RESEARCH COMMUNICATION

Molecular cloning of higher-plant 3-oxoacyl-(acyl carrier protein) reductase

Sequence identities with the *nodG*-gene product of the nitrogen-fixing soil bacterium *Rhizobium meliloti*

Antoni R. SLABAS,* Diana CHASE,* Ikuo NISHIDA,† Norio MURATA,† Christopher SIDEBOTTOM,‡ Richard SAFFORD,‡ Philip S. SHELDON,§¶ Roy G. O. KEKWICK,§ D. Grahame HARDIE|| and Robert W. MACKINTOSH||**

*School of Biological Sciences, University of Durham, Durham DH1 3LE, U.K., †National Institute for Basic Biology, Myodaiji, Okazaki 444, Japan, ‡Biosciences Division, Unilever Research, Bedford MK44 1LQ, U.K., §Department of Biochemistry, University of Birmingham, Birmingham B15 2TT, U.K., and ||Protein Phosphorylation Group, Department of Biochemistry, University of Dundee, Dundee DD1 4HN, Scotland, U.K.

cDNA clones encoding the fatty-acid- biosynthetic enzyme NADPH-linked 3-oxoacyl-(acyl carrier protein) (ACP) reductase were isolated from a *Brassica napus* (rape) developing seed library and from an *Arabidopsis thaliana* (thale cress) leaf library. The *N*-terminal end of the coding region shows features typical of a stromal-targeting plastid-transit peptide. The deduced amino acid sequences have 41 % and 55 % identity respectively with the *nodG*-gene product of *Rhizobium meliloti*, one of the host-specific genes that restrict infectivity of this bacterium to a small range of host plants. The probability that the *nodG*-gene product is a oxoreductase strengthens the hypothesis that some of the host-specific *nod*-gene products are enzymes which synthesize polyketides that uniquely modify the *Rhizobium* nodulation signal molecule.

INTRODUCTION

The overall pathway of fatty acid biosynthesis from malonyl-CoA is the same in all organisms. However, the structural organization of the participating enzymes varies widely among different Kingdoms. A single multifunctional fatty acid synthase polypeptide is found in mammals, whereas two interacting multifunctional polypeptides function in yeast [1]. In contrast, in prokaryotes and plants, each step in the pathway is catalysed by a separate monofunctional enzyme [2]. In plants, evidence from cellular-fractionation and immunolocalization studies suggests that fatty acid biosynthesis takes place in the plastid organelles [3–5]. This biochemical evidence is supported by analysis of cDNA clones of acyl carrier protein (ACP), 3-oxoacyl-ACP synthase I and enoyl-ACP reductase from *Brassica napus* (rape), *Arabidopsis thaliana* (thale cress) and *Hordeum vulgare* (barley), which show these enzymes to be nuclear-encoded and synthesized with a transit peptide for targetting to the stromal compartment of plastids [4,6–8].

3-Oxoacyl-ACP reductase catalyses the first reduction step in fatty acid biosynthesis [2]. Two isoforms of the enzyme, one NADPH-linked and the other NADH-linked, have been reported to be localized in plastids [3]. The function of the NADH-linked enzyme is unknown. It is the NADPH-linked β -oxoacyl-ACP reductase that functions in fatty acid biosynthesis. For example, only the NADPH-linked enzyme is co-ordinately induced with the other fatty-acid-biosynthetic enzymes, just before and during lipid accumulation in developing oilseeds [9].

In order to understand the regulatory mechanisms involved in

the co-ordinated gene expression of the fatty-acid-biosynthetic pathway, the specific targetting of the precursor enzymes to the plastids, the substrate specificity of the enzymes, and to address the evolutionary relationships between the monofunctional and multifunctional fatty acid synthases, knowledge of the molecular structure of the enzymes and their genes is required.

The NADPH-linked isoform of 3-oxoacyl-ACP reductase has been partially purified from barley leaf [10], safflower (*Carthamus tinctorius*) [11] and *Euglena* [12] and to homogeneity from spinach (*Spinacia oleracea*) leaf [13], avocado (*Persea americana*) fruit [5] and the developing seeds of *Brassica napus* [9]. Limited amino acid sequencing has shown the avocado fruit and *Brassica napus* seed enzymes to be closely related.

