

Functional identification of a fatty acid Δ^5 desaturase gene from *Caenorhabditis elegans*

Louise V. Michaelson^{a,*}, Johnathan A. Napier^b, Mervyn Lewis^b, Gareth Griffiths^c,
Colin M. Lazarus^a, A. Keith Stobart^a

^a*School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, UK*

^b*IACR-Long Ashton Research Station, Department of Agricultural Sciences, University of Bristol, Long Ashton, Bristol BS41 9AF, UK*

^c*Horticulture Research International, Wellesbourne, Warwick CV35 9EF, UK*

Received 24 July 1998; received in revised form 8 October 1998

Abstract We have identified a cDNA from the nematode worm *Caenorhabditis elegans* that encodes a fatty acid Δ^5 desaturase. *Saccharomyces cerevisiae* expressing the full-length cDNA was able to convert di-homo- γ -linolenic acid to arachidonic acid, thus confirming Δ^5 desaturation. The 1341 bp Δ^5 desaturase sequence contained an N-terminal cytochrome b_5 domain and was located within a kilobase of the *C. elegans* Δ^6 desaturase on chromosome IV. With an amino acid identity of 45% it is possible that one of these genes arose from the other by gene duplication. This is the first example of a Δ^5 desaturase gene isolated from an animal.

© 1998 Federation of European Biochemical Societies.

Key words: Fatty acid desaturase; Delta five; *Caenorhabditis elegans*

1. Introduction

The Δ^5 desaturase is responsible for the conversion of di-homo- γ -linolenic acid (C20:3 $\Delta^{8,11,14}$) to arachidonic acid (C20:4 $\Delta^{5,8,11,14}$) as well as the synthesis of eicosapentaenoic acid (20:5 $\Delta^{5,8,11,14,17}$) from 20:4 $\Delta^{8,11,14,17}$ substrates [1,2]. In animals, arachidonic acid is a precursor of the short-lived regulatory molecules the eicosanoids, which comprise the prostaglandins, the leukotrienes and the thromboxanes. The eicosanoids have many functions, and are particularly important in the inflammatory response, reproductive function and the regulation of blood pressure [2]. Essential fatty acids may alleviate certain conditions such as eczema and mastalgia [2] and some studies suggest a possible application in the treatment of cancer and tumour growth [3,4].

The Δ^5 desaturase in *Caenorhabditis elegans* most likely functions as other aerobic microsomal desaturases requiring NAD(P)H and an electron transport chain involving cytochrome b_5 [5]. Recently a new class of fusion proteins has been identified in which the protein encoded by certain desaturase cDNAs contain a sequence related to cytochrome b_5 at the N-terminus [6].

Molecular genetic approaches, particularly using *Arabidopsis* mutants, have provided much information on desaturation reactions [7,8]. More recently data emanating from the genome sequencing projects have made it possible to identify

more membrane-bound desaturases [9,10]. These searches make use of the presence of characteristic histidine motifs common to all membrane-bound desaturases [11]. No information is available yet for an animal Δ^5 desaturase.

Here we report the identification of a putative *C. elegans* Δ^5 desaturase gene. When the nematode coding sequence was expressed in yeast grown in the presence of di-homo- γ -linolenic acid, arachidonic acid was produced. This confirmed the identity of the *C. elegans* ORF T13F2.1 as a Δ^5 desaturase gene.

2. Materials and methods

The University of Wisconsin GCG software package [12] was used for initial sequence identification and any further sequence analysis.

2.1. PCR-based cloning

Primers were synthesised using the EMBL sequence as a template. The forward primer was CEFOR, 5'-ATGGTATTACGAGAGCAAGA-3' and the reverse primer was CEREV, 5'-TCTGGGATCTCTGGTTCTTG-3'. The primers were used for PCR amplification of a mixed-stage *C. elegans* cDNA library constructed in λ ZapII. The PCR protocol followed was as [13]. Amplification products were fractionated on 1% agarose gels from which the DNA band was purified. It was ligated directly into pGEM-T (Promega) and the recombinant plasmids used to transform *Escherichia coli* DH5 α cells. Plasmid DNA was purified for sequencing using the Qiagen QIAprep miniprep kit. Nucleotide sequences were determined using a Perkin Elmer ABI-377 DNA sequencer.

2.2. Library screening

The mixed-stage *C. elegans* cDNA library was screened by standard techniques [14] using the cloned PCR product as a probe. DNA fragments were labelled with α^{32} P]dCTP using the Ready To Go DNA-Labelling reaction mix (Amersham Pharmacia Biotech). The full-length cDNA clone L4 was purified by successive rounds of plating and hybridisation before it was subjected to plasmid excision in vivo and the insert was sequenced on both strands.

