Proteome Analysis of Grain Filling and Seed Maturation in Barley

Christine Finnie, Sabrina Melchior, Peter Roepstorff, and Birte Svensson*

Department of Chemistry, Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK–2500 Valby, Copenhagen, Denmark (C.F., B.S.); and Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK–5230 Odense M, Denmark (S.M., P.R.)

In monocotyledonous plants, the process of seed development involves the deposition of reserves in the starchy endosperm and development of the embryo and aleurone layer. The final stages of seed development are accompanied by an increase in desiccation tolerance and drying out of the mature seed. We have used two-dimensional gel electrophoresis for a time-resolved study of the changes in proteins that occur during seed development in barley (Hordeum vulgare). About 1,000 low-salt extractable protein spots could be resolved on the two-dimensional gels. Protein spots were divided into six categories according to the timing of appearance or disappearance during the 5-week period of comparison. Nineteen different proteins or protein fragments in 36 selected spots were identified by matrix-assisted laser-desorption ionization time of flight mass spectrometry (MS) or nano-electrospray tandem MS/MS. Some proteins were present throughout development (for example, cytosolic malate dehydrogenase), whereas others were associated with the early grain filling (ascorbate peroxidase) or desiccation (Cor14b) stages. Most noticeably, the development process is characterized by an accumulation of low-M<sub>ᵦ</sub> α-amylase/trypsin inhibitors, serine protease inhibitors, and enzymes involved in protection against oxidative stress. We present examples of proteins not previously experimentally observed, differential extractability of thiol-bound proteins, and possible allele-specific spot variation. Our results both confirm and expand on knowledge gained from previous analyses of individual proteins involved in grain filling and maturation.

Barley (Hordeum vulgare) is an important crop in northern Europe, where it is used both for feed and in the malting industry. The characteristics of a good malting barley cultivar are quite different from those of a feed barley cultivar and there is an interest to develop a greater knowledge of the determinants of malting quality at the molecular level, to improve the evaluation of new varieties. Therefore, an understanding of the processes occurring during seed development and the proteins involved is required.

Several complex and interlinked processes are involved in cereal seed development (for review, see Bewley and Black, 1978). The starchy endosperm and aleurone layer are formed (for review, see Olsen, 2001) and storage proteins, lipids, and polysaccharides are deposited in the endosperm. These molecules will be broken down later by hydrolases as part of the germination process, to supply the growing embryo with sugars and amino acids. Development of the embryo is accompanied by a buildup of desiccation tolerance (Ingram and Bartels, 1996). The maturation process is marked by desiccation of the seed and the onset of dormancy. The mature seeds contain only about 10% to 15% water (Olsen et al., 1999). The starchy endosperm cells do not survive this desiccation and undergo programmed cell death (Young et al., 1997), whereas the aleurone and embryo stay alive but maintain a basal level of metabolic activity (Leprince et al., 1993).

Proteome analysis is a tool that can be used both to visualize and compare complex mixtures of proteins and to gain a large amount of information about the individual proteins involved in specific biological responses. Recently, a number of plant proteome studies have been published, some of them more descriptive, providing an overview of proteins present in a given tissue, organelle, or stage of development (Kristoffersen and Flengsrud, 2000; Peltier et al., 2000; Prime et al., 2000; Skylas et al., 2000; Gallardo et al., 2001; Haebel and Kehr, 2001; Kruft et al., 2001; Millar et al., 2001; Porubleva et al., 2001), whereas some have been more directed, using the techniques of proteome analysis to address a specific biological question such as the role of jasmonate in defense signaling (Rakwal and Komatsu, 2000) or to identify proteins phosphorylated in response to bacterial or fungal elicitors (Peck et al., 2001). The plant species studied include barley, maize (Zea mays), and rice (Oryza sativa), as well as the model plant Arabidopsis. In fact, two-dimensional electrophoresis techniques were established early on for barley (Görg et al., 1988; Hurkman and Tanaka, 1988; Flengsrud and
Kobro, 1989). Two-dimensional gel analysis has also been used previously to analyze differences in malt-
ing quality of barley cultivars (Görg et al., 1992a, 1992b); however, these analyses focused mainly on the hordeins, which are the major storage proteins in barley. Skylas et al. (2000) used proteome analysis to study wheat (Triticum aestivum) endosperm, but again, because a total protein extract was used, the protein patterns observed were dominated by the wheat storage proteins, glutenins, and gliadins.

Identification of proteins from two-dimensional gels using mass spectrometry (MS) relies on sequence information in the databases. This is clearly a limitation for barley, for which the genome has not been sequenced, although closely related proteins can often be identified from the corresponding rice sequence. The SWISSPROT, TREMBL, and TREMBL NEW databases currently (released October 25, 2001) contain 960 sequences from barley and 8,870 sequences from rice. Despite the limited amount of barley sequence information available, we decided to use two-dimensional electrophoresis and MS techniques to conduct a time-resolved analysis of the seed development process in barley, with emphasis on the water-soluble protein fraction. By avoiding extraction of hordeins, a greater number of proteins involved in the seed development process can be addressed.

By observing the way in which different proteins show changing patterns of appearance, a clearer overview is obtained of the events of seed development than by a comparison of specific developmental stages. Our aim was to visualize the changes in the protein spot pattern, identify some of the proteins involved, and relate them to the processes known to be taking place in the seed. The cv Barke, a good malting barley, was used as a standard. In addition, protein patterns from Barke were compared with three other cultivars with varying properties, to see how individual characteristics might be reflected by the presence of specific protein spots. To our knowledge, this is the first time-resolved proteome study of seed development.