Here we report the molecular cloning of plant NADPH-linked 3-oxoacyl-ACP reductase from developing seed of the oil crop *Brassica napus* and from another member of the Cruciferae family, *Arabidopsis thaliana*. A striking finding is that the entire coding region of plant NADPH-linked 3-oxoacyl-ACP reductase shows sequence identity with the *nodG*-gene product of the soil bacterium *Rhizobium meliloti*, which fixes nitrogen in root nodules of leguminous plants [14]. However, there is no resemblance to the sequence of the 3-oxoreductase domain of the multifunctional mammalian fatty acid synthase polypeptide.

MATERIALS AND METHODS

Enzyme purification

3-Oxoacyl-ACP reductase was purified from avocado and *Brassica napus* seed as described in [5,9]. In outline, extracts of

Abbreviations used: ACP, acyl carrier protein; DTT, dithiothreitol; ORF, open reading frame.

¶ Present address: Department of Biochemistry and Molecular Biology, University of Leeds, Leeds LS2 9JT, U.K.

** To whom correspondence should be addressed.

The nucleotide sequences reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers X64463 (*Brassica napus* reductase) and X64464 (*Arabidopsis* reductase).

Fig. 1. (a) Reverse-phase h.p.l.c. profile of peptides resulting from the tryptic digestion of *Brassica napus* seed 3-oxoacyl-ACP reductase and (b) amino acid sequences of *Brassica napus* peptides purified by reverse-phase chromatography and the amino acid sequences of peptides previously obtained from avocado fruit 3-oxoacyl-ACP reductase [5]

(a) The digested proteins were applied to a 5 C₈ micro-bore column equilibrated in 0.1% trifluoroacetic acid, and the peptides eluted by an increasing gradient of acetonitrile in 0.1% trifluoroacetic acid. (b) Some of the amino acid sequences have also been reported in [9]. The N-terminal sequences of the undigested *Brassica napus* and avocado fruit 3-oxoacyl-ACP reductase are also shown. The N-terminal sequence given for the *Brassica napus* enzyme contained minor contaminating sequences; however, the major sequence shown here is the most similar to the avocado and *Arabidopsis* N-termini.

developing seeds from *Brassica napus* or avocado fruit were subjected to (NH₄)₂SO₄ fractionation, then Procion Red H-E3B affinity chromatography, gel filtration and, finally, high-performance hydroxyapatite chromatography to give pure enzyme.

Peptide purification and amino acid sequencing

The purified protein (up to 30 µg; approx. 900 pmol) was incubated with trypsin at 37 °C at a proteinase/protein ratio of 1:100 (w/w); incubation was for a total of 24 h, the trypsin being added at 0 h and 12 h. The digested proteins were then applied to an ODS-102 spheri 5 C₈ micro-bore h.p.l.c. column (2.1 mm × 10 cm) equilibrated in 0.1% trifluoroacetic acid, and the peptides were eluted by an increasing gradient of acetonitrile in 0.1% trifluoroacetic acid. Fractions were collected manually to maximize separation and purification of the eluted peptides. All the fractions corresponding to absorbance peaks were subjected to amino acid sequence analysis on an Applied Biosystems gas-phase amino acid sequencer. The total volume of the fraction was applied to the sequencing disc in 0.03 ml aliquots, the disc being allowed to dry between additions. For N-terminal sequencing, 10 µg (approx. 300 pmol) of undigested enzyme was denatured by the addition of SDS (2%, v/v) and dithiothreitol (DTT) (20 mM final concn.) and heating to 100 °C for 2 min. After centrifugation the sample was injected on to a Brownlee C 03 032 C₈ micro-bore h.p.l.c. column (2.1 mm × 10 cm) equilibrated in 0.1% trifluoroacetic acid, and then eluted with an increasing gradient of 100% propan-2-ol/acetonitrile (2:1, v/v) in 0.1% trifluoroacetic acid. The eluted sample was dried and redissolved in 0.1 ml of 0.2 M-Tris/HCl (pH 8.0)/20 mM-DTT/2 mM-EDTA/2% SDS and incubated under N₂ at 37 °C for 60 min. The sample was then reductively alkylated by incubating with iodoacetamide (50 mM final concn.) in the dark at 0 °C for 40 min. The sample was then reapplied to, and eluted from, the Brownlee C 03 032 C₈ column as described above. The eluted sample was then applied to the sequencer disc.

