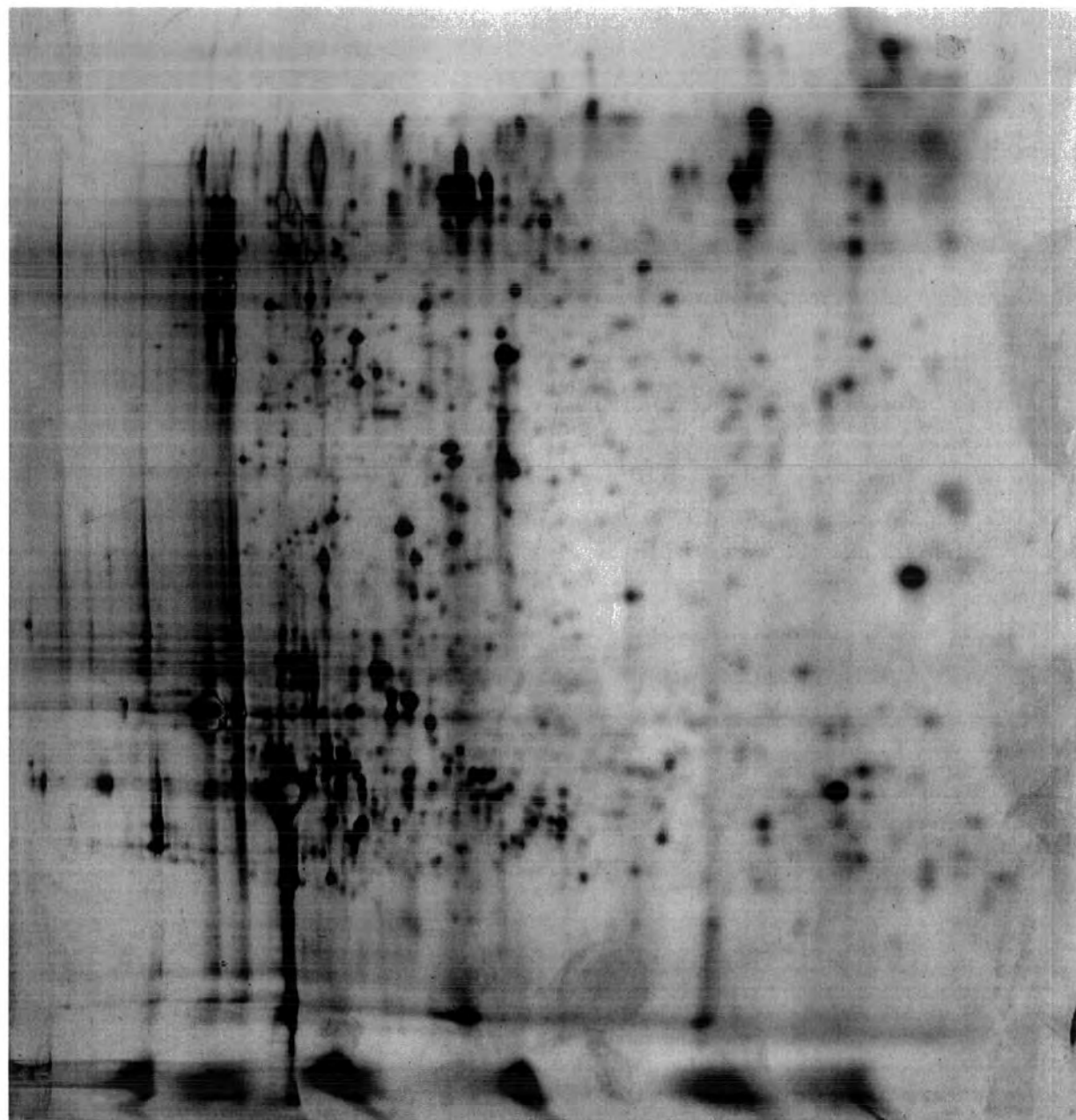
12% SDS-PAGE (20 cm)



**Figure 5.19:**  
Developing ER  
50µg  
Standard silver stain



IPG 3-10 NL (18 cm)



97.4 kDa

66

46

30

21.5

14.3

12% SDS-PAGE (20 cm)

**Figure 5.20:**  
Developing ER  
200µg  
PlusOne silver

lysis/rehydration solutions. Gel profiles were compared using transparent print-outs, as reported in chapter 4, to aid in the labelling of stage-specific spots. The large-scale gels were not compared in any detail with conventional small-scale profiles because of gel compositional differences (non-linear IEF/homogeneous SDS-PAGE versus linear IEF/gradient SDS-PAGE, respectively).

Comparison between analytical and preparative large-scale profiles of the same sample revealed a small but significant number of differences in the exact spots seen, i.e. both systems exhibited exclusive spots (compare figures 5.17 and 5.18, figures 5.19 and 5.20). These differences were attributed to the four-fold change in protein loading and compositional differences in staining materials. It was also possible that TCA/acetone precipitation (50 µg loadings only) resulted in certain protein losses.

The larger format system proved highly effective for the resolution of ER proteins giving profiles of up to more than 200 distinguishable spots. Proteins were separated along the full length of the IEF axis, further confirming the suitability of broad pH gradients (3 – 10) for focussing of these samples. The low levels of aggregation occurring in the smaller format system were no longer evident. This was attributed to the fact that sample application was incorporated into the gel rehydration process. This avoids the severe concentration and pH shock of proteins induced during the use of sample cups.

The aims of mass spectroscopic analysis from preparative gels were to generate novel peptide mass fingerprint data for ER proteins from readily available germinating

endosperm, and to select mainly stage-specific spots. It was also intended to study several common components likely to include PDI and calreticulin as positive controls. Other common predominant spots were included to minimise interference of contaminating peaks. Contamination was expected to be a more significant issue than with Coomassie-stained gels, due to the generally lower levels of individual proteins.

### **5.5.1 MALDI analysis of 2-D protein spots and database analysis**

Nineteen protein spots were excised from the gel shown in Fig. 5.18 (annotated in figure 5.21), plus a blank sample for parallel processing. Tryptic peptides were gel-extracted in acetonitrile and cleaned using ZipTips (Millipore). Samples were dried down and reconstituted in 3  $\mu$ l TFA.

Table 5.3 includes all of the mass data obtained from the processed 2-D samples. Corresponding spectra are shown (figures 5.22-5.34). The lack of processable spectra from the majority of samples necessitated alterations to machine parameters. Most effective was the application of voltage along the guide wire (0.005%) which was at zero for previous characterization.

Increased laser intensities were also examined until optimum results were obtained. However, a relatively high setting also had the effect of raising background noise peaks at the low molecular weight end arising from matrix ionization. Peaks appearing at higher molecular weight were therefore more readily decipherable.

**Figure 5.21: Annotated large scale 2-D gel profile of germinating ER.** 200  $\mu$ g total ER protein was incorporated into IPG rehydration solution (see figure 5.19). 19 spots were selected for tryptic digestion and characterized by MALDI mass spectrometry. A blank sample was also processed as a control. Data was edited as shown in table 5.3. Gel was silver stained (PlusOne, Pharmacia). Spots 1-12: germinating-specific; spots 13-19: major common components; spots 13-15 were clustered in the 50 kDa region and therefore likely to include PDI and calreticulin.

IPG 3-10 NL (18 cm)

981

12% SDS-PAGE (20 cm)

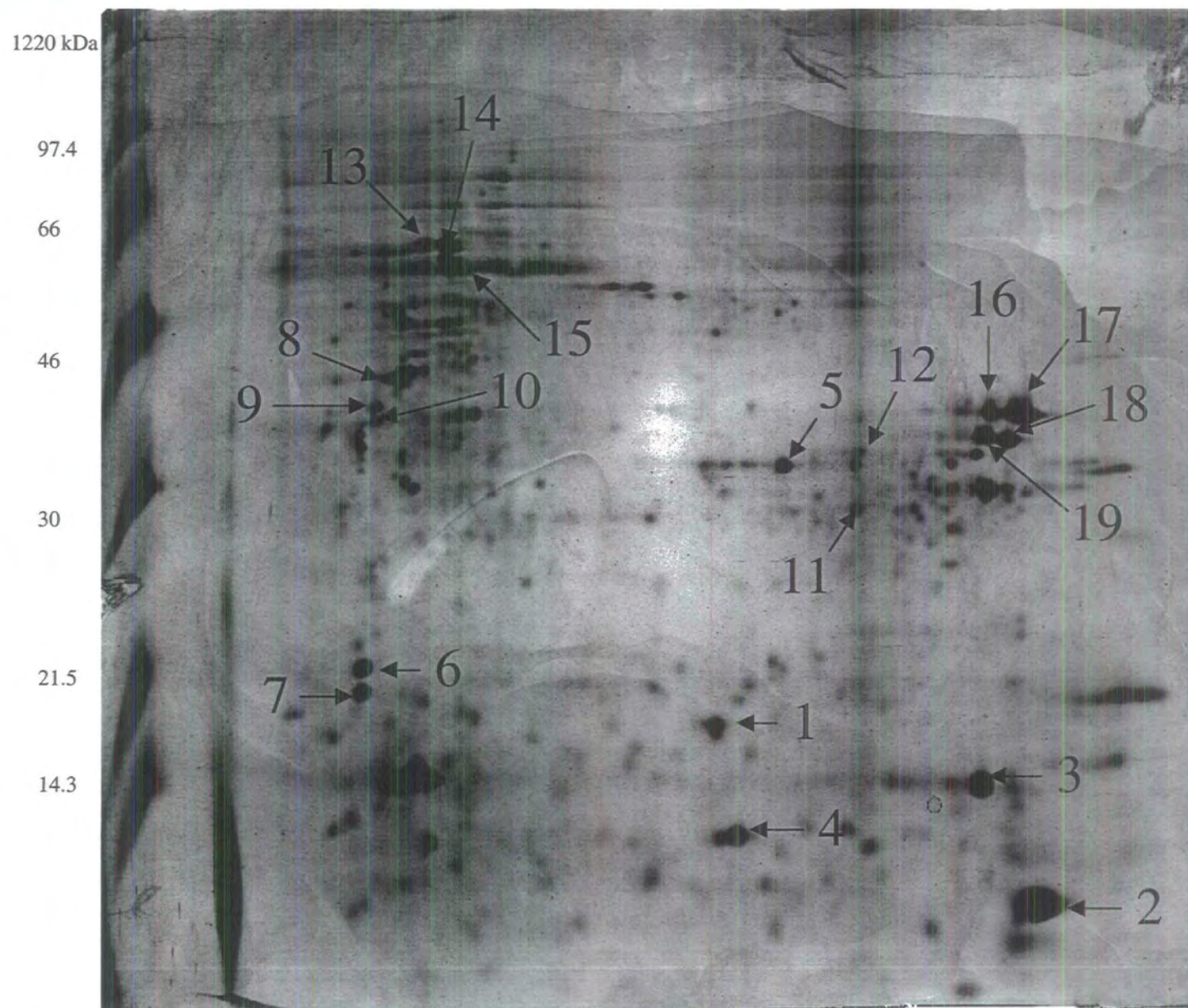


Figure 5.21:  
Germinating ER  
targets for MALDI  
ms  
200µg

**Figures 5.22-5.34: MALDI-TOF ms spectra from 2-D separated ER proteins.**

Tryptic digests of 2-D separated proteins were loaded onto a sample plate with matrix ( $\alpha$ -cyano 4-hydroxy cinnamic acid) following acetonitrile extraction/concentration and Zip Tip (Millipore) clean-up. A guide wire voltage of 0.005% was applied and a laser intensity range of 2170-2330 was used. The mass spectrometer was run in reflector mode, mass values are monoisotopic.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\2D11.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\2D11.MS  
Comment:

Method: BSLOWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.005 %  
Delay: 200 ON  
Sample: 24

Laser: 2170  
Scans Averaged: 110  
Pressure: 1.57e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected 11/23/99 9:57 AM

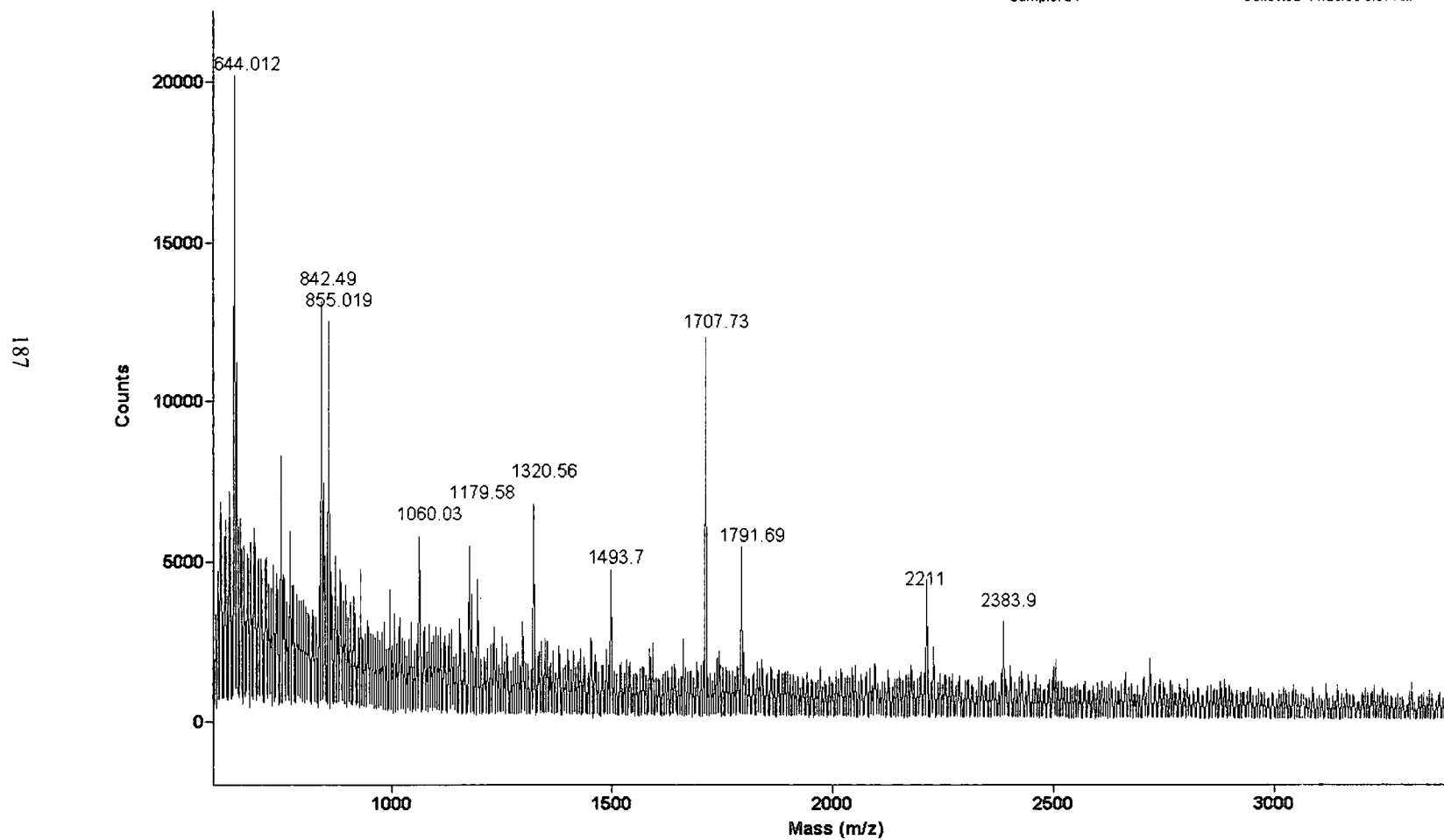


Figure 5.22: MALDI ms spectrum from germinating ER 2-D gel, spot 1.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\2D12.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\2D12.MS  
Comment:

Method: BSLWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.005 %  
Delay: 200 ON  
Sample: 25

Laser: 2170  
Scans Averaged: 233  
Pressure: 1.59e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 11/23/99 9:59 AM

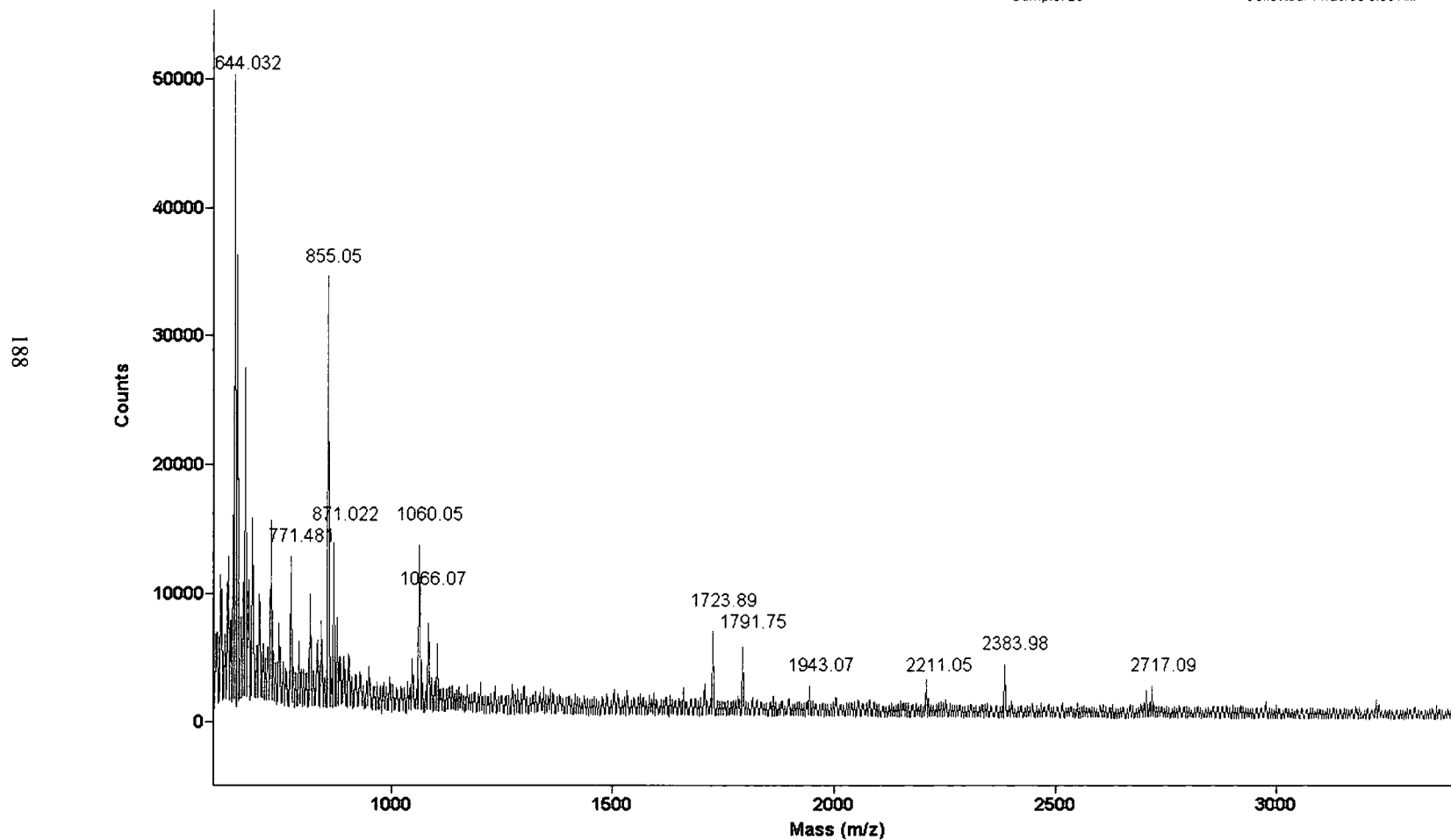


Figure 5.23: MALDI ms spectrum from germinating ER 2-D gel, spot 2.



# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\2D13.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\2D13.MS  
Comment:

Method: BSLOWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.005 %  
Delay: 200 ON  
Sample: 26

Laser: 2300  
Scans Averaged: 121  
Pressure: 1.56e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 11/23/99 10:03 AM

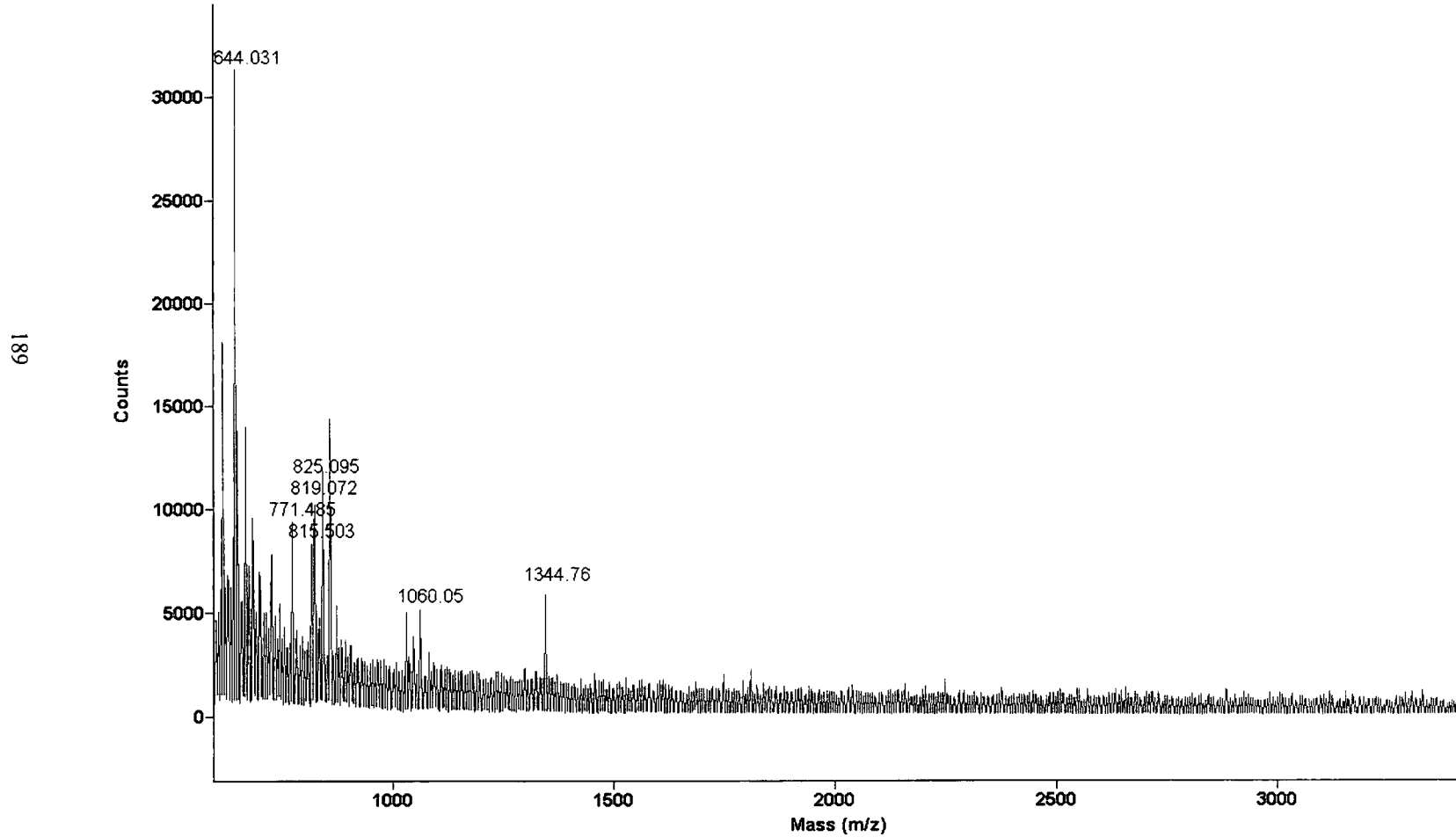


Figure 5.24: MALDI ms spectrum from germinating ER 2-D gel, spot 3.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\2D15.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\2D15.MS  
Comment:

Method: BSLOWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.005 %  
Delay: 200 ON  
Sample: 28

Laser: 2220  
Scans Averaged: 80  
Pressure: 1.56e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 11/23/99 10:07 AM