RESULTS AND DISCUSSION

Seed Collection

Field-grown seeds were collected at weekly intervals over a period of 5 weeks during the grain filling and maturation stage. According to the Zadoks scale of cereal grain development (Zadoks et al., 1974), the five samples of barley cv Barke corresponded to the following stages, respectively: 80 (start of dough development), 82, 85 (soft dough, onset of drying), 86, and 87 (hard dough). Seeds from three other barley cultivars with varying properties were likewise collected for comparison because they would be expected to differ in the protein expression patterns visible by two-dimensional gel electrophoresis. Barke is a good malting barley currently popular in Denmark. barley cv Mentor is a moderate malting barley that has low \( \alpha \)-amylase activity, whereas cv Meltan is a feed barley with unfavorable malting properties. These cultivars are two-rowed, whereas the fourth cultivar, Morex, is a six-rowed North American malting barley.

Changes in Protein Patterns during Seed Development

The progress of barley cv Barke seed development was accompanied by a change in the pattern of soluble proteins as visualized by two-dimensional gel electrophoresis (Fig. 1). Approximately 1,000 well-defined spots could be resolved on the stage 80 gel, and about 900 spots could be resolved on the stage 87 gel in the pI 4 to 7 region. Conventional standard size

![Figure 1. Two-dimensional gel protein patterns of developing barley seeds.](image-url)
gels are capable of separating about 1,500 to 2,000 protein spots (Rabilloud, 2002). The decrease in number of protein spots during seed development correlated with that seen when comparing protein patterns from developing (689 spots; 17 DPA) and mature (651 spots) wheat endosperm in the same pI interval (Skylas et al., 2000). It may be in part attributable to the large increase in amounts of protein in some spots (for example, see those in Fig. 1, box 3), causing some of the less abundant proteins to disappear from the stained spot pattern because essentially the same amount of protein was loaded on the gels. Alternatively, the decrease in number of protein spots could reflect the greater metabolic activity of the developing seed in comparison with the mature seed. The fact that many more protein spots can be resolved under our conditions than those used for wheat by Skylas et al. (2000) is probably partly due to the presence of large amounts of glutenins and gliadins in the total protein extracts used in that study, which may have obscured less abundant spots. Our extraction procedure was designed to minimize extraction of the abundant barley storage proteins.

Many spots in the four cultivars changed in intensity during the 5-week development period of the present survey. About one-half of the most abundant spots at stage 87 were absent at stage 80, whereas 80% of the most abundant protein spots at stage 80 were also present at stage 87. This apparent discrepancy arises from the fact that some of the proteins that appear during the development process are extremely abundant at stage 87.

The amount of salt-extractable protein in the seeds also changed slightly during development, in accordance with the idea that the proportion of proteins in an inaccessible stored form should increase during development. Typically, the amount of protein extracted from seeds at the latest stage of development was one-third of that obtained from developing seeds under identical extraction conditions (data not shown). This could also influence the protein patterns seen on the two-dimensional gels; however, we decided to compare gels with equal protein load to visualize as many protein spots as possible at the late stage of development without losing resolution at the earlier stage.

Different patterns of spot variation could be observed and correlated with the early, middle, and late stages of dough development. Spots were classified into the following groups: O, present over the entire 5-week interval; I, early, present during the first weeks and decreasing during development; II, increasing gradually throughout development; III, transient, absent during the first and last weeks; IV, mid, appearance coinciding with the onset of drying and continuing during the desiccation period; V, late, present only at the final stage of development; and VI, variations that did not fit into the other groups. Proteins in selected spots from each group have been identified (Table I).

### Proteins Identified by MS

It is becoming clear, as the number of proteome studies in plants and other organisms increases, that a single gene can give rise to several different protein products (e.g. Porubleva et al., 2001). Proteins that occur as multiple spots on the two-dimensional gel, presumably either due to differing posttranslational modifications or to expression of highly related gene sequences, have also been identified in this study (Table I). Seven spots have thus been shown to contain cytosolic triose phosphate isomerase (spots 7, 9, 95, 120, 121, 122, and 138; Table I; Figs. 1 and 2A). All have a similar M<sub>r</sub>, close to the theoretical value, but have pI values varying from 5.16 to 5.87 around the theoretical value of 5.39. Spot 95 may correspond to the non-modified protein because it has a pI close to the theoretical value. Occurrence of some of these forms is relatively constant during seed development (e.g. spots 95 and 120; group O), whereas the appearance of the others varies in different ways (e.g. spots 121 and 122; group VI; Fig. 2A). We do not yet know the nature of the difference between these forms. Triose phosphate isomerase has been identified in a number of plant proteome studies so far (Flengsrud, 1993; Skylas et al., 2000; Gallardo et al., 2001). This is unsurprising given its relative abundance and central "housekeeping" role in glycolysis. Remarkably, however, it has not previously been identified in multiple protein spots. The different spots may be due to posttranslational modification, or they may represent sequence variants of the protein resulting from expression of related genes. Currently, only one triose phosphate isomerase sequence from barley is present in the SWISSPROT database, although a search of the National Center for Biotechnology Information EST database with the known sequence reveals barley ESTs encoding potential amino acid changes that could result in higher or lower pI values. However, peaks corresponding to tryptic peptides containing these amino acid changes were not found in the mass spectra for any of the spots identified as triose phosphate isomerase, which all matched accession number P34937 (data not shown).