Oligonucleotide synthesis

From the amino acid sequence EDMEKK derived from *Brassica napus* 3-oxoacyl-ACP reductase, a complementary degenerate oligonucleotide, TT(T/C)TT(T/C)CAT(A/G)TC-(T/C)TC, was prepared on a Applied Biosystems oligonucleotide synthesizer.

cDNA library construction and screening

A *Brassica napus* λgt10 cDNA library was constructed by

using a cDNA synthesis kit (Pharmacia) and phage packaging kit (Amersham International), starting with 1–2 µg of mRNA isolated from developing seeds. The library was amplified and screened on duplicate filters. Prehybridization was carried out in 0.9 M-NaCl/0.09 M-Tris/HCl (pH 7.4)/6 mM-EDTA/0.5% (v/v) Nonidet P40/0.2% (w/v) sodium pyrophosphate/0.04% (w/v) Ficoll/0.04% (w/v) polyvinylpyrrolidone/0.04% (w/v) BSA/denatured calf thymus DNA (100 µg/ml) at 37 °C for 14 h. The oligonucleotide was labelled with [γ-³²P]ATP by T₄ polynucleotide kinase [15] to a specific radioactivity of approx. 1 × 10⁶ c.p.m./pmol. Hybridization was performed under the same conditions for 14 h with the addition of [γ-³²P]ATP-5'-end-labelled oligonucleotide to a concentration of 0.8 pmol/ml [16]. Filters were washed in 0.9 M-NaCl/0.09 M-sodium citrate, pH 7.0, at 37 °C. Positive clones were plaque-purified, and finally the phages were purified by CsCl-density-gradient centrifugation to allow DNA isolation by formamide extraction [17]. One clone was initially purified which contained three *Eco*R1 fragments (0.6, 0.32 and 0.1 kb). The fragments were subcloned into Bluescript pKS-M13⁺, and plasmid DNA was purified from alkali-lysed cells by CsCl-density-gradient centrifugation [15]. DNA sequencing was performed on the double-stranded plasmid DNA by using the dideoxy-chain-termination methods [16] with deoxyadenosine 5'-[α-³²S]thiotriphosphate. The 0.32 and 0.1 kb fragments together coded for the 3' non-coding region and a truncated open reading frame (ORF) for 3-oxoacyl-ACP reductase. The 0.6 kb fragment was an unrelated sequence and may have been present as a result of the cloning procedures. The *Brassica napus* λgt10 cDNA library was rescreened using the 0.32 kb fragment labelled with [γ-³²P]dATP using the random-oligonucleotide-priming method [17], but no other clones of increased length were found. The 0.32 kb cDNA fragment obtained from *Brassica napus* was used to screen an *Arabidopsis thaliana* leaf λ ZAP II cDNA library. A 1.2 kb clone was isolated and sequenced and found to contain a complete ORF for 3-oxoacyl-ACP reductase.

RESULTS

Fig. 1 shows the C₈ reverse-phase h.p.l.c. profile of the peptides derived from the tryptic digestion of *Brassica napus* seed 3-oxoacyl-ACP reductase and the amino acid sequences of the purified peptides. Also shown are peptide sequences that were derived previously from avocado fruit 3-oxoacyl-ACP reductase [5]. From analysis of these sequence data it was possible to choose

(a)

