Folate biosynthesis in higher plants: purification and molecular cloning of a bifunctional 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase/7,8-dihydropteroate synthase localized in mitochondria

Fabrice Rébeillé¹, David Macherel, Jean-Marie Mouillon, Jérôme Garin² and Roland Douce

Laboratoire de Physiologie Cellulaire Végétale, CNRS URA no. 576, Département de Biologie Moléculaire et Structurale, CEA-Grenoble, F-38054 Grenoble Cédex 9, France

²Present address: Laboratoire de Chimie des Protéines, Département de Biologie Moléculaire et Structurale, CEA-Grenoble, F-38054 Grenoble Cédex 9, France

¹Corresponding author

In pea leaves, the synthesis of 7,8-dihydropteroate, a primary step in folate synthesis, was only detected in mitochondria. This reaction is catalyzed by a bifunctional 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase/7,8-dihydropteroate synthase enzyme, which represented 0.04–0.06% of the matrix proteins. The enzyme had a native mol. wt of 280–300 kDa and was made up of identical subunits of 53 kDa. The reaction catalyzed by the 7,8-dihydropteroate synthase domain of the protein was Mg^{2+} -dependent and behaved like a random bireactant system. The related cDNA contained an open reading frame of 1545 bp and the deduced amino acid sequence corresponded to a polypeptide of 515 residues with a calculated M_r of 56 454 Da. Comparison of the deduced amino acid sequence with the N-terminal sequence of the purified protein indicated that the plant enzyme is synthesized with a putative mitochondrial transit peptide of 28 amino acids. The calculated M_r of the mature protein was 53 450 Da. Southern blot experiments suggested that a single-copy gene codes for the enzyme. This result, together with the facts that the protein is synthesized with a mitochondrial transit peptide and that the activity was only detected in mitochondria, strongly supports the view that mitochondria is the major (unique?) site of 7,8-dihydropteroate synthesis in higher plant cells.

Keywords: dihydropteroate synthase/folate synthesis/plant mitochondria

Introduction

One-carbon metabolism in cells is mediated by a variety of tetrahydrofolate polyglutamate derivatives (Cossins, 1984; McGuire and Coward, 1984). As a result, a number of pathways such as those involved in the metabolisms of methionine, serine, glycine, purine or thymidylate are dependent on an endogenous supply of these coenzymes (McGuire and Bertino, 1981; Appling, 1991). Plants and microorganisms, in contrast to animals, are able to synthesize tetrahydrofolate from 6-hydroxymethyl-7,8-

dihydropterin *de novo*. This pathway requires the sequential operation of five enzymes: a 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK), a 7,8-dihydropteroate synthase (DHPS) (EC 2.5.1.15), a dihydrofolate synthetase (DHFS) (EC 6.3.2.12), a dihydrofolate reductase (DHFR) (EC 1.5.1.3) and a folylpolyglutamate synthetase (FPGS) (EC 6.3.2.17). In microorganisms, a lot of attention has been given to the first steps of this synthesis because they represent potential targets for antimicrobial agents. This is the case for DHPS, the target of sulfonamide drugs which are *p*-aminobenzoic acid (*p*-ABA) analogs that are recognized by DHPS as alternate substrates (Shiota, 1984; Hong *et al.*, 1995). In bacteria, such as *Bacillus subtilis* and *Streptococcus pneumoniae*, HPPK and DHPS are part of a cluster of genes involved in folate biosynthesis (Slock *et al.*, 1990; Lacks *et al.*, 1995). In *Escherichia coli*, HPPK is a monofunctional protein with a compositional mol. wt of ~18 kDa (Talarico *et al.*, 1992). In *S.pneumoniae*, the situation is different and HPPK is part of a bifunctional protein also supporting dihydroneopterin aldolase (DHNA) activity (Lopez and Lacks, 1993). DHNA catalyzes the conversion of 7,8-dihydroneopterin into 6-hydroxymethyl-7,8-dihydropterin, substrate of the HPPK activity. In this organism, HPPK/DHNA is a tetrameric protein made up of 31 kDa subunits. DHPS in prokaryotes is a monofunctional protein, distinct from HPPK, with a native mol. wt of ~70–100 kDa (Lopez *et al.*, 1987) and is made up of identical subunits of 30–34 kDa (Lopez *et al.*, 1987; Dallas *et al.*, 1992; Kelam *et al.*, 1995). In contrast, HPPK and DHPS activities were always associated in the eukaryotes studied so far. In the protozoa *Plasmodium falciparum* and *Toxoplasma gondii*, the DHPS enzyme is bifunctional, also supporting HPPK activity (Allegra *et al.*, 1990; Triglia and Cowman, 1994). In *P.falciparum*, the corresponding gene encodes a protein with an apparent M_r of 83 kDa but the observed size on SDS-PAGE gels is 68 kDa (Triglia and Cowman, 1994). In *T.gondii*, the primary structure of the protein was not determined, but the reported molecular weight of the native enzyme [125 kDa (Allegra *et al.*, 1990)] is different from that of *P.falciparum* [200 kDa (Walter and König, 1974)]. The situation is even more complex in the sporozoa *Pneumocystis carinii*, where the enzyme is at least a trifunctional polypeptide supporting DHNA, HPPK and DHPS activities (Volpe *et al.*, 1993). Surprisingly, the size of this trifunctional protein was similar to the size of the bifunctional enzyme of *P.falciparum* (Volpe *et al.*, 1993). Taken as a whole, these data indicate that the proteins supporting HPPK and DHPS activities may vary greatly from one species to another.

Folate metabolism in plants has been studied intensively by several authors, and much information regarding the nature of folates and the regulation of enzymes involved

Table I. DHPS and HPPK activities in the different pea leaf cell compartments

Enzyme	Mitochondria	Chloroplasts	Cytosol
Fumarase	51 000 ± 3000	ND	420 ± 120
PRK	ND	27 000 ± 2400	10 200 ± 2400
PEPc	ND	ND	6000 ± 800
DHPS	16 ± 3	ND	0.08 ± 0.04
HPPK + DHPS	1.8 ± 0.3	ND	ND

HPPK activity was estimated in association with DHPS activity (see Materials and methods). The results are the average of five determinations and are expressed as nmol/h/mg protein. Fumarase, phosphoribulokinase (PRK) and phosphoenolpyruvate carboxylase (PEPc) were marker enzymes for mitochondria, chloroplasts and cytosol respectively. ND: not detected.

in C1 metabolism is now available (for reviews, see Cossins, 1984, 1987). However, there is much less data concerning the initial steps of folate synthesis. In an early work, Okinaka and Iwai (1970a) purified the DHPS from whole pea seedlings. The enzyme appeared as a bifunctional protein, also supporting HPPK activity, with an apparent native M_r of 180 kDa. However, the protein was not characterized further and its primary structure was not determined. Fractionation studies suggested that a large part (73–78%) of the DHPS activity was present in mitochondria, but the cross-contamination of the various fractions was not examined (Okinaka and Iwai, 1970b). In a recent study, we presented evidence indicating that mitochondria contain all the enzymes involved in 6-hydroxymethyl-7,8-dihydropterin to tetrahydrofolate conversion, thus confirming the important role that these organelles play in folate synthesis (Neuburger *et al.*, 1996). In the present study, we have purified the HPPK/DHPS from pea leaf mitochondria and determined some of its kinetic properties. In a second step, we have isolated a cDNA clone encoding the entire protein to allow the comparison of plant HPPK/DHPS with those from other sources and as a prerequisite to study the regulation of the protein synthesis.

Results

Cellular distribution of DHPS

Cell fractionation experiments reported in the early work of Okinaka and Iwai (1970b) indicated that 73% of the DHPS activity was found in a mitochondrial fraction and 25% in a soluble fraction. However, as stated by these authors, it was not certain that the activity found in the soluble fraction was dependent on a DHPS isoenzyme because of a possible contamination from the mitochondrial fraction. Therefore, we measured DHPS and HPPK activities in highly purified cell organelles and in a cytosol-enriched fraction where contamination from the other cell compartments was estimated through the activities of marker enzymes (Table I). The purpose of this approach is not to calculate the intracellular distribution of HPPK and DHPS because the methods used to obtain the different fractions are only qualitative, but simply to assess whether or not these activities are present in a given compartment. As previously reported (Neuburger *et al.*, 1996), and shown in Table I, significant HPPK and DHPS activities were detected only in mitochondria. Indeed, the very low DHPS activity found in the cytosol fraction could be accounted for totally by the small mitochondrial contamination which represented ~0.8–1% of the proteins. It must

be pointed out that the HPPK activity reported here was estimated through DHPS activity, according to several authors (Okinaka and Iwai, 1970c; Talarico *et al.*, 1992; Lopez and Lacks, 1993). Separate control experiments indicate that this reaction, in contrast to the DHPS-catalyzed reaction, has an absolute requirement for ATP. Thus, estimations of the ATP-dependent global activity (HPPK + DHPS) do reflect the presence of a HPPK activity and not simply the ability of DHPS to convert 6-hydroxymethyl-7,8-dihydropterin into 7,8-dihydropteroate. The results presented in Table I suggest that higher plant mitochondria are a major site for 7,8-dihydropteroate synthesis. In order to strengthen this hypothesis, the mitochondrial DHPS protein was purified and the corresponding cDNA clone was isolated.