A group of proteins was observed that appeared transiently during seed development, showing a type III pattern of appearance (spots 84–87, Fig. 2B). These were identified as C-terminal fragments of PDI. Tryptic peptides were identified in the mass spectrum that covered the C-terminal third of the 513-residue protein (Fig. 3A; from residue 269 in spots 84 and 85, and residue 293 in spots 86 and 87). Another, more N-terminal fragment of PDI was identified in spot 124 (Fig. 2B). Tryptic peptides covering residues 103 through 234 could be identified in the mass spectrum from this spot (Fig. 3A), suggesting that the
cleavage may occur between residues 234 and 269 of PDI, separating the two thioredoxin domains predicted on the basis of sequence homology to be present in the protein. However, spot 124 shows a slightly different pattern of appearance than spots 84 through 87 and it is not certain that the fragments in spots 84 through 87 and spot 124 originate from the same proteolytic cleavage event. Although no peptides from the N terminus were detected in the MALDI-TOF spectrum of tryptic fragments from spot 124, the apparent Mr of the spot on the two-dimensional gel suggests that the N terminus is intact. In addition to peptides matching accession number P80284 (the only barley PDI sequence in the database), a peptide matching accession number P52589 (wheat PDI) was seen in the mass spectrum (Fig. 3A). This peptide differs from the barley sequence by a Y to D substitution at residue 196 (P80284 numbering). A BLAST search for barley EST sequences confirmed that transcribed barley sequences exist encoding this sequence variation, and in fact no barley EST sequences were found that encoded Y at position 196. MS/MS sequencing data would be required, however, to confirm the sequence of the protein in spot 124. Spots corresponding to full-length PDI have been identified on the two-dimensional gel.

Table 1. Protein identification by MS

<table>
<thead>
<tr>
<th>Spot</th>
<th>Groupa</th>
<th>Matched Protein</th>
<th>Exp. Size (kD)/pl</th>
<th>Calculated Size kD/pl</th>
<th>Sequence Coverage</th>
<th>Accession No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>O</td>
<td>UDP-Glc pyrophosphorylaseb</td>
<td>59.5/5.42</td>
<td>51.6/5.20</td>
<td>23</td>
<td>Q43772</td>
</tr>
<tr>
<td>5</td>
<td>O</td>
<td>Enolaseb</td>
<td>59.5/5.70</td>
<td>47.9/5.42</td>
<td>32</td>
<td>Q42971</td>
</tr>
<tr>
<td>77</td>
<td>O</td>
<td>Cytosolic triose phosphate isomerase (maize)</td>
<td>42.3/6.11</td>
<td>35.9/5.8</td>
<td>26</td>
<td>T02935</td>
</tr>
<tr>
<td>9</td>
<td>O</td>
<td>Cytosolic triose phosphate isomerase</td>
<td>26.8/5.87</td>
<td>26.6/5.39</td>
<td>17</td>
<td>P34937</td>
</tr>
<tr>
<td>120</td>
<td>O</td>
<td>Cytosolic triose phosphate isomerase</td>
<td>26.8/5.16</td>
<td>26.6/5.39</td>
<td>23</td>
<td>P34937</td>
</tr>
<tr>
<td>7</td>
<td>O</td>
<td>Cytosolic triose phosphate isomerase</td>
<td>26.8/5.33</td>
<td>26.6/5.39</td>
<td>22</td>
<td>P34937</td>
</tr>
<tr>
<td>6</td>
<td>O</td>
<td>Glyoxalase I (rice)b</td>
<td>37.6/5.53</td>
<td>32.5/5.51</td>
<td>42c</td>
<td>Q9ZWJ2</td>
</tr>
<tr>
<td>149</td>
<td>O</td>
<td>Glyoxalase I (rice; C-terminal fragment)</td>
<td>30.0/5.30</td>
<td>32.5/5.51</td>
<td>47c</td>
<td>Q9ZWJ2</td>
</tr>
<tr>
<td>79</td>
<td>I</td>
<td>Cytosolic ascorbate peroxidase</td>
<td>30.0/6.37</td>
<td>27.4/5.85</td>
<td>46</td>
<td>O23983</td>
</tr>
<tr>
<td>139</td>
<td>I</td>
<td>Ribulose bisphosphate carboxylase small subunitb</td>
<td>15.1/6.29</td>
<td>19.4/8.98</td>
<td>40</td>
<td>Q40004</td>
</tr>
<tr>
<td>140</td>
<td>I</td>
<td>Ribulose bisphosphate carboxylase small chain (fragment)</td>
<td>14.4/6.29</td>
<td>19.4/8.98</td>
<td>55</td>
<td>Q40004</td>
</tr>
<tr>
<td>22</td>
<td>II</td>
<td>Serpin-protein Z4</td>
<td>42.2/5.48</td>
<td>43.3/5.72</td>
<td>33</td>
<td>P06293</td>
</tr>
<tr>
<td>24</td>
<td>II</td>
<td>Serpin-protein Z4</td>
<td>42.8/5.51</td>
<td>43.5/5.72</td>
<td>16</td>
<td>P06293</td>
</tr>
<tr>
<td>111</td>
<td>II</td>
<td>α-Amylase/trypsin inhibitor IAA1</td>
<td>15.8/5.56</td>
<td>14.6/5.16</td>
<td>37</td>
<td>P16968</td>
</tr>
<tr>
<td>112</td>
<td>II</td>
<td>α-Amylase/trypsin inhibitor IAA1</td>
<td>13.2/5.56</td>
<td>14.6/5.16</td>
<td>40</td>
<td>P16968</td>
</tr>
<tr>
<td>113</td>
<td>III</td>
<td>α-Amylase/trypsin inhibitor pUP13</td>
<td>14.6/5.19</td>
<td>14.7/5.35</td>
<td>32</td>
<td>225102</td>
</tr>
<tr>
<td>115</td>
<td>III</td>
<td>α-Amylase/trypsin inhibitor IAAD</td>
<td>16.5/5.46</td>
<td>16.1/5.24</td>
<td>50</td>
<td>P11643</td>
</tr>
<tr>
<td>116</td>
<td>III</td>
<td>α-Amylase/trypsin inhibitor IA2</td>
<td>12.1/5.60</td>
<td>13.1/5.06</td>
<td>47</td>
<td>P13691</td>
</tr>
<tr>
<td>117</td>
<td>III</td>
<td>α-Amylase/trypsin inhibitor IAAB</td>
<td>14.4/5.52</td>
<td>14.2/5.78</td>
<td>27</td>
<td>P02936</td>
</tr>
<tr>
<td>84</td>
<td>III</td>
<td>Protein disulfide isomerase (C-terminal fragment)</td>
<td>37.2/4.76</td>
<td>53.9/4.49</td>
<td>24</td>
<td>P80284</td>
</tr>
<tr>
<td>85</td>
<td>III</td>
<td>Protein disulfide isomerase (C-terminal fragment)</td>
<td>36.1/4.83</td>
<td>53.9/4.49</td>
<td>25</td>
<td>P80284</td>
</tr>
<tr>
<td>86</td>
<td>III</td>
<td>Protein disulfide isomerase (C-terminal fragment)</td>
<td>37.2/4.87</td>
<td>53.9/4.49</td>
<td>29</td>
<td>P80284</td>
</tr>
<tr>
<td>87</td>
<td>III</td>
<td>Protein disulfide isomerase (C-terminal fragment)</td>
<td>36.1/4.93</td>
<td>53.9/4.49</td>
<td>33</td>
<td>P80284</td>
</tr>
<tr>
<td>80</td>
<td>III</td>
<td>β-Amylase (fragment)</td>
<td>21.3/4.61</td>
<td>59.3/5.85</td>
<td>26</td>
<td>Q9FUJK6</td>
</tr>
<tr>
<td>81</td>
<td>III</td>
<td>β-Amylase (fragment)</td>
<td>21.6/4.68</td>
<td>59.3/5.66</td>
<td>21</td>
<td>Q9FUJK7</td>
</tr>
<tr>
<td>114</td>
<td>IV</td>
<td>α-Amylase/trypsin inhibitor IAA1</td>
<td>16.0/5.08</td>
<td>14.6/5.16</td>
<td>27</td>
<td>P16968</td>
</tr>
<tr>
<td>118</td>
<td>IV</td>
<td>α-Amylase/trypsin inhibitor IAA1</td>
<td>15.1/5.89</td>
<td>13.1/5.51</td>
<td>23</td>
<td>P28041</td>
</tr>
<tr>
<td>119</td>
<td>IV</td>
<td>α-Amylase/trypsin inhibitor IAA1</td>
<td>13.2/5.08</td>
<td>14.6/5.16</td>
<td>40</td>
<td>P16968</td>
</tr>
<tr>
<td>17</td>
<td>IV</td>
<td>Cold-regulated protein Cor14b</td>
<td>20.9/4.95</td>
<td>17.6/4.93</td>
<td>84</td>
<td>Q9FS5B</td>
</tr>
<tr>
<td>20</td>
<td>IV</td>
<td>Expressed sequence tag (EST) encoding homolog of unknown protein from Arabidopsis</td>
<td>15.1/6.26</td>
<td>15.3/6.13</td>
<td>59e</td>
<td>AF332448</td>
</tr>
<tr>
<td>8</td>
<td>V</td>
<td>Embryo-specific protein (rice)b</td>
<td>29.5/8.25</td>
<td>26.4/5.58</td>
<td>40e</td>
<td>Q9ZN59</td>
</tr>
<tr>
<td>121</td>
<td>VI</td>
<td>Cytosolic triose phosphate isomerase</td>
<td>27.2/5.21</td>
<td>26.6/5.39</td>
<td>21</td>
<td>P34937</td>
</tr>
<tr>
<td>122</td>
<td>VI</td>
<td>Cytosolic triose phosphate isomerase</td>
<td>27.6/5.56</td>
<td>26.6/5.39</td>
<td>27</td>
<td>P34937</td>
</tr>
<tr>
<td>138</td>
<td>VI</td>
<td>Cytosolic triose phosphate isomerase</td>
<td>27.6/5.51</td>
<td>26.6/5.39</td>
<td>31</td>
<td>P34937</td>
</tr>
<tr>
<td>124</td>
<td>VI</td>
<td>Protein disulfide isomerase (fragment)</td>
<td>33.9/5.14</td>
<td>53.9/4.49</td>
<td>26c</td>
<td>P80284/P52589</td>
</tr>
</tbody>
</table>