Ar-n	CTTTTCTCTTGTCTTCT	CCACTGGGTCGATCTTTACGAGGTTCACTT	GTACGGATTCACTTGAACAACAGATTCGCC	80
Ar-n	ATGGCCGCCGCCGTCGAGCACCAGACTC	ATCTCCTTGAAGCCGTCGGAAGCTAGGT	TTCCGTGAGATCTCTCAAAATCCGACAATTG	170
Ar-aa	M A A A V A A P R L	I S L K A V G K L G	F R E I S Q I R Q L	
Ar-n	GCTCCGCTTCACTCTGCGATACCTCATTTC	GGAATGCTGCGATGTAGATCGCGGCAGCCG	TTTTCTACCTCTGTTGTGAAAGCTCAAGCT	260
Ar-aa	A P I H S A I P H F	G M L A C R S R Q P	F S T S V V K A Q A	
Ar-n	ACTGCTACTGAGCAATCACCAGGGGAAGTT	GTTCAAAAAGTGGAAATCTCCAGTGGTCGTT	ATCACTGGTGCCTCGAGAGGGATTGGTAAA	350
Ar-aa	T A T E Q S P G E V	V Q K V E S P V V V	I T G A S R G I G K	
Ar-n	GCAATTGCTCTAGCCTTGGGTAAAGCTGGC	TGCAAGGTCCTGGTGAATTATGCTAGGTCA	GCAAAGGAGGCTGAAGAAGTTGCTAAACAG	440
Ar-aa	A I A L A L G K A G	C K V I V N W A R S	A K E A E E V A K Q	
Ar-n	ATTGAAGAATATGGTGGCCAGGCTATCACT	TTTGGGGTGATGTCTCGAAAGCGACTGAT	GTGGATGCGATGATGAAAACCGCTCTTGAC	530
Ar-aa	I E E W G G Q A I T	F G G D V S K A T D	V D A M M K T A L D	
Ar-n	AAATGGGGAACCATGTGATGTTGTGGTTAAC	AATGCAGGAATTACTCGGGATACCTTGTTA	ATACGAATGAAGCAATCACAATGGGACGAA	620
Ar-aa	K W G T I D V V V N	N A G I T R D T L L	I R M K N S Q W D E	
Ar-n	GTGATTGCTTTGAATCTCACTGGTGTATTT	CTCTGTACTCAGGCAGCAGTAAAGATCATG	ATGAAGAAAAAGGGGAAGAATCATCAAT	710
Br-n	A..G...	T.C.....AC.....	G.G..A.....G.....C	71
Ar-aa	V I A L N L T G V F	L C T Q A A V K I M	M K K K R G R I I N	
Br-aa	* * *	* * S * * * T * * *	* * * R K * * * *	
Ar-n	ATCTCGTCAGTTGTTGGTCTCATTGGTAAT	ATTGGTCAAGCAAACCTACACAGCGACTAAA	GGTGGAGTTATTTTCGTTTCCGAGACTCCT	800
Br-n	..CG.A.....T.....C	G.....G.T..AG.....	..C.....GG.....A.....G.C	161
Ar-aa	I S S V V G L I G N	I G Q A N Y T A T K	G G V I S F S E T P	
Br-aa	* A * * * * *	* * * * * A * A *	A * * * G * * K * A	
Ar-n	GCCAGAGAAGGTGCGAGTAGGAATATAAAT	GTGAATGTGGTTTGCCCGGATTTCATTGCA	TCTGACATGACTGCAGAGCTTGGAGAAGAC	890
Br-n	G.....T.....C	C.....T.....	G.....CA.....	251
Ar-aa	A R E G A S R N I N	V N V V C A G F I A	S D M T A E L G E D	
Br-aa	* * * * * * *	* * * * * P * * * *	* E * * * K * * * *	
Ar-n	ATGGAAAAGAAAATCCTGGGAACAATCCCA	TTAGGACGGTATGGAAAAGCTGAAGAAGTA	GCTGGCTTGGTAGAATCTTGGCTCTCAGT	980
Br-n	T.....A.....C..C.A.....T..G	C.....	C.....	341
Ar-aa	M E K K I L G T I P	L G R Y G R A E E V	A G L V E F L A L S	
Br-aa	* * * * * * *	* * * * * Q P * D V	* * * * * * * *	
Ar-n	CCTGCAGCCAGTTACATTACCGGACAGGCA	TTCACCATGTGAGGATTTGCTATCTAG	ACATTGGTATGAGTCGGCTTTTGTITAGT	1070
Br-n	..G.....A..T.....A..	C.....C.....	G.C...CTCG..CCG	416
Ar-aa	D A A S Y I T G Q A	F T I D G G I A I		
Br-aa	P * * * * * T	* * * * * * *		
Ar-n	TAAAACAGTTTGGTAAAACAGACAAAGTTG	AGTTCTTTCTAGCACTTGGTCGGATTGCTA	TTTTCGGGTTTCTGTTAAGAAATATGTAAA	1160
Ar-n	ACTTACTGCTTAAACTACGTATTGGTCAC	TTTCAGAACTAGAAATC		1208

Fig. 2. (a) Nucleotide (Ar-n, Br-n) and derived amino acid (Ar-aa, Br-aa) sequences of the *Arabidopsis thaliana* leaf and *Brassica napus* seed 3-oxoacyl-ACP reductase clones and (b) sequencing strategy of the *Arabidopsis thaliana* leaf (1208 bp) and *Brassica napus* seed (326 bp and 90 bp) clones