Purification and biochemical properties of the HPPK/DHPS

The DHPS from higher plant mitochondria was purified in three steps (Table II). During the course of purification, the HPPK activity always co-eluted with the DHPS activity, which strongly suggests that these two activities are supported by a common protein. Thus, this enzyme catalyzes the first two reactions involved in folate synthesis:

As shown in Table II, the yield of recovery of the HPPK/DHPS was ~30–40%, but this protein represented only 0.04–0.06% of the soluble proteins of the mitochondrial matrix. The global (HPPK + DHPS) activity was always lower than DHPS activity alone, a result previously observed with the pea seedling enzyme (Okinaka and Iwai, 1970a). It is not certain whether the global activity reported here reflects the maximal HPPK activity, which prevents any comparison between the two maximal rates. The measurement of HPPK activity alone requires a new assay method, the development of which is currently in progress. The apparent molecular weight of the native protein, determined either by gel filtration (Figure 1) or by PAGE in non-denaturing conditions (Figure 2A), was ~280–300 kDa. This result is not in agreement with the previous work of Okinaka and Iwai (1970a) who reported an M_r of 180 kDa. SDS-PAGE analysis (Figure 2B)

Table II. Purification steps for HPPK/DHPS from pea leaf mitochondria

Fractions	Total protein (mg)	HPPK + DHPS		DHPS		Yield (%)
		Total act. (nmol/h)	Sp. act. (nmol/h/mg)	Total act. (nmol/h)	Sp. act. (nmol/h/mg)	
Matrix	630	1010	1.6	9200	14.6	100
Superdex 200	160	768	4.8	8000	50	87
Mono Q	7.4	525	71	5500	743	60
Folate agarose	0.137	383	2800	3500	25 500	38

indicated that HPPK/DHPS was made up of identical subunits of ~53 kDa. Mass spectrometry analysis of the protein, using a laser desorption technique, indicated a single peak of 53.65 kDa (result not shown), thus confirming the purity of our protein and the SDS-PAGE molecular mass determination.

Preliminary studies of the kinetic properties of HPPK/DHPS are shown in Table III. The K_m values for 6-hydroxymethyl-7,8-dihydropterin and *p*-ABA were very low, <1 μ M, but the K_m value for 6-hydroxymethyl-7,8-dihydropterin-pyrophosphate was much higher (30–40 μ M). The maximal DHPS activity was obtained for temperatures close to 50°C and the optimum pH, as often observed for folate enzymes, was ~9. The catalytic properties of the DHPS domain of the bifunctional protein (Equation 2) were studied in more detail. As shown in Figure 3, when the initial rate of the reaction was measured in the presence of variable and limiting concentrations of each of the two substrates, we obtained a family of reciprocal plots intersecting at a unique point on the *x*-axis. On the basis of the rapid equilibrium hypothesis (Segel, 1975), these results are indicative of a random bireactant system where binding of one substrate does not change the affinity of the other. Interestingly, 7,8-dihydropteroate, the product of this reaction, appeared as a potent competitive inhibitor for either 6-hydroxymethyl-7,8-dihydropterin-pyrophosphate ($K_i = 7$ –11 μ M) or *p*-ABA ($K_i = 5$ –8 μ M) (Figure 4), suggesting a strong feed-back inhibition. The presence of Mg^{2+} was required for maximal DHPS activity (result not shown). Indeed, omission of Mg^{2+} from the buffer medium resulted in a 50–60% decrease in the activity, and a complete loss of activity was obtained in the presence of 1 mM EDTA. At the moment, we still do not know whether Mg^{2+} -bound substrates, such as Mg^{2+} -6-hydroxymethyl-7,8-dihydropterin-pyrophosphate, are the true substrates of the reaction or if Mg^{2+} is required at the level of the catalytic pocket. This question is under investigation.

Primary structure of the HPPK/DHPS

From the purified HPPK/DHPS enzyme, the N-terminal part (12 amino acids) and three internal sequences were determined by microsequencing (see Materials and methods and Figure 5). These sequences were compared with those present in protein data banks, which led to the prediction that internal sequences 1 and 2 belonged respectively to HPPK and DHPS. We selected the amino acid sequences DMGRTDG (internal sequence 1) and AHIINDV (internal sequence 2) to design degenerate oligonucleotides which were tailed at their 5' end with unrelated GC-rich sequences. cDNAs obtained by reverse

Fig. 1. Separation by gel filtration of the HPPK and DHPS activities. Soluble proteins from purified mitochondria were loaded on a Superdex 200 (Pharmacia) column as described in Materials and methods. DHPS activity (□); HPPK activity (○). To facilitate the graph reading, the DHPS activity was reduced three times. The arrows mark the position of the peaks of the P-protein (200 kDa), the L-protein (120 kDa), the T-protein (45 kDa) and the H-protein (15 kDa) of the glycine cleavage system (Bourguignon *et al.*, 1988) and the DHFR (140 kDa) (Neuburger *et al.*, 1996). The inset is the plot of the molecular weight of the markers (logarithm scale) versus the fraction number. The arrow indicates the position of the HPPK/DHPS. The calculated value from this curve is 280 kDa.

transcription of pea leaf mRNAs were then combined with these oligonucleotides in a two-step PCR protocol (see Materials and methods and Figure 6). This procedure allowed the correct amplification of the desired sequence in spite of the degeneracy and low T_m of the primers. A 700 bp PCR product was obtained and further amplified using the non-degenerate GC-rich sequences as primers. This PCR product was sequenced, and it appeared to encode the HPPK/DHPS region boarded by the internal sequences 1 and 2 (see Figure 5). Thus, our RT-PCR approach using GC-rich tailed oligonucleotides appeared to be straightforward because it limited the non-specific amplifications generally observed with degenerate oligonucleotides, and allowed a direct sequencing of the PCR product without tedious subcloning. Specific oligonucleotides derived from this first nucleotide sequence were used to amplify a 430 bp DNA fragment which was, in turn, used as a probe to screen a λ gt10 cDNA library prepared from green leaf mRNA. A single positive clone, containing a 2.1 kb insert, was picked out of 120 000 phages and purified. The digestion of phage DNA with *Eco*RI yielded

Fig. 2. Electrophoretic analysis of purified HPPK/DHPS from pea leaf mitochondria. (A) Electrophoresis in non-denaturing conditions. Lane 1, 6 µg of the purified enzyme recovered from the folate agarose column; lanes 2 and 3, molecular weight standards. (B) SDS-PAGE. Lane 1, molecular weight standards; lane 2, 10 µg of the purified enzyme recovered from the folate agarose column.

two fragments of ~1.8 and 0.3 kbp which were subcloned in pBluescript to yield pBSDH1 and pBSDH2 phagemids. The sequence analysis of pBSDH1 and pBSDH2 revealed that the isolated cDNA had a total length of 2159 nucleotides and contained an open reading frame (ORF) of 1545 bp starting at the ATG translation initiation start site at position 156 and ending at the TGA stop codon at position 1701 (Figure 5). The sequence analysis of pBSDH2 indicated that it contained the downstream region of the cDNA ending with a 56 nucleotide poly(A) tail. We do not know at present whether the poly(T) motif observed at the 5' end of the cDNA was representative of a non-coding region of the mRNA or whether it was the result of an artefact during the library construction.

