RESEARCH COMMUNICATION

Identification of a *Caenorhabditis elegans* Δ^6-fatty-acid-desaturase by heterologous expression in *Saccharomyces cerevisiae*

Johnathan A. NAPIER[1], Sandra J. HEY, Dominic J. LACEY and Peter R. SHEWRY

IACR-Long Ashton Research Station, Department of Agricultural Sciences, University of Bristol, Long Ashton, Bristol BS18 9AF, U.K.

We identified a cDNA expressed sequence tag from an animal (the nematode worm *Caenorhabditis elegans*) that showed weak similarity to a higher-plant microsomal Δ^6-desaturase. A full-length cDNA clone was isolated and expressed in the yeast *Saccharomyces cerevisiae*. This demonstrated that the protein encoded by the *C. elegans* cDNA was that of a fatty acid Δ^6-desaturase, as determined by the accumulation of γ-linolenic acid. The *C. elegans* Δ^6-desaturase contained an N-terminal cytochrome b_5 domain, indicating that it had a similar structure to that of the higher-plant Δ^6-desaturase. The *C. elegans* Δ^6-desaturase mapped to cosmid W08D2, a region of chromosome III. This is the first example of a Δ^6-desaturase isolated from an animal and also the first example of an animal desaturase containing a cytochrome b_5 domain.

INTRODUCTION

Over the last few years, a number of microsomal and soluble fatty acid desaturases have been isolated from higher plants, most notably *Arabidopsis thaliana* (thale cress). This has been achieved by a combined genetic and biochemical approach to the generation and complementation of mutant *Arabidopsis* lines defective in fatty acid desaturation or elongation [1]. The importance of this approach has been clearly validated by the isolation and characterization of genes encoding microsomal desaturases such the Δ^{12} [2] and Δ^{15} [3] (encoded by the *FAD2* and *FAD3* genes respectively) enzymes, which had previously proved intractable to classical purification techniques on account of their hydrophobicity. The isolation of these and related genes, such as the Δ^{12}-hydroxylase from *Ricinus communis* (castor bean) [4], has allowed the identification of a number of conserved motifs in plant microsomal desaturases, most notably the so called 'histidine boxes' [5]. These short motifs appear to be required for enzyme function and also allow the proteins containing these motifs to be classified as di-iron-centre-containing enzymes [6].

Recently we isolated a cDNA clone from borage (*Borago officinalis*), using highly degenerate PCR against these histidine motifs, which was shown by heterologous expression in transgenic tobacco (*Nicotiana tabacum*) to encode a microsomal Δ^6-desaturase [7]. Desaturation at the Δ^6 position is an unusual modification in higher plants, occurring only in a small number of species such as borage, evening primrose (*Oenothera* spp.) and redcurrant (*Ribes* spp.), which accumulate the Δ^6-unsaturated fatty acids γ-linolenic acid (GLA) and octadecatetraenoic acid in the seeds and/or leaves. GLA is a high-value plant fatty acid and is widely used in the treatment of a number of medical conditions, including eczema and mastalgia. It has been postulated that the application of GLA replaces the loss of endogenous Δ^6-unsaturated fatty acids [7]. The sequence of the borage microsomal Δ^6-desaturase differed from previously characterized plant microsomal desaturases/hydroxylases in that it contained an N-terminal extension which showed sequence similarity to cytochrome b_5, and also in that the third (most C-terminal) histidine box varied from the consensus [6] H-X-X-H-H, with a glutamine residue replacing the first histidine one.

Although Δ^6-fatty-acid desaturation is an unusual modification in higher plants, it is a common reaction in animals. The essential fatty acid linoleic acid ($C_{18:2,\Delta9,12}$) is desaturated to GLA by a Δ^6-desaturase as the first step on the biosynthetic pathway of the eicosanoids, which includes prostaglandins and leukotrienes. This results in the rapid metabolism of GLA [to dihomo-GLA ($C_{20:3,\Delta8,11,14}$) and arachidonic acid ($C_{20:4,\Delta5,8,11,14}$)], so accumulation of this fatty acid is not usually observed. For example, in the model animal system, the nematode *Caenorhabditis elegans*, polyunsaturated fatty acids which have been Δ^6-desaturated (in the form of arachidonic and eicosapentanoic acids) make up over 20% of the fatty acids of the total lipids, but no GLA is observed [8]. This is presumably due to its rapid elongation to C_{20} fatty acid derivatives.

