ELSEVIER

Gene 209 (1998) 167–174

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

# Gene structure, expression in *Escherichia coli* and biochemical properties of the NAD$^+$-dependent glyceraldehyde-3-phosphate dehydrogenase from *Pinus sylvestris* chloroplasts[1]

Gilbert Meyer-Gauen [2], Heike Herbrand, Jens Pahnke, Rüdiger Cerff, William Martin *

*Institut für Genetik, Technische Universität Braunschweig, Spielmannstrasse 7, D-38023 Braunschweig, Germany*

Received 17 October 1997; accepted 23 December 1997

**Abstract**

Photosynthetic eukaryotes typically possess two distinct glyceraldehyde-3-phosphate dehydrogenases, an NAD$^+$-specific enzyme in the cytosol (GapC: EC 1.2.1.12) and an NADP$^+$-dependent enzyme in the chloroplast (GapAB: EC 1.2.1.13). The gymnosperm *Pinus sylvestris* is an exception in that it is known to express a gene encoding a transit peptide-bearing GapC-like subunit that is imported into chloroplasts (GapCp), but the enzymatic properties of this novel GAPDH have not been described from any source. We have expressed the mature GapCp subunit from *Pinus* in *Escherichia coli* and have characterized the active enzyme. GapCp has a specific activity of 89 units per milligram and is strictly NAD$^+$-dependent, showing no detectable activity with NADP$^+$. Values of the apparent $K_m$ for NAD$^+$ and glyceraldehyde-3-phosphate were determined as 62 and 344 μM, respectively. The *Pinus GapCp1* gene possesses 12 introns, two in the region encoding the transit peptide and ten in the region encoding the mature subunit, all of which are found at positions strictly conserved across genes for higher plant *GapC*. A cDNA encoding a homologue of GapCp was isolated from the heterosporous fern *Marsilea quadrifolia*, indicating that NAD$^+$-dependent chloroplast GAPDH also occurs in other higher plants. © 1998 Elsevier Science B.V.

*Keywords:* GAPDH; Chloroplasts; Introns; Gymnosperms

## 1. Introduction

GAPDH is a ubiquitous enzyme integral to carbohydrate metabolism in prokaryotes and eukaryotes. The Calvin cycle enzyme from higher plant chloroplasts has been characterized from many sources. It is an NADP$^+$-dependent $A_2B_2$ heterotetramer of GapA and GapB subunits that undergoes aggregation in the dark to a 4-fold less active oligomeric form — probably a tetramer of tetramers — and thioredoxin-modulated reactivation in the light (Cerff, 1978a, 1982; Brinkmann et al., 1989; Scagliarini et al., 1993; Baalmann et al., 1994, 1996). By contrast, the cytosolic enzyme of glycolysis/gluconeogenesis is a strictly NAD$^+$-dependent homotetramer of GapC subunits and is not light-regulated at the protein level (Cerff and Quail, 1974; Cerff and Kloppstech, 1982). In a previous study, we reported cDNAs from *Pinus sylvestris* seedlings encoding a novel protein that is closely related to higher plant cytosolic GapC, but is synthesized as a cytosolic, transit peptide-bearing precursor which is imported into pea chloroplasts in vitro under cleavage of the transit peptide (Meyer-Gauen et al., 1994). This plastid-specific GapC protein was thus designated GapCp.

The presence of specific motifs in the primary sequence that are known to be involved in NAD$^+$/NADP$^+$ cosubstrate specificity suggested that the *Pinus* GapCp mRNA should encode an NAD$^+$-specific GAPDH enzyme. However, NAD$^+$-specific GAPDH enzymes have not been isolated from any plastid source. More recently, cDNAs for GapC-related proteins were cloned from several photosynthetic protists that possess N-terminal transit peptides, but, curiously, also possess primary structure motifs, which suggest that they may encode NADP$^+$-dependent GAPDH enzymes (Liaud et al., 1997).

These findings indicate that the classical dichotomy

* Corresponding author. Tel.: +49 531 391 5785; Fax +49 531 391 5785; E-mail: w.martin@tu-bs.de

[1] Sequences reported in this paper have been deposited with GenBank (XAJ001706, *Pinus GapCp1* gene; AJ003783, *Marsilea quadrifolia* GapCp cDNA).