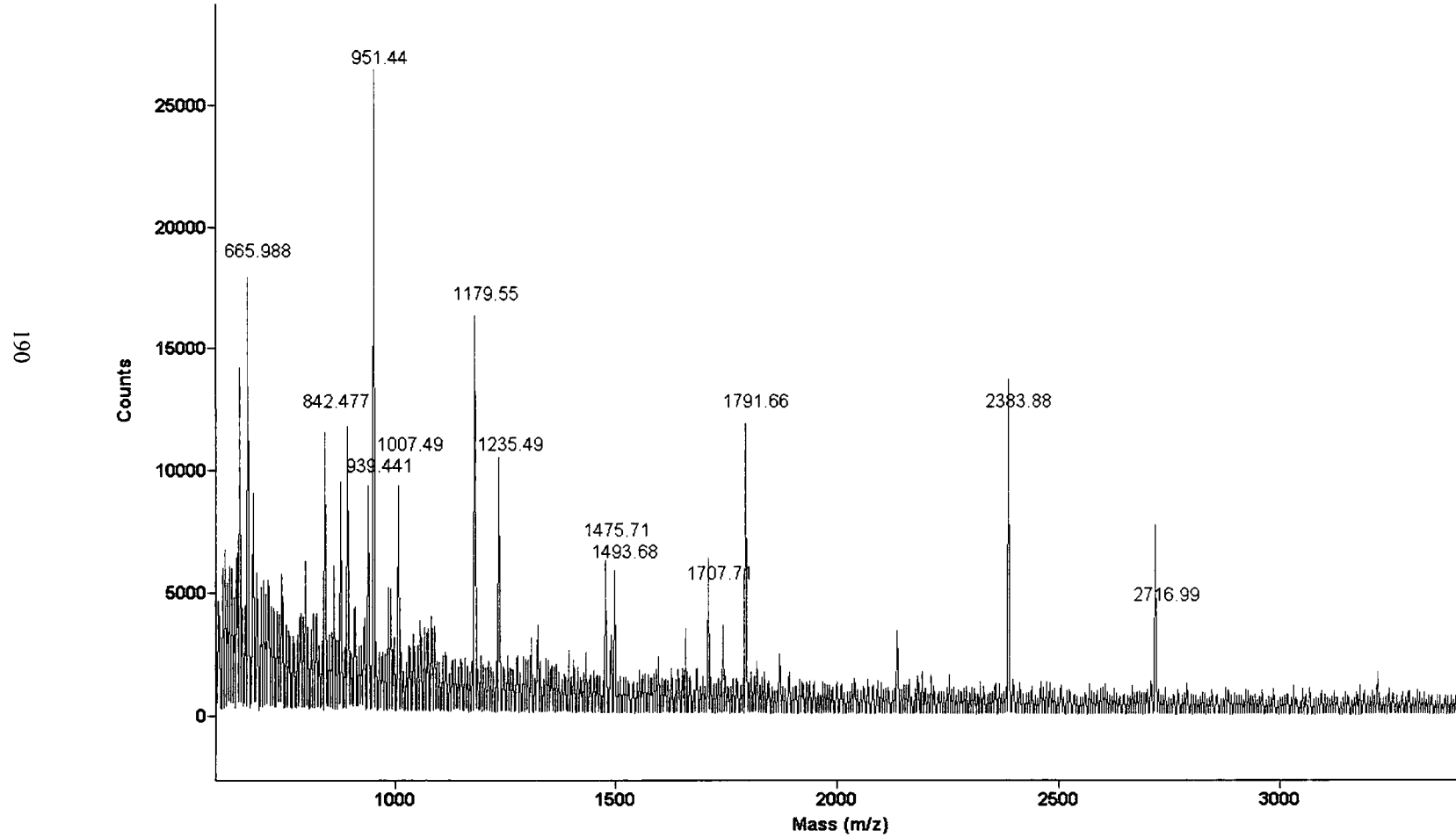


Figure 5.25: MALDI ms spectrum from germinating ER 2-D gel, spot 5.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\2D16.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\2D16.MS  
Comment:

Method: BSLOWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.005 %  
Delay: 200 ON  
Sample: 29

Laser: 2300  
Scans Averaged: 181  
Pressure: 1.57e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 11/23/99 10:11 AM

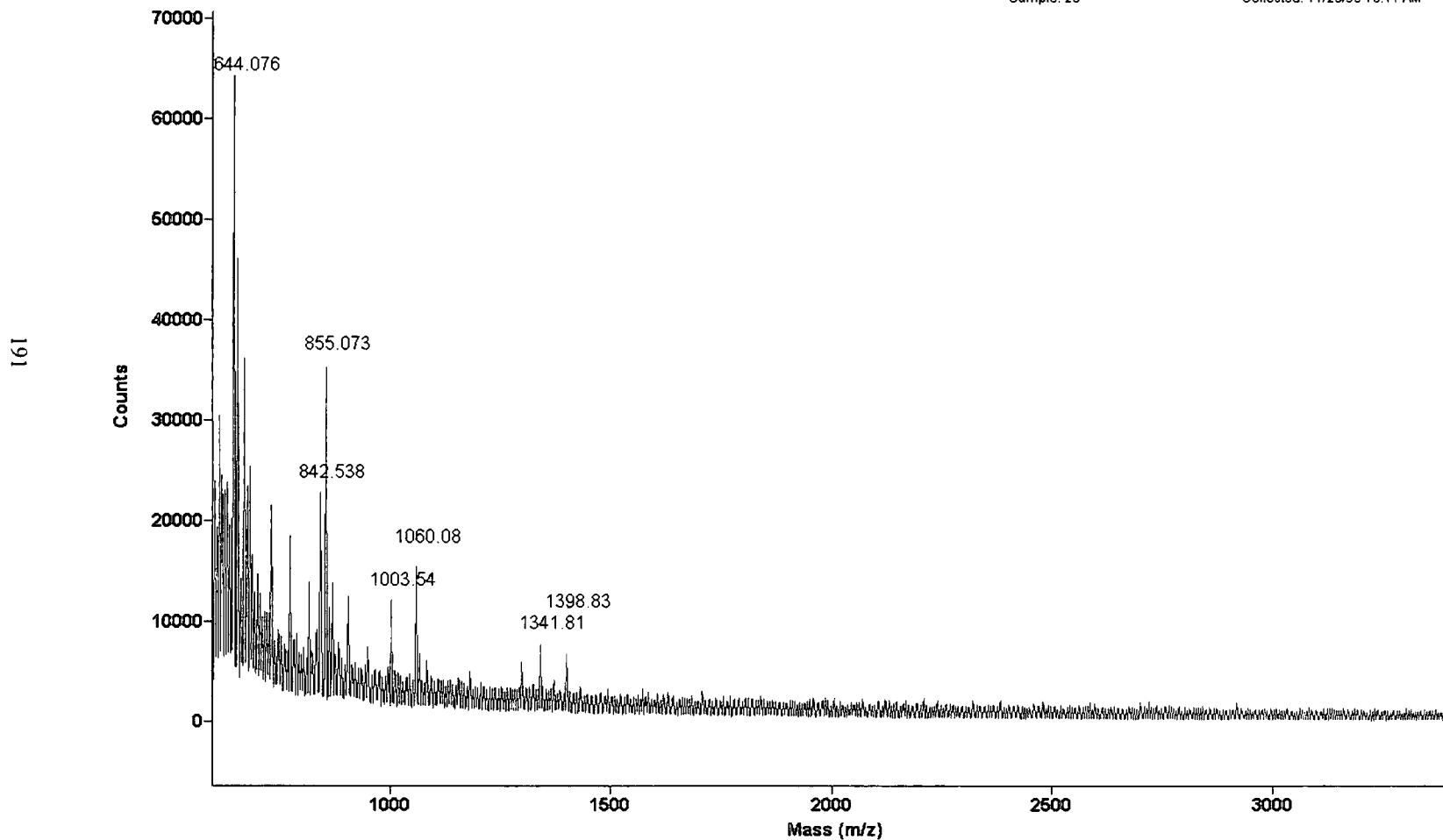


Figure 5.26: MALDI ms spectrum from germinating ER 2-D gel, spot 6.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\2D19.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\2D19.MS  
Comment:

Method: BSLOWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.005 %  
Delay: 200 ON  
Sample: 34

Laser: 2330  
Scans Averaged: 196  
Pressure: 1.55e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected 11/23/99 10:21 AM

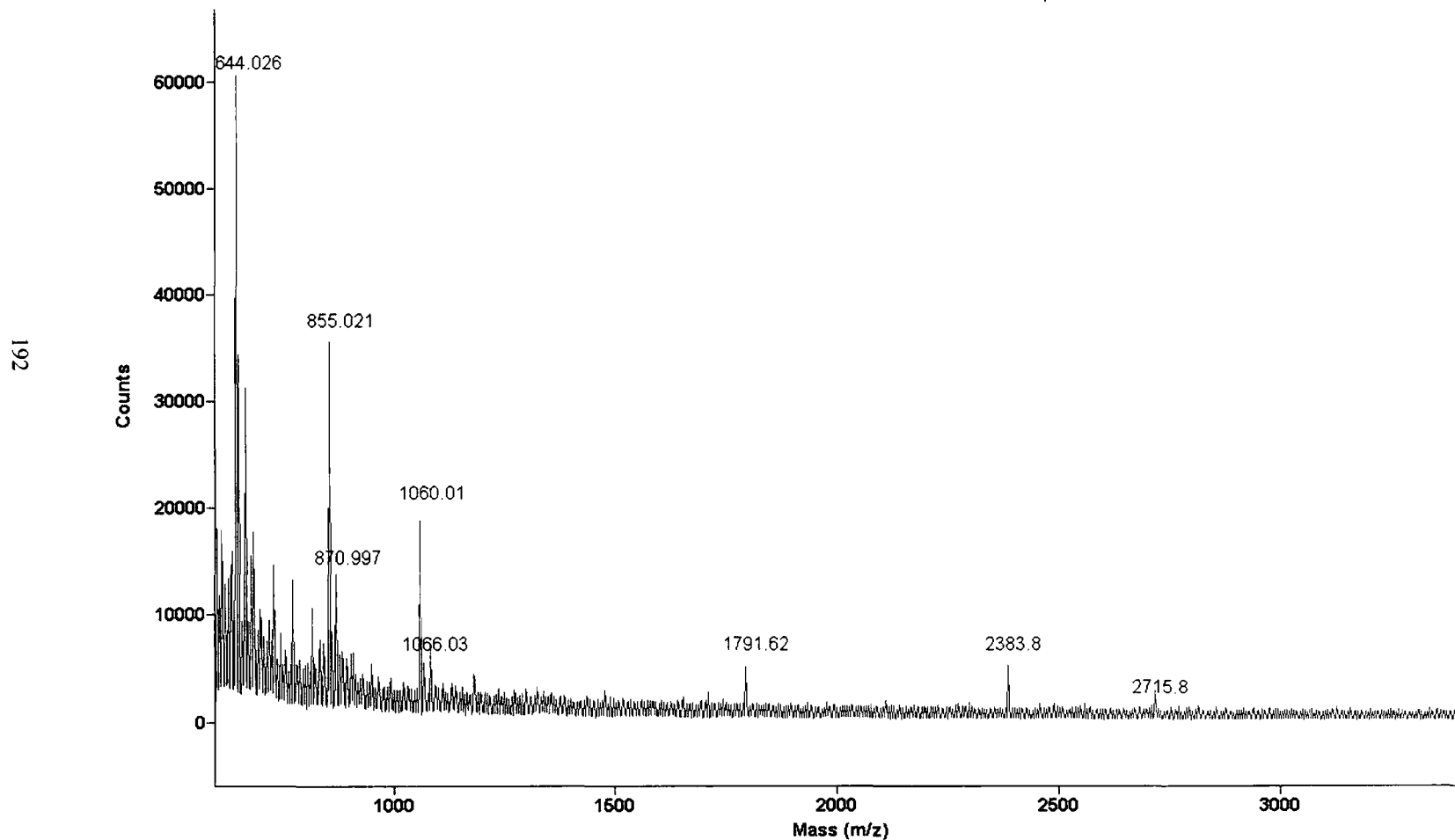


Figure 5.27: MALDI ms spectrum from germinating ER 2-D gel, spot 9.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\2D20.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\2D20.MS  
Comment:

Method: BSLOWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.005 %  
Delay: 200 ON  
Sample: 35

Laser: 2330  
Scans Averaged: 202  
Pressure: 1.56e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 11/23/99 10:25 AM

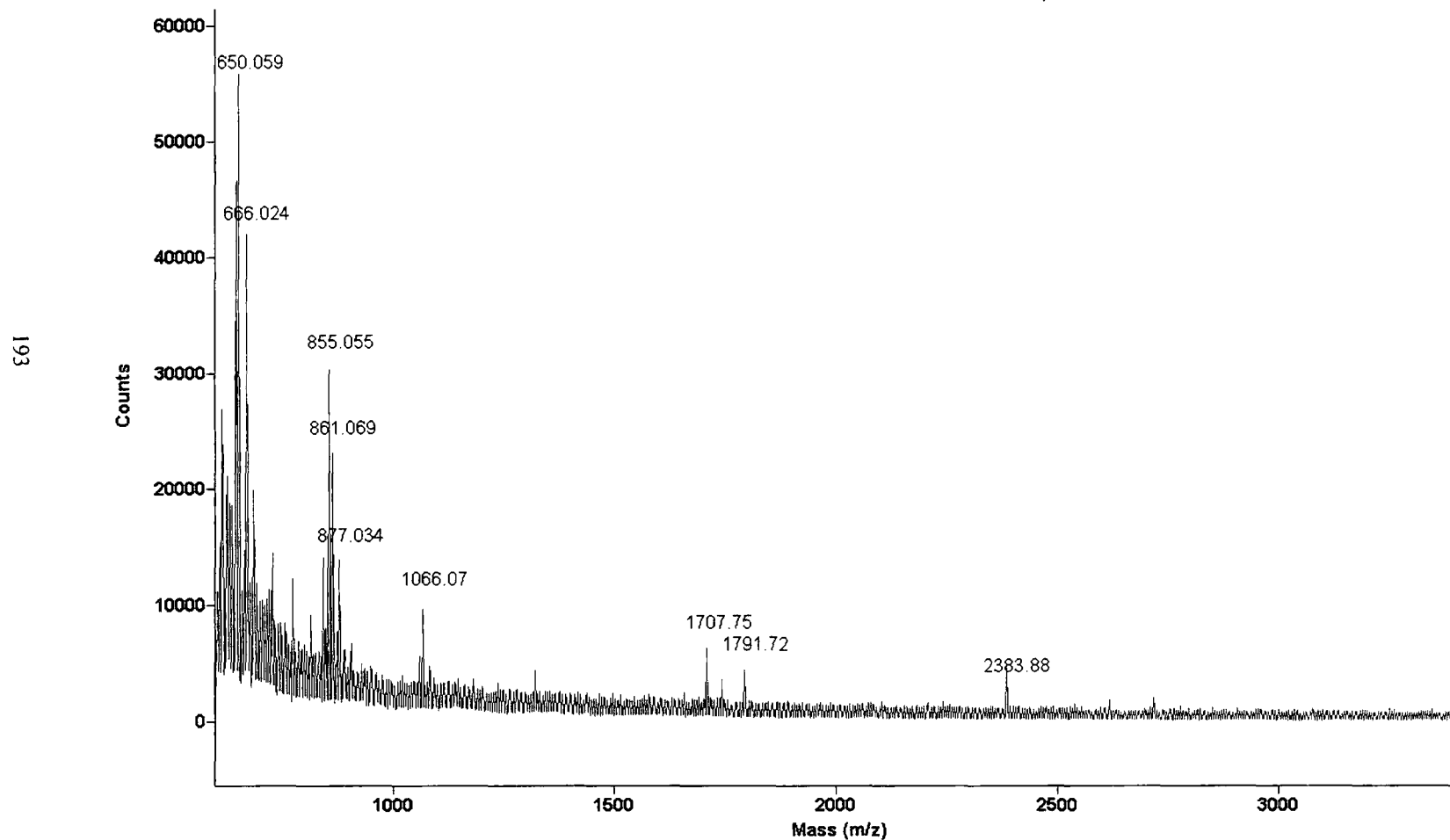


Figure 5.28: MALDI ms spectrum from germinating ER 2-D gel, spot 10.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\2D21.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\2D21.MS  
Comment:

Method: BSLWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.005 %  
Delay: 200 ON  
Sample: 36

Laser: 2330  
Scans Averaged: 169  
Pressure: 1.56e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 11/23/99 10:27 AM

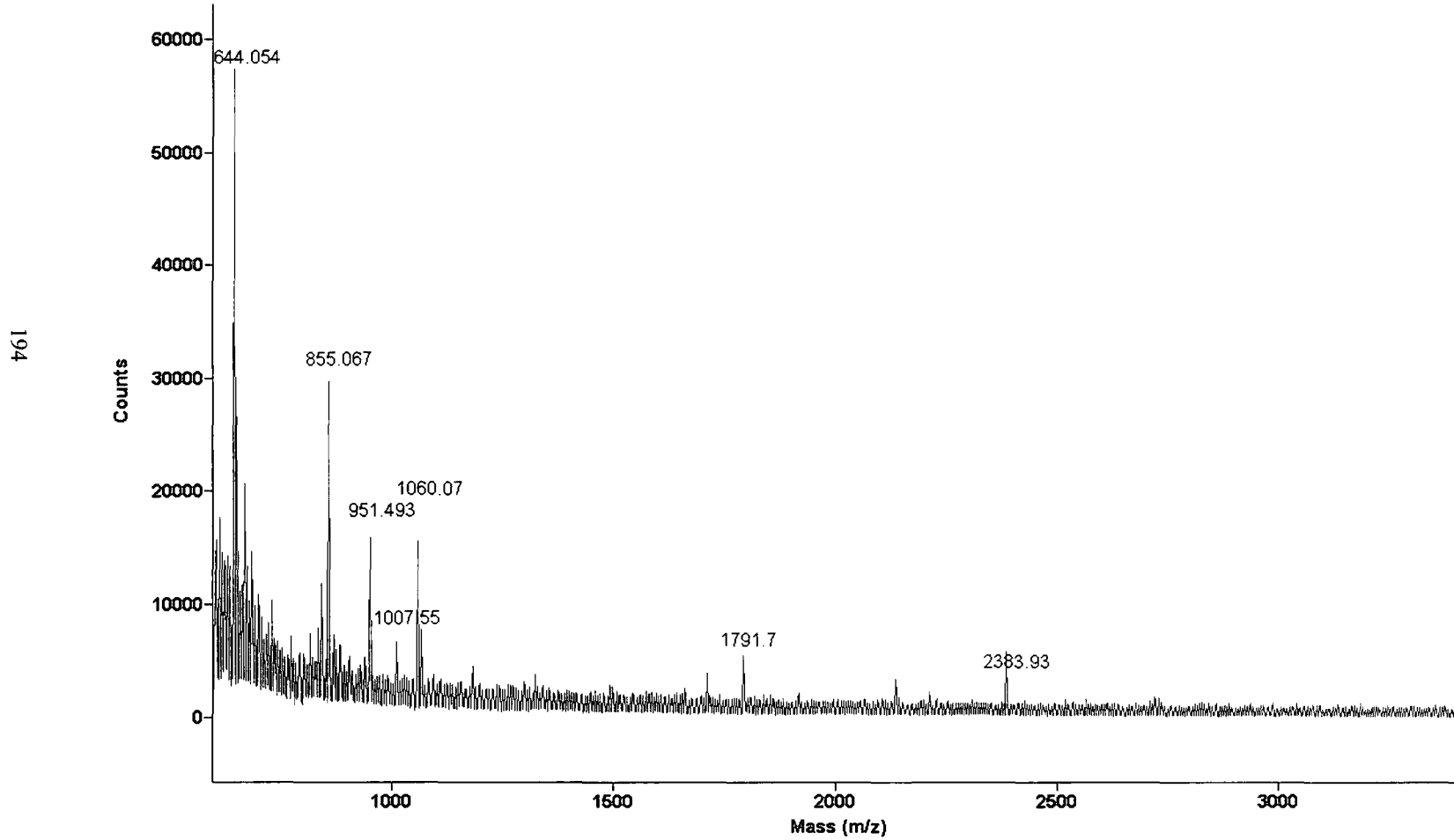


Figure 5.29: MALDI ms spectrum from germinating ER 2-D gel, spot 11.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\2D23.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\2D23.MS  
Comment:

Method: BSLOWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.005 %  
Delay: 200 ON  
Sample: 38

Laser: 2300  
Scans Averaged: 98  
Pressure: 1.49e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 11/23/99 11:31 AM

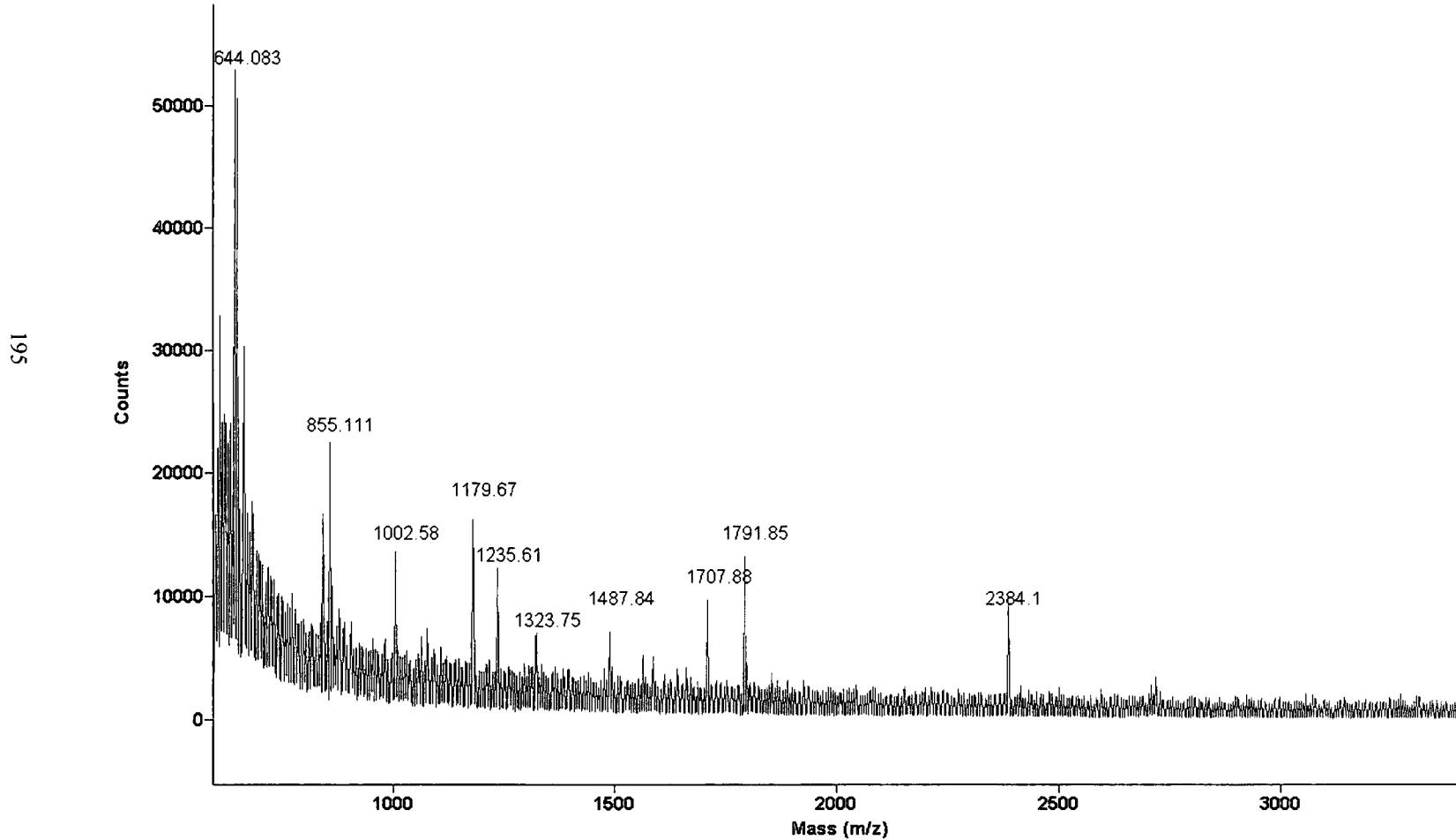


Figure 5.30: MALDI ms spectrum from germinating ER 2-D gel, spot 13.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\2D24.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\2D24.MS  
Comment:

Method: BSLOWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.005 %  
Delay: 200 ON  
Sample: 39

Laser: 2300  
Scans Averaged: 90  
Pressure: 1.51e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 392.2 OFF  
Negative Ions: OFF  
Collected: 11/23/99 11:33 AM

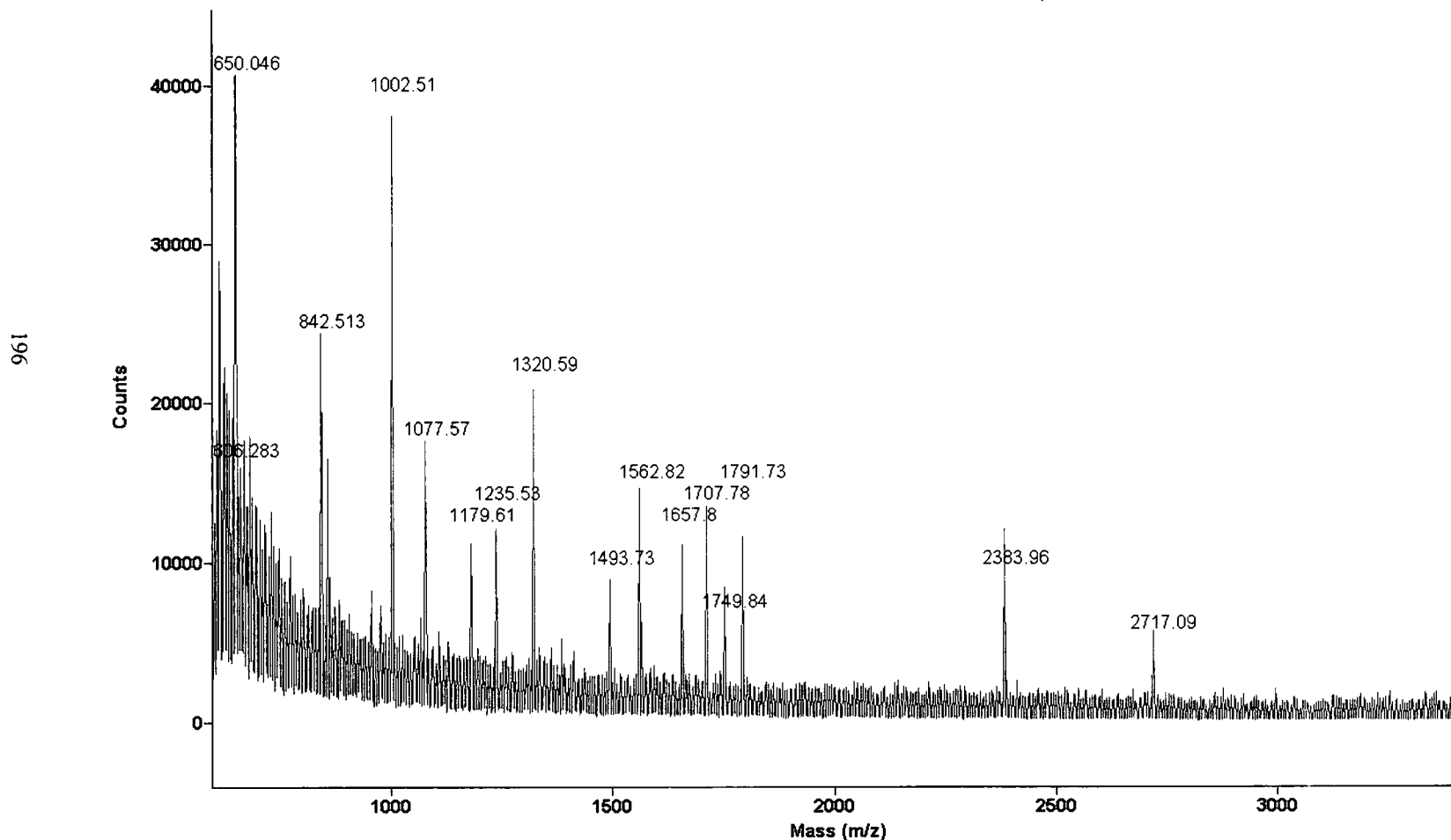


Figure 5.31: MALDI ms spectrum from germinating ER 2-D gel, spot 14.



# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\2D27.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\2D27.MS  
Comment:

Method: BSLOWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.005 %  
Delay: 200 ON  
Sample: 44

Laser: 2300  
Scans Averaged: 152  
Pressure: 1.50e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 11/23/99 11:46 AM

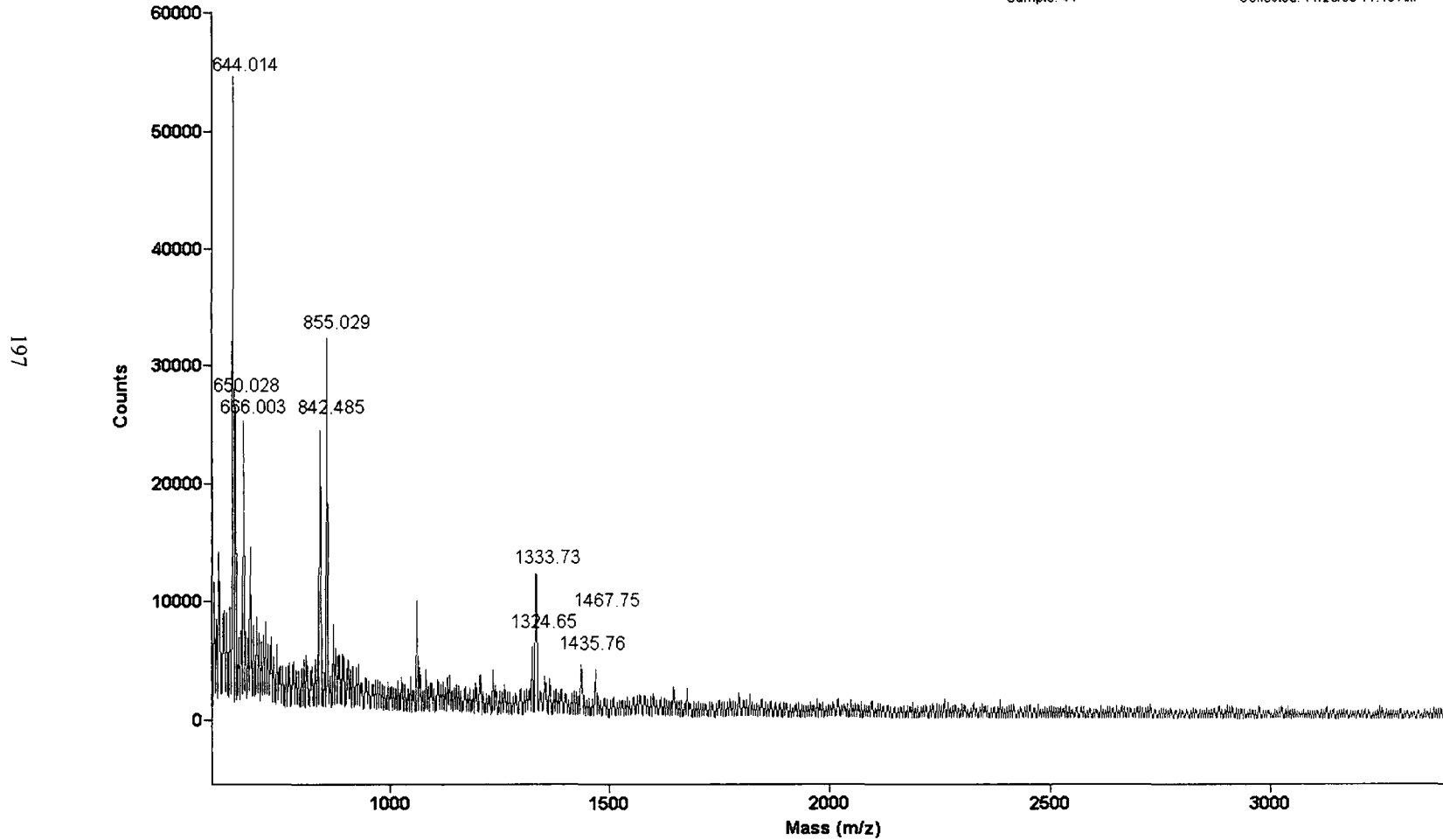


Figure 5.32: MALDI ms spectrum from germinating ER 2-D gel, spot 17.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\2D28.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\2D28.MS  
Comment:

Method: BSLOWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.005 %  
Delay: 200 ON  
Sample: 45

Laser: 2300  
Scans Averaged: 113  
Pressure: 1.52e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 11/23/99 11:49 AM

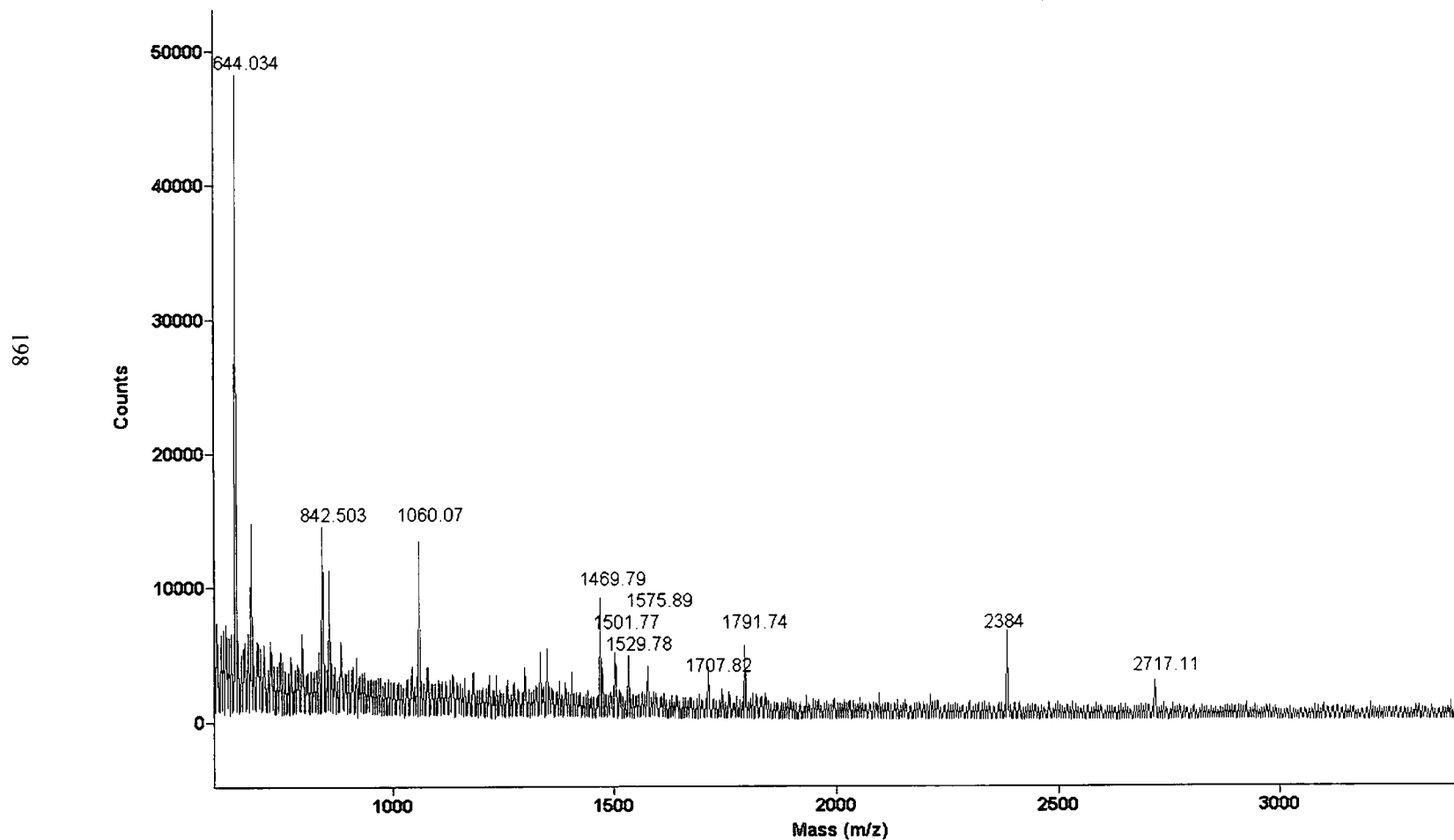


Figure 5.33: MALDI ms spectrum from germinating ER 2-D gel, spot 18.

# PerSeptive Biosystems

Original Filename: c:\voyager\data\bill\2D30.ms  
This File # 1 : C:\VOYAGER\DATA\BILL\2D30.MS  
Comment:

Method: BSLOWRDE  
Mode: Reflector  
Accelerating Voltage: 20000  
Grid Voltage: 74.500 %  
Guide Wire Voltage: 0.005 %  
Delay: 200 ON  
Sample: 47

Laser: 2300  
Scans Averaged: 147  
Pressure: 1.54e-07  
Low Mass Gate: 600.0  
Timed Ion Selector: 382.2 OFF  
Negative Ions: OFF  
Collected: 11/23/99 11:52 AM

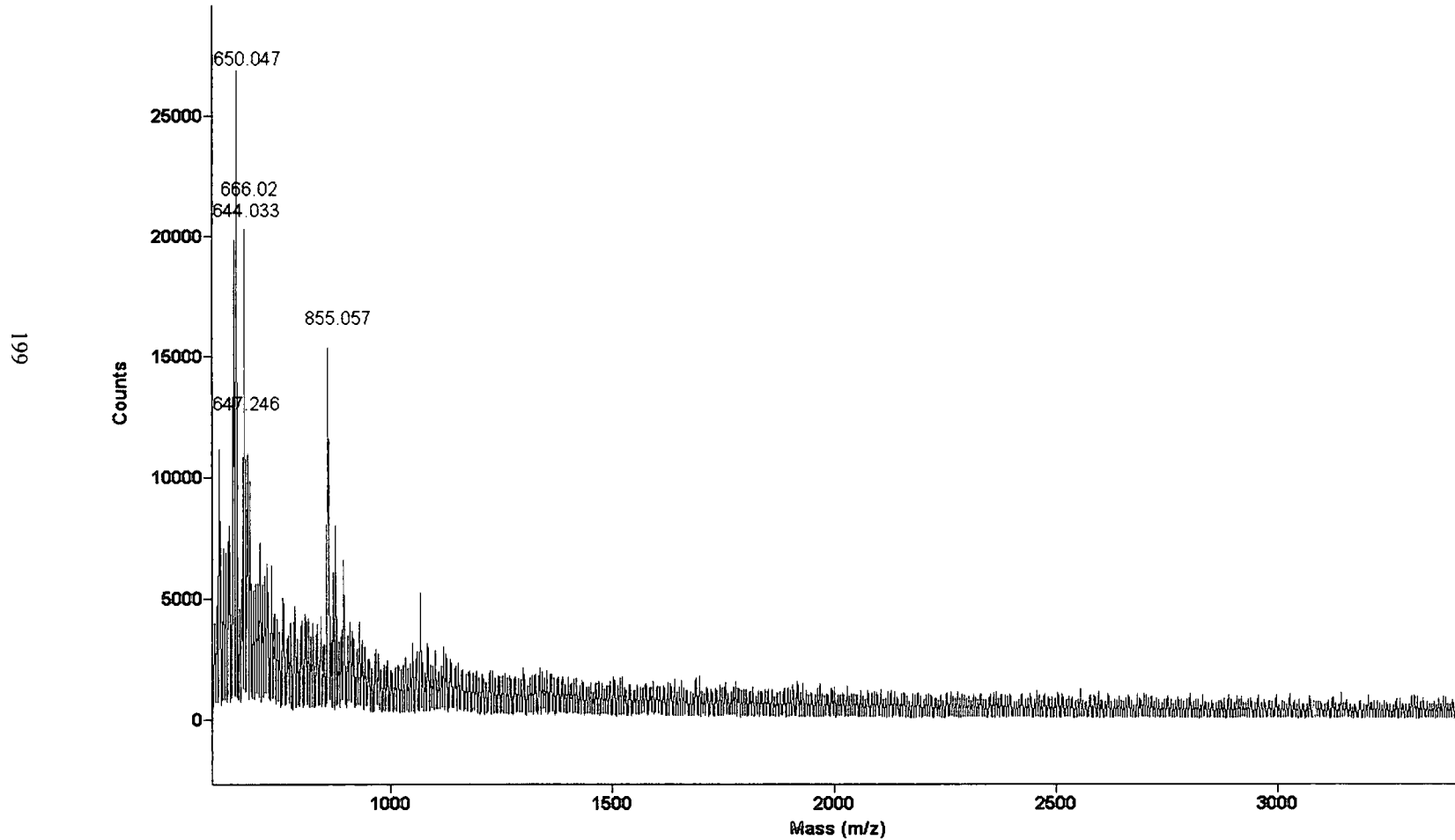


Figure 5.34: MALDI ms spectrum from germinating ER 2-D gel, blank spot .

**Table 5.3: MALDI-TOF ms data from 2-D gel spots.** Silver-stained protein spots were excised from a preparative large format 2-D gel (figure 5.19). Data required extensive editing due to a prevalence of contaminating and common peaks. Edited values include peaks likely to derive from trypsin autolysis and keratins. Remaining data was read as sample-derived. Corresponding spectra are presented (in order, for samples 1,2,3,5,6,9,10,11,13,14,17,18,blank) in figures 5.24-5.36.

| <b>1</b> | <b>2</b> | <b>3</b> | <b>4</b> | <b>5</b> | <b>6</b> | <b>7</b> | <b>8</b> | <b>9</b> | <b>10</b> |
|----------|----------|----------|----------|----------|----------|----------|----------|----------|-----------|
| 1320.56  | 871.022  | 815.503  |          | 939.441  | 1003.54  |          |          | 870.997  | 861.069   |
| 1493.7   | 1723.89  | 819.072  |          | 951.44   | 1341.81  |          |          |          | 877.034   |
| 2211     | 1943.07  | 825.095  |          | 1007.49  | 1398.83  |          |          |          |           |
|          |          | 1344.76  |          | 1475.71  |          |          |          |          |           |
|          |          |          |          | 1493.68  |          |          |          |          |           |

| <b>11</b> | <b>12</b> | <b>13</b> | <b>14</b> | <b>15</b> | <b>16</b> | <b>17</b> | <b>18</b> | <b>19</b> | <b>Blank</b> |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|--------------|
| 951.493   |           | 1323.75   | 1077.57   |           |           | 1324.65   | 1469.79   |           |              |
| 1007.55   |           | 1487.84   | 1320.59   |           |           | 1333.73   | 1501.77   |           |              |
|           |           |           | 1493.73   |           |           | 1435.76   | 1529.78   |           |              |
|           |           |           | 1562.82   |           |           | 1467.75   | 1575.89   |           |              |
|           |           |           | 1657.8    |           |           |           |           |           |              |
|           |           |           | 1749.84   |           |           |           |           |           |              |

Many of the samples yielded very few detectable ions, and no data was returned from sample 7. Cross-contamination appeared to be prevalent, and following standard measures of editing, several more samples were left with no representative spectra (table 5.3). Only 8 of the samples following editing gave 3 or more potentially useful peptide masses. The relatively small number and size of many of these peaks in many of the spectra returned suggested inefficient tryptic digestion and/or low levels of starting materials. This made complication of spectra by contaminant ions a more significant problem. It is expected that a higher loading of total protein (e.g. 400 µg) might alleviate the problem of contamination. In addition efficiency of digestion is likely to play an important part in the generation of an adequate data set. Despite the high incidence of contamination, spectra are included for all samples which rendered data, despite only one database assignment having been made.

All sample data was compared against the OWL protein database with a green plant search restriction, again using the MASCOT engine.

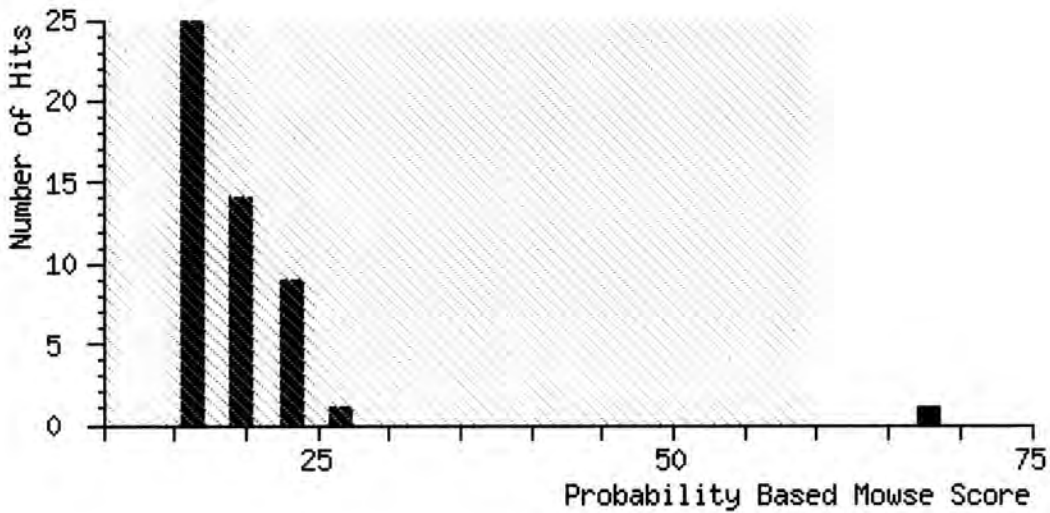
Spot 14 was identified as protein disulphide-isomerase, 4 of the 5 post-editing peptide masses (1002.51, 1077.57, 1562.82, 1749.84) returning the castor database entry with a significant score under a 10 ppm mass tolerance restriction (figure 5.35).

The remaining data searches did not generate any significant protein matches. It is likely that most of the proteins represented by this data have not been sequenced and deposited in public access databases.

**Figure 5.35: Database report returned for four query peptides from germinating ER 2-D gel, spot 14.** Data was edited to remove contaminating and non-sample mass values (see table 5.3). The query peptides (observed column) retrieved castor PDI under a 10 ppm mass accuracy restriction. This represents a significant match with a MOWSE (MOlecular Weight SEarch) score of 67. The search was restricted to green plant (viridiplantae) entries in the NCBI nr protein database. A maximum of 2 missed cleavages was stipulated.

Database : NCBI nr 20000301 (457798 sequences; 140871481 residues)  
 Taxonomy : Viridiplantae (Green Plants) (51513 sequences)  
 Top Score : 67 for gi|2146797, protein disulfide-isomerase (EC 5.3.4.1) - Castor bean

Score is  $-10 \cdot \log(P)$ , where P is the probability that the observed match is a random event.  
 Protein scores greater than 60 are significant ( $p < 0.05$ ).



1. gi|2146797 55606 67 protein disulfide-isomerase (EC 5.3.4.1) - Castor bean

| Observed | Mr(expt) | Mr(calc) | Delta | Start | End | Miss | Peptide          |
|----------|----------|----------|-------|-------|-----|------|------------------|
| 1002.51  | 1001.50  | 1001.50  | 0.01  | 450 - | 457 | 0    | GYPTVYFR         |
| 1077.57  | 1076.56  | 1076.56  | 0.00  | 121 - | 129 | 1    | SIQEYKGPR        |
| 1562.82  | 1561.81  | 1561.82  | -0.01 | 202 - | 216 | 0    | HLPQGESSVTGPVVR  |
| 1749.84  | 1748.83  | 1748.82  | 0.01  | 434 - | 449 | 0    | LDATANDIPSDTFDVR |



In general, spectra derived from silver-stained 2-D gels were characterized by an abundance of persistent non-sample peaks. Such fragments were identified on the basis of their presence in two or more different spectra, and are as follows: 842.5, 855.05, 1060.05, 1066.05, 1179.06, 1235.5, 1320.6, 1493.7, 1707.8, 1791.7, 2383.9, 2717.1. All except one of these peaks were prevalent amongst data from completely different biological samples, run on the same machine, and prepared in the same way (data not included). The two major sources of contamination are keratin, and trypsin itself (see appendix 2a and b, respectively). Indeed, two common tryptic autolysis product peptides, 841.5 and 2210.1, are often used for internal calibration of MALDI instruments.

## 5.6 Discussion

A selection of castor ER proteins has been extracted from 1-D and 2-D gels and processed for analysis by MALDI TOF mass spectrometry. Raw mass data required editing to separate genuine sample-derived peaks from contaminants. Database analysis with results from Coomassie-stained proteins from 1-D gels indicated resolving inadequacy, as revealed by the presence of peptides from two different proteins, PDI and calreticulin, in a single sample. Both of these proteins were easily identified owing to their full sequence entry in the database searched. Similarly, distinction between the two sets of peptides proved simple. However, for the majority of data sets which remain unassigned, elucidation of the complexity of a protein mixture within a single band would prove impossible without a more comprehensive *Ricinus* database. Alternatively, amino acid sequencing of individual peptides could permit homology-based database searches.

Two sets of large-scale 2-D gels were run of total protein from germinating and developing ER. Sample-inclusive IPG rehydration appeared to be effective in reducing in-gel protein precipitation, normally visible as a vertical streak in the second dimension, at the point of sample application. Essentially, total sample protein is applied in a more dilute state across the whole length of the IPG strip rather than at a precise point in a more smaller volume. The results of large-scale IEF and SDS-PAGE was a collection of well-resolved broadly-distributed spot profiles. Although these maps were readily interpreted in terms of protein differences, comparison between analytical and preparative loading gels from identical samples revealed a small number of spots in each, some prominent, that were not seen in the other. The reason for protein absences is presumably related to certain characteristics of the gel system (e.g. loading, detection) and also the method of sample preparation (200  $\mu\text{g}$  incorporated directly, 50  $\mu\text{g}$  via TCA/acetone precipitation).

Non-appearance of spots on the higher preparative loading gel suggests a tendency towards selective aggregation (and consequent loss from the profile) of a subset of proteins at higher concentrations. It is also evident that the less sensitive mass spectroscopy/compatible stain used on the preparative gel is prone to negative staining and high background colouration, factors which may mask certain proteins.

Absence of spots in the 50 $\mu\text{g}$  analytical gel, including some which are prominent in the 200  $\mu\text{g}$  gel, is more difficult to explain and may be the consequence of TCA/acetone

precipitation. Selective degradation during IPG rehydration is unlikely since both sets of gels were processed in the same manner. In fact, the lower loading (4-fold) is almost compensated by the greater sensitivity ( $\sim \times 3$ ) of the conventional stain, so prominent spots on the 200  $\mu\text{g}$  gel should be plainly visible. In summary, an otherwise representative 2-D profile may be incomplete, and caution is necessary when comparing gel images generated from different sample loading, sample preparation, and staining methods. In summary, TCA/acetone precipitation of sample followed by application within the rehydrating solution is recommended.