As also shown in Figure 5, the three amino acid sequences, previously determined by microsequencing of the purified HPPK/DHPS, were recognized on the translated sequence, thus confirming the identity of this cDNA. The deduced amino acid sequence corresponds to a polypeptide of 515 residues with a calculated M_r of 56 454 Da. Comparison of the deduced amino acid sequence with the N-terminal sequence of the purified protein indicates that the HPPK/DHPS is synthesized as a cytosolic precursor containing a 28 amino acid putative transit peptide. As there is only one methionine before the first amino acid (Phe) of the mature protein, and because the reading frame is interrupted upward in the nucleotide sequence, it is likely that the ATG 156 encodes the real initiation codon. The transit peptide was classified as mitochondrial as it exhibits characteristics shared with other mitochondrial transit peptides such as charge transition, a high content of serine and basic residues and no acidic residues (Von Heijne, 1986; Gavel and Von Heijne, 1990). This is in agreement with the mitochondrial origin of the purified protein. The mature protein is predicted to consist of 487 amino acid residues, giving an M_r of 53.43 kDa.

The 1836 bp fragment obtained by digestion of the phage DNA with *Eco*RI (pBSDH1) was used in a Northern hybridization experiment carried out with total RNA

Table III. Kinetic constants, optimal temperature and optimal pH for the HPPK/DHPS of pea leaf mitochondria

	Kinetic parameters
K_m for 6-hydroxymethyl-7,8-dihydropterin (µM)	0.7
K_m for ATP (µM)	70
K_m for 6-hydroxymethyl-7,8-dihydropterin-pyrophosphate (µM)	30
K_m for <i>p</i> -ABA (µM)	0.6
Optimal temperature (°C)	50
Optimal pH	9

Fig. 3. Kinetic studies of the DHPS reaction. (A) Reciprocal plots showing the effect of varying *p*-ABA concentrations at different fixed concentrations of 6-hydroxymethyl-7,8-dihydropterin-pyrophosphate. The fixed 6-hydroxymethyl-7,8-dihydropterin-pyrophosphate concentrations were: ○ 300 µM, □ 50 µM, ▽ 20 µM, ● 10 µM. (B) Reciprocal plots showing the effect of varying 6-hydroxymethyl-7,8-dihydropterin-pyrophosphate concentrations at different fixed concentrations of *p*-ABA. The fixed *p*-ABA concentrations were: ■ 4 µM, ● 2 µM, ▽ 1.25 µM, □ 0.7 µM, ○ 0.6 µM.

Fig. 4. Inhibitory effect of 7,8-dihydropteroate upon DHPS activity. (A) Competitive inhibition of 7,8-dihydropteroate versus 6-hydroxymethyl-7,8-dihydropterin-pyrophosphate in the presence of a saturating concentration of p-ABA (10 μM). 7,8-dihydropteroate concentrations were: □ 50 μM, ▽ 10 μM, ○ 0 μM. (B) Competitive inhibition of 7,8-dihydropteroate versus p-ABA in the presence of a saturating concentration of 6-hydroxymethyl-7,8-dihydropterin-pyrophosphate (100 μM). 7,8-dihydropteroate concentrations were: ▽ 50 μM, □ 5 μM, ○ 0 μM.

prepared from dark- or light-grown leaves. A single transcript of 2.2 kb was detected after several days of exposure in both cases, with a higher intensity for etiolated leaves (Figure 7A). This shows that the length of the cDNA which we isolated fitted very well with that of the transcript. In addition, this figure also indicates that the steady-state level of the transcript encoding HPPK/DHPS was low and not promoted by light, at least at the developmental stage used in this analysis.

The 1836 bp ³²P-labeled probe was used in a Southern hybridization experiment with pea nuclear DNA digested with several restriction endonucleases (Figure 7B). After washing at moderate stringency (see Materials and methods), only one band was seen with *Eco*RI fragments whereas two bands could be detected with *Hind*III and *Eco*RV fragments. *Hind*III and *Eco*RV enzymes had restriction sites on the HPPK/DHPS cDNA at positions 1647 and 943, respectively. Therefore, each of these enzymes produced two restriction fragments which hybridized with the ³²P-labeled probe. These results strongly suggest that a single-copy gene codes for the HPPK/DHPS protein in pea, unless the mitochondrial protein had poor homology with isoenzymes possibly present in the other cell compartments. This last hypothesis is unlikely, however, since strong homologies were always observed among the various HPPK and DHPS reported so far. In addition, it is unlikely that two forms of HPPK/DHPS derive from the same gene because there is no start codon in the region coding for the N-terminal part of the mature protein, as is the case in human folylpolyglutamate synthetase (Freemantle *et al.*, 1995).

Comparison of the plant mitochondria HPPK/DHPS with other HPPK and DHPS

The predicted amino acid sequence of the plant mitochondria HPPK/DHPS was aligned with HPPK and DHPS of either bacteria (*E. coli*) (Dallas *et al.*, 1992; Talarico *et al.*, 1992), protozoa (*P. falciparum*) (Triglia and Cowman, 1994) or sporozoa (*P. carinii*) (Volpe *et al.*, 1992) (Figure

CTTTTCTTCTTCTCAGGATTTTGCTCCTCTAACCCTATGAGAAAGGGTTGTTA	60
ATTATCAGCTGATTGATATGGTAGCTTGAAAAAAGTGAATACAAAGGGTTAGGAT	120
ATTGACTTCAATACACTTCTGTAAGCCTCTCTCAGATGAGTATACTTAAGTCTGGGCG	180
MSIILKCLG	
TGCGCGGAATCAGCTCTGTGCTGCCAGAACTATCTTAAGTCTGGGTTTCTCTCT	240
V R G N Q L C A A R N Y L K V L G F S S	
TTACACAGCTCCAACTCTCTATGAAATCAAACTCAAGATGAAGTAGTGATTG	300
F H T A P N S S I E I Q T Q D E E V V I	
N-ter	
CTTTGGGAAGTAATGTAGGTGATAGACTACATACTTCAAGGAAGCTTGAATATGATGA	360
A L G S N V G D R L H N F K E A L K L M	
GGAAGTCAGCATACACATCACAAGACATGCAAGTCTGTATGAGACAGCAGCGATG	420
R K S G I H I T R H A S L Y E T A P A Y	
TTACTGACCAACCTCGCTTCTCACTCTGCAGTAAGAGCGGATACGAACTCGGGCCAC	480
V T D Q P R F L N S A V R A D T K L G P	
ATGAATATTGGCTCACTCAACGAATCGAGAAGGATATGGCGCTACTGATGGTATAA	540
H E L L A A L K R I E K D M G R T D G I	
I-1	
GGTATGTGCAAGGCCAATGACTTAGACATTTTGTCTATGGTAAATTTAAAGTCAGAT	600
R Y G P R P I D L D I L F Y G K F K V R	
CTGATATTCTCAGTACCTCAGAAAGAAATTTGGGAACGACCGTTTGTCTATGGCCCTT	660
S D I L T V P H E R I W E R P F V M A P	
TGATGGATTTGCTGGGAACAGCTATTGACAGTATGATAGTCTGCTAGCTGGCATTCTTT	720
L M D L L G T A I D S D T V A S W H S F	
CAGGCCATTTCTGGTGGACTAAATGCATTATGGGAAAGTAGTGAGGAATCCCTTATTG	780
S G H S G G L N A L W E K L G G E S L I	
GAGAGGAAGTATGATAGGTAATCGCTTGGCAATGGCTTACTGATGGTCTCGGAA	840
G E E G M Y R V M P V A N G L D W S R	
GAACATTGGTCATGGGATCTTAAATTTGACTCGAGATGTTTCACTGATGGAGGAATT	900
R T L V M G I L N L T P D S F S D G N	
TTCACTCTGGAAGTCTGCTGTTTTCGAGGACGGTAAATGATATCAGAGGGTCTGATA	960
F Q S V K S A V S Q A R L M I S E G A D	
TAATTGATATTGGTCTCAGTCTACTCGGCAATGGCATCAAGGATCTTCTCGGAGAG	1020
I I D I G A Q S T R P M A S R I S A E E	
AATTAGGTAGATTAATCCCTGCTTGGAGCTGTAATGTCAATACCTGAGTAGAGGAA	1080
E L G R L I P V L E A V M S I P E V E G	
AACATCATCTGTTGATCTTCTACTCTGAAGTTCATTAGAAGCAGTACGTAAGGGG	1140
K L I S V D T F Y S E V A L E A V R K G	
CTCATATTATAATGATGATCTGCGGAAAGTATGATGCAAGTATGTTTAAGGTCATGG	1200
A H I I N D V S A G K L D A S M F K V M	
I-2	
CAGAGCTTGATGTTCTTATGTCGCAATGCACATGAGGGGTGACCGAGTACAATGCAGG	1260
A E L D V P Y V A M H M R G D P S T M Q	
ATAGTGAACCTGAAATATGATATGTTTTCGAGGATATGCTCGGAATTATATCTCGC	1320
D S E N L K Y D N V C K D I S S E L Y S	
GGGTAGAGAGGCAGAAATATCGGGAATCCCGGATGAGGAGTATTATGAGACCTGGAA	1380
R V R E A E I S G I P A W R I M D P G	
TTGGATTCTCAAGAAACCGAAGCAATTTAGCGGCATCAACGGGAATACCTGATATTA	1440
I G S F K K T E D N L A A L T G I P D I	
I-3	
GAGAAGAGATTTCAAAAAGAGTTTGCCATCTCTCATGCTCTCTACTAATTTGGACCGT	1500
R E E I S K R S L A I S H A P I L I G P	
CAAGAAAGCGATTTTATGAGTGAATTTGCTCTCGCCCTCTGCGGTTGATAGAGATCCCG	1560
S R K R F L G E I C S R P S A V D R D P	
CTACCATGCTCTCTGTCACCGCAGGTGTGTGTGTCGCGCAATATTTGTCAGTGCATA	1620
A T I A S V T A G V L C G A N I V R V H	
ATGTTAAAGATAATCTAGATGCGGTGAAGCTTTGTCAGTCAATTTCTGAACAAAGAGTT	1680
N V K D N L D A V K L C D A I L K Q K S	
CTCCATAAAATTTAAACAGTGAATTCATCTTGCTATGTTGATGTCGACTGTTTCT	1740
S P I K F K Q -	
TCATATGCTTATGCTACCTCTTCTAGTGGCTTCATTCAACAAATCTATTACTGTTCTCTT	1800
TACCATGCTCACTGCTACTCTCAATGTCATATAGAATTTCTGTCATGCTGCGAGTGCA	1860
GAGCAAGAGCGTGAAGTGCATTAAGATGATGACACATGCAAAATGAGAATAAGTTG	1920
ACACTTCCTGCTCAATTAATTAAGCATACATTTTGAATAAATTTACGATCGTATGTA	1980
TTTCATTTCTTCCCATGTTATTGATTCAACAAATAGATTATTTCTAAGTTCAACATTT	2040
CACTTGTAATCTATTGATTCACAGATTTTGTTCATATGCAAAATGCAACTACTGTTCAAT	2100
CCTAAAAAAG	2159