2.3. Functional analysis: yeast transformation

PCR with the primers YCEDFOR, 5'-GCGAAGCTTAAAATGTATTACGAGAGCAAGAGC-3' (annealing to the initiating methionine indicated by the bold type face), and primer YCEDREV, 5'-GCGGGATCCAATCTAGGCAATCTTTTAGTCAA-3' (annealing to the complement of the stop codon, indicated in bold type face), was used to amplify the pL4 coding region and to flank it with *Hind*III and *Bam*HI restriction sites. The amplified PCR product was ligated into the vector pYES2 (Invitrogen) to generate the plasmid pYES2/L4. The fidelity of the cloned PCR product was checked by in vitro transcription and translation using the TNT system (Promega). The plasmid was transformed into *Saccharomyces cerevisiae* DBY746 by the lithium acetate method [15] and expression of the transgene was induced by the addition of galactose to 1% (w/v). The yeast culture medium was supplemented with 0.5 mM di-homo- γ -linolenic acid in the presence of 1% tergitol, as described by Napier et al. [10].

*Corresponding author. Fax: (44) (117) 925 7374.
E-mail: louise.michaelson@bristol.ac.uk

Abbreviations: PCR, polymerase chain reaction; GC, gas chromatography; GCMS, gas chromatography mass spectroscopy; FAME, fatty acid methyl ester

2.4. Fatty acid analysis

Total fatty acids extracted from yeast cultures were analysed by GC of methyl esters, as described previously [13]. Fatty acids were identified by comparison with retention times of FAME standards (Sigma) and arachidonic acid was further characterised by GC-MS using a Krats MS80RFA operating at an ionisation voltage of 70 eV, with a scan range of 500–40 Da.

3. Results

Recently we identified a fungal Δ^5 desaturase [13] and Δ^6 desaturases from both plant [16] and animal species [10] which were distinct from previously identified microsomal desaturases. This difference was due to the presence of an N-terminal extension which showed homology to the electron donor protein, cytochrome b_5 . In the course of the characterisation of the fungal Δ^5 desaturase [13] and the *C. elegans* Δ^6 desaturase (present on cosmid W08D2; accession number Z70271) [10], we noted the presence of a related sequence on cosmid T13F2 (accession number Z81122). Cosmids W08D2 and T13F2 were found to be overlapping. Closer examination of this ORF (T13F2.1) as predicted by the Genefinder program [12] revealed a number of similarities, in that the ORF contained an N-terminal cytochrome b_5 domain, as defined by the diagnostic HPGG motif [17], as well as three 'histidine boxes' characteristic of all microsomal desaturases [11]. Of particular interest was the observation that this putative desaturase contained a variant third histidine box, with an H→Q substitution for the first histidine in the HXXHH motif. This glutamine substitution had previously been observed by us in the plant and animal Δ^6 desaturases [16,10] and in the fungal Δ^5 desaturase [13]. The overlap between cosmids T13F2 and W08D2 allowed us to determine the proximity of the putative desaturase ORF T13F2.1 to the Δ^6 desaturase. This indicated that the two sequences were arranged in tandem on chromosome IV, separated by 990 bases from the predicted stop codon of T13F2.1 to the initiating methionine triplet of the Δ^6 desaturase.

We wished to determine the function of ORF T13F2.1 and

confirm if this was indeed a fatty acid desaturase. *C. elegans* produces several different unsaturated fatty acids and hence may carry out Δ^5 , Δ^6 , Δ^9 , Δ^{12} and Δ^{15} desaturation reactions [18]. Although it remained possible that T13F2.1 represented a second, possibly redundant, form of the Δ^6 desaturase, we considered it more likely that it encoded the Δ^5 fatty acid desaturase, since the predicted protein showed the greatest homology with a fungal microsomal Δ^5 desaturase [13].

The T13F2.1 ORF is predicted to be interspersed with a number of introns, making heterologous functional expression unfeasible. A PCR fragment representing a large predicted exon at the 5' end (1–233 bp) was amplified from a *C. elegans* cDNA library using the primers CEFOR and CEREV. A DNA fragment of the predicted size was generated and isolated from a gel for cloning and sequencing.