a Patterns of spot appearance (see main text). O, Present at all stages; I, early, decreasing during development; II, increasing during development; III, transient; IV, mid, coincident with dessication; V, late; VI, other patterns. b Identified or confirmed by nano-electrospray ionization quadrupole (ESI) MS/MS. c Based on translated barley EST sequence identified by Mascot search. d Molecular mass/pl values for protein after removal of theoretical targeting sequence. e Values are given for the full-length protein.
Figure 2. Variations in protein spots during seed development. Close-up views of A, region 2 (22–30 kD, pl 5.0–5.9); B, region 3 (30–45 kD, pl 4.7–5.7); C, region 4 (10–20 kD, pl 6.2–7.0); D, region 5 (20–30 kD, pl 5.1–5.6); and E, region 6 (22–30 kD, pl 6.25–6.6). Numbered spots correspond to Table I. For ease of comparison in C through E, some reference spots have been circled. A. This region shows a group of spots that is abundant at the early and late stages of dough formation but less abundant at the middle stage (upwards arrows). Another spot remains constant throughout the process (horizontal arrow) and yet another spot is present at all stages of development, but most abundant at the final stage (downwards arrow). Spots 7, 9 (Fig. 1), 95, 120 through 122, and 138 have been identified as triose phosphate isomerase. An as-yet unidentified spot has expression pattern III is marked (a). B. Group of spots (spots 84–87) accumulates transiently at the middle stages of development. These have been identified as C-terminal fragments of protein disulfide isomerase (PDI). Another group of spots, including spots 22 and 24, increases gradually both in intensity and number throughout development and are identified as protein Z4 (a Ser protease inhibitor [serpin]). Below the serpin spots, glyoxalase I (spot 6) is present throughout and is particularly abundant at the latest stage. A fragment of this protein is also present (spot 149). C. Spot 140 is probably a degradation product of spot 139. These spots decrease gradually in intensity during development (group I). Both have been identified as the small subunit of ribulose bisphosphate carboxylase. Spot 20 (group IV, desiccation related) is a glyoxalase I-related protein (see text). Another, unidentified protein (spot b) has the same pattern of appearance. D. Spot 149 (glyoxalase I fragment) appears to alter position during development. An as-yet unidentified spot with expression pattern I (c) is marked. E. Spot 79, also with expression pattern I, has been identified as cytosolic ascorbate peroxidase (APX). Spots with expression pattern IV (d) and O (e) are marked. Spot d has been identified as a 1cys peroxiredoxin (O. Østergaard, C. Finnie, S. Melchior, P. Roepstorff, and B. Svensson, unpublished data), whereas spot e is not yet identified.
This suggests that the N terminus of the mature protein is present but is modified. The peptide with 
\([\text{M} + \text{H}] = 1,214.92\) was sequenced by MS/MS fragmentation and had the expected sequence of the N
terminus of Rubisco small subunit after removal of the targeting sequence. The masses of the fragment ions showed that the 14-D modification was present on the N-terminal Met residue. This mass difference can be accounted for by the addition of a methyl group. In fact, N-methylation of the N-terminal Met of Rubisco small subunit has been observed previ-
ously (Grimm et al., 1997). The mass spectrum from spot 139 also contained a peak corresponding to the C-terminal peptide of the protein. This was absent in spot 140, supporting the evidence from the two-dimensional gel pattern that the protein in spot 140 is a degradation product of that in spot 139.