[2] Present address: University of Texas at Houston, School of Public Health, P.O. Box 20334, Houston, TX 77225, USA.


*PII* S0378-1119(98)00034-1

of plant GAPDH activities cytosol = $NAD^+$ and chloroplast = $NADP^+$ is not generally valid outside the angiosperms, where the vast majority of biochemical studies on plant GAPDH have been conducted. We wished to further investigate the biochemical properties of GapCp from *Pinus* seedlings, in particular with respect to cosubstrate specificity, but attempts to purify sufficient amounts of active GapCp for characterization from this source were unsuccessful. As an alternative approach to characterize this protein, we have overexpressed it in *E. coli*. Here we report the structure of the gene for GapCp from *Pinus sylvestris*, the expression of the mature GapCp subunit in *E. coli* and the properties of the active enzyme.

## 2. Materials and methods

### 2.1. Genomic cloning

DNA was obtained from 2-week-old light-grown *Pinus sylvestris* seedlings as described (Schwarz-Sommer et al., 1984). *Mbo* I fragments with a size of 15–20 kb were purified on sucrose gradients (Sambrook et al., 1989), ligated into *Bam*HI λEMBL4 arms and packaged with Gigapack II extracts (Stratagene, La Jolla, CA). Filters ($10^6$ recombinants) were screened by plaque hybridization with the randomly labelled *Not*I insert of pGapCp17 from *Pinus sylvestris* (Meyer-Gauen et al., 1994) as previously described (Henze et al., 1994) and washed at 68°C in 2 × SSPE, 0.1% sodium dodecyl sulfate (SDS). Ten positively hybridizing phages were sequenced by the dideoxy method using the primer 5′-CAYTCRTTRTCRTCCA-3′ constructed against the conserved GAPDH amino acid sequence WYDNE (Martin et al., 1993). One clone (λPS47) revealed a sequence 100% identical to that of the GapCp17 cDNA. *Eco*RI and *Bam*HI restriction fragments of λPS47 were subcloned into pSK plasmids (Stratagene), nested deletion series from λPS47 subclones were sequenced on both strands using radioactive techniques.

### 2.2. mRNA isolation and cDNA cloning

Twenty grams of young leaves of *Marsilea quadrifolia* were ground in liquid nitrogen and vigorously shaken for 5 min in 100 ml of 50 mM Tris–HCl (pH 9.0), 100 mM NaCl, 10 mM EDTA, 2% (w/v) SDS, 20 mM 2-mercaptoethanol, 10% w/v Polyklar AT (Sigma, St Louis, MO), 200 μg/ml proteinase K (Merck) and centrifuged for 5 min at 5000 × *g*. The supernatant was extracted vigorously with one volume of phenol/chloroform/isoamyl alcohol (25/24/1, v/v) and centrifuged as above. The aqueous phase was stirred while 0.3 vol of ethanol were slowly added, this mixture was vigorously extracted with 1 vol of chloroform/isoamyl alcohol (24/1, v/v) and centrifuged for 10 min at 15 000 × *g* (copious amounts of polysaccharides accumulate in the interphase). The aqueous phase was precipitated with one volume of isopropanol for 1 h at −20°C. The pellet was recovered by centrifugation for 30 min at 15 000 × *g*, and rapidly suspended in 10 ml of 20 mM 2-mercaptoethanol, 10 μg/ml proteinase K. LiCl was added to 2.5 M, RNA was precipitated for 3 h at 0°C and collected by centrifugation for 30 min at 15 000 × *g* and 4°C. Poly(A)$^+$ mRNA was prepared as described (Henze et al., 1994), except that all 20 mM 2-mercaptoethanol was added to all solutions. Due to the presence of various secondary metabolites, isolation of mRNA suitable for cDNA from ferns is often difficult. We have found this procedure suitable for many species. cDNA was prepared and cloned as described (Henze et al., 1994), screening was performed with the *Not*I insert of pGapCp17 (Meyer-Gauen et al., 1994). Hybridization and washing was performed as above, but at 55°C. Pea and maize leaf cDNA libraries were prepared from mRNA isolated from 1-week-old light-grown seedlings as described (Henze et al., 1994) and screened as for *Marsilea*. Under these conditions, GapC and GapCp cDNAs from *Pinus* strongly cross-hybridize. The longest of the positives obtained from the *Marsilea* library was designated pCpMQ38 and was sequenced using nested deletions.