To isolate the coding region corresponding to the T13F2.1 ORF the PCR product of 233 bp was used to probe the *C. elegans* cDNA library that had been constructed in λ ZapII. The positive plaques were checked for the presence of T13F2.1 sequence by PCR with CEFOR and CEREV. One clone, L4, was purified by additional rounds of plating and hybridisation screening. The plasmid pL4 was released from lambda clone L4 by in vivo excision and the cDNA insert was sequenced on both strands.

The deduced amino acid sequence was identical with that predicted for T13F2.1 over most of the protein. However, DNA sequences encoding residues 198–204 predicted for T13F2.1 were not present in the cDNA clone; this is likely to be due to a misprediction of intron/exon borders by the Genefinder program. This means that the deduced amino acid sequence for L4 is 447 amino acids long compared to the 454 amino acids predicted for T13F2.1. Other than this the sequences were identical at the DNA level.

The deduced sequence of L4 is shown in Fig. 1 and is compared with the previously characterised *C. elegans* Δ^6 desaturase. The amino acid sequences of these two genes show 45% identity and 68% similarity. At the nucleic acid level L4 is

CED 6	1	MVVDKN-ASGLRMKDGKWL YLS EELVKKHPGGAVIEQYRNSDATHIFHAFHEGSSQAYK
CED 5	1	MVLRQEHEPFFFKIDGKWCQIDDAVLRESHPGGSALT TYKNMDATTVFHTFHTGSKEAYQ
CED 6	60	QLDLKKKHG--EHDEFLKQLEKRLDKVDINVSAYDYVVAQEKKMVESEFEKLRQKIHDDG
CED 5	61	WLTLEKKKECPTQEP EIPDIKDDPIK GIDDVNMGTFTNISEKRSAQINKSFTDLRMKVRANG
CED 6	118	LMKANET YFLFKAIS TLSIMAFAFYLQYLGVYITSACLLALAWQFGWLTHEFCHQOPTK
CED 5	121	LMDGSP LFYIRKILETIFTILFAFYLQYHTTYVLP SAILMGVAVQQLGWL IHEFAHQLFK
CED 6	178	NRPLNDTISLFFCNFLQGFSDRWKDKHNTHHAATNVIDHDGDIIDLAPLFAFIPGDLCKY
CED 5	181	NRYYNDLASIFVGNFLQGFSSGGWKEQHNVHHAATNVVGRDGDIDLVPFATVAEHLNNY
CED 6	238	KASF EKAILKIVPYQHLYFTAMLPLMRFSWTGQSVQWVFKENQMEYKVYQRNAPFEQAT
CED 5	241	SQ--DSWVMTLFRMQHLYHMTFMPLPFLRLSLWLLQSIIVSQMPHTHYDYRNTATIEQVGL
CED 6	298	VGHWA WVFYQLFLPTWPLRVAYFIISQMGGLLIAHVTFNHNVSVDKY PANSRIENNA
CED 5	299	SLHWAWSLGQLFLLPDWSTRIMFFLVSHLVGGFLLSHVVTFNHYSVEKFA LSSNIMSNA
CED 6	358	ALQILTTTRNMTESPFIDWLWGGLNYQIEHHLFPTMPRCNLNACVKYVKECKENNLPLYV
CED 5	359	CLQILTTTRNMRPGRFIDWLWGGLNYQIEHHLFPTMPRHNLNTVMPLVKEFAAANGLPYV
CED 6	418	DDYFDGYAMNLQQLKNNMAEHIQAKAA
CED 5	419	DDYETGWLETEQFFRNIAANVA AKLTCKIA

Fig. 1. Comparison of *C. elegans* Δ^5 and Δ^6 sequences. The translation of the coding sequence of L4 (GenBank accession number AF078796) is aligned with the deduced amino acid sequence from a *C. elegans* Δ^6 desaturase cDNA. Identical or conserved residues are shaded and the conserved histidine boxes and cytochrome b_5 domain are underlined.