We presume that the ribulose bisphosphate carboxylase in spots 139 and 140 originates from the pericarp, which is green and photosynthetic before maturation of the seed. The amount of spot 139 gradually declines during development, in parallel with seed
maturation. As spot 139 decreases in intensity, two spots (spot 20 and spot b in Fig. 2C) appear adjacent to it and become very abundant in stage 87 seeds. The mass spectrum from spot b did not contain peaks arising from Rubisco, and no identification resulted from peptide mass mapping.

Peptide mass mapping of spot 20 did not indicate the presence of Rubisco and did not result in a positive identification in routine searches against National Center for Biotechnology Information and SWISSPROT databases. However, using the Mascot server to search against the EST database resulted in a significant match to a translated barley EST sequence (accession no. BI780314). Because the EST sequence was apparently not full length, it was used in a BLAST search of The Institute for Genomic Research Barley Gene Index (http://www.tigr.org/tdb/hvgi) and a tentative consensus sequence was obtained (TC8995) that contained an open reading frame encoding a protein with predicted mass of 15.3 kD and pI 6.13. This was in agreement with the experimental values for spot 20 (Table I). A previously unassigned peak in the MALDI spectrum for spot 20 was now found to correspond to the N terminus of the protein that was not within the region encoded by the originally identified EST, strongly supporting the identification.

The protein sequence was 59% identical to a protein of unknown function from Arabidopsis (accession no. AF332448; Table I). Both proteins contained a single signature motif for glyoxalase I, and were 20% to 30% identical to several bacterial glyoxalase sequences. The glyoxalase system, comprising glyoxalases I and II, may be involved in detoxification of the cytotoxic metabolite methylglyoxal that can be produced by increased levels of glycolysis under conditions of stress (Iyengar and Rose, 1981). Glyoxalase I gene expression in plants has been shown to be up-regulated in response to stress (Espartero et al., 1995; Romo et al., 1998) and overexpression of the gene confers stress tolerance to transgenic tobacco (Nicotiana tabacum; Veena et al., 1999).

Bacterial and many eukaryotic glyoxalase I proteins function as dimers, each monomer containing a single glyoxalase domain. However, the yeast (Saccharomyces cerevisiae) and most plant enzymes described to date have two tandem glyoxalase domains in a single polypeptide, and are thought to result from a gene duplication event (Clugston et al., 1998). Glyoxalase I from tomato (Lycopersicon esculentum; accession no. Q42891) has a single domain but is longer than the Arabidopsis and barley sequences and only shares about 20% sequence identity. Therefore, the barley and Arabidopsis proteins may belong to a different subset of single-domain enzymes. Glyoxalases are metalloproteins that require zinc for activity. A sequence alignment of single-domain glyoxalases showed that the residues involved in Zn$^{2+}$ binding in the human enzyme (for which the three-dimensional structure is known; Cameron et al., 1997) were conserved in tomato but not in the barley and Arabidopsis sequences. This suggests that these proteins may not have glyoxalase activity. In this context, it is of interest that a glyoxalase I with the duplicated domain structure has also been identified by Mascot searching of the EST database and confirmed by MS/MS sequencing (spots 6 and 149, group O; Fig. 2B). The identified protein is highly homologous to glyoxalase I from rice. The abundance of this protein in the two-dimensional gel pattern probably reflects the high metabolic activity of the developing seeds.

Proteins Involved in Oxidative Stress

In addition to homologs of glyoxalase I, several proteins have been identified that have potential roles in responses to oxidative stress. These proteins are expressed at varying times throughout the development process, reflecting the importance of protection against desiccation-induced injury due to active oxygen species produced during seed development (for review, see Noctor and Foyer, 1998). Ascorbate peroxidase is the major enzyme involved in the ascorbate-glutathione cycle for detoxification of hydrogen peroxide (Asada, 1992). The cytosolic form of the enzyme is present early on in grain filling (group I) and cannot be detected at later stages (spot 79, Fig. 2E). A protein with expression pattern IV (spot d, Fig. 2E) has been identified as 1cys-peroxiredoxin (Ø. Østergaard, C. Finnie, S. Melchior, P. Roepstorff, and B. Svensson, unpublished data). The 1cys-peroxiredoxin proteins have a single conserved Cys residue and display peroxidase activity when coupled to a thiol-reducing system. It has been proposed that 1cys-peroxiredoxin might be involved in dormancy because it is expressed during imbibition in dormant seeds, whereas its expression decreases in nondormant seeds (Stacy et al., 1999). However, overexpression of rice 1cys-peroxiredoxin in transgenic tobacco enhanced oxidative stress tolerance, but did not affect dormancy (Lee et al., 2000). We have also identified a 2cys-peroxiredoxin, a thioredoxin-dependent antioxidant, from developing barley endosperm (C. Finnie, unpublished data). This protein has also been identified from wheat endosperm (Skylas et al., 2000).