We wished to determine whether the Δ^6-desaturase isolated from borage was representative of Δ^6-desaturases as a whole. Since most higher plants do not contain this enzyme [7], we decided to take advantage of the large amount of animal sequences available on public databases. To this end we identified a putative *C. elegans* Δ^6-desaturase expressed sequence tag (EST) and verified its function by expressing the corresponding cDNA in yeast. When the nematode coding sequence was expressed in yeast supplemented by the addition of linoleic acid, GLA was produced. This was confirmed by GC–MS, identifying the coding sequence similar to the *C. elegans* predicted open reading frame (ORF) W08D2.4 as a Δ^6-desaturase.

MATERIALS AND METHODS

The National Center for Biotechnology Information (NCBI) EST sequence database was searched for polypeptide sequences which were related to the higher-plant Δ^6-fatty-acid desaturase [7] and contained the variant histidine box Q-X-X-H-H. Putative positive *C. elegans* ESTs were further characterized by searching the *C. elegans* EST project database (http://www.ddbj.nig.ac.jp/

Abbreviations used: EST, expressed sequence tag; GLA, γ-linolenic acid; NCBI, National Center for Biotechnology Information; ORF, open reading frame.

[1] To whom correspondence should be sent.

htmls/c-elegans/html/ce-index.html) in order to identify related cosmid clones.

A partial cDNA clone identified by these searches was obtained from the *C. elegans* EST project (kindly supplied by Professor Y. Kohara, National Institute of Genetics, Mishima, Japan), and this was used to screen a *C. elegans* cDNA library (mixed stage; also supplied by Professor Y. Kohara) constructed in λZAPII. A number of positives were identified and further purified, and full-length clones were confirmed by sequencing to encode a transcript likely to have been transcribed from the gene designated W08D2.4, on cosmid W08D2, as determined by database searching of the genes sequenced by the *C. elegans* genome project.

The coding sequence of W08D2.4 was introduced into the yeast expression vector pYES2 by PCR. Oligonucleotides with 5′ overhangs were used to introduce *Kpn*I and *Sac*I sites at the 5′ and 3′ ends respectively. The fidelity of the construct was checked by *in vitro* transcription and translation using the *TnT s*ystem (Promega).

The resulting plasmid was introduced into yeast (*Saccharomyces cerevisiae*) by the lithium acetate method [9], and expression of the transgene was induced by addition of galactose. The yeast was supplemented by the addition of 0.2 mM linoleate in the presence of 1 % tergitol, following the method of [10].

Yeast total fatty acids were analysed by GC of methyl esters, exactly as described previously [7]. Confirmation of the presence of GLA was carried out by GC–MS using a Kratos MS80RFA instrument operating at an ionization voltage of 70 eV, with a scan range of 500–40 Da. The mass spectrum of the novel peak resolved by GC was compared with that of an authentic GLA standard (Sigma).

RESULTS

The sequence of the borage Δ^6-desaturase was used to search databases for related sequences in species which, although they do not accumulate GLA, might be expected to perform Δ^6-desaturation. The simplest organism which fulfilled this criterion was the free-living nematode *C. elegans*. This small animal has been subject to both random cDNA (EST) sequencing programs and large-scale genome sequencing. Our searches of EST databases identified a high-scoring nematode EST, namely yk436b12. This partial sequence of 448 bases was used to search for related cosmid clones sequenced by the *C. elegans* genome project, using the DNA database of the Japan *C. elegans* EST project server. This indicated that the clone yk436b12 showed sequence similarity to part of a gene present on cosmid W08D2 (GenBank accession number Z70271), which forms part of chromosome III [11]. Bases 21–2957 of cosmid W08D2 are predicted by the protein prediction program Genefinder [11] to encode an ORF of 473 residues which is interrupted by five introns. Examination of this predicted protein sequence (designated W08D2.4 by the Sanger Centre Nematode Sequencing Project, Hinxton, Saffron Walden, Essex, U.K.) revealed that it had a number of characteristics reminiscent of a microsomal fatty acid desaturase, including three histidine boxes. However, the predicated protein sequence indicated the presence of an N-terminal domain similar to that of cytochrome b_5, containing the diagnostic H-P-G-G motif found in cytochrome b_5 proteins [12]. Since the Δ^6-desaturase isolated by us from borage [7] also contained an N-terminal b_5 domain, this indicated that W08D2.4 may encode a Δ^6-desaturase. Closer examination of the sequence revealed the presence of the variant third histidine box, with a H → Q substitution (again as observed in the borage Δ^6-desaturase). However, the similarity between W08D2.4 and the borage Δ^6-desaturase is low (51.7 %), as is the value of 31.0 % for identity. Since W08D2.4 was encoded by a gene containing many introns, it was necessary to isolate a full-length cDNA to verify the sequence predicated by the Genefinder program [11] and also to allow the expression of the ORF to define the encoded function.