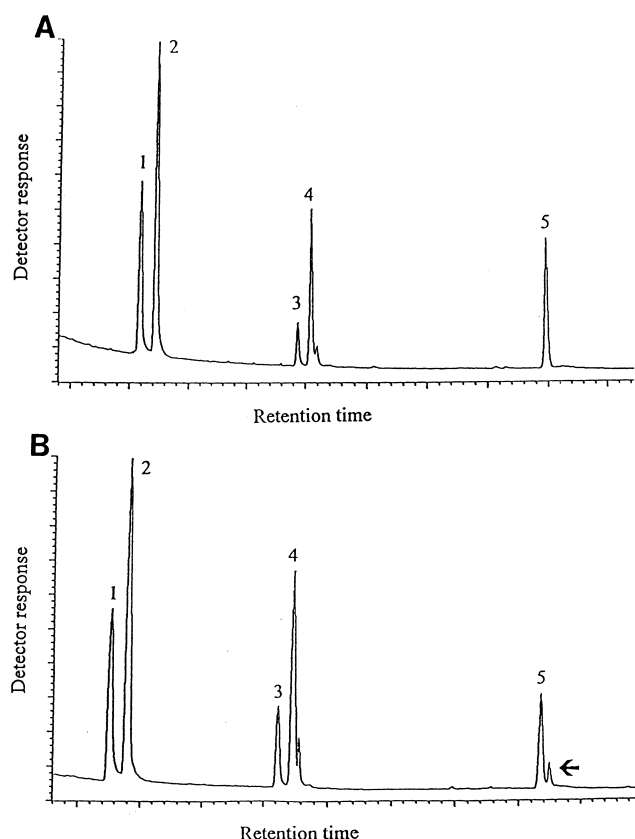


Fig. 2. Identification of arachidonic acid synthesised in transgenic yeast. Fatty acid methyl esters of total lipids of *S. cerevisiae* grown under inducing conditions in the presence of di-homo- γ -linolenic acid were analysed by GC, using flame ionisation detection. A: FAMES extracted from yeast transformed with control (empty) vector pYES2. B: FAMES extracted from yeast transformed with pYES2/L4. The common peaks were identified as C16:0 (peak 1), C16:1 (peak 2), C18:0 (peak 3), C18:1 (peak 4) and C20:3 (peak 5 supplied exogenously). The additional peak, which corresponds to the retention time of arachidonic acid, is indicated by the arrow.

60% identical to the *C. elegans* Δ^6 desaturase gene. It is also interesting to note that the final two intron/exon junctions in both the *C. elegans* Δ^6 desaturase gene and T13F2.1 are in identical positions when the two amino acid sequences are aligned. The intron/exon junction for intron number 6 in T13F2.1 occurs after position 348 and this aligns with the intron/exon junction for intron 5 in the *C. elegans* Δ^6 desaturase which occurs after position 347. The intron/exon junctions for intron 7 in T13F2.1 and 6 in the *C. elegans* Δ^6 desaturase are similarly aligned, occurring after positions 384 and 383 respectively.

3.1. Functional analysis of L4 in yeast

The complete coding region (447 amino acids plus a stop codon) of L4 was amplified by PCR using the primers YCED-For and YCEDRev and inserted into the yeast expression vector pYES2 downstream of the *GAL1* promoter. This construct was transformed into *E. coli*. The fidelity of the PCR-generated insert in plasmid pYES2/L4 was confirmed in vitro by coupled transcription/translation under the control of the T7 RNA polymerase promoter in the vector. The translation products were analysed by SDS-PAGE and autoradiography. A product of $M_r \sim 56000$ was produced from the plasmid

pYES2/L4 whereas the empty vector control (pYES2) failed to yield any protein products (data not shown).

For functional analysis of the L4 coding region the recombinant plasmid was transferred to yeast. Cells were cultured overnight in a medium containing raffinose as sole carbon source, and supplemented by the addition of either linoleic acid (18:2 $\Delta^{9,12}$) or di-homo- γ -linolenic acid (C20:3 $\Delta^{8,11,14}$). These fatty acids are not present in *S. cerevisiae* but serve as the specific substrates for either Δ^6 or Δ^5 desaturases, respectively. Expression of the L4 coding region from the *GAL1* promoter of the vector was induced by the addition of galactose to 1%. Growth of the cultures was continued for 16 h, when aliquots were removed for the analysis of fatty acids by GC. Fig. 2 shows the result of GC analysis of the fatty acid methyl esters of transformed yeast strains. An additional peak is apparent in the trace obtained from induced pYES2/L4 grown in the presence of di-homo- γ -linolenic acid compared to an empty-vector control. This peak was also absent from uninduced cultures grown on di-homo- γ -linolenic acid and it should be noted also that pYES2/L4 grown in the presence of linoleic acid failed to accumulate any novel fatty acids indicating that this fatty acid is not a substrate for the enzyme encoded by the clone pYES2/L4. The retention time of the additional peak is identical to that of authentic methyl-arachidonic acid standard. The fatty acid produced from di-homo- γ -linolenic acid was further characterised by GCMS and identified as arachidonic acid (Fig. 3). The results show, therefore, that yeast cells transformed with the plasmid pYES2/L4 had acquired functional Δ^5 desaturase activity and were now capable of synthesising arachidonic acid from the substrate di-homo- γ -linolenic acid. The Δ^5 desaturase in the transformed yeast converted some 15% of the substrate under the conditions of the experiment.