MALDI data, particularly that originating from the preparative germinating ER 2-D gel, required considerable editing to remove contaminating values. Editing was therefore crucial for effective database searching since raw data often led to low-scoring false positive matches. In addition to data editing, factors such as peptide mass accuracy, restricting search parameters (eg. protein modifications, number of missed cleavages), mass distribution of query fragments, and size of database all affect the likelihood of reliable protein identification. Several known proteins were identified from 1-D SDS-PAGE including the major chaperones PDI and calreticulin (through a combination of theoretical digest comparison and database searching) whose presence in the same Coomassie-stained band further indicated a requirement for the greatly superior resolving power of 2-DE. The 1-D analyses also identified precursors to the seed storage lectins *Ricinus communis* agglutinin (RCA) and ricin. The passage of these precursors through the ER has been documented (Lord, 1985). These findings provide further evidence of the purity of the membrane preparations since contamination from the seed abundant

processed forms of these proteins were not detected. From preparative large-scale 2-DE, the protein disulphide-isomerase spot was unambiguously identified from only four peptides with a 10 ppm mass error margin. This testifies to the power of resolution of the 2-D technique over 1-D SDS-PAGE, and to the accuracy at which the mass spectrometer was operating/calibrated.

The overall scarcity of data from the 2-D samples is likely to have resulted from inefficient tryptic digestion, in combination with (possibly even a consequence of) lower than optimal protein loading, indicated that higher loading might be preferable.

Results have confirmed that, for a sufficient quantity of material, 2-D electrophoresis and MALDI-TOF analysis, together with practiced data editing and database searching, high accuracy peptide mass fingerprints can be generated from individual castor ER proteins. However, as this system is presented at present, routine identification of ER proteins from 2-D spots is severely impeded by a combination of insufficient starting protein quantity in the majority of spots analysed, and also by the absence of significant genomic data from castor in the public access databases. It is expected that improvements in mass accuracy will result from the use of internal calibration standards. However, since over-representation of separate standards in the final spectrum can obscure sample data, internal calibration from trypsin autolysis values is suggested as an alternative. It is important to note that routine generation of comprehensive MALDI data from 2-D spots will require experimental improvement which is most likely to involve increased sample loadings.

A sound single organelle-based system has been established for more comprehensive proteomic mapping. The opportunity now exists to target specific groups of ER components, such as complex lipid biosynthetic enzymes (eg. by radiolabelling with photoactive acyl-CoA analogues). The application of advanced 2-D image analysis software and automated database searching will further add to the power of this approach. While the continued growth of plant databases will undoubtedly aid protein identification it is anticipated that a combination of C-terminal sequencing and electrospray peptide mass spectrometry amino acid sequencing (using triple quadrupole or ion-trap technology) will overcome many of the difficulties associated with protein identification from peptide mass data alone.

## CHAPTER 6

### **Analysis of a putative plant phosphatidic acid phosphatase gene: from EST towards function**

#### **6.1 Introduction**

The aim of any proteomic study is the identification of specific proteins through the combined use of protein analytical methods and bioinformatics. Identification of novel proteins is limited by the absence in the databases of functionally assigned corresponding sequence. In such cases preliminary information may be rapidly obtained on the basis of homology with known proteins from other species although functional assignment in its strictest terms has to involve some form of biochemical analysis. Much of the DNA data present in the database is unassigned with regard to protein function. A particularly useful resource is the database of partially sequenced cDNA data or expressed sequence tags (ESTs) since this data represents actively transcribed mRNA. The identification of a cognate EST is therefore a rapid means of obtaining additional sequence information for a specific protein. However, care must be taken when interpreting EST data since it is derived from single-pass high throughput DNA sequencing and therefore likely to contain errors in the form of base changes and frame shifts.

This chapter presents work undertaken on an Arabidopsis EST clone identified in this laboratory by homology-based database searches as a putative phosphatidic acid phosphatase (PAP). PAP is of special interest in this group as an enzyme involved in the biosynthesis of triacylglycerol. It has been purified and biochemically characterized from avocado (Pearce and Slabas, 1998). The particular importance of

PAP lies in the fact that it occupies a pivotal point in complex lipid biosynthesis and as such represents a possible major control point for the flux of lipid intermediates into either storage triglyceride or membrane phospholipid. Two PAP isoforms have been described from a variety of non-plant sources and these are primarily distinguished on the basis of magnesium ion dependency. PAP1 (magnesium-dependent) isoforms are localised to the endoplasmic reticulum where a role in TAG biosynthesis is performed. PAP2 isoforms are magnesium-independent and are integral proteins of the plasma membrane where they are implicated in lipid signaling pathways.

The aims of this section of work were to obtain and fully sequence the EST clone with a view to performing further sequence analysis, expression analysis and functional biochemistry via bacterial overexpression.

## **6.2 Methods specific to this chapter**

### **6.2.1 Identification of an *Arabidopsis* EST encoding a putative PAP**

A PCR fragment was subcloned from a rape cDNA library in this laboratory following the identification of a rape putative PAP EST (Brown, unpublished). The EST had been identified during a series of gapped BLAST homology searches against mammalian PAP conserved domains (Slabas, unpublished). *The Arabidopsis* EST 1036240 was retrieved from the databases on the basis of sequence similarity to 3' RACE (rapid amplification of cDNA ends) data from the rape library. The *Arabidopsis* entry was 452 bp in length. It was ordered directly from the Arabidopsis Biological Resource Center (ABRC).

### **6.2.2 Ordering of EST 158J20T7**

The EST was obtained free of charge via the internet from the AIMS (Arabidopsis Information Management System) database at the ABRC, Ohio State University (order # 4902, stock # 158J20T7). Subsequent reference to 158J20T7 describes the recombinant plasmid supplied. The database entry resulted from sequencing from a lambda PRL2 cDNA library (Newman *et al*, 1994). The clone was received as bacterial stab in agar.

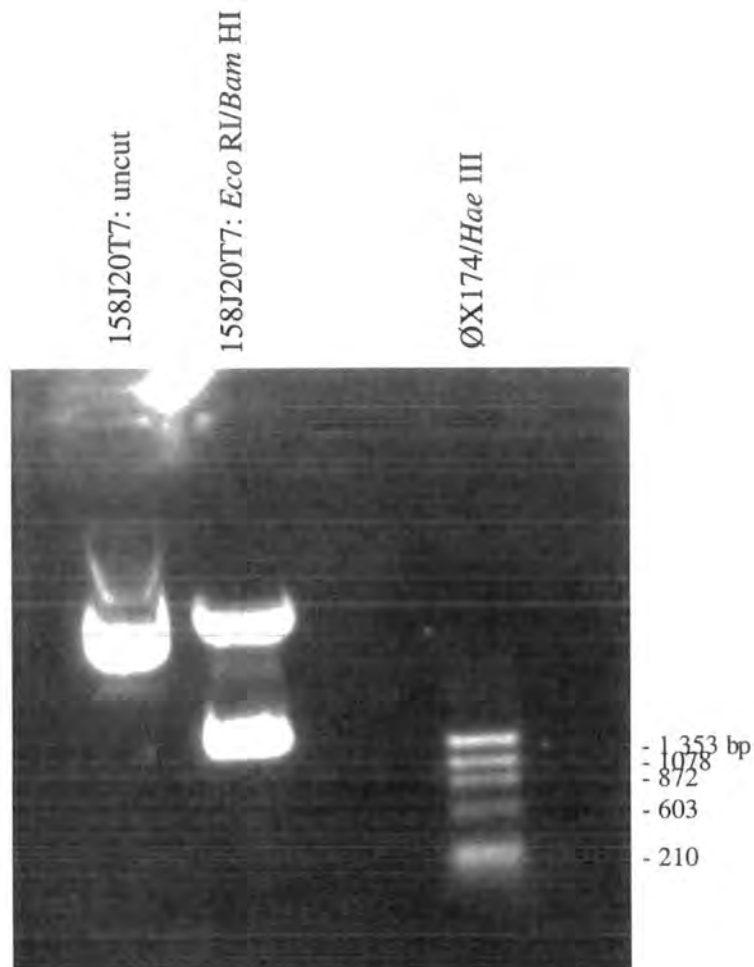
### **6.2.3 Preparation of 158J20T7 and restriction digest analysis**

The plasmid was isolated from an overnight 5 ml culture (LB/ampicillin). Test digests were performed with *Bam* HI and *Eco* R1, sites for which were present in the multiple cloning site either side of the cDNA insert. A single fragment approximately 1300 bp in length was excised (figure 6.1).

### **6.2.4 DNA sequencing of the 158J20T7 insert**

DNA sequencing was performed to obtain complete data for the cDNA insert to establish whether a full cDNA or at least an intact open reading frame (ORF) was present. A total of four different oligonucleotide primers were employed for this exercise (see figure 6.2).





**Figure 6.1: Enzymatic excision of 158J20T7 from plasmid vector pZL 1.**

Recombinant plasmid was prepared (Hybaid Recovery kit) and incubated with *Eco* RI and *Bam* HI. Digested DNA was analysed alongside a sample of undigested plasmid on a 0.8% (w/v) agarose gel. Excision of a fragment approximately 1300 bp in length was demonstrated.

**ARA1:**

5' - GTTTATGACCTGCATCATGC - 3'

**ARA2:**

5' - TCATCAACTCTGGATACACC - 3'

**ZipLox T7 promoter:**

5' - TAATACGACTCACTATAGGG - 3'

**ZipLox SP6 promoter:**

5' - ATTTAGGTGACACTATAG - 3'

**Figure 6.2: Oligonucleotide primers used for sequencing of the 158J20T7 insert.**

ARA1 and ARA2 were designed from internal insert sequence obtained via sequencing from the ZipLox primers (Gibco-BRL).

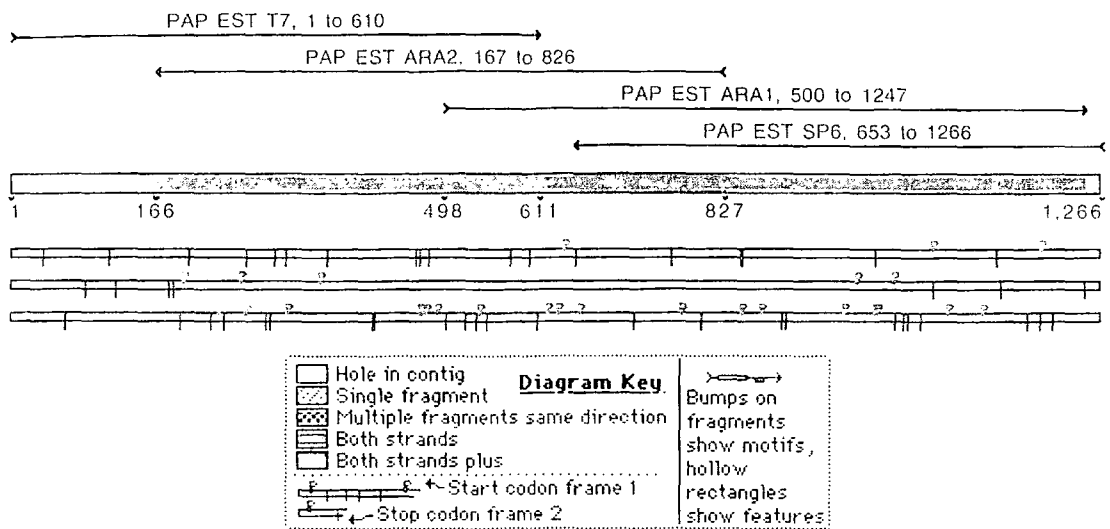
The direction of the insert was surmised from the presence of a poly-A 'tail' (possibly truncated). Clear sequence data was obtained as far as base pair 610 from the 5' end (ZipLox T7 promoter primer) and backwards as far as base pair 653 (SP6 terminator primer). Primers ARA1 and ARA2 were used to obtain overlapping internal data. The Sequencher™ program was used to generate a contiguous sequence from the assembled information (see figure 6.3). The insert was found to be 1266 bp long which is in agreement with the restriction digest results. Back comparison with the database revealed 23 base errors as well as 8 erroneous base insertions in the EST entry. The contiguous sequence was compiled for open reading frame analysis.

### **6.3 Sequence analysis of the 158J20T7 insert**

This second reading frame was found to contain an ORF of 870 bp which conceptually translates into a 290 amino acid protein as shown in figure 6.4. Kyte-Doolittle hydropathy analysis was performed using the DNA Strider program. Six hydrophobic stretches were highlighted which are potentially involved in membrane attachment (figure 6.5). Subcellular localization prediction using the PSORT program suggested plasma membrane and ER as possible sites of localization.

### **6.4 Database searches with the putative PAP amino acid sequence**

The non redundant protein database at NCBI (National Center for Biotechnology Information) was queried with the putative PAP sequence using the BLASTP algorithm.



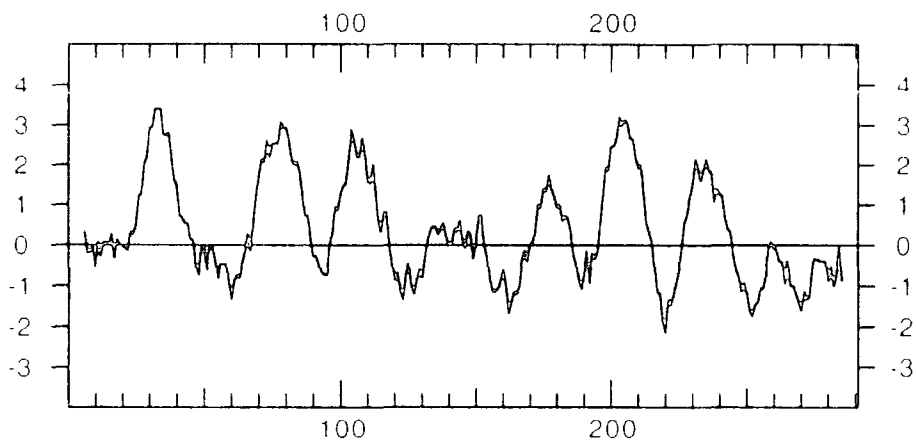
**Figure 6.3: Schematic representation of the 158J20T7 insert sequence.** Edited data obtained from the 158J20T7 plasmid (see figure 6.2 for primers used) was assembled into a 1266 bp contiguous sequence using the sequencer program. The position of an 870 bp open reading frame was discovered in the second reading frame.

```

1/1
CCA CGC GTC CGC AAA AAA CTT TAT CTT TCC TTC CTT TGA AAT CTC CCG GAG AAA AAC TAT AGA GAT TTT CCG TTT CCC GCT TTA ATA CAG
P R V R K K L Y L S F L * N L P E K N Y R D F P F P A L I Q
H A S A K N F I F P S F E I S R R K T I E I F R F P L * Y S
T R P Q K T L S F L P L K S P G E K L * R F S V S R F N T V
91/31 121/41 151/51
TGC CCC AAT TCG CGC GAC ACA TAG AGT GTA GAG ACG CTT TCA CGA GCG TTT CCG ACG TCG GAC TTT CAG CTC ATC ATC TCC ACA TCT TTA
C P N S R D T * S V E T L S R A F P T S D F Q L I I S T S L
A P I R A T H R V * R R F H E R F R R R T F S S S S P H L *
P Q F A R H I E C R D A F T S V S D V G L S A H H L H I F N
181/61 211/71 241/81
ACG GTA AAG ATT AAT CAT GCC TGA AAT TCA TTT GGG TGC TCA TAC AAT AAG ATC CCA TGG AGT AAC AGT CGC GAG GTT CCA CAT GCA TGA
T V K I N H A * N S F G C S Y N K I P W S N S R E V P H A *
R * R L I *
G K D * S C L K F I W V L I Q * D P M E * Q S R G S T C M T
271/91 301/101 331/111
CTG GCT CAT TCT TCT GCT AAT AGT CAT TGA AAT TGT TCT TAA TGT CAT CGA ACC CTT TCA TCG TTT TGT TGG AGA AGA TAT GCT CAC
L A H S S A A N S H * N C S * C H R T L S S F C W R R Y A H
G S F F C C * * S L K L F L M S S N P F I V L L E K I C S L
361/121 391/131 421/141
TGA TCT CAG ATA CCC TCT GCA GGA CAA CAC AAT TCC TTT CTG GGC TGT CCC GTT GAT AGC TGT TGT GCT ACC TTT TGC TGT CAT ITG TGT
* S Q I P S A G Q H N S F L G C P V D S C C A T F C C H L C
I S D T L C R T T Q F L S G L S R * * L L C Y L L L S F V F
451/151 481/161 511/171
TTA CTA CTT CAT TAG AAA TGA TGT TTA TGA CCT GCA TCA TGC AAT ACT AGG GCT TTT GTT CTC TGT ACT TAT AAC CGG TGT CAT AAC CGA
L L L H * K * C L * P A S C N T R A F V L C T Y N R C H N R
T T S L E M M F M T C I M Q Y * G F C S L Y L * P V S * P M
541/181 571/191 601/201
TGC TAT AAA GGA CGC TGT TGG TCG ACC TCG TCC TGA TTT CTT TTG GCG TTG TTT CCC TGA CGG TAT AGG GAT CTT TCA CAA TGT CAC GAA
C Y K G R C W S T S S * F L L A L F P * R Y R D L S Q C H E
L * R T L L V D L V L I S F G V V S L T V * G S F T M S R R
631/211 661/221 691/231
GAA TGT TCT ATG TAC TGG AGC TAA GGA TGT GGT CAA AGA GGG ACA CAA GAG CTT CCC CAG CGG CCA CAC ATC TTG GTC GTT TGC TGG TCT
E C S M Y W S * G C G Q R G T Q E L P V C W S
M F Y V L E L R M W S K R D T R A S P A A T H L G R L L V *
721/241 751/251 781/261
AGG ATT TCT ATC GTT ATA CTT GTC TGG GAA AAT CAG GGT GTT TGA CCA GAG AGG GCA TGT TGC AAA GCT CTG CAT TGT GAT TTT ACC TCT
R I S I V I L V W E N Q G V * P E R A C C K A L H C D F T S
D F Y R Y T C L G K S G C L T R E G M L Q S S A L * F Y L Y
811/271 841/281 871/291
ACT GGT TGC AGC ATT GGT TGG TGT ATC CAG AGT TGA TGA CTA TTG GCA TCA CTG GCA AGA TGT TTT TGG AGG AGC TAT CAT AGG ATT GAC
T G C S I G W C I Q S * * L L A S L A R C F W R S Y H R I D
W L Q H W L V Y P E L M T I G I T G K M F L E E L S * D * L
901/301 931/311 961/321
TGT GGC CAC ATT TTG TTA TCT GCA ATT TTT CCC TCC ATA CGA TCC AGA CGG TTG GGG ACC TCA TGC CTA CTT CCA GAT GCT GGC AGA
C G H I L L S A I F P S S I R S R R L G T S C L L P D A G R
W P H F V I C N F S L L L H T I Q T V G D L M P T S R C W Q T
991/331 1021/341 1051/351
CTC AAG AAA TGA TGT CCA AGA TTC AGC AGG AAT GAA TCA TCT AAG CGT GAG GCA AAC AGA GCT AGA GAG CGT ACG TTG ATG GAG AAG AGA
L K K * C P R F S R N E S S K R E A N R A R E R T L M E K R
Q E M M S K I Q Q E * I I * A * G K Q S * R A Y V D G E E T
1081/361 1111/371 1141/381
CGT CCA TGG AAA TAT CAA GAA GCA ACA CGC GGG ACA CCA CCC GTA TGC TTC AGA ACC CGC TAA GTG AAG TCT TTG TAC TCG TTA TCT ATC
R P W K Y Q E A T R G T P P V C F R T R * V K S L Y S L S I
V H G N I K K Q H A G H H P Y A S E P A K * S L C T R Y L S
S M E I S R S N T R D T T R M L Q N P L S E V F V L V I Y Q
1171/391 1201/401 1231/411
AAT CTT AGG CAT TGC GCA TTG ATA TGT ATT GGC TTA ATC ACA AGG CCC AAT ATT GGT TGG AAC CCA TTC GCT AAA AAA AAA AAA AAA
N L R H C A L I C I G L I T R P N I G W N P F A K K K K K K
I L G I A H * Y V L A * S Q G P I L V G T H S L K K K K K K
S * A L R I D M Y W L N H K A Q Y W L E P I R * K K K K K K
1261/421
AAA A
K
K

```

**Figure 6.4: Three frame conceptual translation of the full 158J20T7 insert DNA sequence.** Translation was performed using Strider 1.3. The second reading frame contains an 870 bp ORF encoding a putative 290 amino acid protein (highlighted)



**Figure 6.5 : Kyte-Doolittle hydropathy plot of the ARAPAP primary Sequence.** The full ARAPAP amino acid sequence was analysed using DNA Strider 1.2. The plot indicayes 6 discrete and continuous hydrophobic Segments (positive values).

The highest-scoring match in the database was a *Saccharomyces cerevisiae* probable membrane protein called YDR284c (accession pir||S70114), characterized as the *DPP1* gene product (Toke *et al*, 1998a). This entry shared 31% (63/199) identity and 49% (100/199) positive similarity with the *Arabidopsis* sequence. Remaining high-scoring database hits were type 2 phosphatidic acid phosphatases from various species including rat, mouse and human. A full amino acid alignment between the yeast and *Arabidopsis* sequences is shown in figure 6.6 which also highlights the position of a conserved novel phosphatase motif (Stukey & Carman, 1997). These proteins differ in length by only 1 amino acid, the yeast protein being 289 long. These findings provide further evidence for the existence in 158J20T7 of a full-length open reading frame encoding phosphatidic acid phosphatase activity in *Arabidopsis*. From here on this protein is referred to as *ARAPAP*.

The observed similarity between *ARAPAP* and the yeast *DPP1* product, together with hydropathy analysis product highlighted the possibility of *ARAPAP* falling into the type 2 PAP (magnesium-independent) isoform category (see table 6.1).

|               | <b>No. amino acids</b> | <b>% identity</b> | <b>% positive similarity</b> | <b>Novel phosphatase motif</b> | <b>Predicted membrane domains</b> |
|---------------|------------------------|-------------------|------------------------------|--------------------------------|-----------------------------------|
| <i>ARAPAP</i> | 290                    | 30                | 52                           | yes                            | 6                                 |
| <i>DPP1</i>   | 289                    |                   |                              | yes                            | 6                                 |

**Table 6.1: Comparison of yeast *DPP1* and *ARAPAP* products.**

CLUSTAL W(1.4) multiple sequence alignment

```

Arabidopsi      MPEIHLGAHTIRSHGVTVARFHMHDWLILLLLIVIEIVLNVIEPFHRFVGEDMLTDLRYP
YeastDPP1      ----MNRVSFIKTPFNIGAKWRLEDVFLLIIMILLNYPVYYQQPFERQFYINDLTISHPY
                *..      * . . * . * . . . . . . . . . . . . . . . . . . . . . . . . .
Arabidopsi      LQDNTIPFWAVPLIAVVLPPFAVICVYYFIRND----VYDLHHAILGLLFSVLITGVITDA
YeastDPP1      ATTERVNNMMLFVYSFVPSLTIIGSILADRRHLIFILYTSLLGLSLAWFSTSFFTNE
                . . . . * . * . * * . . * . . . . . . . . . . . . . . . . . . . .
Arabidopsi      IKDAVGRPRPDFFWRCFP-DGIGIFHNVTKNVLCGTG-AKDVVKEGHKSFPSGHTSWSFAG
YeastDPP1      IKNWIGRLRPDFLDRCQPVEGLPLDTLFTAKDVCTTKNHERLLDGFRTTPSGHSSESFAG
                ** . ** * ** * * * . * . . . * . . . . . . . . . . . . . . . . . . . .
Arabidopsi      LGFLSLYLSGKIRVFDQRGHVAKLCIVILPLLVAALVGVS RVDDYWHHWQDVFGGAIIGL
YeastDPP1      LGYLYFWLCGQLLTESPLMPLWRKMVAFLPLLGAALIALSRTQDYRHHFVDVILGSMLGY
                ** . * * * . * . . . . . . . . . . . . . . . . . . . . . . . . . .
Arabidopsi      TVATFCYLQFFPPPYDPDGWGPWAYFQMLADSRNDVQDSAGMNHLSVRQTELESVR----
YeastDPP1      IMAHFFYRRIFPPIDP-----LPEKPLMDD-SDVTLEEAVTHQRIPDEELHPLSDEGM
                . * * * . * ** * * . * . * * * * * . . . . . . . . . . . . . . .

```

**Figure 6.6: Full alignment between the *Arabidopsis* (ARAPAP) and *Saccharomyces* amino acid sequences using the Clustal program. A conserved novel phosphatase motif (Stukey & Carman, 1997) is present and highlighted in each sequence.**



Subsequent additions to the *Arabidopsis* database were found to include further putative PAP data.

### **6.5 Other plant PAPs**

Continuing database searches using the *ARAPAP* sequence revealed two further *Arabidopsis* entries during the course of this work. Both were contained within genomic BAC clones. It was apparent that the *ARAPAP* protein is derived from one of these clones, AC007591, which covers the whole of chromosome 1. This sequence reveals the presence of 6 introns in the *ARAPAP* gene. The other genomic clone, AC006200, contains an intronless homologue of *ARAPAP* (see alignment, figure 6.7). The identification of these sequences offers the potential to investigate likely promoter sequences controlling the expression of PAPS, predominantly in embryonic tissue.