A cDNA library and EST yk436b12 were generously provided by Professor Y. Kohara, and a number of positive plaques were identified by screening with the EST insert. These were further purified to homogeneity, excised, and the largest inserts (∼

```
Boofd6  MAAQIKKYIT  SDELKNHDKP  GDLWISIQGK  AYDVS.DWVK  DHPGGSFPLK  SLAGQEVTDA  59
Ceeld6  ..........  ...MVVDKNA  SGLRMKVDGK  WLYLSEELVK  KHPGGAV.IE  QYRNSDATHI  46

Boofd6  FVAFHPAS..  TWKNLDKF..  ...FTGYYLK  DY........  ...SVSEVS.  KDYRKLVFEF  100
Ceeld6  FHAFHEGSSQ  AYKQLDLLKK  HGEHDEFLEK  QLEKRLDKVD  INVSAYDVSV  AQEKKMVESF  106

Boofd6  SKMGLYDKKG  HIMFA..TLC  FIAMLFAMSV  YGVLFCEGVL  VHLF.SGCLM  GFLWIQSGWI  157
Ceeld6  EKLRQKLHDD  GLMKANETYF  LFKAISTLSI  MAFAFYLQYL  GWYITSACLL  ALAWQQFGWL  166

Boofd6  GHDAGHYMVV  SDSRLNKFMG  IFAANCLSGI  SIGWWKWNHN  AHHIACNSLE  YDPDLQYIPF  217
Ceeld6  THEFCHQQPT  KNRPLNDTIS  LFFGNFLQGF  SRDWWKDKHN  THHAATNVID  HDGDIDLAP.  225

Boofd6  LVVSSKFFGS  LTSHFYEKRL  TFDSLSRFFV  SYQHWTFYPI  MCAARLNMYV  QSLIMLLTKR  277
Ceeld6  ......LFAF  IPGDLCKYKA  SFEKAILKIV  PYQHLYFTAM  LPMLRFSWTG  QSVQWVFKEN  279

Boofd6  NVSYRAHE..  .......LLG  CLVFSIWYPL  LVSCLPNWGE  RIMFVIASLS  VTGMQQVQ.F  327
Ceeld6  QMEYKVYQRN  AFWEQATIVG  HWAW.VFYQL  FL..LPTWPL  RVAYFIISQM  GGGLLIAHVV  336

Boofd6  SLNHFSSSVY  VGKPKG.NNW  FEKQTDGTLD  ISCPPWMDWF  HGGLQFQIEH  HLFPKMPRCN  386
Ceeld6  TFNHNSVDKY  PANSRILNNF  AALQILTTRN  MTPSPFIDWL  WGGLNYQIEH  HLFPTMPRCN  396

Boofd6  LRKISPYVIE  LCKKHNLPYN  YASFSKANEM  TLRTLRNTA.  .LQARDITKP  LPKNLVWEAL  444
Ceeld6  LNACVKYVKE  WCKENNLPYL  VDDYFDGYAM  NLQQLKNMAE  HIQAKAA*..  ..........  443

Boofd6  HTHG*448
Ceeld6  .....443
```

Figure 1 A comparison of the deduced amino acid sequences of the borage (*B. officinalis*) Δ^6-desaturase [7] and the *C. elegans* cDNA CeD6.1

Abbreviations: Ceeld6, CeD6.1; Boofd6, borage Δ^6-desaturase.