### 2.3. Expression of active GapCp in E. coli

The pET-3a expression vector (Studier et al., 1990) was used. The region of the pGapCp17 from *Pinus* encoding the mature subunit was amplified using the primers 5′-ATATTCATATGAAGATAGGAATAAATGGATTTGG-3′ (forward) and 5′-ATATTTGATCACTATTTGCGTGAAGCAACTAAAG-3′ (reverse) with Pwo polymerase (Boehringer Mannheim, Germany) according to the manufacturer's protocol. The *Nde*I–*Bcl*I sites (underlined) of the amplification product were cut, products were purified on Microcon 30 devices (Amicon, Beverley, MA), ligated into *Nde*I–*Bam*HI-digested pET-3a and transformed into *E. coli* NM522. Transformant plasmids were sequenced to verify the reading frame, this plasmid pETCp1 was transformed into *E. coli* BL21. Cultures were grown in 200 ml LB with 50 μg/ml ampicillin overnight, 50 ml of fresh medium was added, after 1 h expression was induced by addition of 1 mM IPTG. Cells were harvested, resuspended in 0.5 vol of buffer A [50 mM Tris–HCl, 2 mM EDTA (pH 8.0)] with 0.5 μg/μl lysozyme, incubated for 15 min at 30°C, vortexed and centrifuged for 2 min in a microcentrifuge. *E. coli* BL21 harbouring pET3a (control) was handled in the same manner. For induction kinetics, activity in the supernatant was measured. For purification of the overexpressed protein, 10 ml of supernatant was passed through a Blue Sepharose Cl-6B

column (1 ml bed volume) equilibrated with buffer A. The column was washed with 20 ml of buffer A, proteins were eluted with 5 ml 1 M NaCl in buffer A, 0.5 ml fractions were collected, desalted on PD-10 columns (Pharmacia, Uppsala, Sweden) and assayed for GAPDH activity.

### 2.4. Enzyme assay

For the reverse reaction (reduction of 1,3-bisphosphoglycerate), GAPDH was analysed at 20°C as described (Heber et al., 1963). The 1 ml assay mixtures contained 0.1 M Tricine (pH 7.5), 8 mM $MgCl_2$, 5 mM dithiothreitol, 1 mM ATP, 1 unit each of phosphoglycerate kinase and triosephosphate isomerase (Boehringer), 250 μM NADH (or NADPH) and 2.5 mM 3-phosphoglycerate. The forward reaction was measured as described (Corbier et al., 1990) in 40 mM triethanolamine, 50 mM $Na_2HPO_4$, 0.2 mM EDTA, 250 μM $NAD^+$ (or $NADP^+$) and 250 μM glyceraldehyde-3-phosphate (GA3P) prepared from the D,L-diethylacetal (Sigma) according to the supplier's protocol. The concentration of D-GA3P was determined enzymatically with yeast GAPDH (Boehringer) prior to assays. Reactions were pre-incubated for 1 min to assay endogenous nicotinamide cosubstrate consumption and initiated by the addition of either 3-phosphoglycerate or GA3P. One unit converts 1 μmol of substrate in 1 min at 25°C.

### 2.5. Other methods

Protein concentration was determined according to Bradford (1976) using BSA as a standard. Protein electrophoresis and standard molecular techniques were performed as described (Sambrook et al., 1989). Sequence handling and alignment was performed with the Genetics Computer Group package (Genetics Computer Group, 1994). Phylogenetic analysis was performed by the neighbour-joining method (Saitou and Nei, 1987) for distances estimated with the Dayhoff matrix using the PROTDIST, SEQBOOT and NEIGHBOR programs of the PHYLIP package (Felsenstein, 1993).