Fig. 5. Nucleotide sequence and deduced amino acid sequence of the HPPK/DHPS cDNA. Initiation (ATG) and stop (TGA) translation codons are indicated in bold. Annealing regions of primers DM and AH (see Materials and methods) are overlined with dotted lines. The N-terminal and the three internal sequences (named I-1, I-2, I-3) determined by microsequencing of the purified HPPK/DHPS are marked in bold and underlined. Numbers indicate nucleotide residues. The sequence was registered in the EMBL Nucleotide Sequence Database under the accession No. Y08611.

8A and B). As shown in Figure 8A, the N-terminal part of the pea enzyme showed a good identity with HPPK from *E. coli* (~34% amino acid identity), whereas the C-terminal part of the protein presented numerous domains identical to the *E. coli* DHPS (~36% amino acid identity). Most of these common regions were also found in all HPPK and DHPS studied so far (Lopez *et al.*, 1987, 1990; Slock *et al.*, 1990; Dallas *et al.*, 1992; Talarico *et al.*, 1992; Volpe *et al.*, 1992; Ballantine *et al.*, 1994; Brooks *et al.*, 1994; Triglia and Cowman, 1994), indicating highly conserved domains which are possibly involved in substrate binding and catalysis. Unfortunately, there is

Fig. 6. Schematic representation of the cloning strategy to obtain the HPPK/DHPS cDNA. R.T., reverse transcriptase; DM and AH were the degenerated primers constructed from the internal sequences (black boxes) of the protein; X₁ and X₂ were the unrelated GC-rich sequences intended to increase the T_m of the whole primer. See Materials and methods for details.

presently no data available for the corresponding proteins in algae or fungi. Interestingly, the plant enzyme resembles a fusion of the two bacterial enzymes with a small spacer of 10–15 amino acids (Figure 8A). Indeed, addition of the molecular mass of the two bacterial enzymes gives a value of 48 kDa, which is roughly similar to the 53 kDa of the mature pea protein. This situation is not seen with HPPK/DHPS of other eucaryotes. In *P. carinii*, the central and C-terminal part of the enzyme, containing the HPPK and DHPS activities respectively, presented a strong homology with the pea enzyme (~36% amino acid identity) and the bacterial enzymes (Volpe *et al.*, 1993). These two regions were also comparable in size with either the plant protein or their bacterial counterparts (Figure 8B). However, as shown in Figure 8B, the N-terminal part of the *P. carinii* HPPK/DHPS is responsible for DHNA activity, an apparently unique feature among the known HPPK/DHPS. The pea leaf mitochondria HPPK/DHPS is also very different from the *P. falciparum* enzyme. In this latter case, the N-terminal part of the protein (HPPK domain) contains two large insertions of 92 amino acids and the C-terminal part (DHPS domain) a smaller insertion of 32 amino acids (Triglia and Cowman, 1994), thus conferring on the *P. falciparum* protein a relatively high molecular weight.

Discussion

Our preliminary studies indicate that plant DHPS catalyses a Mg^{2+} -dependent reaction that can be described as a random bireactant system, according to the rapid equilibrium hypothesis of Michaelis–Menten (Segel, 1975). Interestingly, 7,8-dihydropteroate was a competitive inhibitor of the two substrates of the pea leaf mitochondrial enzyme, suggesting that it interacts with both 6-hydroxymethyl-7,8-dihydropterin-pyrophosphate- and *p*-ABA-binding sites. Furthermore, the low inhibitory constant measured for 7,8-dihydropteroate [the K_i value for 7,8-dihydropteroate was, in fact, in the same range as many K_i values reported for sulfonamide compounds (Shiota, 1984; Allegra *et al.*, 1990; Zhang and Meshnick, 1991)] indicates that DHPS is tightly controlled by the product of the reaction. Similar results were also reported for *E. coli* DHPS (Roland *et al.*, 1979). Clearly, 7,8-dihydropteroate cannot accumulate in the matrix space of mitochondria and appears as an important regulatory point of the folate biosynthesis pathway.

The proteins involved in the first two steps of folate synthesis are very different from one species to another. In the prokaryotes *B. subtilis* (Slock *et al.*, 1990) and *E. coli* (Dallas *et al.*, 1992; Talarico *et al.*, 1992), HPPK and

Fig. 7. (A) Northern blot analysis of pea leaf RNA. Samples of denatured total RNA (10 μ g) from light-grown (L) and dark-grown (D) leaves were separated on a 1.2% agarose gel, transferred to a nitrocellulose membrane and hybridized with the 1.8 kbp 32 P-labeled cDNA probe (see Materials and methods). The blot was washed with $0.5 \times$ SSC/0.1% SDS at 65°C . Transcript sizes are indicated on the left. (B) Southern blot analysis of pea genomic DNA. A genomic DNA gel blot of pea DNA (10 μ g) digested with *Eco*RI (EI), *Hind*III (H) and *Eco*RV (EV) was hybridized with the 1.8 kbp 32 P-labeled cDNA probe as described in Materials and methods. The blot was washed with $2 \times$ SSC/0.1% SDS at 60°C . Arrowheads indicate the position of the two minor bands observed with *Hind*III and *Eco*RV. Approximate fragment sizes are indicated on the right.

DHPS are separate enzymes. In *S. pneumoniae* (Lopez and Lacks, 1993), these two enzymes are also distinct, but HPPK is part of a bifunctional enzyme with DHNA located at the N-terminal part of the protein. In eukaryotes, the related proteins studied so far were always multifunctional, containing HPPK and DHPS activities in *P. falciparum* and *T. gondii* (Allegra *et al.*, 1990; Triglia and Cowman, 1994) and DHNA, HPPK and DHPS activities in *P. carinii* (Volpe *et al.*, 1993). Our results indicate that in pea leaf mitochondria, as in *P. falciparum* and *T. gondii*, HPPK and DHPS activities are supported by a single protein. This conclusion is based on the analysis of the primary structure of the enzyme, reported here for the first time, and the presence of both activities in the purified protein. Because HPPK activity was not assayed directly, one may question whether this enzyme could also exist, to some extent, as a monofunctional protein. However, the existence of such a monofunctional enzyme was not revealed by Northern and Southern blot experiments, despite the fact that the cDNA probe used in these studies exhibited strong similarities with all the HPPKs known so far. On the contrary, these experiments indicated only one mRNA family and one gene corresponding to the bifunctional enzyme. The plant mitochondria protein appeared different from the bifunctional HPPK/DHPS of *P. falciparum*. Compared with the protozoan enzyme, which contains several amino acid insertions in its N-terminal part, the pea protein is much smaller and appears like a mere fusion of the two bacterial enzymes. In contrast, the M_r of the native bifunctional pea protein is much larger (300 kDa, possibly a hexamer) than the *P. falciparum* enzyme [190 kDa (Walter and König, 1974), possibly a dimer] or the *T. gondii* enzyme [125 kDa (Allegra *et al.*, 1990)].