(a) Identities in the nucleotide sequences are represented by , and in the derived amino acid sequence by ****. The positions of similar or identical peptides obtained from purified *Brassica napus* (—) and avocado fruit (---) 3-oxoacyl-ACP reductase are also indicated. Numbering of the nucleotide sequences is from the 5' end of the cDNA clones. The position of the putative leader-peptide cleavage site is indicated by the arrow. (b) The coding sequence is denoted by the thick filled bar and the non-coding regions by the thin bars. The scale shows the nucleotide position from the 5' end of the cDNA insert. The arrows indicate the direction and length of DNA sequences obtained. Sequencing was initiated with specific oligonucleotide primers (open circles) or Bluescript primers (filled circles).

a peptide, EDMEKK, that has a complementary oligonucleotide sequence of low degeneracy, thus allowing the initial screening of the *Brassica napus* cDNA library. The sequencing strategies, nucleotide sequences and derived amino acid sequences of the *Brassica napus* seed and *Arabidopsis thaliana* leaf 3-oxoacyl-ACP reductase cDNA clones are presented in Fig. 2. The *Brassica napus* clone has an ORF of 398 bp, which starts at nucleotide 1 and codes for 132 amino acids. Since there is no initiating methionine codon, this clone does not contain a complete coding sequence. This truncation was not immediately apparent from the size of the DNA insert because the clone also carried an additional 600 bp *EcoRI* DNA insert which turned out to be unrelated to 3-oxoacyl-ACP reductase (see the Materials and methods section). The nucleotide sequence of the *Arabidopsis thaliana* clone contains an ORF of 957 bp corresponding to a polypeptide of 319 amino acids. The 3'-untranslated region is 171 bp in length, carried a possible polyadenylation signal (AAATA), but does not include a polyadenylated tail. At the 5' end of the clone there is an untranslated region of 80 bp. By using the peptide sequences derived from the avocado fruit and *Brassica napus* seed it was possible to confirm that the cDNA clones encode 3-oxoacyl-ACP reductase amino acid sequences. The *Brassica napus* peptide sequences are identical with the derived amino acid sequence in the *Brassica napus* clone (peptides 1, 2, 5 and 7 from Fig. 1b) and also show a high degree of similarity to the *Arabidopsis thaliana*-clone-derived sequence. The avocado fruit peptide sequences are also very similar to the derived amino acid sequences of both the *Brassica napus* and *Arabidopsis thaliana* cDNA clones. The degree of identity between the *Brassica napus* seed and *Arabidopsis thaliana* leaf clones of 3-oxoacyl-ACP reductase is extremely high: 86% at the nucleotide level and 85% at the amino acid level. A sequence of high identity in the deduced amino acid sequence of the *Arabidopsis thaliana* clone was identified by aligning it with the N-terminal sequences of purified 3-oxoacyl-ACP reductase from *Brassica napus* seed and avocado fruit (Fig. 1), which suggests that the protein is synthesized as a precursor containing an N-terminal extension (Fig. 2a). This putative extension sequence contains two possible initiating methionine residues which would result in either a 12- or a 59-amino-acid leader peptide. Only the 59-amino-acid extension sequence reveals the three distinct features characteristic of stromal-targeting plastid transit peptides [18], namely an N-terminal domain that is uncharged (MAAFAA), a central domain with few or no acidic residues, and a C-terminal domain (residues 34–57) able to form an amphiphilic β -strand. The molecular mass of the protein encoded by the *Arabidopsis thaliana* 3-oxoacyl-ACP reductase clone is 27 208 Da, excluding the putative leader sequence. This is in good agreement with the subunit molecular masses of 28 kDa and 30 kDa determined by SDS/PAGE for the avocado and *Brassica napus* enzymes respectively [5,9].

Fig. 3 shows that the deduced amino acid sequences of the *Arabidopsis thaliana* 3-oxoacyl-ACP reductase and the *nodG*-gene product of *Rhizobium meliloti* [14] exhibit 41% overall identity, which rises to 60% if conservative substitutions are allowed for. The first 11 amino acids of the *nodG*-gene product show sequence identity with the N-terminal region of the transit sequence of the 3-oxoacyl-ACP reductase, but the remainder of the transit sequence is not present in the *nodG*-product sequence and the sequence identity re-commences at the proposed start of the mature plant protein. As the *nodG*-gene product is expressed within the bacterium, it would not need a stromal-transit-peptide sequence. The identity between the *Brassica napus* clone and the *nodG*-gene product is 51% over the region able to be compared, rising to 66% if conservative substitutions are allowed for. No significant similarity between plant 3-oxoacyl-ACP reductase