These findings indicated the possible existence of at least two PAP genes in *Arabidopsis*. Gene copy number was therefore investigated further by Southern blotting.

### **6.6 Southern hybridization analysis using ARAPAP ORF as probe**

The open reading frame of *ARAPAP* was used as a template for a radiolabelled probe to investigate the potential copy number of probe-related gene sequences in *Arabidopsis* and Rape (*Brassica napus*). The aim was to establish the likelihood of the existence of different PAP isoforms in either species.

```

ARAPAP      MPEIHLGAHTIRSHGVTVARFHMHDWLILLLLIVIEIVLNVIEPFHRFVGEDMLTDLRYP
AC006200    MQEIDLSVHTIKSHGGRVASKHKHDWIILVILIAIEIGLNLISPFYRYVGKDMMTDLKYP
            * * * * * . * * * * * * * * * * * . * * * * * * * * * * * . * * * * * . * * * * *
ARAPAP      LQDNTIPFWAVPLIAVVLPFAVICVYYFIRNDVYDLHHAAILGLLFSLITGVITDAIKDA
AC006200    FKDNTVPIWSVPVYAVLLPIIVFVCFYLRKTCVYDLHHSILGLLFAVLITGVITDSIKVA
            . * * * * * . * * * * * * * * * * * . * * * * * . * * * * * . * * * * *
ARAPAP      VGRPRPDFFWRCFPDGIGIFHNVTKNVLTGAKDVVKEGHKSFPSGHTSWSFAGLGFLSL
AC006200    TGRPRPNFYWRCFPDGK-ELYDALGGVVCHGKAAEVKEGHKSFPSGHTSWSFAGLTFLSL
            * * * * * . * * * * * * * * * * * . * * * * * * * * * * * * * * * * * * * * *
ARAPAP      YLSGKIRVFDQRGHVAKLCIVILPLLVAALVGVSRVDDYWHHWQDVFGGAIIGLTVATFC
AC006200    YLSGKIKAFNNEGHVAKLCLVIFPLLAACLVGISRVDDYWHHWQDVFGALIGTLVA AFC
            * * * * * . * * * * * . * * * * * * * * * * * . * * * * * . * * * * * . * * * * *
ARAPAP      YLQFFPPPYDPPDGWGPYAYFQMLAD-----SRN-DVQDSAGMNHLSVRQTELESVR--
AC006200    YRQFYPNPYHEEGWGPYAYFKAAQERGVPVTSSQNGDALRAMSLQMDSTSLENMESGTST
            * * * * * . * * * * * * * * * * * . * * * * * . * * * * * . * * * * *
ARAPAP      ---
AC006200    APR

```

**Figure 6.7: Clustal amino acid alignment between ARAPAP and a putative Arabidopsis PAP from genomic clone AC006200.** The coding sequence for the AC006200 protein is uninterrupted whereas the ARAPAP genomic sequence contains 6 introns. The conserved novel phosphatase motif (Stukey & Carman, 1997) is highlighted.

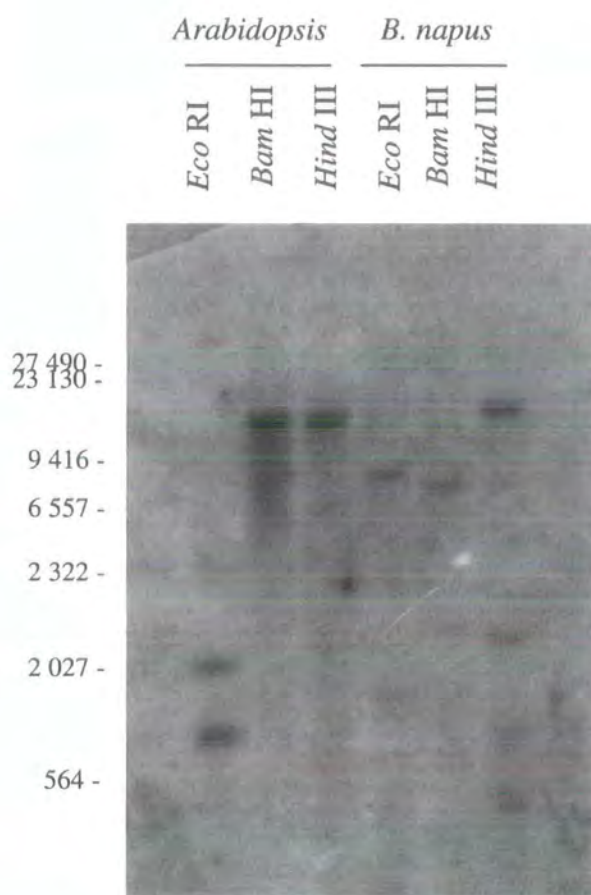
Separate samples of *Arabidopsis* and *Brassica* genomic DNA were digested with *Eco* RI, *Bam* HI and *Hind* III and separated on a 0.5% agarose gel. A probe was prepared by random priming, incorporating  $^{32}\text{P}$  as label. Figure 6.8 shows the autoradiogram resulting from hybridization. Single bands were observed in the *Bam* HI and *Hind* III digest lanes although two bands were seen in the *Eco* RI lane. Therefore under the low stringency conditions used results suggest that potentially two ARAPAP related sequences were detected.

Single band *Eco* RI and *Bam* HI profiles infer single copy existence in *Brassica*. A more complex pattern of at least 4 distinct bands is obtained following *Hind* III digestion. Southern hybridization had shown that the ARAPAP ORF derived probe was capable of detecting a homologous sequence from rape genomic DNA. This probe was subsequently used to investigate patterns of expression of the ARAPAP transcript via Northern blotting.

## **6.7 Expression studies of the ARAPAP transcript and protein**

### **6.7.1 Northern analysis**

The tissue distribution of the ARAPAP transcript was investigated by heterologous Northern hybridization using a [ $^{32}\text{P}$ ]-radiolabelled probe generated from a gel-purified ARAPAP PCR fragment. Initially Northern blots of *Arabidopsis* total RNA were unsuccessful in detecting the ARAPAP transcript. This was attributed to very low transcript levels. Subsequently readily available and high quality poly-A<sup>+</sup> RNA from

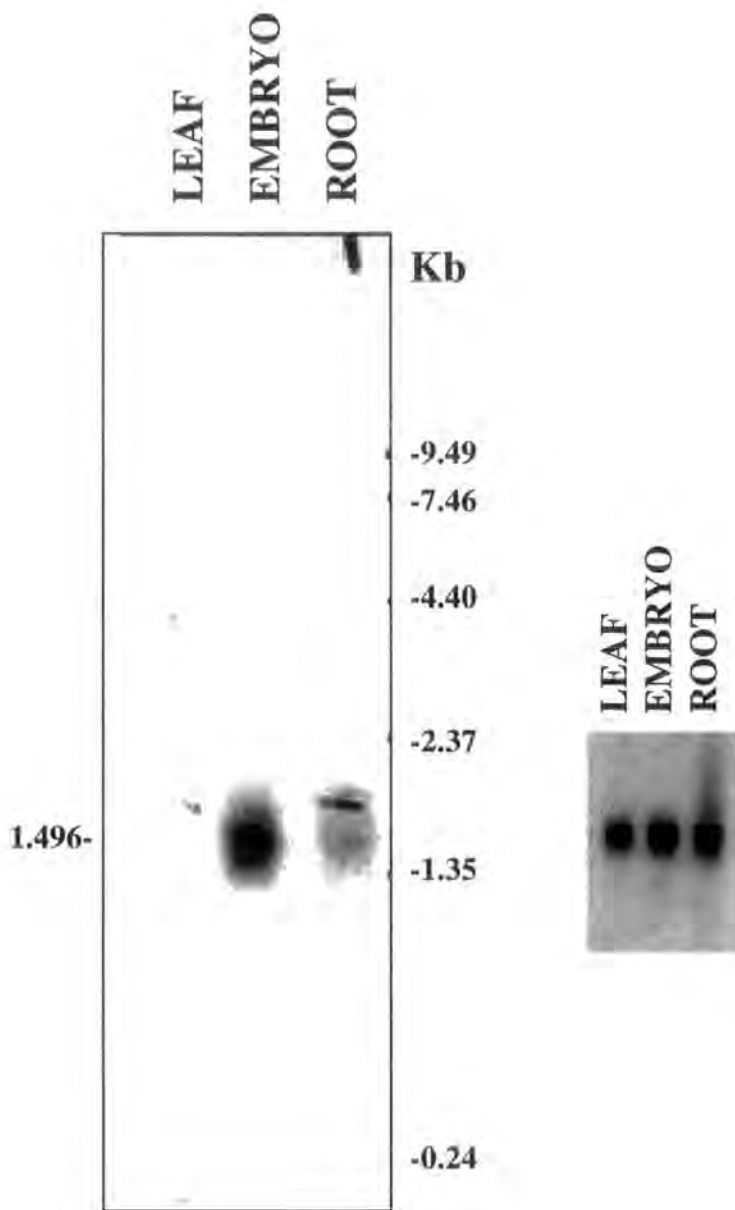


**Figure 6.8: Southern blot autoradiogram showing hybridization of the ARAPAP ORF to *Arabidopsis* and *Brassica* genomic DNA.** <sup>32</sup>P-labeled ARAPAP ORF was generated by a random priming reaction (Redi Prime). Genomic DNA from *Arabidopsis* and *Brassica napus* was digested with *Eco* RI, *Bam* HI, and *Hind*III. Digested DNA was electrophoresed and blotted onto nitrocellulose membrane. Following hybridization, X-ray film was exposed to the blot for 3 weeks.

*Brassica napus* seed embryo, root and leaf was incubated with the same probe under conditions of very low stringency.

A replica blot was also incubated with an actin probe as a means of confirming equal RNA loading in each lane (figure 6.9). The results suggest a relative predominance in embryonic tissue of the *ARAPAP* messenger. The target was also detected in the root sample, but was absent from leaf. Alignment of the developed X-ray film with the membrane, together with calibration using an RNA marker ladder facilitated size estimation of the target mRNA and provided a value of approximately 1500 bases. This value is in excess of the 158J20T7 insert by more than 200 nucleotides and infers one of two possibilities. Firstly the 158J20T7 clone is derived from a truncated mRNA in the original source *Arabidopsis* tissue, the likelihood of which can be supported by the apparent absence of 5' untranslated motifs. Alternatively the estimate could reflect a natural size difference between the *Arabidopsis* subject and the rape target.

The tissue distribution pattern of the *ARAPAP* messenger homologue would be consistent with a PAP involved in complex lipid biosynthesis. In embryonic material the primary function of this enzyme is likely to be the provision of DAG for storage TAG assembly. Conversely in elongating root zones, metabolic emphasis on membrane phosphoglycerolipid synthesis is likely, although the possibility of a role in lipid signaling is not discounted.



**Figure 6.9:** Northern blot autoradiogram showing expression patterns of an ARAPAP homolog in *Brassica napus*.  $^{32}$ -labeled ARAPAP-derived probe was generated by the Redi Prime method. Electrophoretically separated poly A+ RNA (2 $\mu$ g) from leaf, embryo, and young root was blotted onto nylon membrane (Hybond N). Following hybridization, X-ray film was exposed to the blot for 1 week. Inset: Actin-probed replica blot demonstrates equal loading of RNA from each tissue.

It is expected that cross hybridization with other PAP isoforms/homologues is possible. Northern analysis demonstrated the tissue selectivity of a rape *ARAPAP* homologue.

However, definitive proof of function required demonstration of function. It was therefore decided to subclone the *ARAPAP* ORF into an expression vector for transformation into bacterial cells.

## **6.7.2 PCR subcloning of the *ARAPAP* ORF into a bacterial expression system**

### **6.7.2.1 PCR amplification**

Two oligonucleotide primers were designed for the amplification of the *ARAPAP* ORF and its subsequent insertion into Stratagene's pET-24a(+) vector (figure 6.10). The forward primer, ARAP1b includes an *Nde* I restriction site with a 7 nucleotide extension for efficient digestion at the 5' end of the PCR product. This restriction site allows insertion of the target fragment precisely at the starter methionine (ATG) codon. The 3' primer, ARAP2, incorporated an *Eco* RI restriction site and is located 63bp downstream of the *ARAPAP* stop (TGA) codon.

PCR conditions were optimized using a Stratagene RoboCycler<sup>®</sup> (table 6.2) and Vent<sub>r</sub><sup>®</sup> DNA polymerase (New England BioLabs<sup>®</sup> Inc.) was chosen due to its reported high fidelity (5-15 fold higher than that of Taq). Amplification reactions were analysed on agarose gels and a product of approximately 900 bp was observed (figure 6.11).

**ARAP1b:**

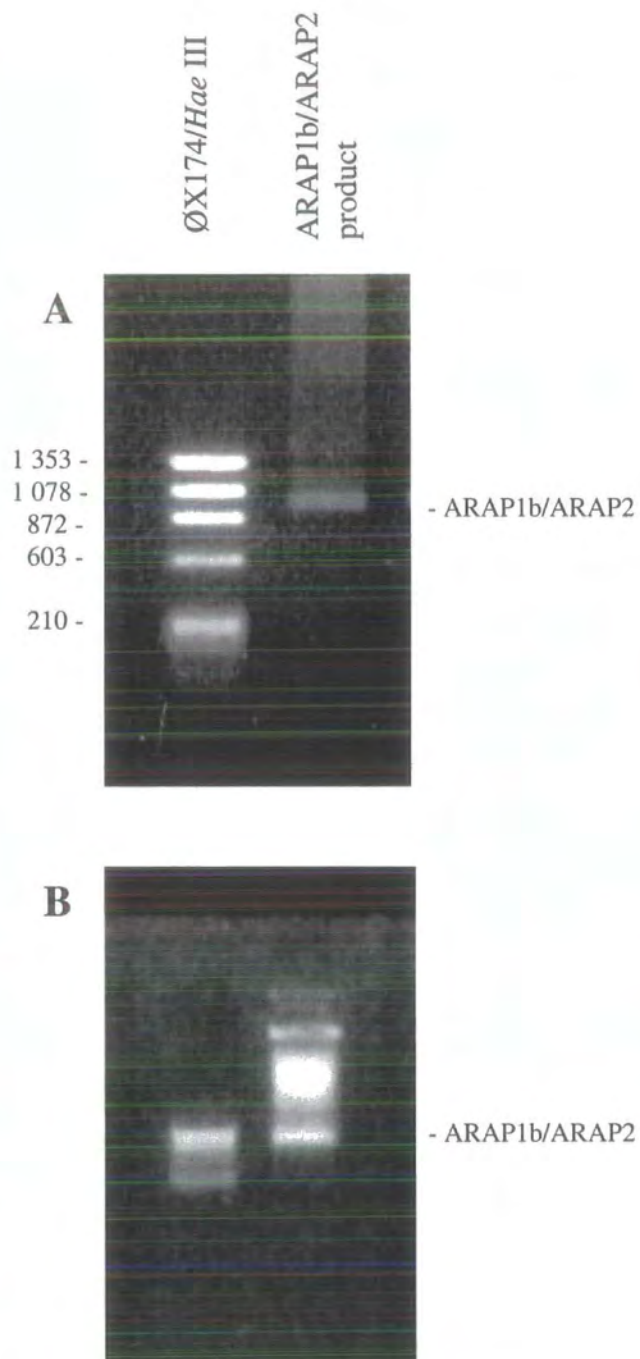
5' - GGAATTCCATATGCCTGAAATTCATTTGGG - 3'  
*Nde* 1

**ARAP2:**

5' - GCGCGAATTCCTTCACTTAGCGGGTTC - 3'  
*Eco* R1

**Figure 6.10: Oligonucleotide primers designed for the PCR subcloning of the ARAPAP open reading frame into pET-24(a) (Stratagene).** The primers include engineered *Nde* 1 and *Eco* R1 restriction sites (highlighted) for directional insertion into the vector.





**Figure 6.11: PCR amplified ARAPAP ORF from primers ARAP1b and ARAP2.** A. sample aliquot of PCR reaction product. B: Pooled and concentrated product.

| Phase | Conditions  | Number of cycles |
|-------|---|------------------|
| 1     | 94°C, 3 minutes                                       | 1                |
| 2     | 94°C, 45 seconds<br>46°C, 1 minute<br>72°C, 2 minutes | 35               |
| 3     | 72°C, 6 minutes                                       | 1                |

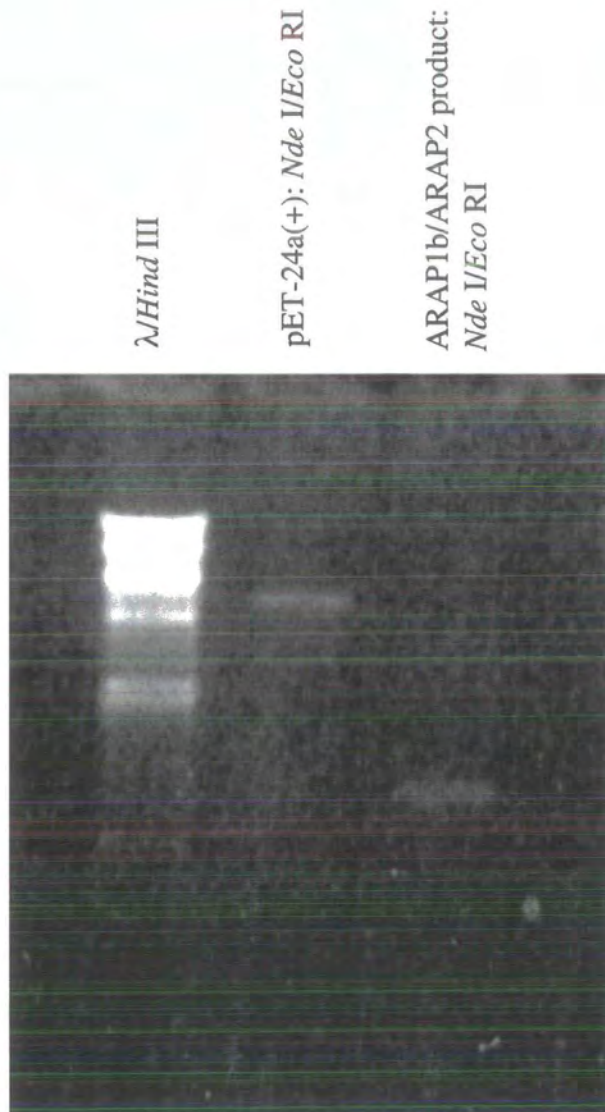
**Table 6.2: Optimized reaction conditions for PCR amplification of the ARAPAP ORF using Vent polymerase (New England BioLabs). Reactions were performed in a Stratagene RoboCycler.**

#### **6.7.2.2 Preparation of PCR insert and pET-24a(+) for ligation**

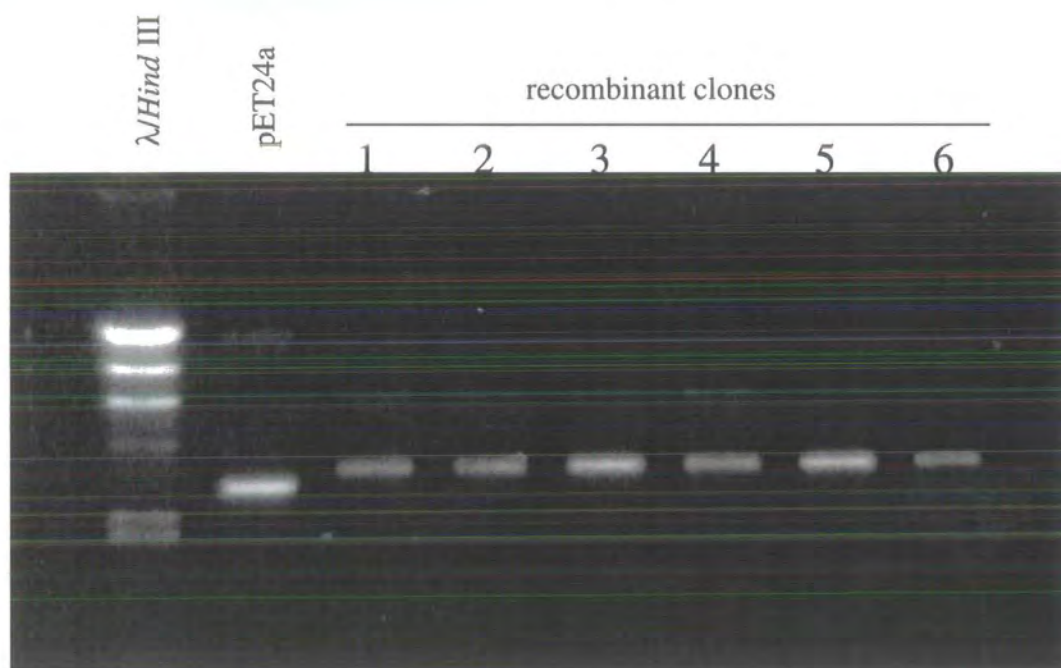
Pooled PCR product was precipitated by ethanol/sodium acetate. Digestion with *Nde* I and *Eco* R1 was followed by preparative agarose gel electrophoresis (figure 6.12).

#### **6.7.2.3 Ligation reactions**

Ligations were performed at a vector-to-insert ratio of 3:1 (50 and 30 ng respectively) with T4 DNA ligase and corresponding buffer from Promega. Competent XL1-Blue *E.coli* cells were transformed and plated on LB-agar/kanamycin. Six positive colonies were selected at random for plasmid isolation. Each clone was of a size expected for successful insertion (figure 6.13). Clone number 1 was prepared for single pass DNA sequencing with the T7 promoter primer which verified in-frame insertion at the 5' end. Sequencing in the reverse direction (T7 terminator primer) confirmed a full length insert but highlighted a single base substitution resulting in an amino acid alteration at residue 237 (Ala->Asp), assumed to be the result of PCR misincorporation. This substitution



**Figure 6.12: Preparative agarose gel for ligation of ARAPAP ORF into pET-24a(+).** Samples of *Nde I/Eco RI*-digested ARAPAP ORF and pET-24a(+) (Novagen) DNA were run into a 1.5% (w/v) agarose gel. DNA bands were excised and extracted (QIAGEN QiAquick Gel Extraction Kit).



**Figure 6.13: Plasmid preparations from 6 positively selected recombinant pET clones.** PCR amplified ARAPAP ORF was ligated into pET-24(a), and 6 positive clones were randomly selected. Plasmids were isolated and samples run alongside pET-24(a) on an agarose gel. Apparent size increase is consistent with the insertion of an 890 bp fragment.

was confirmed by further DNA sequencing. From hereon this construct was used as a parent stock from which all transformations for bacterial expression were performed.

Subsequent sequencing of clones 2 and 3 revealed different single base substitutions in each (table 6.3). None of the mutations observed disturbed the novel conserved phosphatase motif. It is possible that negative selection against error-free construct transformants was taking place due to toxicity of expressed *ARAPAP*.

| Recombinant clone | Observed amino acid substitution |
|-------------------|----------------------------------|
| 1                 | residue 237: Ala -> Asp          |
| 2                 | residue 235: Thr -> Ile          |
| 3                 | residue 205: Leu -> Ile          |

**Table 6.3: Amino acid substitutions in *ARAPAP* ORF fragments PCR subcloned into pET-24a.**

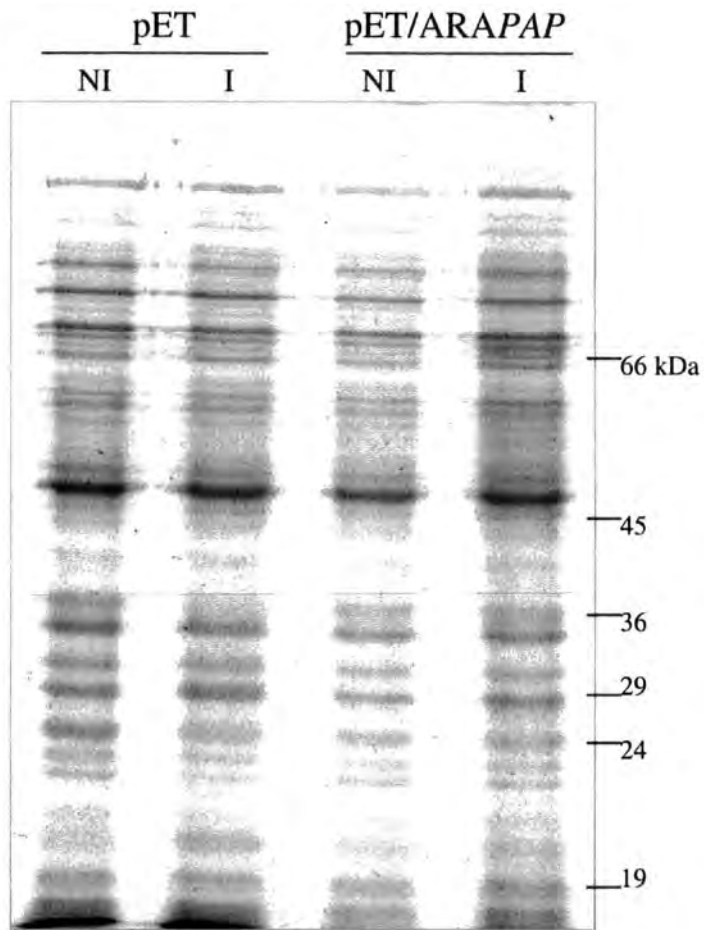
### 6.7.3 Attempted overexpression in *E.coli*

Having established that the 158J20T7 insert contained an open reading frame encoding a putative phosphatidic acid phosphatase, for which a Rape homologue was shown by Northern analysis to be predominantly expressed in seed embryo, attempts were made to demonstrate biochemical function, following expression in *E. coli*.