The differences observed among the HPPKs and DHPSs

from various species strongly suggest that there were considerable variations in their gene evolution. Multifunctional enzymes are thought to have evolved by the fusion of smaller monofunctional proteins, each functional domain being connected by protease-sensitive linker regions (Schmincke-Ott and Bisswanger, 1980). In connection with this, it has been shown recently that the HPPK domain of the multifunctional protein of *P. carinii* can be overexpressed as an independent enzyme in *E. coli* (Ballantine *et al.*, 1994), and refolding of the recombinant protein yielded enzymatically active HPPK in a monomeric form. Although the function of such an evolution is not clearly understood, a possible advantage of the bifunctional nature of HPPK/DHPS could be to bring the two catalytic sites closer. Because these two activities are sequential in folate synthesis, such a situation would reduce the diffusional pathway in the mitochondrial matrix space containing a high protein concentration (0.4 mg/ml). Very little is known about the kinetic mechanisms of multifunctional enzymes, and it would be of interest to study the coordination of these two reactions either with the bifunctional enzyme or after splitting the protein into its two components. This point is currently under investigation.

In plants, HPPK and DHPS activities could be detected only in mitochondria. This observation, together with the fact that the bifunctional protein was synthesized with a putative mitochondrial transit peptide and that a single-copy gene was observed in nuclear DNA, strongly suggests that the mitochondria are, indeed, the unique site of 7,8-dihydropteroate synthesis in plants. In other eukaryotes, the subcellular compartmentalization of HPPK/DHPS was not investigated. It is interesting to note, however, that reported ORF sequences coding for *P. falciparum* (Triglia and Cowman, 1994) or *P. carinii* (Volpe *et al.*, 1992) HPPK/DHPS do not indicate the presence of such a transit peptide, although this assumption should be confirmed by the analysis of the N-terminal parts of the mature proteins. In contrast to higher plants where the pathway involved in 6-hydroxymethyl-7,8-dihydropterin to tetrahydrofolate conversion is present in mitochondria (Neuburger *et al.*, 1996), the subcellular distribution of the corresponding enzymes in the other eukaryotes remains obscure. For example, the DNA sequence coding for the bifunctional DHFR/TS in the protozoa *Leishmania major* predicts an amino acid extension from the N-terminal part that resembles the import sequence of mitochondrial proteins (Beverley *et al.*, 1986), but immunogold experiments indicated a cytosolic distribution of the protein (Swafford *et al.*, 1990). If this holds true, the mitochondrial localization of tetrahydrofolate synthesis in plants appears as a unique feature. This localization raises the problem of folate transport through the inner mitochondrial membrane and its distribution in the cytosol and plastids.

Finally, these results raise the question of the origin of the HPPK/DHPS substrates. There is actually no data available concerning the subcellular localization of 6-hydroxymethyl-7,8-dihydropterin synthesis in higher plants. *p*-ABA is synthesized from chorismate, a branch point in the aromatic amino acid pathway which has been localized in plastids (Herrmann, 1995). The synthesis of these metabolites and their transport across the mitochon-

A

HPPK <i>E.coli</i>	F H T A P N S S I E I Q T Q D E E V V	M T V A Y	I A I G S N L A S P	L E Q V N A	A L K A L	G D I P E S	H I L T	36
HPPK/DHPS <i>pea</i>	F H T A P N S S I E I Q T Q D E E V V	I A L G S N V G D R	L H N F K E	A L K	L	M R K S G I	H I T R	49
HPPK <i>E.coli</i>	V S S F Y R T P P L G P Q D Q P D Y L N A A V	G L E T S L A P E	E L L N H T Q	R I E L Q Q	G R V R K			86
HPPK/DHPS <i>pea</i>	H A S L Y E T A P A Y V T D Q P R F L N S A V	R A D T K L G P H	E L L A A L K	R I E K D M	G R T D G			99
HPPK <i>E.coli</i>	A E R W G P R T L D L D I M L F	G N E V I N T E R	L T V P H Y D M K N	R G F M L W P L F Q I A P E L				136
HPPK/DHPS <i>pea</i>	I R Y G P R P I D L D I L F Y G K F K V R S D I	L T V P H E R I W E R P F V M A P L M D L L G T A						148
HPPK <i>E.coli</i>	V F P D G E M L R Q I L H T R A F D K L N K W							159
HPPK/DHPS <i>pea</i>	I D S D T V A S W H S F S G H S G G L N A L W	E K L G G E S L I G E E G M Y R V M P V	A N G L L D W					198
DHPS <i>E.coli</i>						M K L F A Q G T S L D		11
HPPK/DHPS <i>pea</i>	S R R T L V M G I L N L T P D S F S D G G N F Q	S V K S A V S Q A R L M I S E G A D	I I D I G A Q S					248
DHPS <i>E.coli</i> L	L S H P H V M G I L N V T P D S F S D G G T H N	S L I D A V K H A N L M I N A G A T	I I D V G G E S					61
HPPK/DHPS <i>pea</i>	T R P M A S R I S A E E E L G R L I F V L E A V M S I P E V E G K L	I S V D T F Y S E V A L E A V R						298
DHPS <i>E.coli</i>	T R P G A A E V S V E E E L Q R V I F V V E A I A Q R F E V W . . .	I S V D T S K P E V I R E S A K						108
HPPK/DHPS <i>pea</i>	K G A H I I N D V S A G K L D A S M F K V M A E L D V P Y V A M H M R G D P S T M Q D S E N L K Y D							348
DHPS <i>E.coli</i>	V G A H I I N D I R S L S E P G A L E A A E T G . L E V C L M H M Q G N P K T M Q E A P K Y D D V							157
HPPK/DHPS <i>pea</i>	N V C K D I S S E L Y S R V R E A E I S G I P A W R I I M D P G I G F S K K T E D N L A A L T G I P							398
DHPS <i>E.coli</i>	F A E V N R Y F I E Q I A R C E Q . . A G I A K E K L L L D P G F G F G K N L S H N Y S L L A R L A							205
HPPK/DHPS <i>pea</i>	D I R E E I S K R S L A I S H A P I L I G P S R K R F L G E I C S R P S A V D R D P A T I A S V T A							448
DHPS <i>E.coli</i>	E F H H F N L P L L V G M S R K S M T G Q L L N V G P S E R L S G S L A C A V I A							246
HPPK/DHPS <i>pea</i>	G V L C G A N I V R V H N V K D N L D A V K L C D A I L K Q K S S P I K F K Q							487
DHPS <i>E.coli</i>	A M Q . G A H I I R V H D V K E T V E A M R V V E A T L S A K E N K R Y E							282

B

Fig. 8. Comparison of the predicted amino acid sequence of pea leaf mitochondria HPPK/DHPS with HPPK and DHPS primary sequences from several organisms. (A) Comparison between the pea protein and the two monofunctional HPPK (Talarico *et al.*, 1992) and DHPS (Dallas *et al.*, 1992) proteins of *E.coli*. Boxed areas indicate amino acid identity, dots indicate spaces inserted into the sequence to provide optimal homology, and numbers indicate amino acid residues. (B) Schematic comparison of the size of HPPK and DHPS from various organisms. Numbers indicate the first and the last amino acid of the proteins, black boxes indicate the main conserved regions of the different catalytic domains.

drial membranes to sustain the HPPK/DHPS activity might also be important regulatory points for folate synthesis.

Materials and methods

Materials

Pea (*Pisum sativum* L. Var. Douce Provence) plants were grown from seeds in vermiculite for 15 days under a 12 h photoperiod at 26°C (day) or 20°C (night). Mitochondria and chloroplasts were isolated and purified as previously described (Mourioux and Douce, 1981; Douce *et al.*, 1987) using self-generating gradients of Percoll. With these experimental procedures, mitochondria and chloroplasts were usually devoid of contamination from the other compartments. Soluble proteins from these

organelles were obtained as previously described (Neuburger *et al.*, 1986). The cytosol-enriched fraction was prepared as previously described (Neuburger *et al.*, 1996). Marker enzymes (fumarase for mitochondria, phosphoribulokinase for chloroplasts and phosphoenolpyruvate carb-oxylase for cytosol) were assayed in the various cell fractions as described in Neuburger *et al.* (1996).