<i>nodG</i>	M A A A V A A P R L I S L K A L V - - - - -	11
Ar-R	M A A A V A A P R L I S L K A L V G K L G F R E I S Q I R Q L A P	32
<i>nodG</i>	- - - - -	64
Ar-R	I H S A I P H F G M L A C R S R Q P F S T S V V K A Q A T A T E	
<i>nodG</i>	Q S P G E V V V Q K V E S P V V V I T G A S G A I G G A I A R V L	26
Ar-R	Q S P G E V V V Q K V E S P V V V I T G A S G A I G G A I A R V L	96
<i>nodG</i>	- - H A Q G A I V G L H G T Q I E K L E T L - - A T E L G D R V	54
Ar-R	G K A G C K V I V N M A R S A K E A E V A K Q I E E W G G Q A	128
<i>nodG</i>	K L F P A N L A N R D E V K A L G Q R A E A D L E G V D I L V N	86
Ar-R	I T F G G D V S K A T D V D A M M K T A L D K M G T I D V V V N	160
<i>nodG</i>	N A G I T K D G T E L H M A D P D M D I L E V N L T A M F R L	118
Ar-R	N A G I T K D G T E L H M A D P D M D I L E V N L T A M F R L	192
Br-R	N A G I T K D G T E L H M A D P D M D I L E V N L T A M F R L	5
<i>nodG</i>	T R E I T Q Q M I R R R N G R I N V T S V A G A I G N I P G Q T	150
Ar-R	T Q S A A V K M M K K R G R I N I T S V V G L I G N I G O A	224
Br-R	S Q A A T K I M M K K R K G R I N I T A S V V G L I G N I G O A	37
<i>nodG</i>	N V I G A S K A C M I T G F S K S L A Q S I A R N I N V N C V A P	182
Ar-R	N V A A K G V I E F S K T A A R E G A S R N I N V N V C P	256
Br-R	N V A A K A G V I G F S K T A A R E G A S R N I N V N V C P	69
<i>nodG</i>	G F I E S A M T D K L N H K O K E K T M V A T P T H R M G T G T	214
Ar-R	G F I A S N M T A V L G E D M E K K I L G T I P L G R Y G E A G	288
Br-R	G F I A S E M T A K L G E D M E K K I L G T I P L G R Y G Q P E	101
<i>nodG</i>	E V A S A V A Y L A S D H A A L Y V T G Q T I H V N G C M A M I	245
Ar-R	G V A G L V E F L A L S P A A S Y I T G Q A F T I D G G I A I	319
Br-R	R V A G L V E F L A L S P A A S Y I T G Q T F T I D G G I A I	132

Fig. 3. Comparison of the deduced amino acid sequences of the *Brassica napus* (Br-R) and *Arabidopsis thaliana* (Ar-R) β -oxoacyl-ACP reductase clones with that of the *nodG*-gene product (*nodG*) of the bacterium, *Rhizobium meliloti*

Identical amino acid are boxed, and conservative substitutions are underlined individually. The continuous line indicates the position of the proposed transit peptide. Gaps to aid sequence comparison are indicated (---).

and the sequences of the mammalian and yeast fatty acid synthase multifunctional polypeptides was observed, and no bacterial 3-oxoacyl-ACP reductase sequences are yet available for comparison.

DISCUSSION

cDNA sequencing of plant 3-oxoacyl-ACP reductase has confirmed it to be a nuclear-encoded enzyme that is targeted to the plasmid stroma, being synthesized from cytoplasmic mRNA and possessing a transit peptide (Fig. 2a). The transit peptides of 3-oxoacyl-ACP reductase, ACP [4,6], 3-oxoacyl-ACP synthase I [7] and enoyl-ACP reductase [8], show no significant sequence identity with each other. However, analysis of the transit sequences reveals the characteristic features of stromal-targeting plastid-transit peptides [18]. The degree of identity between the *Brassica napus* seed and *Arabidopsis thaliana* leaf cDNA clones of 3-oxoacyl-ACP reductase is extremely high, and the minor variation in the 3-oxoacyl-ACP reductase sequences most probably reflects species differences rather than the existence of tissue-specific isoenzymes. This sequence similarity contrasts with plant ACP, where the degree of sequence diversity of ACP from different tissue types suggests a multigene family [4,6]. As the coding sequences of the seed and leaf 3-oxoacyl-ACP reductase are very similar, tissue-specific changes in gene expression which occur during development [9] must depend on tissue-specific control elements in the non-coding regions of the mRNA and/or the promoter of the genes.