As shown in preceding sections the *ARAPAP* open reading frame was subcloned via PCR amplification, error free, into the expression vector pET-24(a) (see appendix 4). This construct was used to transform competent cells of Stratagene's *E. coli* BL21

(DE3) strain, which is suitable for T7 promoter-driven expression upon IPTG induction.

Initial induction experiments were performed with standard BL21 (DE3) strain *E. coli* grown in LB media and under kanamycin selection as 50ml cultures. Cultures were grown to an OD<sub>600</sub> of approximately 0.6. Aliquots removed immediately before and 3 hours after addition of IPTG (to 0.5mM) were prepared for SDS-PAGE by harvesting and uptake of cells in SDS loading buffer followed by boiling for 5 minutes. The Coomassie-stained SDS-PAGE profiles revealed no visible induction (figure 6.14). Separate analyses of distinct subcellular fractions (high-speed microsome pellet and corresponding supernatants similarly failed to highlight expression of the recombinant protein. Induction conditions were manipulated (IPTG concentration, incubation period) and a positive control in-house clone provided successful expression. The failure to detect visual induction was thought to result from a range of possible causes. PCR-derived errors in the open reading frame were not disruptive to the reading frame of the insert. It was however possible that the recombinant protein was toxic to the bacterial cells and therefore never expressed beyond a very low level; the possibility of pre-induction background expression of ARAPAP could perhaps be brought under tighter control by the use of a pLysS bacterial strain. Freeze/thaw extracts of induced and pre-induced cells were analysed by SDS-PAGE to examine the possibility that recombinant ARAPAP had been held in inclusion bodies but without detection of a recombinant protein band.



**Figure 6.14 : SDS-PAGE whole cell induction profiles from BL21 DE3 (pLysS)empty pET24a(+) and recombinant pET24a(+)/ARAPAP transformants. NI: non-induced (t = 0) cultures; I: cultures at 3 hrs after induction with 0.4 mM IPTG.**

Different codon usage between *Arabidopsis* and *E. coli* was also considered. The forced high-level expression of a foreign gene commonly results in depletion of bacterial tRNA pools and this in turn leads to delayed translation of the recombinant RNA and consequent enhanced degradation. The issue of rare codon usage was returned to following failed biochemical assay analyses and is discussed in section 6.6.4.2.

The possibility that expression levels were too low level for visualization was not ruled out and it was decided that a biochemical assay might be able to detect any recombinant PAP activity.

#### **6.7.4 Biochemical assay analysis of ARAPAP induced bacterial cultures**

Subsequent to repeated induction experiments performed under a variety of conditions and corresponding SDS-PAGE assessment, the detection of ARAPAP expression was attempted by assay for biochemical activity.

Two assays were employed to investigate the induction of ARAPAP. The first, a radioactive assay based on the liberation of inorganic [<sup>32</sup>P] phosphate from [<sup>32</sup>P]-labelled phosphatidic acid (PA) is specific for PAP activity. The second is a non-specific phosphatase assay which colourmetrically monitors the hydrolysis of paranitrophenyl phosphate (*p*-NPP).



#### 6.7.4.1 The PAP assay

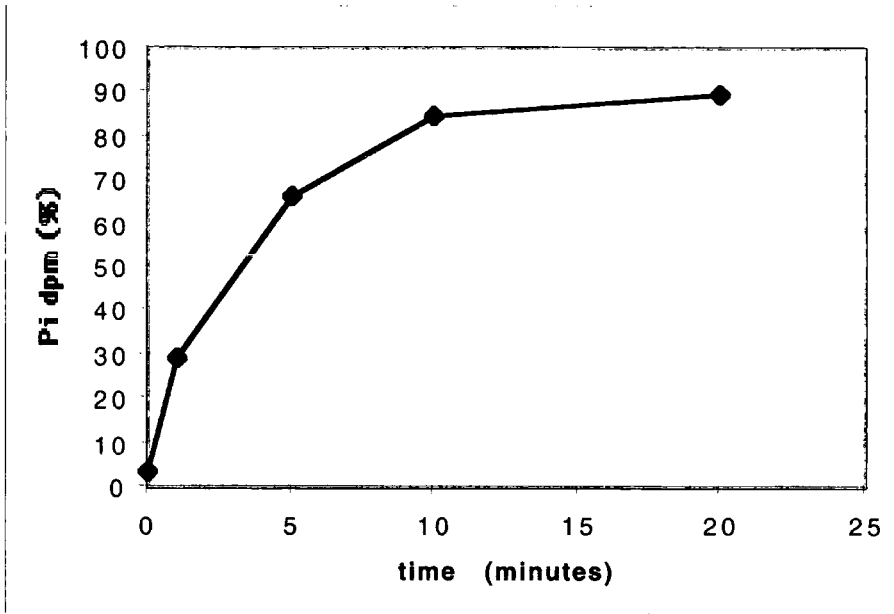
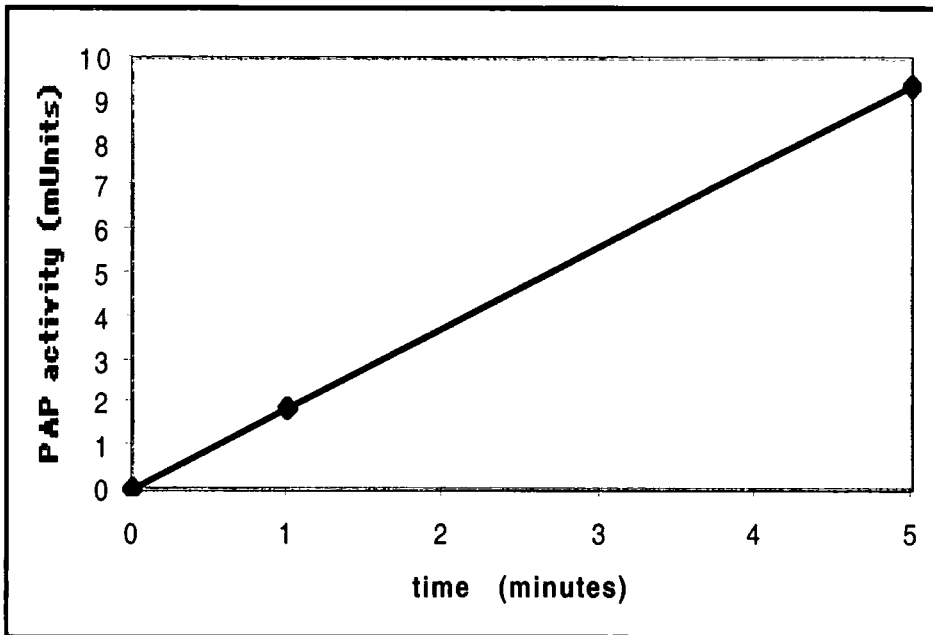
The PAP assay used was as that refined in this laboratory for avocado PAP (Pearce PhD thesis, 1997) from the method of Lin and Carman (1989). Before the assay could be performed, [ $^{32}\text{P}$ ]-labelled PA had to be synthesized from DAG and [ $\gamma$ - $^{32}\text{P}$ ]ATP in a DAG kinase catalysed reaction (Pearce and Slabas, 1998). Figure 6.15 shows a TLC autoradiogram of this substrate which was found to be approximately 95% pure in terms of total [ $^{32}\text{P}$ ] present in the final phase-extracted sample. The substrate and reaction conditions were tested using specially isolated avocado microsomes. Figure 6.16 shows the results of this assay as performed at set time intervals. Under the conditions chosen, the reaction proceeded in a linear fashion up to 5 minutes, and that the substrate was suitable for detecting PAP activity. The assay was next applied to bacterial cultures as previously introduced.

Subcellular fractions were prepared from induced and non-induced *E. coli* BL21(DE3) cells. These consisted of a high-speed microsome pellet and the corresponding supernatant from sonicated cells (a lower speed pellet was discarded). SDS-PAGE analysis of the fractions revealed highlighted no induced protein band. Furthermore, the assays were not able to detect any elevation of aqueous inorganic phosphate from induced samples, and radioactivity remained in the chloroform (substrate) phase (data not shown). The results indicated the absence of detectable PAP activity from *ARAPAP* induced cells, either due to non-expression or to problems with the assay (ie. reaction conditions, specific substrate) for use with this system.



**Figure 6.15: Autoradiogram showing the radiolabelled components of the lipid fraction of the Phosphatidic acid synthesis.** An aliquot of the lipid phase was separated by TLC with a chloroform/methanol/water (12:6:1) solvent. The phosphatidic acid product ( $R_f + 0.35$ ) was **95% pure in terms of radioactive counts.**

**Figure 6.16: Phosphatidic acid phosphatase assay performed on avocado microsomes.** (A) Total microsomal protein (5 $\mu$ g) was incubated in the PAP assay reaction mixture and aliquots removed for phase partitioning and liquid scintillation counting at set time points. Activity was detected by the release of radioactive inorganic phosphate (which enters the aqueous fraction) from the phosphatidate substrate. (B) Aqueous phase data was used to calculate units for PAP activity. Under the conditions used PAP activity was linear up to 5 minutes after reaction initiation.

**A****B**

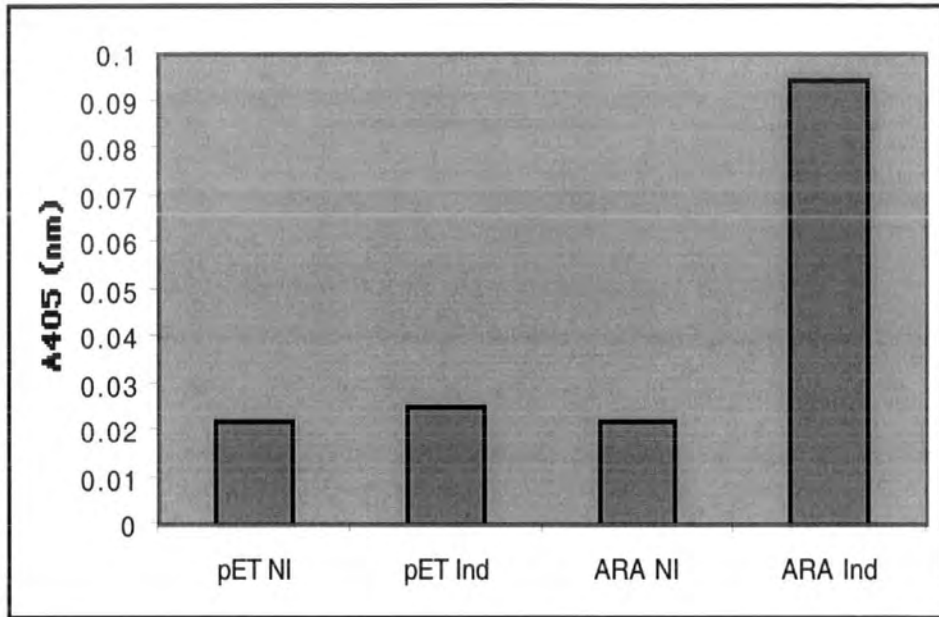
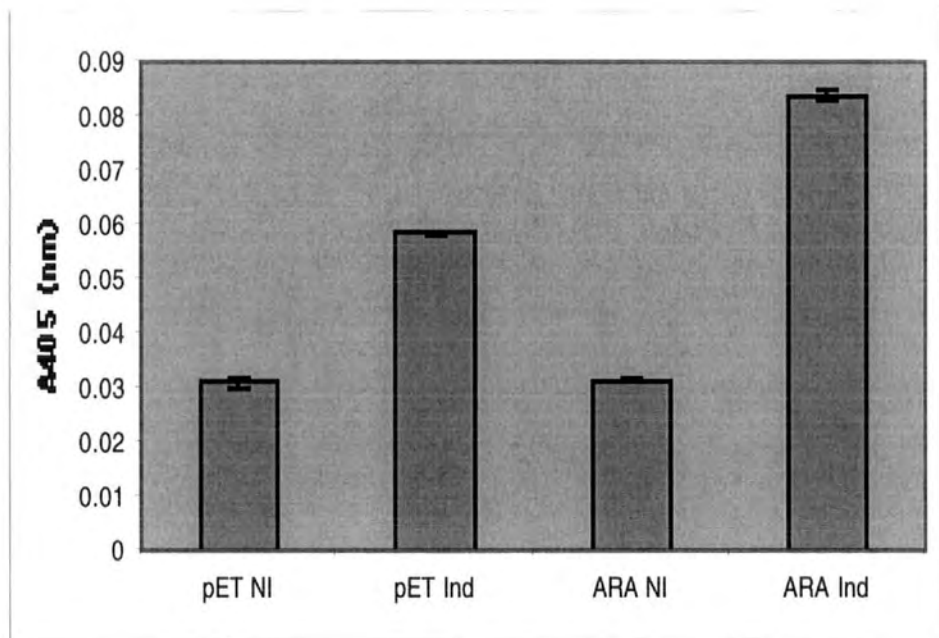
Alternatively the recombinant protein may be lethal to the expressing cells, in which case alternative proof of function would be required (eg. expression in eukaryotic cells, gene knockout in Arabidopsis, functional complementation).

#### **6.7.4.2 Non-specific phosphatase assay**

Paranitrophenyl phosphate (*p*-NPP) was investigated as a general phosphatase substrate. As with the PAP assay, the *p*-NPP assay was tested using avocado microsomes as the protein (enzyme) sample, and similarly using an *E. coli* microsome fraction. Use of the *p*-NPP assay coincided with the availability of a codon bias-adjusted BL21 derivative from Stratagene (BL21-CodonPlus(DE3)-RIL). The strain contains extra copies (plasmid-based) of rare *E. coli* tRNA genes, the products of which recognize AGA/AGG, AUA, and CUA codons.

The possibility that codon usage was causing an impediment to expression prompted investigation into the extent of *E. coli* rare codons in the ARAPAP ORF. The four most problematic codons with regard to usage between Arabidopsis and *E. coli* are AGG (Arg), AGA (Arg), CUA (Leu) and AUA (Ile). Analysis of the ORF revealed the widespread presence (27 in total) of these codons. The pET24A(+)/ARAPAP construct was transformed into BL21(DE3)-RIL cells for induction experiments. As previously, SDS-PAGE profiles did not reveal induction in these cultures. The *p*-NPP assay was used to investigate phosphatase activity between induced and non-induced whole cells (sonicated). Figure 6.17a indicates an increase of almost 5-fold proportions in the level

**Figure 6.17: Measurement of whole-cell non-specific phosphatase activity in *E.coli* cultures.** Aliquots of each inoculation were sonicated and analysed using the p-NPP assay. (A) A crude analysis which did not include duplicates was performed which indicated an approximate 5-fold elevation of phosphatase in induced strains containing the ARAPAP construct over control cells containing the pET-24a vector alone. (B) Subsequent tests confirmed elevated phosphatase levels (approximately 2.5-fold over non-induced background, and approximately 1.5-fold over induced control). The control inoculum was found to contain approximately a 2-fold increase over non-induced background. Experimental range is shown.

**A****B**

of paranitrophenol product from the induced cells containing the recombinant pET construct. Non-induced samples together with empty pET transformants (IPTG-induced and non-induced) all displayed similar low levels of product. All samples were normalized for protein content (via Bradford assay) prior to the experiment. This crude assay was repeated, in duplicate, to confirm results (figure 6.17b). Approximately 3-fold increase in phosphatase was detected in the recombinant cultures. An increase in phosphatase activity was also detected in the induced pET control culture, although at a lower level (approximately 2-fold).

Further [<sup>32</sup>P]PA PAP assays were unsuccessful in measuring PAP-specific phosphatase activity, despite use of the BL21(DE3)-RIL strain and in the absence and presence of β-mercaptoethanol.

## 6.8 Discussion

A putative *Arabidopsis* PAP EST clone was subjected to full DNA sequence analysis which revealed the presence of 1266 bp open reading frame encoding a conceptual 290 amino acid sequence. Database comparisons revealed homology (30% identity) with the yeast *dpp1* protein which displays phosphatidic acid phosphatase activity and has been implicated in signal transduction, as well as complex lipid biosynthesis.

Southern hybridization experiments provided evidence for a single genomic copy of the *ARAPAP* gene in both *Arabidopsis* and *Brassica napus*, although a close homologue in *Arabidopsis* has recently appeared in the databases.



Heterologous Northern hybridization of the *ARAPAP* open reading frame with *Brassica* mRNA from three tissues revealed highest expression levels in seed embryo. Expression was also detected in root, but not in leaf. This expression pattern would be consistent with that of an enzyme involved in storage triglyceride biosynthesis in the endoplasmic reticulum. This analysis required conditions of very low stringency, and previous attempts to hybridize against total RNA blots had been unsuccessful. These findings suggest that the transcript for the enzyme naturally exists at a low level, perhaps reflecting a high degree of regulation associated with this enzyme.

The *ARAPAP* open reading frame was subcloned into a bacterial expression vector via PCR amplification. The recombinant construct was found to contain a single amino acid substitution which was deemed to be the result of PCR error. Overexpression in bacteria under standard conditions of IPTG induction could not be detected on SDS-PAGE profiles. The apparent lack/very low level of expression could be attributed to one or more of a range of factors including cellular toxicity, membrane inclusion, and codon preference differences between *Arabidopsis* and *E. coli*. During the time course of this research detection of the *ARAPAP* protein with anti-(his)<sub>6</sub> antibodies was not attempted but it is suggested that this method may be effective if in fact expression was occurring at a very low level.

Specific PAP activity was not demonstrable in *E. coli* cultures using a radioactive PA substrate. Modifications to reaction conditions, including inclusion of the reducing agent  $\beta$ -mercaptoethanol, and also the use of a codon bias engineered *E. coli* strain proved ineffective. Further modifications may be necessary for detection with this system. Alternatively, overexpression in a eukaryotic overexpression system may

prove successful. Elevated phosphatase activity was however detected, via a non-specific phosphatase assay, in induced cultures containing the recombinant *ARAPAP* plasmid. It is therefore believed that the *ARAPAP* gene encodes a membrane bound phosphatase enzyme, and that based upon homology with characterised yeast this enzyme is likely to display a PAP activity localised to either the plasma membrane or endoplasmic reticulum. The expression pattern observed by heterologous Northern hybridization supports the latter in that a proposed involvement in complex lipid assembly would take place at this organelle.

With the ever increasing availability of EST information from a variety of plant sources the identification of clones of interest, directly through the generation of protein data through specific proteomics approaches, should provide a powerful route to the molecular characterization and manipulation of biochemical pathways.

## General Discussion

An established method for the preparation of biochemically pure endoplasmic reticulum from castor endosperm has been exploited. The method was used to generate bulk quantities of material for in depth electrophoretic investigation and protein characterization. Close attention to particular details of the procedure including prolonged homogenization and sucrose gradient sample loading has resulted in high yields of ER. Visual inspection of SDS-PAGE profiles and verification of ER markers by N-terminal sequencing was taken as evidence of fraction purity.

ER samples were subfractionated by sodium chloride and chloroform/methanol solutions. The sodium chloride procedure represents an effective protein extraction resulting in distinct subfractions as viewed on 1-D SDS gels. One-dimensional studies indicated a requirement for improved resolution of subfraction protein constituents. Hydrophobic chloroform/methanol extracts are on the other hand simple in protein composition and provide a means of enriching hydrophobic proteins for characterization from 1-D gels.

Establishing a workable and reproducible 2-DE methodology for ER samples constituted a major part of this work. Difficulties in attaining gel reproducibility have been overcome as far as possible largely through increased understanding and confidence at each stage of the process and through the use of precast IEF and SDS gels. The membranous nature of the sample presented aggregation concerns which were most notable with membrane-enriched salt-generated subfractions. In-gel aggregation has been greatly reduced by improved sample preparation methods and

sample handling. ER subfraction samples have also been successfully separated by 2-D electrophoresis. The problematic membrane subfraction has been isolated by ultracentrifugation onto a dense barrier sucrose solution, a procedure which greatly reduces aggregation by avoiding ultracentrifugal pelleting.

Reproducible 2-D profiles have been used to compare developing and germinating ER. This analysis required the use of enlarged transparent gel colour images. With no available dedicated image analysis software, comparison relied particularly heavily on experimental reproducibility. Analyses revealed considerable protein differentiation. 2-D experiments have emphasized the importance not only of the 2-DE technique itself but of methods of sample generation and preparation. The work has demonstrated the establishment of a process for the reliable separation of the proteins of this single, high purity organelle. The effective enrichment of proteins during such organellar isolation further adds to the power of examining specific groups of proteins by a combined 2-DE and mass spectrometry approach.

MALDI mass spectrometry has been applied to characterizing selected ER components, both from 1 and 2-D gels. Results have shown that Coomassie-stained bands taken from 1-D gels represent protein levels amenable to peptide mass fingerprinting with limited interference from contaminating sources. However, this approach is severely limited by electrophoretic resolving power, an issue initially recognised during N-terminal sequencing and confirmed during MALDI TOF by the presence of mixed peptide data in a single band. Nonetheless, several proteins were recognised/putatively identified during this exercise. Characterization from 2-D gels made use of a larger scale preparative electrophoresis system. Germinating ER

proteins were visualized by a mass spectrometry-compatible silver stain and selected components were processed for MALDI. The lower quantities of protein analysed at this level resulted in greater interference from contaminant peptides (notably keratins). Data editing of one of the more abundant candidates led to its identification as PDI at only 10 ppm mass range tolerance. Clearly 2-DE offers an unrivalled method for obtaining single proteins from a complex mix. However, for high throughput proteomic characterization from this system it is believed that sample loadings would need to be significantly increased. It has been found that any requirement to precipitate proteins prior to IEF (eg. in order to reduce volume) is undesirable as significant losses to total protein can be expected. Direct uptake of the final ER pellet during isolation in a suitable volume of 2-D lysis buffer is suggested here as an alternative. While the skills acquired during this study are applicable to a wide range of biological problems, it has been shown that a specific sample presents its own challenges. In addition to sample preparation and gel reproducibility, this study has highlighted gel loading quantity and MALDI data editing as key to the success of 2-DE and subsequent MALDI characterization. Of major importance for protein identification from MALDI data is the presence of an extensive corresponding species DNA (genomic) or EST database. Retrieval of an interesting cognate EST is likely to be followed by demonstration of function (see below). The present study represents a sound basis for powerful proteomic research into germination or maturing seed specific protein components of a single organelle. Use of biochemical tools such as photoaffinity probes and substrate analogs immobilized substrate analogs may provide a means of identifying/purifying proteins in these samples involved in complex lipid metabolism. The organellar approach to proteomics is expected to provide an additional level of focus to the understanding of cellular protein expression

patterns, especially when combined with the enriching power of membrane subfractionation techniques.

An unassigned EST encoding a putative PAP had been identified from *Arabidopsis thaliana*, and this clone was ordered and sequence analysed prior to investigation of tissue expression and overexpression attempts. This work involved the combined use of molecular and biochemical techniques and presents an example of functional research downstream of the identification of an EST of interest. The work presented here outlines the identification of a protein likely to be irresponsible for PAP hydrolysis. This inference is primarily based on close sequence homology with known PAPs. Despite a probable transmembrane topology (hydropathy analysis, homologous proteins) which might suggest a more likely role in complex lipid signalling pathways, at the plasma membrane, apparent tissue predominance in seed embryo of *Brassica napus* implies a function in complex lipid assembly. Overexpression of the Arabidopsis protein in *E. coli* was problematic and could not be visualized in SDS-PAGE profiles from fractions of induced cells. Although specific PAP activity was not detected by an established radioactive assay, elevated non-specific activity detected in induced recombinant bacteria supports the functional assignment of phosphatase to this clone.