6-Hydroxymethyl-pterin was obtained from Sigma and 6-hydroxy-methyl-pterin pyrophosphate was obtained from Schircks Laboratory, Jona, Switzerland. These products were reduced to dihydro-compounds as described by Scrimgeour (1980).

Determination of HPPK and DHPS activities

For these measurements, the media and all the solutions were maintained under a stream of argon to minimize the oxidation of the pterin substrates. The reactions were assayed at 30°C.

The HPPK activity was estimated in association with the DHPS activity according to Okinaka and Iwai (1970c). The standard reaction medium (medium A) contained, in a total volume of 120 μ l: 20 mM Tris, 20 mM K_2HPO_4 (pH 8), 20 mM β -mercaptoethanol, 15 mM $MgCl_2$, 10 mM ATP and various amounts of soluble proteins. Two μ l of 2 mM [carboxyl- ^{14}C]-ABA (1.85 Gbq/mmol) (ICN Biomedicals) were added to the assay medium, then the reaction was started by the addition of 100 μ M 6-hydroxymethyl-7,8-dihydropterin. After 20 min of incubation (the reaction is linear for at least 30 min), the reaction was stopped by heating the samples at 100°C for 5 min. The samples were centrifuged to remove the precipitated proteins and the [^{14}C]7,8-dihydropteroate formed was determined with a reverse phase HPLC system (Waters, Nova Pack C18 column) coupled with a Berthold (LB 506D) scintillation counter. The HPLC conditions were: solvent A, 0.1 M sodium acetate, pH 6; solvent B, acetonitrile; B increased linearly by 1% every minute; the flow-rate was 1 ml/min. With these experimental conditions, excess [^{14}C]-p-ABA was not retained in the column whereas [^{14}C]7,8-dihydropteroate, the product of the reaction, was eluted after 10 min of chromatography. The recovery of the enzyme-catalyzed product, as judged from the recovery of calibrated solutions of 7,8-dihydropteroate, was >90%.

The DHPS activity was measured in medium A (final volume 120 μ l) devoid of ATP. Two μ l of 2 mM [carboxyl- ^{14}C]-p-ABA (1.85 Gbq/mmol) were added to the assay medium, then the reaction was started by the addition of 100 μ M 6-hydroxymethyl-7,8-dihydropterin-pyrophosphate. After various times of incubation, the reaction was stopped and the [^{14}C]7,8-dihydropteroate formed was estimated as described above.

Purification of the mitochondrial HPPK/DHPS

The entire purification procedure was conducted at 4°C. Mitochondria from ~20 kg of leaves were required for this purification. Soluble proteins obtained from purified mitochondria (Neuburger *et al.*, 1986) were loaded on a Superdex 200 (Pharmacia) column (60 cm $\times$ 1 cm) previously equilibrated with the following buffer (buffer B): 10 mM K_2HPO_4/KH_2PO_4 , 10 mM Tris, pH 7.2, 1 mM EDTA, 1 mM dithiothreitol, 10 mM β -mercaptoethanol, 10% (w/v) glycerol. The proteins were eluted with the same buffer (rate of elution: 0.3 ml/min; fraction size: 2.4 ml). To estimate the molecular weight of eluted enzymes, the proteins P, H, T and L of the glycine cleavage system present in the matrix extract (Bourguignon *et al.*, 1988) were used as internal standards to calibrate the column.

In a second step, fractions containing the DHPS activity were combined, concentrated to a final volume of 2 ml by ultrafiltration on a 10 kDa cutoff membrane (Filtron), and loaded on a Mono Q HR 5/5 (Pharmacia) anion exchange column (0.5 cm $\times$ 5 cm) previously equilibrated with buffer C [25 mM KH_2PO_4/K_2HPO_4 (pH 7.2); 10% (w/v) glycerol]. The column was eluted with a linear KCl gradient in buffer C (0–500 mM; flow rate: 0.5 ml/min). DHPS emerged as a sharp peak at ~250 mM KCl. In a final step, fractions containing DHPS activity were pooled, concentrated to a final volume of 2 ml by ultrafiltration on a 10 kDa cutoff membrane, and loaded on a folate agarose affinity column (Sigma; 0.5 cm $\times$ 5 cm) previously equilibrated with buffer C. The flow rate was 0.3 ml/min. The column was washed for 40 min with buffer C, then eluted with a linear folate (pteroylmonoglutamate) gradient (0–1 mM) in buffer C. Fractions containing the purified DHPS were dialyzed against medium C and concentrated to a final volume of 200 μ l.

Electrophoretic methods

In denaturing conditions, the electrophoresis was performed in a SDS–polyacrylamide slab gel, 1 mm thick, containing a 7.5–15% linear acrylamide gradient. The procedures used were those described by Laemmli (1970). In non-denaturing conditions, the electrophoresis was performed without SDS on a linear acrylamide gradient (3.5–27%) as described by Clarke and Critchley (1992). Gradient polyacrylamide slab gels were stained with Coomassie Brilliant Blue R-250 as described by Chua (1980).

Determinations of N-terminal and internal sequences by microsequencing

Nine μ g of purified DHPS were analyzed by SDS–PAGE in a 7.5–15% acrylamide gradient (as described above). For N-terminal microsequencing, proteins on the gel were electrophoretically transferred onto PVDF (polyvinylidene difluoride) membrane (Problott, Applied Biosystems), using a Bio-Rad electrotransfer apparatus and a transfer medium containing 10 mM CAPS (3-cyclohexylamino-1-propanesulfonic acid), pH 11, in 10% methanol. PVDF membrane was stained in Coomassie Blue R-250 and destained in 50% methanol. The DHPS spot was

visualized and excised. Sequencing of the DHPS N-terminus was performed on an automated Applied Biosystems sequenator (477A) equipped with an on-line PTH-amino acid analyzer (120A). For internal microsequencing, the gel was stained for 20 min with 0.1% Coomassie Blue R-250 in 20% methanol, 0.5% acetic acid and the DHPS band was visualized and cut out from the gel for endolysin digestion. The stained protein band was then destained with 30% ethanol for 1 h at room temperature, washed twice with 20 ml of 50% acetonitrile, 100 mM ammonium hydrogenocarbonate for 15 min at 30°C, equilibrated twice with 20 ml of cleavage buffer (25 mM Tris–HCl, 1 mM EDTA, 10% acetonitrile, pH 8.5) and partially dehydrated in a Speed Vac concentrator (Savant). The gel slice was then rehydrated in 10 μ l of cleavage buffer. Then Lys endoprotease (Boehringer) was added as an enzyme to protein ratio of 1/50 (w/w). The digestion was carried out for 36 h at 37°C. Peptides were extracted as reported by Rosenfield *et al.* (1992) and were separated by reverse phase high performance chromatography using a VYDAC TP C18 column (5 μ m, 300 Å, 250 $\times$ 2.1 mm) as described in Bourmeyster *et al.* (1992). Peptide-containing fractions were spotted onto a glass fiber disk coated with polybrene, and amino acid sequencing was performed on an automated Applied Biosystems sequenator (477A) equipped with an on-line PTH-amino acid analyzer (120A).

Cloning of the cDNA encoding HPPK/DHPS

Two internal sequences of HPPK/DHPS were determined (see above) and their approximate location was deduced from their alignments with known sequences of several HPPKs and DHPSs. From these amino acid sequences, we designed several degenerate primers to clone the cDNA by RT–PCR. To increase PCR sensitivity, we applied the efficient method described by Weighard *et al.* (1993) to amplify genomic sequences. Thus we tailored our degenerate primers at their 5' ends with an unrelated GC-rich 17 nucleotide sequence, in order to increase the T_m of the whole primer. This was combined with an amplification protocol designed to amplify, in a first step, the degenerate primers and their complementary products for a few cycles, then, in a second step, after increasing the annealing temperature, to amplify the initial PCR products selectively (see Figure 6). The sense primers were DM1: 5'-**GGCATGCGGC-CGCTTCGGA**(TC)ATGGG(GTAC)CG(GATC)AC(GATC)GA(TC)GG or DM2: 5'-**GGCATGCGGC-CGCTTCGGA**(TC)ATGGG(GATC)AG-(GA)AC(GATC)GA(TC)GG. The sequence in bold is the GC-rich unrelated sequence called X1. The degenerate sequences derive from the amino acid sequence DMGRTDG, and the two primers correspond to the presence of arginine which is encoded by six codons. The antisense primer, named AH, is: 5'-**GGCAGCGTCCGTGACGGAC**(GA)TC(G-A)TT(GAT)AT(GAT)AT(GA)TG(GATC)GC. The sequence in bold is the unrelated GC-rich sequence called X2. The degenerate sequence corresponds to the inverse complemented strand encoding the amino acid sequence AHIINDV.