## 3. Results

### 3.1. Structure of the GapCp gene: intron conservation

A map of the region of the genomic clone λPS47 encoding *GapCp* is given in Fig. 1. In exon regions, the sequence of λPS47 is identical to that of the expressed cDNA pGapC17, indicating that the GapCp mRNA previously described (Meyer-Gauen et al., 1994) is encoded by the *GapCp* gene contained in λPS47. However, the first exon of the *GapCp* gene is not

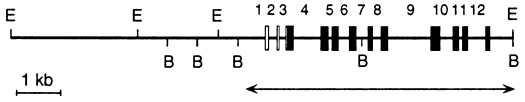


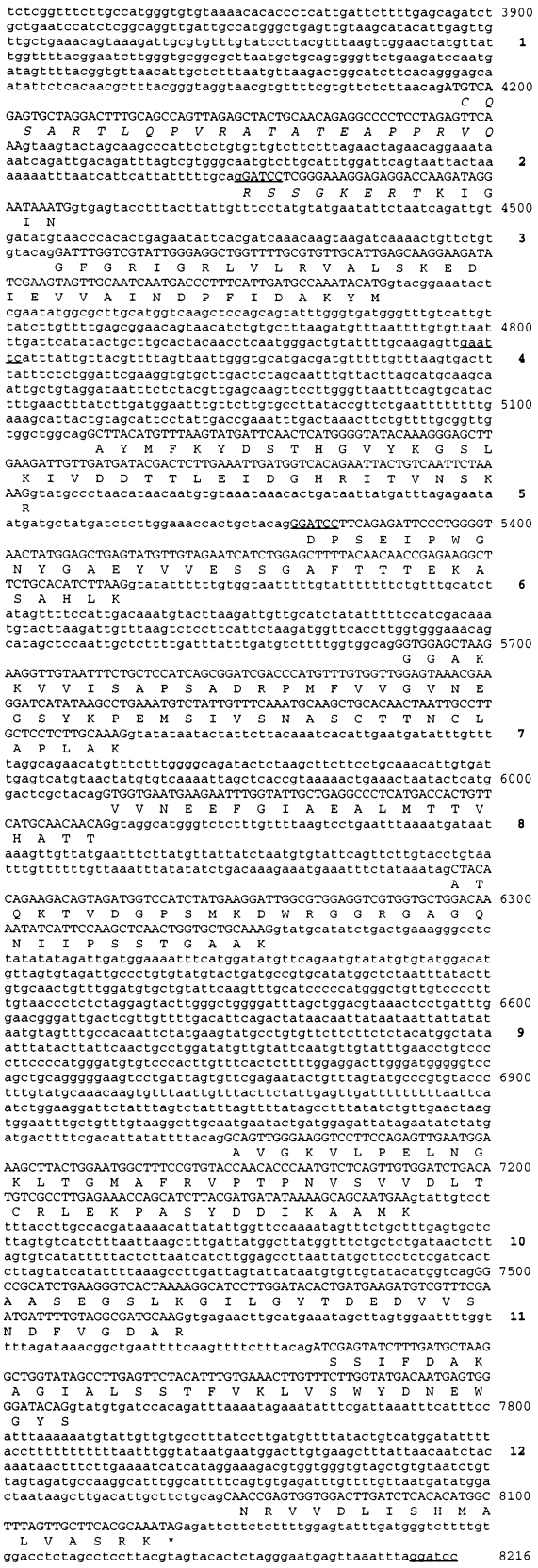

Fig. 1. Structure of λPS47. Open boxes indicate exons encoding the transit peptide region, solid boxes indicate exons encoding the mature subunit. E, *Eco*RI; *B*, *Bam*HI. Introns are numbered. The sequenced region is depicted, the double arrow indicates the segment given in Fig. 2.

contained in λPS47. Rescreening of the genomic library with a 180 bp *Hin*dIII fragment from the 5′ end of pGapCp17 detected 30 independent positively hybridizing clones, none of which, however, possessed *Eco*RI fragments that overlapped with λPS47, indicating that the promoter and/or the first exon may harbour stretches of DNA that are difficult to clone. Thus, although the promoter region and the first exon are missing, the intron–exon structure for the region of the gene encoding the mature subunit can be deduced, as shown in Fig. 2. All introns in *GapCp* conform to the GT–AG rule.