## Appendix 1a

Theoretical tryptic digest: castor calreticulin (800-2000 Da, max # missed cleavages = 2)

| Residues | MH <sup>+</sup> (mi) | MH <sup>+</sup> (av) | Sequence                 |
|----------|----------------------|----------------------|--------------------------|
| 209-215  | 800.4518             | 800.9336             | (K)IKDPEAK(K)            |
| 284-290  | 820.4317             | 820.9269             | (K)NPNYKGK(W)            |
| 275-281  | 840.5096             | 841.0479             | (K)GPWKPKK(I)            |
| 282-288  | 876.4943             | 877.0349             | (K)IKNPNYK(G)            |
| 367373   | 878.4107             | 878.9142             | (K)KEEEESK(D)            |
| 243-250  | 884.4365             | 884.9647             | (K)EIPDPAK(K)            |
| 358-365  | 894.4209             | 894.9597             | (K)AAFEEAEK(K)           |
| 104-112  | 943.4017             | 944.1206             | (K)LDCGGGYMK(L)          |
| 58-65    | 945.4066             | 945.9671             | (K)WNGDPNDK(G)           |
| 36-42    | 948.4943             | 949.1011             | (R)WVKSDWK(K)            |
| 113-121  | 978.5108             | 979.0787             | (K)LLSSSTDQK(K)          |
| 358-366  | 1022.5159            | 1023.1345            | (K)AAFEEAEKK(K)          |
| 216-223  | 1046.4795            | 1047.1130            | (K)KPEDWDEK(E)           |
| 28-35    | 1052.4437            | 1053.0793            | (R)FEDGWENR(W)           |
| 343-351  | 1061.5267            | 1062.1713            | (K)QLAEETWGK(N)          |
| 66-74    | 1068.4962            | 1069.1197            | (K)GIQTSYR(F)            |
| 91-99    | 1068.6094            | 1069.2936            | (K)TLVFQFSVK(H)          |
| 224-232  | 1105.5053            | 1106.1779            | (K)EYIPDPEDK(K)          |
| 113-122  | 1106.6057            | 1107.2534            | (K)LLSSSTDQKK(F)         |
| 293-302  | 1147.5458            | 1148.3277            | (K)APMIDNPDFK(D)         |
| 233-242  | 1161.5792            | 1162.2890            | (K)KPEGYDDIPK(E)         |
| 315-324  | 1234.6836            | 1235.4738            | (K)YVGIELWQVK(S)         |
| 343-353  | 1303.6646            | 1304.4503            | (K)QLAEETWGKKNK(D)       |
| 89-99    | 311.7313             | 1312.5571            | (K)DKTLVFQFSVK(H)        |
| 354-365  | 1337.6225            | 1338.4181            | (K)DAEKAAFEEAEK(K)       |
| 291-302  | 1461.7200            | 1462.7165            | (K)WKAPMIDNPDFK(D)       |
| 100-112  | 1465.6568            | 1466.6839            | (K)HEQKLDCGGGYMK(L)      |
| 28-38    | 1465.6864            | 1466.6012            | (R)FEDGWENRWVK(S)        |
| 303-314  | 1465.7215            | 1466.6359            | (K)DDPEIYVYPNLK(Y)       |
| 44-57    | 1571.6614            | 1572.5897            | (K)DENTAGEWNYTSGK(W)     |
| 211-223  | 1586.7338            | 1587.6885            | (K)DPEAKKPEDWDEK(E)      |
| 91-103   | 1590.8644            | 1591.8569            | (K)TLVFQFSVKHEQK(L)      |
| 401-415  | 1602.6407            | 1603.5544            | (K)SDAAEDSAEDVHDEL(-)    |
| 75-88    | 1649.7851            | 1650.8340            | (R)FYAISAEFPEFSNK(D)     |
| 43-57 1  | 699.7564             | 1700.7644            | (K)KDENTAGEWNYTSGK(W)    |
| 193-207  | 1733.8750            | 1734.9529            | (K)QTGSLYTDWDLPPK(K)     |
| 145-159  | 1764.9397            | 1766.0169            | (K)VHAILNYNDTNHLIK(K)    |
| 193-208  | 1861.9700            | 1863.1277            | (K)QTGSLYTDWDLPPKK(I)    |
| 75-90    | 1892.9071            | 1894.0975            | (R)FYAISAEFPEFSNKDK(T)   |
| 144-159  | 1893.0347            | 1894.1917            | (K)KVHAILNYNDTNHLIK(K)   |
| 145-160  | 1893.0347            | 1894.1917            | (K)VHAILNYNDTNHLIKK(E)   |
| 104-121  | 1902.8941            | 1904.1760            | (K)LDCGGGYMKLLSSSTDQK(K) |
| 58-74    | 1994.8844            | 1996.0636            | (K)WNGDPNDKGIQTSYR(F)    |
| 325-342  | 1998.9483            | 2000.2421            | (K)SGTLFDNVLICNDPEYAK(Q) |

## Appendix 1b

Theoretical tryptic digest: castor protein disulphide-isomerase (800-2000 Da, max # missed cleavages = 2)

| Residues | MH <sup>+</sup> (mi) | MH <sup>+</sup> (av) | Sequence               |
|----------|----------------------|----------------------|------------------------|
| 264 271  | 925.4420             | 926.0200             | (K)FFNSPDAK(A)         |
| 425 433  | 931.5100             | 932.0653             | (K)SDADIVIAK(L)        |
| 67 73    | 934.4998             | 935.0715             | (K)LRPEYEK(A)          |
| 358 365  | 950.5060             | 951.0740             | (K)EGKVQAYR(K)         |
| 463 470  | 981.4278             | 981.9977             | (K)VEQYDGDR(T)         |
| 90 98    | 989.4540             | 990.0179             | (K)VDANEEANK(E)        |
| 450 457  | 1002.5049            | 1003.1495            | (R)GYPTVYFR(S)         |
| 486 495  | 1003.4809            | 1004.0480            | (K)AAQQESANGK(D)       |
| 141 150  | 1015.5424            | 1016.1431            | (K)QSGPASVEIK(S)       |
| 164 172  | 1019.6294            | 1020.3078            | (K)IFIVGVFPK(F)        |
| 108 116  | 1044.6570            | 1045.3183            | (K)GFPTLKILR(N)        |
| 230 238  | 1050.5108            | 1051.1449            | (K)DFNVDALEK(F)        |
| 66 73    | 1062.5948            | 1063.2463            | (K)KLRPEYEK(A)         |
| 121 129  | 1077.5693            | 1078.2174            | (K)SIQEYKGPR(E)        |
| 80 89    | 1078.6261            | 1079.2893            | (K)SHDIPVVLAK(V)       |
| 473 481  | 1079.5625            | 1080.2271            | (K)DDIISFIEK(N)        |
| 99 107   | 1080.5577            | 1081.2148            | (K)ELATQYDIK(G)        |
| 130 139  | 1108.5526            | 1109.2252            | (R)EADGIAEYLK(K)       |
| 117 126  | 1123.5748            | 1124.2431            | (R)NGGKSIQEYK(G)       |
| 140 150  | 1143.6374            | 1144.3179            | (K)KQSGPASVEIK(S)      |
| 163 172  | 1147.7243            | 1148.4826            | (K)KIFIVGVFPK(F)       |
| 290 298  | 1149.5693            | 1150.2837            | (K)YQEVAHQFK(G)        |
| 463 472  | 1210.5704            | 1211.2779            | (K)VEQYDGDRTK(D)       |
| 130 140  | 1236.6476            | 1237.4000            | (R)EADGIAEYLKK(Q)      |
| 484 495  | 1246.6028            | 1247.3116            | (R)DKAAQQESANGK(D)     |
| 151 162  | 1253.6014            | 1254.3434            | (K)STEAANTFIGDK(K)     |
| 471 481  | 1308.7051            | 1309.5072            | (R)TKDDIISFIEK(N)      |
| 473 483  | 1349.7065            | 1350.5196            | (K)DDIISFIEKNR(D)      |
| 486 498  | 1360.6345            | 1361.4126            | (K)AAQQESANGKDEL(-)    |
| 288 298  | 1364.6963            | 1365.5368            | (K)SKYQEVAHQFK(G)      |
| 151 163  | 1381.6963            | 1382.5182            | (K)STEAANTFIGDKK(I)    |
| 190 201  | 1382.6228            | 1383.4617            | (R)SDYEFGHTLDAK(H)     |
| 458 470  | 1411.6454            | 1412.4604            | (R)SASGKVEQYDGDR(T)    |
| 127 139  | 1418.7280            | 1419.5827            | (K)GPREADGIAEYLK(K)    |
| 450 462  | 1432.7225            | 1433.6122            | (R)GYPTVYFRSASGK(V)    |
| 367 379  | 1451.7382            | 1452.6095            | (K)SEPIPEVNNEPVK(V)    |
| 342 354  | 1489.7803            | 1490.7079            | (K)ANLEPDHIAPWVK(A)    |
| 67 79    | 1517.8692            | 1518.8030            | (K)LRPEYEKAASILK(S)    |
| 411 424  | 1545.8528            | 1546.8102            | (K)QLAPILDEVAVSYK(S)   |
| 202 216  | 1562.8291            | 1563.7600            | (K)HLPQGESSVTGPVVR(L)  |
| 366 379  | 1579.8332            | 1580.7843            | (R)KSEPIPEVNNEPVK(V)   |
| 217 229  | 1644.8678            | 1645.9452            | (R)LFKPFDELFDVDFK(D)   |
| 188 201  | 1651.8080            | 1652.8101            | (K)LRSDYEFGHTLDAK(H)   |
| 74 89    | 1661.9954            | 1663.0207            | (K)AASILKSHDIPVVLAK(V) |



|     |     |           |           |                          |
|-----|-----|-----------|-----------|--------------------------|
| 380 | 395 | 1704.9172 | 1705.9554 | (K)VVVADTLQDIVFNSGK(N)   |
| 99  | 113 | 1723.9271 | 1725.0014 | (K)ELATQYDIKGFPTLK(I)    |
| 434 | 449 | 1749.8295 | 1750.8656 | (K)LDATANDIPSDTFDVR(G)   |
| 272 | 287 | 1758.8195 | 1760.0480 | (K)AMLFMNFNGEAADSIK(S)   |
| 173 | 187 | 1768.7376 | 1769.8870 | (K)FSGEEYENYMSVADK(L)    |
| 396 | 410 | 1779.8351 | 1781.1157 | (K)NVLLEFYAPWCGHCK(Q)    |
| 323 | 338 | 1810.9551 | 1812.0368 | (K)EDQVPLIIQTNDGQK(Y)    |
| 342 | 357 | 1851.9757 | 1853.1383 | (K)ANLEPDHIAPWVKAYK(E)   |
| 339 | 354 | 1894.0227 | 1895.2193 | (K)YLVKANLEPDHIAPWVK(A)  |
| 272 | 289 | 1973.9465 | 1975.3012 | (K)AMLFMNFNGEAADSIKSK(Y) |

## Appendix 2a

### Tryptic peptide mass fingerprints from four human keratins

| M+      | Sequence            | Residues |         |         |         |
|---------|---------------------|----------|---------|---------|---------|
|         |                     | K2C1     | K1C1    | K1CJ    | K22E    |
| 874.50  | SLVNLGGSK           | 65-73    |         |         |         |
| 973.53  | IEISELNR            | 395-402  |         |         | 400-407 |
| 1033.52 | TLLEGEESR           | 483-491  |         |         |         |
| 1041.56 | VDPEIQNVK           |          |         |         | 173-181 |
| 1064.78 | AQYEDIAQK           |          |         |         |         |
| 1090.53 | VTMQNLNDR           |          |         | 148-156 |         |
| 1092.50 | GSGGGSSGGSIGGR      | 602-615  |         |         |         |
| 1107.54 | AQYEEIAQK           |          |         |         | 360-368 |
| 1109.49 | DAEAWFNEK           |          |         | 335-343 |         |
| 1125.54 | AEAESLYQSK          | 366-375  |         |         |         |
| 1157.59 | QGV DADINGLR        |          | 250-260 |         |         |
| 1165.58 | LENEIQTYR           |          |         | 442-453 |         |
| 1179.60 | YEELQITAGR          | 376-385  |         |         |         |
| 1265.64 | TNAENEFVTIK         | 277-287  |         |         |         |
| 1329.73 | NLDLDSIIAEVK        |          |         |         | 348-359 |
| 1357.70 | LNDLEDALQQAK        | 443-454  |         |         |         |
| 1365.64 | SQYEQLAEQNR         |          |         | 323-333 |         |
| 1381.65 | ALEESNYELEGK        |          |         | 166-177 |         |
| 1475.79 | FLEQQNQVLQTK        | 199-210  |         |         |         |
| 1475.75 | WELLQQVDTSTR        | 211-222  |         |         |         |
| 1586.77 | VQALEEANNDLENK      |          | 170-183 |         |         |
| 1657.79 | SGGGFSSGSAGIINYQR   | 12-28    |         |         |         |
| 1707.77 | GSLGGGFSSGGFSGGSFSR |          |         | 41-59   |         |

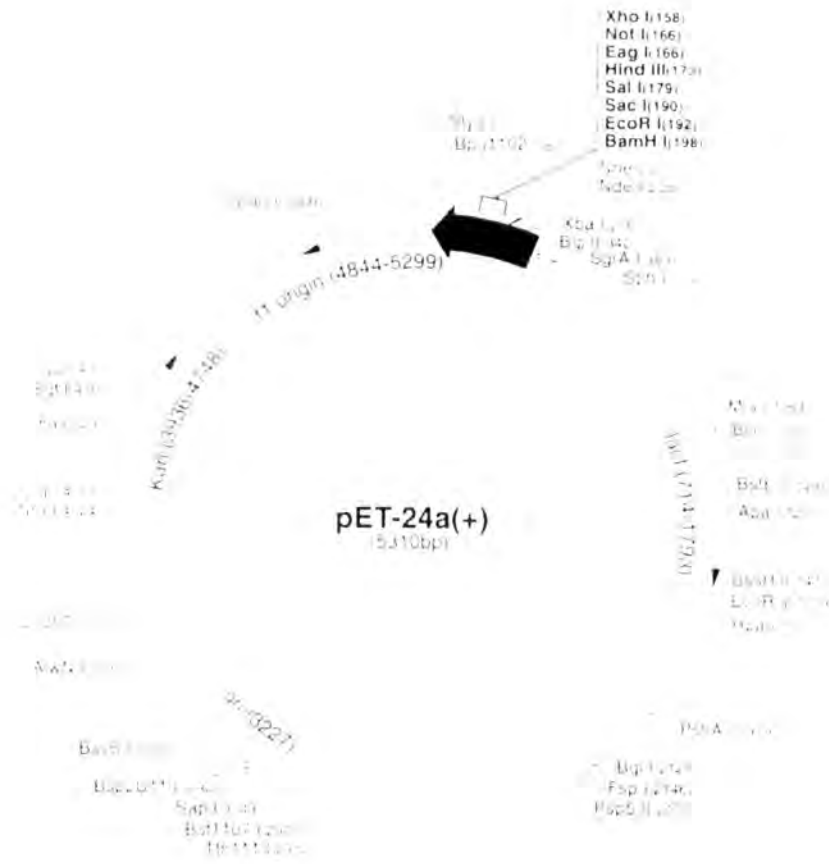
## Appendix 2b

Autolysis products of porcine trypsin. Peptides at 842 and 2210 Da are commonly observed in MALDI spectra

| Residues | M+      | Sequence              |
|----------|---------|-----------------------|
| 52-53    | 261.14  | SR                    |
| 54-57    | 514.32  | IQVR                  |
| 108-115  | 842.50  | VATVSLPR              |
| 209-216  | 905.50  | NKPGVYYK              |
| 148-157  | 1005.48 | APVLSOSSCK            |
| 98-107   | 1044.56 | LSSPATLNSR            |
| 134-147  | 1468.72 | SSGSSYPSLLQCLK        |
| 217-231  | 1735.84 | VCNYVNWIQQYIAAN       |
| 116-133  | 1767.79 | SCAAAGTECLISGWGNTK    |
| 158-178  | 2157.02 | SSYPGQITGNMICVGFLEGGK |
| 58-77    | 2210.10 | LGEHNIDVLEGNEQFINAAK  |
| 78-97    | 2282.17 | IITHPNFNGNTLDNDIMLIK  |
| 179-208  | 3012.32 | DSCQGDSSG...SWGYGCAQK |
| 9-51     | 4488.11 | IVGGYTCAA...VVSAAHCYK |

# Appendix 3

## Bacterial Expression Vector pET-24a(+) (Novagen CN Biosciences UK)



## References

**Adams MD, Dubnick M, Kerlavage AR, Moreno R, Kelley JM, Utterback TR, Nagle JW, Fields C, Venter JC (1992)** Sequence identification of 2,375 human brain genes. *Nature* **355**: 632-634

**Allen NS, Brown DT (1988)** Dynamics of the endoplasmic reticulum in living onion epidermal cells in relation to microtubules, microfilaments and intracellular particle movement. *Cell Motil Cytoskeleton* **10**: 153-163

**Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990)** Basic local alignment search tool. *J Mol Biol* **215**: 403-410

**Altschuler Y, Galili G (1994)** Role of conserved cysteines of a wheat gliadin in its transport and assembly into protein bodies in *Xenopus* oocytes. *J Biol Chem* **269**: 6677-6682

**Amar-Costesec A, Beaufay H, Wibo M, Thines-Sempoux D, Feytmans E, Robbi M, Berthet J (1974)** Analytical study of microsomes and isolated subcellular membranes from rat liver II. Preparation and composition of the microsomal fraction. *J Cell Biol* **61**: 201-212

**Aronel V, Lemieux B, Hwang I, Gibson S, Goodman HM, Somerville CR (1992)** Map-based cloning of a gene controlling omega-3 fatty acid desaturation in *Arabidopsis*. *Science* **258**: 1353-1355

**Barlowe C, Orci L, Yeung T, Hosobuchi M, Hamamoto S, Salama N, Rexach MF, Ravazzolo M, Amherdt M, Schekman R (1994)** COP II: A membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell* **77**: 895-907

**Beaufay H, Amar-Costesec A, Thines-Sempoux D, Wibo M, Robbi M, Berthet J (1974)** Analytical study of microsomes and isolated subcellular membranes from rat liver III. Subfractionation of the microsomal fraction by isopycnic and differential centrifugation in density gradients. *J Cell Biol* **61**: 213-231

**Beaufay H, Amar-Costesec A, Feytmans E, Thines-Sempoux D, Wibo M, Robbi M, Berthet J (1974)** Analytical study of microsomes and isolated subcellular membranes from rat liver I. Biochemical methods. *J Cell Biol* **61**: 188-200

**Bednarek SY, Ravazzola M, Hosobuchi M, Amherdt M, Perrelet A, Schekman R, Orci L (1995)** COPI- and COPII-coated vesicles bud directly from the endoplasmic reticulum. *Cell* **83**: 1183-1196

- Beevers H** (1979) Microbodies in higher plants. *Ann Rev Plant Physiol* **30**: 159-193
- Berger M, Schmidt MF** (1985) Protein fatty acyltransferase is located in the rough endoplasmic reticulum. *FEBS Lett* **187**: 289-294
- Bergeron JJM, Brenner MB, Thomas DY, Williams DB** (1994) Calnexin: a membrane-bound chaperone of the endoplasmic reticulum. *TIBS* **19**: 124-128
- Bille J, Weiser T, Bentrup FW** (1992) The lysolipid sphingosine modulates pyrophosphatase activity in tonoplast vesicles and vacuoles from a heterotrophic cell-suspension culture of *Chenopodium rubrum*. *Physiologia Plantarum* **84**: 250-254
- Bjellqvist B, Ek K, Righetti PG, Gianazza E, Gorg A, Westermeier R, Postel W** (1972) Isoelectric focusing in immobilized pH gradients - principle methodology and some applications. *J Biochem Biophys Methods* **6**: 317
- Boevink P, Oparka K, Cruz SS, Martin B, Betteridge A, Hawes C** (1998) Stacks on tracks: the plant Golgi apparatus traffics on an actin/ER network. *Plant J* **15**: 441-447
- Boevink P, Martin B, Oparka K, Cruz SS, Hawes C** (1999) Transport of virally expressed green fluorescent protein through the secretory pathway in tobacco leaves is inhibited by cold shock and brefeldin A. *Planta* **208**: 392-400
- Borgeson CE, Bowman BJ** (1983) Isolation and characterization of the *Neurospora crassa* endoplasmic reticulum. *J Bacteriology* **156**: 362-368
- Bradford M** (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254
- Bretscher MS, Munro S** (1993) Cholesterol and the Golgi apparatus. *Science* **261**: 1280-1281
- Brindley DN & Waggoner DW** (1996) Phosphatidate phosphohydrolase and signal transduction. *Chem Phys Lipids* **80**: 45-57
- Brough CL, Coventry JM, Christie WW, Kroon JTM, Brown AP, Barsby TL, Slabas AR** (1996) Towards the genetic engineering of triacylglycerols of defined fatty acid composition: major changes in erucic acid content at the *sn*-2 position affected by the introduction of a 1-acyl-*sn*-glycerol-3-phosphate acyltransferase from *Limnanthes douglasii* into oilseed rape. *Molecular Breeding* **2**: 133-142

**Brown AP, Coleman J, Tommey AM, Watson MD, Slabas AR (1994)** Isolation and characterisation of a maize cDNA that complements a 1-acyl *sn*-glycerol-3-phosphate acyltransferase mutant of *Escherichia coli* and encodes a protein which has similarities to other acyltransferases. *Plant Mol Biol* **26**: 211-223

**Cantrill LC, Overall RL, Goodwin PB (1999)** Cell-to-cell communication via plant endomembranes. *Cell Biol Int* **23**: 653-661

**Carman GM, Quinlan JJ (1992)** Phosphatidate phosphatase from yeast mitochondria. *Methods Enzymol* **209**: 219-224

**Chalifour RJ, Spiro RJ (1988)** Effect of phospholipids on thyroid oligosaccharyltransferase activity and orientation evaluation of structural determinants for stimulation of N-glycosylation. *J Cell Biol* **263**: 15673-15680

**Chauvea J, Moule Y, Rouiller C, Schneebeli J (1962)** Isolation of smooth vesicles and free ribosomes from rat liver microsomes. *J Cell Biol* **12**: 17-29

**Chen ZQ, Ulsh LS, DuBois G, Shih TY (1985)** Post translational processing of p21 *ras* proteins involves palmitoylation of the C-terminal tetrapeptide containing cysteine-186. *J Virol* **56**: 607-612

**Chrispeels MJ, Agre P (1994)** Aquaporins: water channel proteins of plant and animal cells. *TIBS* **19**: 421-425

**Claude A (1938)** A fraction from normal chick embryo similar to the tumour producing fraction of chicken tumour I. *Proc Soc Exp Biol Med* **39**: 398

**Conder MJ, Lord TM (1983)** Heterogeneous distribution of glycosyltransferases in the endoplasmic reticulum of castor bean endosperm. *Plant Physiol* **72**: 547-552

**Cosson P, Letourneur F (1994)** Coatamer interaction with di-lysine endoplasmic reticulum retention motifs. *Science* **263**: 1629-1631

**Coughlan SJ, Hastings C, Winfrey Jr RJ (1996)** Molecular characterization of plant endoplasmic reticulum. Identification of protein disulphide-isomerase as the major reticuloplasmin. *Eur J Biochem* **235**: 215-224

**Coughlan SJ, Hastings C, Winfrey Jr R (1997)** Cloning and characterization of the calreticulin gene from *Ricinus communis* L. *Plant Mol Biol* **34**: 897-911

- Craig S, Stahelin LA** (1988) High pressure freezing of intact plant tissues. Evaluation and characterization of novel features of the endoplasmic reticulum and associated membrane systems. *Eur J Cell Biol* **46**: 80-93
- Crofts AJ, Leborgne-Castel N, Pesca M, Vitale A, Denecke J** (1998) BiP and calreticulin form an abundant complex that is independent of endoplasmic reticulum stress. *Plant Cell* **10**: 813-823
- Crowell DN** (2000) Functional implications of protein isoprenylation in plants. *Prog Lipid Res* **39**: 393-408
- Cutler SR, Ehrhardt DW, Griffiths JS, Somerville CR** (2000) Random GFP::cDNA fusions enable visualization of subcellular structures in cells of *Arabidopsis* at a high frequency. *Proc Natl Acad Sci USA* **97**: 3718-3722
- Czichi U, Lennarz WJ** (1977) Localization of the enzyme system for glycosylation of proteins via the lipid-linked pathway in rough endoplasmic reticulum. *J Biol Chem* **252**: 7901-7912
- Dean C, Schmidt R** (1995) Plant genomes: A current molecular description. *Annu Rev Plant Physiol* **46**: 395-418
- Denecke J** (1996) Soluble endoplasmic reticulum resident proteins and their function in protein synthesis and transport. *Plant Physiol Biochem* **34**: 197-205
- D'Enfert C, Gense M, Gaillardin C** (1992) Fission yeast and a plant have functional homologues of the Sar 1 and Sec 12 proteins involved in ER to Golgi traffic in budding yeast. *EMBO J* **11**: 4205-4211
- Frentzen M, Heinz E, McKeon TA, Stumpf PK** (1983) Specificities and selectivities of glycerol-3-phosphate acyltransferase and monoacylglycerol-3-phosphate acyltransferase from pea and spinach chloroplasts. *Eur J Biochem* **129**: 629-636
- Fujiki Y, Hubbard AL, Fowler S, Lazarow PB** (1982) Isolation of intracellular membranes by means of sodium carbonate treatment: application to endoplasmic reticulum. *J Cell Biol* **93**: 97-102
- Galili G, Sengupta-Gopalan C, Ceriotti A** (1998) The endoplasmic reticulum of plant cells and its role in protein maturation and biogenesis of oil bodies. *Plant Mol Biol* **38**:1-29
- Geli MI, Torrent M, Ludevid D** (1994) Two structural domains mediate two sequential events in  $\gamma$ -zein targeting: protein endoplasmic reticulum retention and protein body formation. *Plant Cell* **6**: 1911-1922