Although plant ACP shows a significant sequence identity with the acyl-carrier domain of mammalian multifunctional fatty acid synthase [1], 3-oxoacyl-ACP reductase shows no such identity with the mammalian enzyme. As 3-oxoacyl-ACP synthase I shows only very limited similarity to the mammalian enzyme in the immediate active-site region [7], it appears that the enzyme components of the fatty-acid-synthesis pathway from different Kingdoms are less related than are the ACPs to the mammalian enzyme acyl-carrier domains. There are likely to

Fig. 4. Proposed structures of the nodulation signals of (a) *Rhizobium meliloti* (from Lerouge *et al.* [26]) and (b) *Rhizobium leguminosarum* (from Spaink *et al.* [27])

have been greater constraints on the evolution of the ACPs, as they must interact with all six or seven enzymes of the fatty-acid-biosynthesis pathway. The lack of conservation between the plant 3-oxoacyl-ACP reductase and 3-oxoacyl-ACP synthase I and the mammalian fatty acid synthase domains prevents any clarification of the evolutionary relationships between the mono-functional and multifunctional systems.

There had previously been reports that *nodG*-gene product has some limited similarity to several dehydrogenases, which is confined to the central portion of the *nodG*-gene product amino acid sequence [14,19,20]. These studies suggested that the conserved amino acids formed part of a nucleotide-binding site of a dehydrogenase or reductase. Our finding that plant 3-oxoacyl-ACP reductase has a high degree of identity with the *Rhizobium nodG*-gene product throughout the entire coding region strongly suggests that the *nodG*-gene product functions as an oxo-reductase. This observation was only possible because of the purification and cDNA cloning of the plant 3-oxoacyl-ACP reductase. This finding reinforces previous observations that other so-called 'host-specific' *nod*-gene products are related to enzymes of fatty acid biosynthesis [21], e.g. the *nodE* product has sequence similarity to the active-site region of the enzyme catalysing the condensation reaction of fatty acid biosynthesis (3-oxoacyl-ACP synthase) [22], whereas the *nodF* product carries a sequence similar to the phosphopantetheine-attachment site of ACP from *Escherichia coli* [23] and purified *nodF*-gene protein has a 4'-phosphopantetheine prosthetic group [24]. *Rhizobium* species have plasmids which carry different combinations of *nod* genes that determine which host plant the bacterium can infect symbiotically without eliciting the plant's defence mechanisms. For example, *Rhizobium meliloti* has *nodE*, *nodF* and *nodG* among its 'host-specific' *nod* genes, but *Rhizobium legumino-*

sarum, although retaining *nodE* and *nodF*, lacks *nodG* [25]. The current model for the role of the *nod*-gene family in *Rhizobium*-host signalling is that expression of the so-called 'common' *nod* genes results in the synthesis of the core signal molecule which is an oligosaccharide [21]. This core molecule is modified by other *nod*-gene products to give a 'host-specific' signal. The structures of the specific nodulation signals of *Rhizobium meliloti* and *Rhizobium leguminosarum* have recently been determined (Fig. 4). In the case of *R. meliloti* the key modifications of the oligosaccharide core are acylation by a C₁₆ bis-unsaturated fatty acid and N-sulphation of the oligosaccharide core [26]. For *R. leguminosarum* there is no N-sulphation of the oligosaccharide core, and the acylation modification is a C₁₈ tetra-unsaturated fatty acid [27]. The nodulation signal in *R. meliloti*, which carries the *nodG* gene, is more saturated than in *R. leguminosarum*, which lacks this gene. Intuitively one might expect this pattern if the *nodG*-gene product was an enoyl reductase; however, there is no sequence similarity between *Brassica napus* enoyl-ACP reductase and the *nodG*-gene protein [8]. As the *nodG*-gene protein is very similar to an oxoreductase, its absence presumably results in the increased desaturation of the *R. leguminosarum* fatty acid moiety compared with that of *R. meliloti*. The fatty acid moiety of the nodulation signals is quite unique, and it seems clear that the fatty-acid-biosynthetic genes that are encoded by the bacterial genome are unable to substitute for the *nod* genes carried on the nodulation plasmid. Therefore the substrate specificities of the plasmid- and genomic-encoded enzymes must be very different.