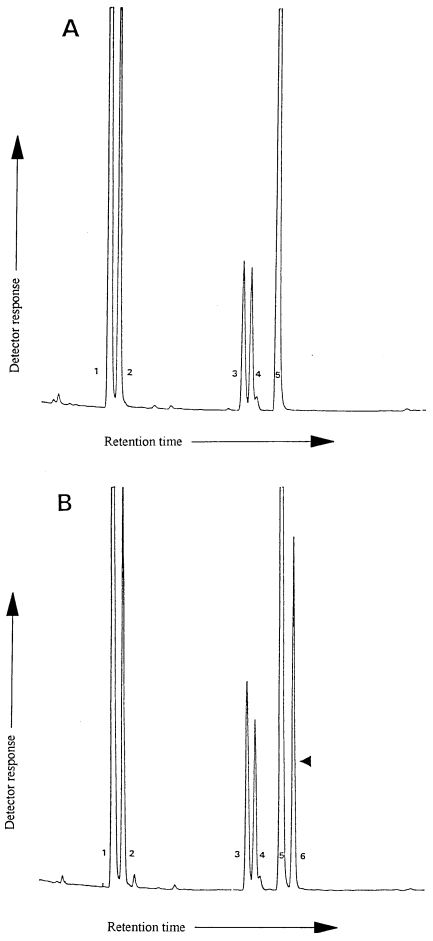

Figure 2 Identification of GLA in transgenic yeast by GC

Fatty acid methyl esters of total lipids of *S. cerevisiae* grown under inducing conditions (linoleate and galactose) were analysed by GC, using flame-ionization detection. (**A**) is yeast transformed with control (empty) vector pYES2 and (**B**) is transformed with pYCeD6.1. The common peaks were identified as $C_{16:0}$ (peak 1), $C_{16:1}$ (peak 2), $C_{18:0}$ (peak 3), $C_{18:1}$ (peak 4) and $C_{18:2}$ (peak 5; supplied exogenously). The additional peak (peak 6 in **B**), which corresponds to the retention time of GLA, is indicated by the arrowhead.

1450 bp) from the resulting rescued phagemids were sequenced. This confirmed that the cDNAs isolated by us did indeed show similarity to W08D2.4, with the 5′ and 3′ ends of the cDNA being equivalent to bases 9 and 3079 of the sequence of cosmid W08D2. Since the ATG initiating coding predicted by the Genefinder program to be the start of gene product W08D2.4 was indeed the first methionine residue in the cDNA clone, we reasoned that we had isolated a *bona fide* full-length cDNA. One representative cDNA clone (termed Cede.1; 1463 bp in length) was sequenced on both strands (Genbank ID: AF031477); the deduced amino acid sequence is identical with that predicted for W08D2.4 over the majority of the protein. However, DNA sequences encoding residues 38–67 (Y-S-I...L-Y-F) predicted for W08D2.4 are not present in the cDNA clone. This means that the deduced amino acid sequence of pCeD6.1 is in fact 443 amino acids long, as opposed to that predicted for W08D2.4, which is 473 residues in length. The only other difference between the two amino acid sequences is an M → V substitution at residue 401, resulting from a A → G base change (base 1211). The deduced amino acid sequence of CeD6.1 is shown in Figure 1, compared with the previously characterized borage Δ^6-desaturase [7]. Note the presence in the *C. elegans* sequence of the H-P-G-G cytochrome b_5 motif in the N-terminus (encoded by bases 96–108) and the H → Q substitution in the third histidine box (encoded by bases 1157–1172).

Clone pCeD6.1 was then used as a template for PCR amplification of the entire predicated coding sequence (443 amino acid residues in length) and cloned into the yeast expression vector pYES2 (Invitrogen) to yield pYCeD6. The fidelity of this PCR-generated sequence was checked by *in vitro* transcription/translation of the plasmid, using the T_7 RNA polymerase promoter present in pYES2. Using the Promega *TnT*-coupled transcription/translation system, translation products were generated and analysed by SDS/PAGE and autoradiography, following the supplier's instructions. This revealed (results not shown) that the plasmid pYCeD6 generated a product of molecular mass 55 kDa, whereas the control (pYES2) failed to yield any protein products, indicating that the construct was correct.

Transformation and selection of yeast able to grow on uracil-deficient medium revealed yeast colonies carrying the recombinant plasmid pYCeD6 by virtue of the URA3-selectable marker carried by pYES2. Expression of pYCeD6 was obtained by inducing the GAL promoter which is present in pYES2. This was carried out after the cells had been grown up overnight with raffinose as a carbon source, and the medium supplemented by the addition of linoleate ($C_{18:2,\Delta9,12}$) in the presence of low concentrations of detergent. This latter addition was required since the normal substrate for Δ^6-desaturation is $C_{18:2}$ fatty acid, which does not normally occur in *S. cerevisiae*. The cultures were then allowed to continue to grow after induction, with aliquots being removed for analysis by GC. When methyl esters of total fatty acids isolated from yeast carrying the plasmid pYCeD6, grown in the presence of galactose and linoleate, were analysed by GC, an additional peak was observed (Figure 2). This had the same retention time as an authentic GLA standard, indicating that the transgenic yeast was capable of desaturating linoleic acid at the Δ^6 position. No such peaks were observed in any of the control samples (transformation with pYES2). The identity of this extra peak was confirmed by GC–MS, which positively identified the compound as GLA (Figure 3). This confirms that Cede.1 encodes a *C. elegans* Δ^6-desaturase, and that this cDNA is likely to be transcribed from the gene predicted to encode ORF W08D2.4, though the deduced amino acid sequence of Cede.1 is 30 residues smaller than that of W08D2.4.