- Gilmore R** (1993) Protein translocation across the endoplasmic reticulum – a tunnel with toll booths at entry and exit. *Cell* **75**: 589-592
- Glaumann H, Dallner G** (1970) Subfractionation of smooth microsomes from rat liver. *J Cell Biol* **47**: 34-48
- Gorlich D, Rapoport TA** (1993) Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum. *Cell* **75**: 615-630
- Greenwood JS, Bewley JD** (1981) Seed development in *Ricinus communis* (castor bean). I. Descriptive morphology. *Can J Bot* **60**: 1751-1760
- Hamburger D, Egerton M, Riezman H** (1995) Yeast galp is required for attachment of a completed GPI anchor onto proteins. *J Cell Biol* **129**: 629-639
- Hammond C, Hellenius A** (1995) Quality control in the secretory pathway. *Curr Opin Cell Biol* **7**: 523-529
- Hanein D, Matlack KES, Jungnickel B, Plath K, Kalies K-U, Miller KR, Rapoport TA, Akey CW** (1996) Oligomeric rings of the Sec61p complex induced by ligands required for protein translocation. *Cell* **87**: 721-732
- Hanover JA, Lennarz WJ** (1980) N-linked glycoprotein assembly. *J Biol Chem* **255**: 3600-3604
- Hardwick KG, Lewis MJ, Semenza J, Dean N, Pelham HRB** (1990) ERD1, a yeast gene required for the retention of luminal endoplasmic reticulum proteins, affects glycoprotein processing in the Golgi apparatus. *EMBO J* **9**: 623-630
- Hardwick KG, Pelham HRB** (1992) SED5 encodes a 39-kD integral membrane protein required for vesicle transport between the ER and the Golgi complex. *J Cell Biol Chem* **119**: 513-521
- Hares W, Frentzen M** (1991) Substrate specificities of the membrane-bound and partially purified microsomal acyl-CoA: 1-acylglycerol-3-phosphate acyltransferase from etiolated shoots of *Pisum sativum* (L.). *Planta* **185**: 124-131
- Harter C, Pavel J, Coccia F, Draken E, Wegehingel S, Tschochner H, Wieland F** (1996) Nonclathrin coat protein  $\gamma$ , a subunit of coatamer, binds to the cytoplasmic dilysine motif of membrane proteins of the early secretory pathway. *Proc Natl Acad Sci USA* **93**: 1902-1906

**Hartinger J, Stenius K, Hogemann D, Jahn R** (1996) 16-BAC/SDS-PAGE: A two-dimensional gel electrophoresis system suitable for the separation of integral membrane proteins. *Anal Biochem* **240**: 126-133

**Hartmann E, Sommer T, Prehn S, Gorlich D, Jantsch S, Rapoport TA** (1994) Evolutionary conservation of components of the protein translocation machinery. *Nature* **367**: 654-657

**Hassiloff J, Siemering KR, Prasher DC, Hodge S** (1997) Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic Arabidopsis plants brightly. *Proc Natl Acad Sci USA* **94**: 2122-2127

**Hein M, Tang Y, McLeod DA, Janda KD, Hiatt A** (1991) Evaluation of immunoglobulins from plant cells. *Biotechnol Prog* **7**: 455-461

**Helenius A, Trombetta ES, Herbert DN, Simons JF** (1997) Calnexin calreticulin and the folding of glycoproteins. *Trans Cell Biol* **7**: 193-200

**Hepler PK** (1982) Endoplasmic reticulum in the formation of the cell plate and plasmodesmata. *Protoplasma* **111**: 121-133

**Hosaka K, Yamashita S** (1984) Partial purification and properties of phosphatidate phosphatase in *Saccharomyces cerevisiae*. *Biochim Biophys Acta* **796**: 102-109

**Huang G, Lee DH, Singh S** (1988) Identification of the thiol ester linked lipids in apolipoprotein B. *Biochemistry* **27**: 1395-1400

**Huang AHC, Qu R, Lai YK, Ratnayake C, Chan KL *et al*** (1990) In *Compartmentalization of Plant Metabolism*, ed. MJ Emes, pp43-58 Cambridge: Cambridge Univ. Press.

**Huang AHC** (1992) Oil bodies and oleosins in seeds. *Annu Rev Plant Physiol Plant Mol Biol* **43**: 177-200

**Huang LQ, Franklin AE, Hoffman NE** (1993) Primary structure and characterization of an *Arabidopsis thaliana* calnexin-like protein. *J Biol Chem* **268**: 6560-6566

**Hurkman WJ, Tanaka CK** (1986) Solubilization of plant membrane proteins for two-dimensional gel electrophoresis. *Plant Physiol* **81**: 802-806

**Israel M, Morel N, Lesbats B, Birman S, Manaranche R** (1986) Purification of a presynaptic membrane-protein that mediates a calcium-

dependent translocation of acetylcholine. Proc Natl Acad Sci USA **83**: 9226-9230

**Jackson MR, Nilsson T, Peterson PA (1990)** Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. EMBO J. **9**: 3153-3162

**James DE, Piper RC (1994)** Insulin resistance, diabetes, and the insulin-regulated trafficking of GLUT-4. J Cell Biol **126**: 1123-1126

**Joyard J, Grossma A, Bartlett SG, Douce R, Chua NH (1982)** Characterization of envelope membrane polypeptides from spinach chloroplasts. J Biol Chem **257**: 1095-1101

**Kalies K-U, Gorlich D, Rapoport TA (1994)** Binding of ribosomes to the rough endoplasmic reticulum mediated by the Sec61p complex. J Cell Biol **126**: 925-934

**Karaoglu D, Kelleher DJ, Gilmore R (1995)** Functional characterization of Ost3p - Loss of the 34-kD subunit of the *Saccharomyces cerevisiae* oligosaccharyltransferase results in biased under glycosylation of acceptor substrates. J Cell Biol **130**: 567-577

**Kelleher DJ, Gilmore R (1994)** The *Saccharomyces cerevisiae* oligosaccharyltransferase is a protein complex composed of Wbp1p, Swp1p, and four additional polypeptides. J Cell Biol **269**: 12908-12917

**Koivu J, Myllyla R, Helaakoski T, Pihlajaniemi T, Tasanen K, Kivirikko KI (1987)** A single polypeptide acts both as the  $\beta$  subunit of prolyl 4-hydroxylase and as a protein disulfide-isomerase. J Biol Chem **262**: 6447-6449

**Kumar V, Heinemann FS, Ozols J (1994)** Purification and characterization of avian oligosaccharyltransferase - complete amino acid sequence of the 50-kDa subunit. J Biol Chem **269**: 13451-13457

**Kumar V, Heinemann FS, Ozols J (1995)** Human oligosaccharyltransferase: Isolation, characterization, and the complete amino acid sequence of the 50-kDa subunit. Archives of Biochem Biophys **320**: 217-223

**Lacey DJ, Hills MJ (1996)** Heterogeneity of the endoplasmic reticulum with respect to lipid synthesis in developing seeds of *Brassica napus* L. Planta **199**: 545-551

**Laemmli UK (1970)** Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**: 680-685

- Larson TJ, Lightner VA, Green PR, Modrich P, Bell RM (1980)** Membrane phospholipid synthesis in *Escherichia coli*. Identification of the *sn*-glycerol-3-phosphate acyltransferase polypeptide as the *plsB* gene product. *J Biol Chem* **255**: 9421-9426
- Letourneur F, Gaynor EC, Hennecke S, Demolliere C, Duden R, Emr SD, Riezman H, Cosson P (1994)** Coatamer is essential for retrieval of dilysine-tagged proteins to the endoplasmic reticulum. *Cell* **79**: 1199-1207
- Li XH et al (1998)** The molecular chaperone calnexin associates with the vacuolar H<sup>+</sup>-ATPase from oat seedlings. *Plant Cell* **10**: 119-130
- Lian JP, Ferro-Novick S (1993)** Bos1p, an integral membrane protein of the endoplasmic reticulum to Golgi transport vesicles, is required for their fusion competence. *Cell* **73**: 735-745
- Lichtscheidl IK, Lancelle SA, Hepler PK (1990)** Actin-endoplasmic reticulum complexes in *Drosera*. Their structural relationship with the plasmalemma, nucleus, and organelles in cells prepared by high pressure freezing. *Protoplasma* **155**: 116-126
- Lightner VA Larson TJ, Tailleux P, Kantor GD, Raetz CRH, Bell RM, Modrich P (1980)** Membrane phospholipid synthesis in *Escherichia coli*. Cloning of a structural gene (*plsB*) of the *sn*-glycerol-3-phosphate acyltransferase. *J Biol Chem* **255**: 9413-9420
- Lin JT, Woodruff CL, Lagouche OJ, KcKeon TA, Stafford AE, Goodrich-Tanrikulu M, Singleton JA, Haney CA (1998)** Biosynthesis of triacylglycerols containing ricinoleate in castor microsomes using 1-acyl-2-oleoyl-*sn*-glycero-3-phosphocholine as the substrate of oleoyl-12-hydroxylase. *Lipids* **33**: 59-69
- Lin YP, Carman GM (1989)** Purification and characterization of phosphatidate phosphatase from *Saccharomyces cerevisiae*. *J Biol Chem* **264**: 8641-8645
- Lodish HF (1988)** Multi-spanning membrane proteins: how accurate are the models? *TIBS* **13**: 332-334
- Lord JM, Kagawa T, Moore TS, Beevers H (1973)** Endoplasmic reticulum as the site of lecithin formation in castor bean endosperm. *J Cell Biol* **57**: 659-667

**Lord JM** (1985) Synthesis and intracellular transport of lectin and storage protein precursors in endosperm from castor bean. *Eur J Biochem* **146**: 403-409

**Lowry OH, Rosebrough NJ, Farr AL, Randall RJ** (1951) Protein measurement with the Folin phenol reagent. *J Biol Chem* **193**: 265

**Lucas WJ, Ding B, Van der Schoot C** (1994) Plasmodesmata and the supracellular nature of plants. *New Phytol* **126**: 719-719

**Lui JY, Hara C, Umeda M, Zhao Y, Okita TW, Uchimiya H** (1995) Analysis of randomly isolated cDNAs from developing endosperm of rice (*Oryza sativa* L.): evaluation of expressed sequence tags, and expression levels of mRNAs. *Plant Mol Biol* **29**: 685-689

**Lund P, Lee RY, Dunsmuir P** (1989) Bacterial chitinase is modified and secreted in transgenic tobacco. *Plant Physiol* **91**: 130-135

**Lund P, Dunsmuir P** (1992) A plant signal sequence enhances the secretion of bacterial ChiA in transgenic tobacco. *Plant Mol Biol* **18**: 47-53

**Lynch DV, Fairfield SR** (1993) Sphingolipid long-chain synthesis in plants - characterization of serine palmitoyltransferase activity in squash fruit microsomes. *Plant Physiol* **103**: 1421-1429

**Menegazzi P, Guzzo F, Baldan B, Mariani P, Treves S** (1993) Purification of calreticulin-like protein(s) from spinach leaves. *Biochem Biophys Res Comm* **190**: 1130-1135

**Meyer DI, Krause E, Dobberstein B** (1982) Secretory protein translocation across membranes – the role of the ‘docking proteins’. *Nature* **297**: 647-650

**Mitchell P** (1979) Keilin’s respiratory chain concept and its chemiosmotic consequences. *Science* **206**: 1148-1159

**Morlock KR, McLaughlin JJ, Lin YP, Carman GM** (1991) Phosphatidate phosphatase from *Saccharomyces cerevisiae*: isolation of 45- and 104-kDa forms of the enzyme that are differentially regulated by inositol. *J Biol Chem* **266**: 3586-3593

**Moore Jr TS** (1982) Phospholipid biosynthesis. *Ann Rev Plant Physiol* **33**: 235-259

**Munro S, Pelham HRB** (1986) An hsp70-like protein in the ER: Identity with the 78 kd glucose-regulated protein and immunoglobulin heavy chain binding protein. *Cell* **46**: 291-300

- Murata N (1983)** Molecular species composition of phosphatidylglycerols from chilling-sensitive and chilling-resistant plants. *Plant Cell Physiol* **24**: 81-86
- Nakano A, Brada D, Schekman R (1988)** A membrane glycoprotein, Sec12p, required for protein transport from the endoplasmic reticulum to the Golgi apparatus in yeast. *J Cell Biol* **107**: 851-863
- Newman T, de Bruijn FJ, Green P, Keegstra K, Kende H, McIntosh L, Ohlrogge J, Raikhel N, Somerville S, Thomashow M, Retzel E, Somerville C (1994)** Genes galore: A summary of methods for accessing results from large-scale partial sequencing of anonymous *Arabidopsis* cDNA clones. *Plant Physiol* **106**: 1241-1255
- Nilsson I, von Heijne G (1993)** Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. *J Biol Chem* **268**: 5798-5801
- Nilsson I, Whitley P, von Heijne G (1994)** The COOH-terminal ends of internal signal and signal-anchor sequences are positioned differently in the ER translocase. *J Cell Biol* **126**: 1127-1132
- Noiva R, Freedman RB, Lennarz WJ (1993)** Peptide binding to protein disulphide isomerase occurs at a site distinct from the active sites. *J Biol Chem* **268**: 19210-19217
- O'Farrel PH (1975)** High resolution, two-dimensional electrophoresis of proteins. *J Biol Chem* **250**: 4007-4021
- Ohlrogge J, Browse J (1995)** Lipid biosynthesis. *Plant Cell* **7**: 957-970
- Okita TW, Rogers JC (1996)** Compartmentation of proteins in the endomembrane system of plant cells. *Annu Rev Plant Physiol Plant Mol Biol* **47**: 327-350
- Okuley J, Lightner J, Feldmann K, Yadav N, Lark E, Browse J (1994)** *Arabidopsis FAD2* gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. *Plant Cell* **6**: 147-158
- Osterman J, Orci L, Tani K, Amherdt M, Ravazzola M, Elazar Z, Rothman JE (1993)** Stepwise assembly of functionally active transport vesicles. *Cell* **75**: 1015-1025
- Oyler GA, Higgins GA, Hart RA, Battenberg E, Billingsley M, Bloom FE, Wilson MC (1989)** The identification of a novel synaptosomal-associated

protein, SNAP-25, differently expressed by neuronal subpopulations. *J Cell Biol* **109**: 3039-3052

**Palade GE, Siekevitz P** (1956) Pancreatic microsomes. An integrated morphological and biochemical study. *J Biophys Biochem Cytol* **2**: 671

**Palade G** (1975) Intracellular aspects of the process of protein synthesis. *Science* **189**: 347-358

**Panzner S, Dreier L, Hartmann E, Kostka S, Rapoport T** (1995) Post-translational protein transport in yeast reconstituted with a purified complex of Sec proteins and Kar2p. *Cell* **81**: 561-570

**Pearce ML** (1997) Ph.D Thesis: The membrane bound phosphatidic acid phosphatase from avocado.

**Pearce M, Slabas AR** (1998) Phosphatidate phosphatase from avocado (*Persea americana*) – purification, substrate specificity and possible metabolic implications for the Kennedy pathway and cell signalling in plants. *The Plant Journal* **14**: 555-564

**Pelham HRB, Hardwick KG, Lewis MJ** (1988) Sorting of soluble ER protein in yeast. *EMBO J.* **7**: 1757-1762

**Pelham HRB** (1995) Sorting and retrieval between the endoplasmic reticulum and Golgi apparatus. *Curr Op Cell Biol* **7**: 530-535

**Peterson GL** (1979) Review of the Folin phenol protein quantitation method of Lowry, Farr, and Randall. *Anal Biochem* **100**: 201-220

**Pfeffer SR, Rothman JE** (1987) Biosynthetic protein transport and sorting by the endoplasmic reticulum and Golgi. *Ann Rev Biochem* **56**: 829-852

**Quadar H, Schnepf E** (1986) Endoplasmic reticulum and cytoplasmic streaming: fluorescent microscopic observations in adaxial epidermis cells of onion bulb scales. *Protoplasma* **131**: 250-252

**Randall SK, Marshall MS, Crowell DN** (1993) Protein isoprenylation in suspension-cultured tobacco cells. *Plant Cell* **5**: 433-442

**Regad F, Bardet C, Tremousaygue D, Moisan A, Lescure B, Axelos M** (1993) cDNA cloning and expression of an *Arabidopsis* GTP-binding protein of the ARF family. *FEBS Lett* **316**: 133-136

**Reinhold L, Kaplan A** (1984) Membrane transport of sugars and amino acids. *Ann Rev Plant Physiol* **35**: 45-83

- Ridge RW, Uozumi Y, Plazinski J, Hurley UA, Williamson RE (1999)** Developmental transitions and dynamics of the cortical ER of Arabidopsis cells seem with green fluorescent protein. *Plant Cell Physiol* **40**: 1253-1261
- Rothman JE, Wieland FT (1996)** Protein sorting by transport vesicles. *Science* **272**: 227-234
- Rothschild J (1963)** The isolation of microsomal membranes. *Biochem Soc Symp* **22**: 4
- Rounsley SD, Glodek A, Sutton G, Adams MD, Somerville CR, Venter JC, Kerlavage AR (1996)** The construction of Arabidopsis expressed sequence tag assemblies. A new resource to facilitate gene identification. *Plant Physiol* **112**: 1177-1183
- Schimmoller F, Singer-Kruger B, Schroder S, Kruger U, Barlowe C, Riezman H (1995)** The absence of Emp24p, a component of ER-derived COPII-coated vesicles, causes a defect in transport of selected proteins to the Golgi. *EMBO J* **14**: 1329-1339
- Schultz AM, Henderson LE, Oroszlan S (1988)** Fatty acylation of proteins. *Ann Rev Cell Biol* **4**: 611-647
- Semenza J, Hardwick KG, Dean N, Pelham HRB (1990)** ERD2, a yeast gene required for the receptor-mediated retrieval of luminal ER protein from the secretory pathway. *Cell* **61**: 1349-1357
- Serafini T, Orci L, Amherdt M, Brunner M, Kahn RA, Rothman JE (1991)** ADP-ribosylation factor is a subunit of the coat of Golgi-derived COP-coated vesicles: a novel role for a GTP-binding protein. *Cell* **67**: 239-253
- Shelness GS, Blobel G (1990)** Two subunits of the canine signal peptidase complex are homologous to yeast SECII protein. *J Biol Chem* **265**: 9512-9519
- Sherrier DJ, Prime TA, Dupree P (1999)** Glycosylphosphatidylinositol-anchored cell-surface proteins from *Arabidopsis*. *Electrophoresis* **20**: 2027-2035
- Shevchenko A, Jensen ON, Podtelejnikov AV, Sagliocco F, Wilm M, Vorm O, Mortensen P, Shevchenko A, Boucherie H, Mann M (1996)** Linking genome to proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. *Proc Natl Acad Sci USA* **93**: 14440-14445



**Shorrosh BS, Subramaniam J, Schubert KR, Dixon RA** (1993) Expression and localization of plant protein disulphide isomerase. *Plant Physiol* **103**: 719-726

**Sieigneurin-Berny D, Rolland N, Garin J, Joyard J** (1999) Differential extraction of hydrophobic proteins from chloroplast envelope membranes: a subcellular-specific proteomic approach to identify rare intrinsic membrane proteins. *Plant J* **19**: 217-228

**Sijmons PC, Dekker BMM, Schrammeijer B, Verwoerd TC, van den Elzen PJM, Hoekema A** (1990) Production of correctly processed human serum albumin in transgenic plants. *Biotechnology* **8**: 217-221

**Silberstein S, Collins PG, Kelleher DJ, Rapiejko PJ, Gilmore R** (1995) The alpha-subunit of the *Saccharomyces cerevisiae* oligosaccharyltransferase complex is essential for vegetative growth of yeast and is homologous to mammalian ribophorin I. *J Cell Biol* **128**: 525-536

**Silberstein S, Collins PG, Kelleher DJ, Gilmore R** (1995) The essential OST2 gene encodes the 16-kD subunit of the yeast oligosaccharyltransferase, a highly conserved protein expressed in diverse eukaryotic organisms. *J Cell Biol* **131**: 371-383

**Singer SJ, Nicolson GL** (1972) The fluid mosaic model of the structure of cell membranes. *Science* **175**: 720-731

**Somerville C, Browse J** (1991) Plant lipids: metabolism, mutants, and membranes. *Science* **252**: 80-87

**Stukey J, Carman GM** (1997) Identification of a novel phosphatase sequence motif. *Protein Sci* **6**: 469-472

**Sturm A, Kindl H** (1983) Fucosyl transferase activity and fucose incorporation in vivo as markers for subfractioning cucumber microsomes. *FEBS Lett* **160**: 165-168

**Swida U, Kreutzfeldt C, Ramezani-Rad M, Kaufer N** (1982) Isolation and characterisation of rough and smooth endoplasmic reticulum from *Saccharomyces cerevisiae*. *FEMS Microbiol Lett* **15**: 313-318

**Te Heesen S, Janetsky B, Lehle L, Aebi M** (1992) The yeast Wbp1 is essential for oligosaccharyltransferase activity *in vivo* and *in vitro*. *EMBO J* **11**: 2071-2075

**Te Heesen S, Knauer R, Lehle L, Aebi M (1993)** Yeast Wbp1p and Swp1p form a protein complex essential for oligosaccharyltransferase activity. *EMBO J* **12**: 279-284

**Toke DA, Bennett WL, Dillon DA, Wu WI, Chen XM, Ostrander DB, Oshiro J, Cremesti A, Voelker DR, Fischl AS & Carman GM (1998a)** Isolation and characterization of the *Saccharomyces cerevisiae* DPP1 gene encoding diacylglycerol pyrophosphate phosphatase. *J Biol Chem* **273**: 3278-3284

**Toke DA, Bennett WL, Oshiro J, Wu WI, Voelker DR, Carman GM (1998b)** Isolation and characterization of the *Saccharomyces cerevisiae* LPP1 gene encoding a Mg<sup>2+</sup>-independent phosphatidate phosphatase. *J Biol Chem* **273**: 14331-14338

**Trimble WS, Cowan DM, Scheller RH (1988)** VAMP-1: A synaptic vesicle-associated integral membrane protein. *Proc Natl Acad Sci USA* **85**: 4538-4542

**Truchet G, Roche P, Leronge P, Vasse J, Camet S, de Billy F, Prome JC, Denarie J (1991)** Sulphated lipo-oligosaccharide signals of *Rhizobium meliloti* elicit root nodule organogenesis in alfalfa. *Nature* **351**: 670-673

**Udenfriend S, Kodukula K (1995)** How glycosylphosphatidylinositol-anchored membrane proteins are made. *Ann Rev Biochem* **64**: 563-591

**Van de Loo FJ, Turner S, Somerville C (1995)** Expressed sequence tags from developing castor seeds. *Plant Physiol* **108**: 1141-1150

**Vidugiriene J, Menon AK (1994)** The GPI anchor of cell-surface proteins is synthesized on the cytoplasmic face of the endoplasmic reticulum. *J Cell Biol* **127**: 333-341

**Vitale A, Ceriotti A, Denecke J (1993)** The role of the endoplasmic reticulum in protein synthesis, modification and intracellular transport. *J Exp Bot* **44**: 1417-1444

**Vitale A, Bielli A, Ceriotti A (1995)** The binding protein associates with monomeric phaseolin. *Plant Physiol* **107**: 1411-1418

**von Heijne G (1985)** Signal sequences. The limits of variation. *J Mol Biol* **184**: 99-105

**Wada I, Rindress D, Cameron PH, Ou WJ, Doherty JJ, Louvard D, Bell AW, Dignard D, Thomas DY, Bergeron JJM (1991)** SSR $\alpha$  and associated calnexin are major calcium binding proteins of the endoplasmic reticulum membrane. *J Biol Chem* **266**: 19599-19610

**Walter P, Blobel G (1981)** Translocation of proteins across the endoplasmic reticulum II. Signal recognition protein (SRP) mediates the selective binding to microsomal membranes of in-vitro-assembled polysomes synthesizing secretory protein. *J Cell Biol* **91**: 551-556

**Wanker EE, Sun Y, Savitz, Meyer DI (1995)** Functional characterization of the 180 kDa ribosome receptor in vivo. *J Cell Biol* **130**: 29-39

**Warren G (1993)** Membrane partitioning during cell division. *Annu Rev Biochem* **62**: 323-348

**Waters MG, Serafini T, Rothman JE (1991)** 'Coatomer': a cytosolic protein complex containing subunits of non-clathrin-coated Golgi transport vesicles. *Nature* **349**: 248-251

**Wessel D, Flugge IU (1984)** A method for the quantitative recovery of protein in a dilute solution in the presence of detergents and lipids. *Anal Biochem* **138**: 141-143

**Wickner WT, Lodish HF (1985)** Multiple mechanisms of protein insertion into and across membranes. *Science* **230**: 400-407

**Wieland FT, Gleason ML, Serafini TA, Rothman JE (1987)** The rate of bulk flow from the endoplasmic reticulum to the cell surface. *Cell* **50**: 289-300

**Xu L, Paulsen AQ, Ryu SB, Wang X (1996)** Intracellular localization of phospholipase D in leaves and seedling tissues of castor bean. *Plant Physiol* **111**: 101-107