α-Amylase/Trypsin Inhibitor Spots

One of the most noticeable differences in the two-dimensional gel patterns when comparing developing Barke stage 80 and stage 87 seeds is the appearance of a group of intense spots of low molecular mass (10–18 kD) and pI 4.5 of 5.5 (Fig. 1, region 1). This group of spots increases gradually throughout seed development (Fig. 4A), showing a type II pattern of appearance. A number of these proteins were identified by MALDI-TOF MS as different α-amy-
lase/trypsin inhibitors (Table I). The endogenous barley \(\alpha\)-amylases, which are de novo synthesized during germination, are not targets for these inhibitors; instead, it has been shown that some of these proteins inhibit the \(\alpha\)-amylase activity of extracts from various insect pests (Gutierrez et al., 1990). Therefore, it is presumed that their role is to defend the starch reserves of the seed against invading insect pathogens and it is expected that these proteins are seen to accumulate in parallel with grain filling.

Barley cv Mentor and cv Meltan were found to differ in the pattern of spots containing \(\alpha\)-amylase/trypsin inhibitors when comparing similar stages of development. In particular, a spot corresponding to the \(\alpha\)-amylase trypsin inhibitor (IAAB inhibitor; spot 117, Fig. 4B, bottom; Table I), present in barley cv Barke and cv Meltan, was absent in cv Mentor (Fig. 4B, top). The mature seed patterns of these cultivars were compared to determine whether this inhibitor appeared later in development in Mentor. Surprisingly, even fewer inhibitor spots were present in the mature seed extract from barley cv Mentor than in the stage 86 to 87 extract (Fig. 4B, middle). In addition to the IAAB protein, spots corresponding to the IAA2 and IAAD inhibitors (Fig. 4B, spots 116 and 115; Table I) were absent in harvested seeds from barley cv Mentor (Fig. 4B, middle). The pattern of spots in stage 87 and mature seeds of barley cv Barke, however, was unchanged.

These small proteins are very Cys rich, containing four or five disulfide bonds, and it is possible that they may be stored in thiol-bound form in the seed, as is the case for other abundant proteins with storage roles in the seed such as serpins and \(\beta\)-amylases.
Therefore, it was speculated whether these proteins were only transiently produced during development in barley cv Mentor, or whether they were produced and then sequestered in a thiol-bound form that was inaccessible to the applied extraction procedure. To examine this question, a sequential extraction procedure was carried out, in which the material remaining after the routine extraction of soluble proteins was re-extracted with the same buffer containing 20 mM DTT to release thiol-bound proteins. This treatment is known to release proteins like β-amylase and serpins from a thiol-bound form in the seed (Rosenkrands et al., 1994). Comparison of the thiol extracts with the standard extracts from the same material showed an increase in intensity of the serpin spots as well as a number of other spots (not shown). The pattern obtained with the thiol extracts from barley cv Mentor showed that the IAA2, IAAB, and IAAD proteins were released into the extracts by this treatment (Fig. 4B, bottom). The thiol extraction procedure did not greatly affect the pattern observed for barley cv Barke, suggesting a cultivar-specific availability of proteins putatively involved in defense against insect pests.

This result suggests that rather than differing in the number and/or amount of α-amylase/trypsin inhibitors synthesized during grain filling, these cultivars differ greatly in the degree to which these proteins are sequestered in a thiol-extractable form in the mature seed. We can determine by comparing the spots present at different stages of development that the event leading to the changed extractability of these proteins in barley cv Mentor apparently occurs late in the development process. We are currently investigating the possible reasons for the difference between the cultivar spot patterns. The fact that other proteins are also released into the extracts by thiol extraction suggests that the α-amylase/trypsin inhibitors may not be the only proteins affected.

Allele-Specific Spots of β-Amylase

It is to be expected that when comparing spot patterns between different barley cultivars that some of the observed differences might be due to different alleles giving rise to proteins with slightly different mobilities on the two-dimensional gel. Previously, Görg et al. (1988) observed spot displacements possibly due to allele differences between barley cultivars. At that time, mass spectrometric techniques were not available for identification of proteins from silver-stained gels, so this interpretation could not easily be confirmed.

A possible example of allele-specific spots is shown in Figure 5A. Comparison of stage 85 and 86 gels shows that barley cv Morex has two spots that correspond to a similar $M_r$ but with pI values differing by approximately 0.1 unit (spots 80 and 81, Fig. 5A). Spot 80 is also present in barley cv Barke and cv Meltan, whereas only spot 81 is present in barley cv Mentor. Both of these spots have been identified as fragments of β-amylase, truncated both at the N and C termini. Almost 100% sequence coverage was obtained by MALDI-TOF peptide mapping for residues 109 through 291 (data not shown). Judging by the apparent molecular mass of the fragments on the two-dimensional gel, it is likely that the cleavage sites are close to these positions. Because the amount of β-amylase is expected to increase during seed development, the transient nature of these β-amylase fragments (group III) may be due to a transient pros...
teolytic activity. These spots showed the same pattern of appearance when examining thiol extractions of the developing seeds (data not shown).