DISCUSSION

Organisms such as *C. elegans* perform Δ^6-desaturation, but unlike plants such as borage or evening primrose, they do not accumulate Δ^6 unsaturated fatty acids such as GLA. We provide evidence that a *C. elegans* cDNA (Cede.1) encodes a Δ^6-desaturase, and that this sequence is similar to the predicted ORF W08D2.4, except for a 30-residue insertion present in the N-terminal region of the latter protein. Whether the deduced amino acid sequence predicted for Cede.1 represents a splicing variant of W08D2.4, or is a result of a misprediction of the intron/exon junctions by the Genefinder program is unclear. However, it is clear that Cede.1 encodes a Δ^6-desaturase. The ORF encoded by this *C. elegans* sequence appears to be related to the higher-plant Δ^6-fatty-acid desaturase previously isolated by us [7], in that they both contain N-terminal domains which show similarity to cytochrome b_5. In contrast, other microsomal fatty acid desaturases from plants do not contain this domain and use free cytochrome b_5 as an electron donor [1,13,14]. Similarly, the domain is absent from the only fatty acid

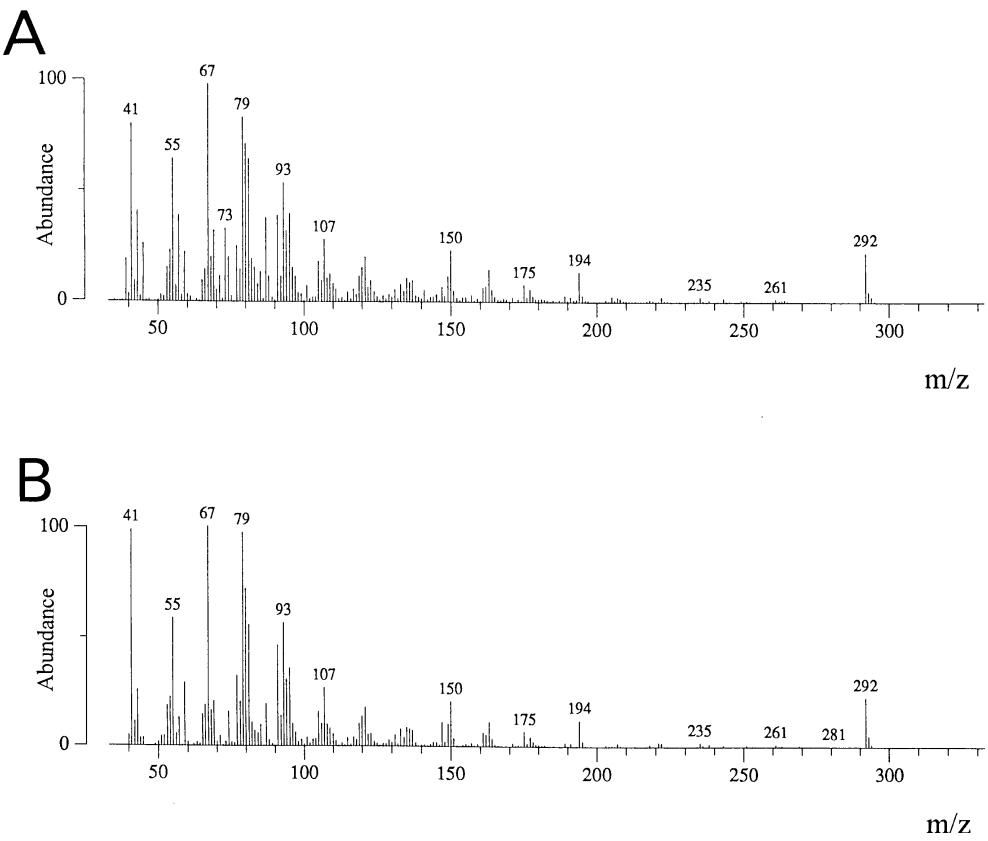

Figure 3 GC–MS analysis of the novel peak identified in yeast carrying pYCeD6.1

The sample was analysed for mass spectra as described previously [7], and the data were used to search a library of profiles. The sample was identified as GLA. A comparison of the mass spectra of the novel peak (**A**) and an authentic GLA standard (**B**) is shown. Visual- and computer-based inspection indicates that the two spectra are identical.

desaturases isolated from animals, a desaturase from *C. elegans* which recognizes a range of C_{18} and $C_{20,\omega-6}$ substrates [15] and a putative fatty acid desaturase from man (*Homo sapiens*) [16]. These animal sequences also differ from the borage and *C. elegans* Δ^6-desaturases in lacking the variant histidine box.