Poly(A)⁺ RNAs were extracted directly with magnetic oligo(dT) Dynabeads beads (Dyna, France) from 100 mg of 7-day-old pea leaves. The mRNA were reverse transcribed at 42°C with M-MuLV reverse transcriptase (New England Biolabs, USA) using oligo(dT) (20mer) primers. One twentieth of the reaction product was used as template for PCR, which was carried out with 2.5 U of *Taq* polymerase (Boehringer) in the presence of 100 pmol of DM1 or DM2 paired with AH in the recommended buffer. The amplification was performed according to the following scheme with a Robocycler gradient 40 (Stratagene, USA): 1 min, 94°C; 2 min, 45°C; 3 min, 72°C; 30 s, 94°C; 2 min 50°C; 3 min, 72°C for four cycles and finally 20 s, 94°C; 1 min, 62°C; 2 min, 72°C for 30 cycles. The analysis by agarose gel electrophoresis revealed the presence of a single product of ~700 bp which was more abundant with DM2 than with DM1, suggesting that the arginine of the DMGRTDG sequence was encoded by the codon AG(GA). This PCR product was in turn used as a template for a PCR carried out with the non-degenerate primers X1 and X2 at an annealing temperature of 60°C. After ethanol precipitation and agarose gel electrophoresis, the 700 bp product was extracted from the gel using Jetsorb (Bioprobe). Its sequence was determined with dye terminators using 100 ng of DNA, 5 pmol of primer X1 or X2, and the Applied Biosystems protocol carried out by Genome Express (Grenoble, France). The translation of the DNA sequence revealed a continuous reading frame including both internal amino acid sequences and exhibiting a strong homology with the corresponding region of other DHPS. We synthesized a pair of specific primers DHF (5'-GGTATGGTCCAAGGCCAA) and DHR (5'-CGAGTAGACTG-AGCACCA) to detect by PCR the presence of the DHPS cDNA in the pea cDNA libraries available in the laboratory. Aliquots of amplified libraries were treated at 95°C for 3 min in 10 μ l of H₂O containing 0.05% (v/v) Tween-20, cooled on ice and used as a template in a PCR

with 100 pmol of primers DHF and DHR and 1.5 U of *Taq* polymerase (Boehringer). The reaction conditions were 30 s, 95°C; 45 s, 62°C; 1 min, 72°C for 30 cycles. The gel electrophoresis revealed the specific amplification of a 430 bp product matching the expected size between DHF and DHR. A λ gt10 cDNA library (Macherel *et al.*, 1990) was chosen for screening. The amplified 430 bp fragment was random labeled with 32 P using a Rediprime kit (Amersham, UK). Approximately 120 000 phage were screened by hybridization of phage DNA on Hybond N⁺ membrane (Amersham) with the labeled probe according to the manufacturer's instructions. A single positive clone (LgDH) was purified by two successive rounds of screening and shown to contain a 2.1 kbp insert by PCR analysis. The phage DNA was extracted and digested by *Eco*RI yielding two DNA inserts of ~1800 and 300 bp which were subcloned in the *Eco*RI site of pBluescript SK⁺ to give the respectively recombinant phagemids pBSDH1 and pBSDH2. The DNA inserts were sequenced entirely on both strands by PCR amplification (Genome Express, France). Computer handling of the sequences was carried out with the program PCGene (Intelligenetics, USA).

Northern and Southern blot analysis

High molecular mass nuclear DNA was prepared according to Watson and Thompson (1986). It was digested overnight with restriction endonucleases then concentrated by ethanol precipitation. Separation of the DNA fragments by agarose (0.8%) electrophoresis and Southern transfer on Hybond N⁺ membrane (Amersham) were done as described by Maniatis *et al.* (1982). Hybridization with the 1.8 kbp (pBSDH1) 32 P-labeled probe (Rediprime kit, Amersham) was conducted overnight at 65°C in the buffer recommended for the membrane. The final washes were at 60°C in 2× SSC/0.1% SDS. For Northern blot experiments, total RNA was prepared from light- or dark-grown leaves (7 days old) according to Macherel *et al.* (1990). RNA (10 µg) was denatured with glyoxal and separated by electrophoresis according to Maniatis *et al.* (1982). Transfer to Hybond N⁺ membrane (Amersham) and hybridization at 65°C were carried out following the manufacturer's recommendations. The final wash was at 65°C in 0.5× SSC, 0.1 SDS.

Acknowledgements

We are grateful to Dr M.Jacquinot (Laboratoire de spectrométrie de masse des protéines, Institut de Biologie Structurale, Grenoble, France) for mass spectrometry analysis of DHPS and to Drs M.Neuburger and J.Bourguignon for helpful discussions.

References

Allegra, C.J., Boarman, D., Kovacs, J.A., Morrison, P., Beaver, J., Chabner, B.A. and Masur, H. (1990) Interaction of sulfonamide and sulfone compounds with *Toxoplasma gondii* dihydropteroate synthase. *J. Clin. Invest.*, **85**, 371–379.

Appling, D.R. (1991) Compartmentation of folate-mediated one-carbon metabolism in eukaryotes. *FASEB J.*, **5**, 2645–2651.

Ballantine, S.P., Volpe, F. and Delves, C.J. (1994) The hydroxymethyl-dihydropterin pyrophosphokinase domain of the multifunctional folic acid synthesis fas protein of *Pneumocystis carinii* expressed as an independent enzyme in *Escherichia coli*: refolding and characterization of the recombinant enzyme. *Protein Expr. Purif.*, **5**, 371–378.

Beverly, S.M., Ellenberg, T.E. and Cordingley, J.S. (1986) Primary structure of the gene encoding the bifunctional dihydrofolate reductase-thymidylate synthase of *Leishmania major*. *Proc. Natl Acad. Sci. USA*, **83**, 2584–2588.

Bourguignon, J., Neuburger, M. and Douce, R. (1988) Resolution and characterization of the glycine-cleavage reaction in pea leaf mitochondria. *Biochem. J.*, **255**, 169–178.

Bourmeyster, N., Stasia, M.J., Garin, J., Boquet, P. and Vignais, P.V. (1992) Copurification of Rho protein and the Rho-GDP dissociation inhibitor from bovine neutrophil cytosol. Effect of phosphoinositides on Rho ADP-ribosylation by the C3 exoenzyme of *Clostridium botulinum*. *Biochemistry*, **31**, 12863–12869.

Brooks, D.R., Wang, P., Read, M., Watkins, W.M., Sims, P.F.G. and Hyde, J.E. (1994) Sequence variation of the hydroxymethyl-dihydropterin pyrophosphokinase: dihydropteroate synthase gene in lines of the human malaria parasite, *Plasmodium falciparum*, with differing resistance to sulfadoxine. *Eur. J. Biochem.*, **224**, 397–405.

Chua, N.H. (1980) Electrophoretic analysis of chloroplast proteins. *Methods Enzymol.*, **69**, 434–446.

Clarke, A.K. and Critchley, C. (1992) The identification of the heat shock protein complex in chloroplasts of barley leaves. *Plant Physiol.*, **100**, 2081–2089.

Cossins, E.A. (1984) Foliates in biological materials. In Blakley, R.L. and Benkovic, S.J. (eds), *Foliates and Pterins*. Wiley (Interscience), New York, Vol. 1, pp. 1–59.

Cossins, E.A. (1987) Folate biochemistry and the metabolism of one-carbon units. In Davies, D.D. (ed), *The Biochemistry of Plants*. Academic Press, San Diego, CA, Vol. 11, pp. 317–353.

Dallas, W.S., Gowen, J., Ray, P.H., Cox, J. and Dev, I.K. (1992) Cloning, sequencing, and enhanced expression of the dihydropteroate synthase gene of *Escherichia coli* MC4100. *J. Bacteriol.*, **174**, 5961–5970.

Douce, R., Bourguignon, J., Brouquisse, R. and Neuburger, M. (1987) Isolation of plant mitochondria: general principles and criteria of integrity. *Methods Enzymol.*, **148**, 403–425.

Freemantle, S.J., Taylor, S.M., Krystal, G. and Moran, R.G. (1995) Upstream organization of and multiple transcripts from the human folylpoly- γ -glutamate synthetase gene. *J. Biol. Chem.*, **270**, 9579–9584.