Detailed inspection of the mass spectra obtained from the two barley cv Morex spots (Fig. 5B) showed that a peak with $[M + H] = 2,025.86$ in spot 80 was replaced by a peak with $[M + H] = 2021.93$ in spot 81. These peaks correspond to tryptic peptides covering amino acids 109 through 126 from allelic \( \beta \)-amylase sequences containing a Cys or an Arg residue, respectively, at position 115. This amino acid substitution is sufficient to explain the pI difference between the two spots. Three alleles of \( \beta \)-amylase have been identified at the \( B my1 \) locus, two of which have Arg at position 115 and one of which has Cys (Ma et al., 2001). It has been demonstrated that the C115R substitution increases the \( K_m \) of \( \beta \)-amylase for soluble starch by 2.5-fold (Ma et al., 2001). Thus, the allele differences of the cultivars could be reflected in the diastatic power of the malt, which is largely determined by the \( \beta \)-amylase activity. The same amino acid substitution also affects the ability of the protein to form intermolecular disulfide bridges that might be important for storage of the protein within the seeds (Ma, 2001). However, it will be necessary to identify the spots corresponding to the full-length protein(s) before the significance of the spot differences shown here can be related to cultivar-specific differences in \( \beta \)-amylase activity. The extraction procedure was designed to minimize the risk of proteolysis; however, we cannot be sure of the biological significance of these fragments until it is clear whether the cleavage occurs during seed development or during the extraction procedure.

Proteins of Unknown Function

Some proteins have been identified in the present work for which the function, or the specific role in seeds, is not known. The barley cold-regulated protein Cor14b (spot 17, Fig. 1) has been shown to be involved in cold tolerance of seedlings (Crosatti et al., 1999). It has not been observed previously in seeds, but here we show that it is seen to accumulate during the desiccation stage of seed development.

A fairly abundant protein accumulating late in development (spot 8, Fig. 2A) has been identified as a homolog of the rice sequence Q9ZNS9. No function has been published for this protein and it shares no homology with proteins of known function.

CONCLUSIONS

Studying changes in individual protein spots over the seed development period enabled classification according to expression patterns. These patterns can be related to the events taking place in the seed, based on the proteins identified within these groups. Group O spots, present throughout the experimental period, included “housekeeping” proteins important in basic metabolism (triose phosphate isomerase, malate dehydrogenase, and enolase). Group I spots, characteristic for the immature seed, included a photosynthetic enzyme not expected to be present in mature seeds (ribulose bisphosphate carboxylase). Group II spots, which increase gradually throughout seed development, included proteins involved in both deposition of storage proteins and in defense of the stored reserves against insect pests (serpins and \( \alpha \)-amylase/trypsin inhibitors). Group III spots appeared transiently, and were identified as proteolytic fragments of PDI and \( \beta \)-amylase, providing evidence for a transient proteolytic activity at this development stage. Group IV spots, which increased at the desiccation stage, included proteins possibly involved in stress responses (putative glyoxalase I and Cor14b), as well as \( \alpha \)-amylase/trypsin inhibitors related to the group II pattern. Only one protein, for which no function is known (spot 8, Fig. 2A; Table I), was identified from the spots with Group V (late) expression.

Notably, different forms of the same protein can be found in different groups (e.g. triose phosphate isomerase), and different proteins with similar functions can be seen to have similar expression patterns (e.g. several \( \alpha \)-amylase/trypsin inhibitors). The fact that of the different proteins identified, several are involved in oxidative stress responses and nine are associated with storage protein deposition can be taken as a measure of the importance of these processes during seed development.

The use of MALDI-TOF MS and ESI MS/MS to identify proteins has enabled us to distinguish between varying forms of proteins (e.g. PDI and \( \beta \)-amylase) that may have consequences for their function. We have also demonstrated the potential for identification of posttranslational modification of proteins on two-dimensional gels (e.g. Rubisco small subunit). It is also possible, based on the available information in EST databases, to identify previously unknown proteins (e.g. glyoxalase-like protein in spot 20).

By comparing different barley cultivars, we have identified variations in spot pattern in \( \alpha \)-amylase/trypsin inhibitors that apparently do not arise directly due to differences in gene expression or obvious posttranslational modification but by a mechanism affecting the accessibility or extractability of the proteins.

In conclusion, the techniques of two-dimensional gel electrophoresis and protein identification by MS can be used to visualize and describe the complex metabolic processes occurring during seed development. Here, we have concentrated on the proteins within the pI range 4 through 7. Addition of the high-pI proteins will provide a more comprehensive analysis. More work is also required to relate the
changes in protein expression that we observe to functional characteristics of barley cultivars.

MATERIALS AND METHODS

Plant Material

Spring barley (Hordeum vulgare cv Barke, cv Meltan, cv Mentor, and cv Morex) was field grown in Fyn, Denmark, in the summer of 2000, under the supervision of Sejet Plantbreeding (Sejet, Horsens, Denmark). At the onset of grain filling, seeds were collected at weekly intervals until maturity. The plants were harvested 2 weeks after collection of the final sample. At each collection stage, the time of development was estimated using the Zadoks scale (Zadoks et al., 1974), and 40 ears were cut and immediately frozen in liquid nitrogen. Samples were stored at −80°C until needed.

Protein Extraction

The barley ears were freeze dried for 48 h before extraction, and awns and stalks were removed. Seeds from 10 ears were milled to flour in a water-cooled mill. Approximately 4 g of flour was added to 20 mL of extraction buffer (5 mM Tris, pH 7.5; and 1 mM CaCl2) at 4°C. From this stage onwards, all manipulations were carried out at or below 4°C. The flour was extracted with stirring for 30 min and insoluble material was removed by centrifugation at 16,000 rpm for 30 min (JA-20 rotor, Beckman Instruments, Fullerton, CA). The supernatant containing the soluble protein fraction was aliquoted and stored at −80°C until required. In some cases, the insoluble pellet was re-extracted for 30 min with 20 mL of extraction buffer containing 20 mM DTT to release thiol-bound proteins. Insoluble material was removed by centrifugation and the supernatant containing thiol-extractable proteins was stored at −80°C until required. Protein concentrations in the extracts were estimated using the Bradford (1976) or Popov et al. (1975) methods, using bovine serum albumin as standard. To enable sufficient protein to be loaded on the two-dimensional gel, thiol extracts were concentrated by precipitation (4 volumes of acetone for 2 h at −20°C).