4. Discussion

We provide evidence that a *C. elegans* cDNA, T13F2.1, encodes a Δ^5 desaturase. The protein encoded by ORF T13F2.1 appears to be related to the Δ^6 desaturase isolated from *C. elegans*, as they both contain N-terminal domains with similarity to cytochrome b_5 . It is uncertain whether the organism uses only this domain to provide the electron donor for desaturation or utilises free cytochrome b_5 similarly to some microsomal fatty acid desaturases from plants [19,20].

It is not known why the Δ^5 desaturases from *M. alpina* and *C. elegans* contain this N-terminal cytochrome b_5 domain which is also observed in the Δ^6 desaturases from *Borago officinalis* (borage) and *C. elegans*. The Δ^9 desaturase from *Saccharomyces cerevisiae* has a C-terminal cytochrome b_5 domain which appears to be essential for its function [21]. A Δ^5 desaturase which catalyses formation of a double bond in a saturated fatty acid from *Bacillus subtilis* has recently been described [22]. This is quite different from the *C. elegans* Δ^5 desaturase showing only 19% identity and having a length of 352 amino acids. The difference in size is mainly due to the lack of a cytochrome b_5 domain.

It is interesting that the *C. elegans* Δ^5 and Δ^6 desaturase genes are only 990 bp apart on chromosome IV, as this is unusually close. Studies on chromosome III have revealed that on average it contains one gene per 5.6 kb of sequence [23]. It is possible that one of these genes arose by duplication of the other. Given that they show only 45% identity and 68%

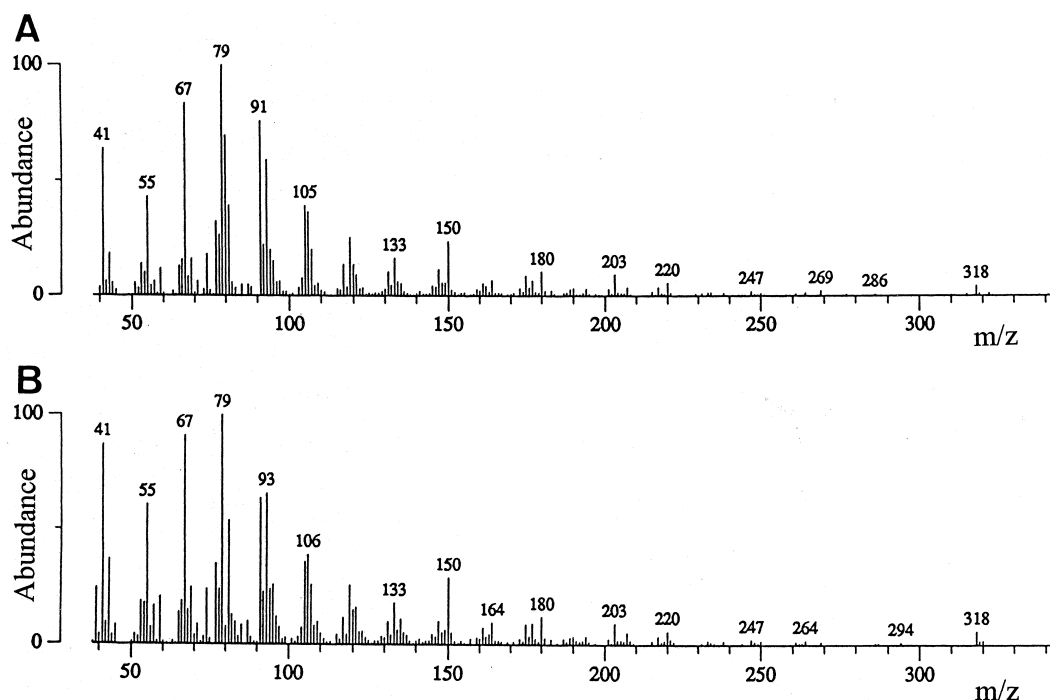


Fig. 3. GCMS analysis of the novel peak identified in yeast carrying pYES2/L4. A comparison of the mass spectra of the novel peak (A) and an authentic arachidonic acid standard (B) is shown. Visual and computer-based inspections indicate that the two spectra are identical.

similarity this must have been an ancient event. Further evidence for gene duplication is provided by the fact that the positions of the final two intron/exon junctions of the two genes are identical.