The reason why the Δ^6-desaturases have a fused cytochrome b_5 domain is not known [17]; the only other examples of desaturases with this extension are fungal microsomal (OLE1) Δ^9-desaturases [10] in which the domain is fused to the C-terminus rather than the N-terminus of the protein. However, the borage Δ^6-desaturase differs from all the other characterized plant microsomal desaturases in carrying out 'front-end' desaturation, which is the introduction of a double bond between C-3 and C-7 of an already unsaturated fatty acid [18]. This means the enzyme desaturates at positions between the carboxy group and pre-existing double bonds, whereas other plant enzymes desaturate sequentially towards the methyl group. It will be of interest to determine whether this feature is shared by other 'front-end' desaturases of plant and animal origin. It is also clear that identification of heterologous fatty acid desaturases will be facilitated by the yeast expression system described in the present study.

We are very grateful to Professor Yuji Kohara for supplying nematode ESTs and cDNA libraries. We thank Mervyn Lewis for carrying out the GC–MS analysis. IACR-Long Ashton Research Station receives grant-aided support from the Biotechnology and Biological Sciences Research Council (U.K.).

Received 5 December 1997/13 January 1998; accepted 13 January 1998

REFERENCES

1 Somerville, C. and Browse, J. (1996) Trends Cell Biol. **6**, 148–1153
2 Okuley, J., Lightner, J., Feldmann, K., Yadav, N., Lark, E. and Browse, J. (1994) Plant Cell **6**, 147–158
3 Arondel, V., Lemieux, B., Hwang, I., Gibson, S., Goodman, H. and Somerville, C. R. (1992) Science **258**, 1353–1355
4 van de Loo, F. N., Broun, P., Turner, S. and Somerville, C. R. (1995) Proc. Natl. Acad. Sci. U.S.A. **92**, 6743–6747
5 Shanklin, J., Whittle, E. and Fox, B. G. (1994) Biochemistry **33**, 12787–12794
6 Shanklin, J., Achim, C., Schmidt, H., Fox, B. G. and Munck, E. (1997) Proc. Natl. Acad. Sci. U.S.A. **94**, 2981–1986
7 Sayanova, O., Smith, M. A., Lapinskas, P., Stobart, A. K., Dobson, G., Christie, W. W., Shewry, P. R. and Napier, J. A. (1997) Proc. Natl. Acad. Sci. U.S.A. **94**, 4211–4216
8 Tanaka, T., Ikita, K., Ashida, T., Motoyama, Y., Yamaguchi, Y. and Satouchi, K. (1996) Lipids **31**, 1173–1178
9 Elble, R. (1992) Biotechniques **13**, 18–20
10 Mitchell, A. G. and Martin, C. E. (1995) J. Biol. Chem. **270**, 29766–29772
11 Wilson, R., Ainscough, R., Anderson, K., Baynes C., Berks, M., Burton, J., Connell, M., Bonfield, J., Copsey, T., Cooper, J. et al. (1994) Nature (London) **368**, 32–38
12 Lederer, F. (1994) Biochimie **76**, 674–692
13 Smith, M. A., Cross, A. R., Jones, O. T. G., Griffiths, W. T., Stymne, S. and Stobart, A. K. (1990) Biochem. J. **272**, 23–29
14 Smith, M. A., Jonsson, L., Stymne, S. and Stobart, A. K. (1992) Biochem. J. **287**, 141–144
15 Spychalla, J. P., Kinney, A. J. and Browse, J. (1997) Proc. Natl. Acad. Sci. U.S.A. **94**, 1142–1147
16 Cadena, D. L., Kurten, R. C. and Gill, G. N. (1997) Biochemistry **36**, 6960–6967
17 Napier, J. A., Sayanova, O., Stobart, A. K. and Shewry, P. R. (1997) Biochem. J. **328**, 717–718
18 Aitzetmuller, K. and Tseegsuren, N. (1994) J. Plant Physiol. **143**, 538–543