Two-Dimensional Gel Electrophoresis

Isoelectric focusing (IEF) of approximately 40 μg of protein in reswelling buffer (8% urea; 2% w/v CHAPS; 0.5% w/v IPG buffer 4–7; 20 mM DTT; and 0.01% w/v bromphenol blue) was run using immobilized pH gradient 18-cm 4–7L IPG strips on an IECphor (Amersham-Pharmacia Biotech, Uppsala; 6 h at 30 V, 6 h at 60 V, 1 h at 200 V, 1 h at 500 V, 30 min at 1,000 V, gradient to 8,000 V, and hold at 8,000 V until a total of at least 63,000 V h−1 was reached). After IEF, IEF strips were equilibrated for 20 min in equilibration buffer (50 mM Tris-HCl, pH 8.8; 6% urea; 30% [v/v] glycerol; 2% [w/v] SDS; and 0.01% w/v bromphenol blue) containing 10% [v/v] DTT, followed by 20 min in equilibration buffer containing 25% [v/v] iodoacetamide. Second dimension SDS-PAGE gels (12%–14%, 18 × 24 cm, Amersham-Pharmacia Biotech) were run on a Pharmacia Multiphor II according to the manufacturer’s recommendations. Gels were stained with silver nitrate in a gel stainer (Hoeffer, San Francisco) according to Shevchenko et al. (1996).

Protein patterns resulting from duplicate protein extractions and duplicate two-dimensional gels were compared with ensure reproducibility. The same variations in protein appearance could be observed in all gels and the spot pattern was found to change in a continuous manner during development. An additional control was provided by comparing equivalent extracts from the four cultivars because most protein spots were common to them all. To avoid estimation of relative spot intensities, assignment of protein spots to groups was based only on the presence or absence of the spot at each stage of development examined.

In-Gel Digestion of Protein Spots

Spots were cut out from silver-stained gels and subjected to in-gel trypsin digestion according to Shevchenko et al. (1996). After soaking trypsin (modified porcine trypsin, sequencing grade, Promega, Madison, WI) into the gel pieces, the supernatant containing excess trypsin was removed and the gel pieces were covered with 60 μL of 50 mM NH4HCO3 and incubated at 37°C overnight. The supernatant containing trypptic peptides was transferred to a clean tube and 10 μL was then used for micropurification of peptides and subsequent analysis by MALDI-TOF MS. Peptides were desalted and concentrated according to Gobom et al. (1999). A homemade 5-mm nanocolumn was packed with POROS R2 resin (Applied Biosystems, Foster City, CA) in a constricted GELoader tip (Eppendorf Scientific, Westbury, NY). For analyses by MALDI-MS, the peptides were eluted with 0.8 μL of matrix solution (20 mg mL−1 α-cyano-hydroxycinnamic acid in 70% [v/v] CH3CN and 0.1% [w/v] trifuluroacetate) and deposited directly onto the MALDI target. For tandem MS, the peptides were eluted from the column with 1 μL of 50% (v/v) methanol/49% (v/v) water/1% (v/v) formic acid directly into a precoated borosilicate nano-electrospray needle.

Protein Identification

Peptide Mass Mapping by MALDI-TOF-MS

A REFLEX MALDI-TOF mass spectrometer (Bruker-Daltonics, Bremen, Germany) in positive ion reflector mode was used to analyze tryptic peptides. The m/z software (Proteometrics, New York) was used to analyze spectra. Spectra were calibrated using trypsin autolysis products (m/z 842.51 and 2,211.10) as internal standards. Protein identification was performed by searching in a nonredundant protein sequence database program using the Protein (http://www.proteomecommons.com) and/or Mascot (http://www.matrixscience.com) servers. The following parameters were used for database searches with MALDI-TOF-peptide mass data: mono-isotopic mass accuracy, 80 ppm; missed cleavages, 1; and allowed modifications, carbamidomethylation of Cys (complete) and oxidation of Met (partial). To qualify as a positive identification, over 15% sequence coverage was required, including at least three independent peptides with a mass deviation of less than 50 ppm.

Amino Acid Sequencing by Nano-Electrospray MS/MS

Tandem MS of peptides was performed on an ESI-Q-TOF mass spectrometer (Micromass, Manchester, UK). The instrument was calibrated using NaI. Database searching was performed using the Mascot program (http://www.matrixscience.com). Data were processed with a Mass Lynx version 3.5 system (Micromass, Manchester, UK). Positively identified spots were given reference numbers in our laboratory database.

ACKNOWLEDGMENTS

We thank Mette Herson Bien (Carlsberg Laboratory, Valby, Denmark) for excellent technical assistance, Kate Rafn (University of Southern Denmark, Odense) for help with electrospray MS/MS analysis, Jørgen Larsen (Carlsberg Research Laboratory, Valby, Denmark) for advice regarding barley cultivars, Kurt Hjortsholm and Birger Eriksen (Sejet Plantbreeding) for field-grown plants, and Kristian Bak-Jensen (Carlsberg Laboratory), Ole Østergaard (Carlsberg Laboratory), and members of the SUE project “The Barley Proteome: Identification of Proteins Important in Malting and Nutrient Uptake” for helpful discussions.

Received February 8, 2002; returned for revision March 11, 2002; accepted March 25, 2002.

LITERATURE CITED


Bewley JD, Black M (1976) Physiology and Biochemistry of Seeds. Springer-Verlag, Berlin


Copyright © 2002 American Society of Plant Biologists. All rights reserved.


Young TE, Gallie DR, DeMason DA (1997) Ethylene-mediatedprogrammed cell death during maize endosperm development of wild-type and shrunk2 genotypes. Plant Physiol 115: 737–751