Other genes identified in *C. elegans* show clustering of function and are transcribed as polycistronic RNAs. Genes within a polycistronic cluster are transcribed sequentially from a single promoter at the 5' end [24]. Approximately 25% of *C. elegans* genes are transcribed in this way [25]. It is possible that the Δ^5 and Δ^6 desaturase genes are part of a single transcription unit, although a 990 bp gap between genes in a polycistronic RNA is larger than previously reported.

Acknowledgements: We would like to thank Prof. Yuji Kohara (National Institute of Genetics, Mishima, Japan) for supplying the nematode cDNA library. L.M. thanks Dr Steve Screen for helpful discussions. A.K.S. is grateful to the Wolfson foundation for laboratory refurbishment. IACR-Long Ashton receives grant aided support from the BBSRC, UK. L.M. is supported by a BBSRC studentship and a CASE award from HRI Wellesbourne.

References

- [1] Leikin, A.I. and Brenner, R.R. (1992) *Biochim. Biophys. Acta* 1165, 189–193.
- [2] Horrobin, D.F. (1992) *Prog. Lipid Res.* 31, 167–194.
- [3] Abouelela, S.H., Prasse, K.W., Carroll, R., Wade, A.E., Dharwadkar, S. and Bunce, O.R. (1988) *Lipids* 23, 948–954.
- [4] Fearon, K.C.H., Falconer, J.S., Ross, J.A., Carter, D.C., Hunter, J.O., Reynolds, P.D. and Tuffnell, Q. (1996) *Anticancer Res.* 16, 867–874.
- [5] Dailey, H.A. and Strittmatter, P. (1980) *J. Biol. Chem.* 255, 5184–5189.
- [6] Napier, J.A., Sayanova, O., Stobart, A.K. and Shewry, P.R. (1997) *Biochem. J.* 328, 717–718.
- [7] Arondel, V., Lemieux, B., Hwang, I., Gibson, S., Goodman, H. and Somerville, C.R. (1992) *Science* 258, 1353–1355.
- [8] Okuley, J., Lightner, J., Feldmann, K., Yadav, N., Lark, E. and Browse, J. (1994) *Plant Cell* 6, 147–158.
- [9] Spychalla, J.P., Kinney, A.J. and Browse, J. (1997) *Proc. Natl. Acad. Sci. USA* 94, 1142–1147.
- [10] Napier, J.A., Hey, S.H., Lacey, D.J. and Shewry, P.R. (1998) *Biochem. J.* 330, 611–614.
- [11] Shanklin, J., Whittle, E. and Fox, D.C. (1994) *Biochemistry* 33, 12787–12794.
- [12] Devereux, J., Haerberli, P. and Smithies, O. (1984) *Nucleic Acids Res.* 12, 387–395.
- [13] Michaelson, L.V., Lazarus, C.M., Griffiths, G., Napier, J.A. and Stobart, A.K. (1998) *J. Biol. Chem.* 273, 19055–19059.
- [14] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., Cold Spring Harbor Laboratory Press, Plainview, NY.
- [15] Elble, R. (1992) *BioTechniques* 13, 18–20.
- [16] 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. USA* 94, 4211–4216.
- [17] Schmidt, H., Dresselhaus, T., Buck, F. and Heinz, E. (1994) *Plant Mol. Biol.* 26, 631–642.
- [18] Tanaka, T., Ikita, K., Ashida, T., Motoyama, Y., Yamaguchi, Y. and Satouchi, K. (1996) *Lipids* 31, 1173–1178.
- [19] Smith, M.A., Jonsson, L., Stymne, S. and Stobart, A.K. (1992) *Biochem. J.* 287, 141–144.
- [20] 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.
- [21] Mitchell, A.G. and Martin, C.E. (1995) *J. Biol. Chem.* 270, 29766–29772.
- [22] Aguilar, P.S., Cronan, J.E. and de Mendoza, D. (1998) *J. Bacteriol.* 180, 2194–2200.
- [23] Wilson, R. et al. (1994) *Nature* 368, 32–38.
- [24] Evans, D., Zorio, D., MacMorris, M., Winter, C.E., Lea, K. and Blumenthal, T. (1997) *Proc. Natl. Acad. Sci. USA* 94, 9751–9756.
- [25] Ross, L.H., Freedman, J.H. and Rubin, C.S. (1995) *J. Biol. Chem.* 270, 22066–22075.