Gavel, Y. and Von Heijne, G. (1990) Cleavage-site motifs in mitochondrial targeting peptides. *Protein Engng.*, **4**, 33–37.

Herrmann, K.M. (1995) The shikimate pathway: early steps in the biosynthesis of aromatic compounds. *The Plant Cell*, **7**, 907–919.

Hong, Y.-L., Hossler, P.A., Calhoun, D.H. and Meshnick, S.R. (1995) Inhibition of recombinant *Pneumocystis carinii* dihydropteroate synthetase by sulfa drugs. *J. Antimicrob. Chemother.*, **39**, 1756–1763.

Kellam, P., Dallas, W., Ballantine, S.P. and Delves, C.J. (1995) Functional cloning of the dihydropteroate synthase gene of *Staphylococcus haemolyticus*. *FEMS Microbiol. Lett.*, **134**, 165–169.

Lacks, S.A., Greenberg, B. and Lopez, P. (1995) A cluster of four genes encoding enzymes for five steps in the folate biosynthetic pathway of *Streptococcus pneumoniae*. *J. Bacteriol.*, **177**, 66–74.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.

Lopez, P. and Lacks, S.A. (1993) A bifunctional protein in the folate biosynthetic pathway of *Streptococcus pneumoniae* with dihydropterin aldolase and hydroxymethyl-dihydropterin pyrophosphokinase activities. *J. Bacteriol.*, **175**, 2214–2200.

Lopez, P., Espinosa, M., Greenberg, B. and Lacks, S.A. (1987) Sulfonamide resistance in *Streptococcus pneumoniae*: DNA sequence of the gene encoding dihydropteroate synthase and characterization of the enzyme. *J. Bacteriol.*, **169**, 4320–4326.

Lopez, P., Greenberg, B. and Lacks, S.A. (1990) DNA sequence of folate biosynthesis gene *sulD*, encoding hydroxymethyl-dihydropterin pyrophosphokinase in *Streptococcus pneumoniae*, and characterization of the enzyme. *J. Bacteriol.*, **172**, 4766–4774.

Macherel, D., Lebrun, M., Gagnon, J., Neuburger, M. and Douce, R. (1990) cDNA cloning, primary structure and gene expression for H-protein, a component of the glycine-cleavage system (glycine decarboxylase) of pea (*Pisum sativum*) leaf mitochondria. *Biochem. J.*, **268**, 783–789.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

McGuire, J.J. and Bertino, J.R. (1981) Enzymatic synthesis and function of polyglutamates. *Mol. Cell. Biochem.*, **38**, 19–48.

McGuire, J.J. and Coward, J.K. (1984) Pteroylpolylglutamates: biosynthesis, degradation and function. In Blakley, R.L. and Benkovic, S.J. (eds), *Foliates and Pterins*. Wiley (Interscience), New York, Vol. 1, pp. 135–190.

Mouriaux, G. and Douce, R. (1981) Slow passive diffusion of orthophosphate between intact isolated chloroplasts and suspending medium. *Plant Physiol.*, **67**, 470–473.

Neuburger, M., Bourguignon, J. and Douce, R. (1986) Isolation of a large complex from the matrix of pea leaf mitochondria involved in the rapid transformation of glycine into serine. *FEBS Lett.*, **207**, 18–22.

Neuburger, M., Rébeillé, F., Jourdain, A., Nakamura, S. and Douce, R. (1996) Mitochondria are a major site for folate and thymidylate synthesis in plants. *J. Biol. Chem.*, **271**, 9466–9472.

Okinaka, O. and Iwai, K. (1970a) The biosynthesis of folic acid compounds in plants. IV. Purification and properties of the dihydropteroate-synthesizing enzyme from pea seedlings. *J. Vitaminol.*, **16**, 201–209.

Okinaka, O. and Iwai, K. (1970b) The biosynthesis of folic acid compounds in plants. III. Distribution of the dihydropteroate synthesizing enzyme in plants. *J. Vitaminol.*, **16**, 196–200.

Okinaka, O. and Iwai, K. (1970c) The biosynthesis of folic acid compounds in plants. V. Reaction mechanism of the dihydropteroate-synthesizing enzyme from pea seedlings. *J. Vitaminol.*, **16**, 210–214.

- Roland, S., Ferone, R., Harvey, R.J., Styles, V.L. and Morrison, R.W. (1979) The characteristics and significance of sulfonamides as substrates for *Escherichia coli* dihydropteroate synthase. *J. Biol. Chem.*, **254**, 10337–10345.
- Rosenfield, J., Capdevielle, J., Guillemot, J.-C. and Ferrara, P. (1992) In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. *Anal. Biochem.*, **203**, 173–179.
- Schmincke-Ott, E. and Bisswanger, H. (1980) Introduction. In *Multifunctional Proteins*. Bisswanger, H. and Schmincke-Ott, E. (eds), Wiley, New York, pp. 1–29.
- Scrimgeour, K.G. (1980) Methods for reduction, stabilization, and analyses of folates. *Methods Enzymol.*, **66**, 517–523.
- Segel, I.H. (1975) *Enzyme Kinetics*. John Wiley and Sons, Inc., New York.
- Shiota, T. (1984) Biosynthesis of folate from pterin precursors. In Blakley, R.L. and Benkovic, S.J. (eds), *Folates and Pterins*. Wiley (Interscience), New York, Vol. 1, pp. 121–134.
- Slock, J., Stahly, D.P., Han, C., Six, E.W. and Crawford, I.P. (1990) An apparent *Bacillus subtilis* folic acid biosynthetic operon containing *pab*, an amphibolic *trpG* gene, a third gene required for synthesis of para-aminobenzoic acid, and the dihydropteroate synthase gene. *J. Bacteriol.*, **172**, 7211–7226.
- Swafford, J.R., Beverley, S.M., Kan, C.C. and Caulfield, J.P. (1990) Distribution of dihydrofolate reductase-thymidylate synthetase polypeptide in methotrexate resistant and wild type *Leishmania major*: evidence for cytoplasmic distribution despite a hydrophilic leader sequence. In Peachey, L.D. and Williams, D.B. (eds), *Proceedings of the XIIIth Congress for Electron Microscopy*. San Francisco Press, San Francisco, CA, pp. 610–611.
- Talarico, T.L., Ray, P.H., Dev, I.K., Merrill, B. and Dallas, W.S. (1992) Cloning, sequence analysis, and overexpression of *Escherichia coli* *folK*, the gene coding for 7,8-dihydro-6-hydroxymethylpterin-pyrophosphokinase. *J. Bacteriol.*, **174**, 5971–5977.
- Triglia, T. and Cowman, A.F. (1994) Primary structure and expression of the dihydropteroate synthase gene of *Plasmodium falciparum*. *Proc. Natl Acad. Sci. USA*, **91**, 7149–7153.
- Volpe, F., Dyer, M., Scaife, J.G., Darby, G., Stammers, D.K. and Delves, C.J. (1992) The multifunctional folic acid synthesis *fas* gene of *Pneumocystis carinii* appears to encode dihydropteroate synthase and hydroxymethyldihydropterin pyrophosphokinase. *Gene*, **112**, 213–218.
- Volpe, F., Ballantine, S.P. and Delves, C.J. (1993) The multifunctional folic acid synthesis *fas* gene of *Pneumocystis carinii* encodes dihydro-neopterin aldolase, hydroxymethyldihydropterin pyrophosphokinase and dihydropteroate synthase. *Eur. J. Biochem.*, **216**, 449–458.
- Von Heijne, G. (1986) Mitochondrial targeting sequences may form amphiphilic helices. *EMBO J.*, **5**, 1335–1342.
- Walter, R.D. and König, E. (1974) Purification and properties of the 7,8-dihydropteroate-synthesizing enzyme from *Plasmodium chabaudi*. *Z. Physiol. Chem.*, **355**, 431–437.
- Watson, J.C. and Thompson, W.F. (1986) Purification and restriction endonuclease analysis of plant nuclear DNA. *Methods Enzymol.*, **118**, 57–75.
- Weighardt, F., Biamonti, G. and Riva, S. (1993) A simple procedure for enhancing PCR specificity. *PCR Methods and Applications*, **3**, 77–80.
- Zhang, Y. and Meshnick, S.R. (1991) Inhibition of *Plasmodium falciparum* dihydropteroate synthetase and growth *in vitro* by sulfa drugs. *Antimicrob. Agents Chemother.*, **35**, 267–271.

Received on October 7, 1996; revised on November 22, 1996