Several questions arise from these observations. How is host specificity obtained via acylation and sulphation of the core molecule? Is the hydrophobic nature of the acylation important? Furthermore, to what extent does the variation in the 'host-specific' *nod* genes result in heterogeneity of the final signal molecules? These questions will only be clarified when the location and molecular structure of the plant-nodulation-signal receptor is known. This receptor is likely to be quite unique to plants, as the signal molecule is unlike anything so far described for other eukaryotes. The function of the other *nod*-gene products will also have to be elucidated.

Finally, an intriguing evolutionary question concerns the origins of the *nod* genes. As several of them are very closely related to plant nuclear-encoded proteins, and the *nod* genes themselves are located on a plasmid, it may be that the *nod*-gene family arose from the plant host genome by horizontal transmission.

This work was supported by the Science and Engineering Research Council and a travel grant from The Royal Society (London).

REFERENCES

- McCarthy, A. D. & Hardie, D. G. (1984) Trends. Biol. Sci. **9**, 60–62
- Harwood, J. L. (1988) Annu. Rev. Plant Physiol. Plant Mol. Biol. **39**, 101–138
- Caughey, I. & Kekwick, G. O. (1982) Eur. J. Biochem. **123**, 553–561
- Safford, R., Windust, J. H. C., Lucas, C., De Silva, J., James, C. M., Hellyer, A., Smith, C. G., Slabas, A. R. & Hughes, S. G. (1988) Eur. J. Biochem. **174**, 287–295
- Sheldon, P. S., Kekwick, R. G. O., Smith, C. G., Sidebottom, C. M. & Slabas, A. R. (1990) Biochem. J. **271**, 713–720
- Post-Beittenmiller, M. A., Hlousek-Radojcic, A. & Ohlrogge, J. B. (1989) Nucleic Acids Res. **17**, 1777
- Siggaard-Andersen, M., Kauppinen, S. & von Wettstein-Knowles, P. (1991) Proc. Natl. Acad. Sci. U.S.A. **88**, 4114–4118
- Kater, M. M., Koningsstein, G. M., Nijkamp, H. J. J. & Stuitje, A. R. (1991) Plant Mol. Biol. **17**, 895–909
- Sheldon, P. S., Kekwick, R. G. O., Smith, C. G., Safford, R. & Slabas, A. R. (1991) Biochem. Biophys. Acta, in the press.
- Høj, P. B. & Mikkesen, J. D. (1982) Carlsberg Res. Commun. **47**, 119–141

11. Shimakata, T. & Stumpf, P. K. (1982) *Arch. Biochem. Biophys.* **217**, 144–148
12. Hendren, R. W. & Bloch, K. (1980) *J. Biol. Chem.* **255**, 1504–1508
13. Shimakata, T. & Stumpf, P. K. (1982) *Arch. Biochem. Biophys.* **218**, 77–91
14. Debelle, F. & Sharma, S. B. (1986) *Nucleic Acids Res.* **14**, 7453–7472
15. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Handbook*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
16. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
17. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13 [and addendum: *Anal. Biochem.* (1984) **137**, 266–267]
18. von Heijne, G., Steppuhn, J. & Herrmann, R. G. (1989) *Eur. J. Biochem.* **180**, 535–545
19. Baker, M. E. (1989) *Mol. Endocrinol.* **3**, 881–884
20. Baker, M. E. (1990) *FASEB J.* **4**, 222–226
21. Hopwood, D. A. & Sherman, D. H. (1990) *Annu. Rev. Genet.* **24**, 37–66
22. Bibb, M., Biro, S., Motamedi, H., Collins, J. F. & Hutchinson, C. R. (1989) *EMBO J.* **8**, 2727–2736
23. Sherman, C. A., Rossen, L., Johnston, A. W. B. & Downie, J. A. (1986) *EMBO J.* **5**, 647–652
24. Geiger, O., Spaink, H. P. & Kennedy, E. P. (1991) *J. Bacteriol.* **173**, 2872–2878
25. Long, S. R. (1989) *Cell (Cambridge, Mass.)* **56**, 203–214
26. Lerouge, P., Roche, P., Faucher, C., Maillet, F., Truchet, G., Prome, J. C. & Denarie, J. (1990) *Nature (London)* **344**, 781–784
27. Spaink, H. P., Sheeley, D. M., van Brussel, A. A. N., Glushka, J., York, W. S., Tak, T., Geiger, O., Kennedy, E. P., Reinhold, V. N. & Lugtenberg, B. J. J. (1991) *Nature (London)* **354**, 125–130

Received 12 December 1991/31 January 1992; accepted 5 February 1992