In the process of rescreening, one phage was identified (λPS20) that possessed two *Eco*RI fragments (2.0 and 4.0 kb) which hybridize to the *Hin*dIII fragment from exon 1. These fragments of λPS20 were subcloned into pSK-plasmids and sequenced with the same primer. Both gave a sequence with only 90% identity to the transit peptide region of pGapCp17, and lacked the introns found in the λPS47 transit peptide region. The sequence of the 4 kb *Eco*RI fragment furthermore revealed a deletion which led to a frameshift mutation. The sequences in λPS20 thus probably represent two processed pseudogenes.

The 12 introns that occur in *GapCp* are found at positions that are conserved to the nucleotide with respect to higher plant genes for cytosolic GAPDH (GapC). This is evident from the intron matrix in Fig. 3. As seen in the topology, the gene duplication which gave rise to *GapCp* from the *GapC* progenitor occurred early in chlorophyte evolution. Since all *GapCp* introns are conserved across the *GapCp*/*GapC* duplication, it can be concluded that they were in place prior to that duplication.

The two introns of *GapCp* not depicted in the matrix are introns 1 and 2, they occur in the transit peptide-coding region. Curiously, higher plant *GapA* and *GapB* genes that encode $NADP^+$-dependent chloroplast GAPDH also possess two introns each in the region of the gene encoding the transit peptide (Liaud et al., 1994), but these cannot be homologized to the GapCp transit peptide introns in terms of position. This is because there is no amino acid sequence similarity in the transit peptide of $NAD^+$-dependent GapCp relative to transit peptides of $NADP^+$-dependent GapAB (Meyer-Gauen et al., 1994). The duplication which gave rise to *GapA* and *GapB* also occurred relatively early in



chlorophyte evolution, but independently of the *GapC*/*GapCp* duplication. Accordingly, the transit peptide for chloroplast localization of GapCp was acquired independently of that in GapAB.

### 3.2. *GapCp exists and is expressed in higher plants other than* Pinus

Despite numerous previous studies of higher plant GAPDH genes, cDNAs encoding proteins corresponding to GapCp have not been previously described from any source other than *Pinus*. To identify such homologues from other species, we screened cDNA libraries prepared from pea and maize seedlings and from leaves of the fern *Marsilea quadrifolia* under heterologous hybridization conditions that permit cross-hybridization of *Pinus* GapCp and GapC cDNAs. Only from the fern were positively hybridizing clones obtained that encoded GapCp homologues. The largest of the *Marsilea* cDNAs is still not full size (Fig. 4), lacking a portion of the transit peptide. However, amino acid sequence conservation in the transit peptides of GapCp from *Pinus* and *Marsilea*, the common branch shared with the *Pinus* counterpart (Fig. 3), and the shared indel at position 52 of the alignment (Fig. 4) indicate that pCpMQ38 encodes a homologue of GapCp. Using RT–PCR, we isolated GapC-related PCR products that possessed the same characteristic indel at position 52 from leaves of three other ferns (*Osmunda regalis, Asplenium nidus*, and *Matteuccia struthiopteris*) (unpublished). Although these were not shown to possess transit peptide-like regions, it appears that they represent GapCp orthologues.

### 3.3. *Substrate specificity and affinity of* Pinus *GapCp expressed in* E. coli

The reading frame corresponding to the putative mature subunit of GapCp from *Pinus* expressed in *E. coli* yields a highly active protein. Prior to IPTG induction, lysates of *E. coli* BL21/pETCp1 possess approx. 40-fold more GAPDH activity than the control, but expressing cells grow vigorously despite this. Two hours after induction, activity increases another 3-fold in *E. coli* BL21/pETCp1, but does not change in the control, which possesses a negligible amount (0.5%) of the GAPDH activity of the overexpressing strain. No activity was detectable using NADPH instead of NADH as the substrate (Table 1). Denaturing polyacrylamide gel

Fig. 2. Sequence of *GapCp* from *Pinus sylvestris*. Numbering of positions was chosen to coincide with that in the accession XAJ001706. Introns and the 3′ UTR are given in lower-case letters, exons are in upper-case letters, the translation is given below the DNA sequence. The translation of the region encoding the transit peptide is given in italics, *Bam*HI and *Eco*RI sites are underlined. Introns are numbered in bold.

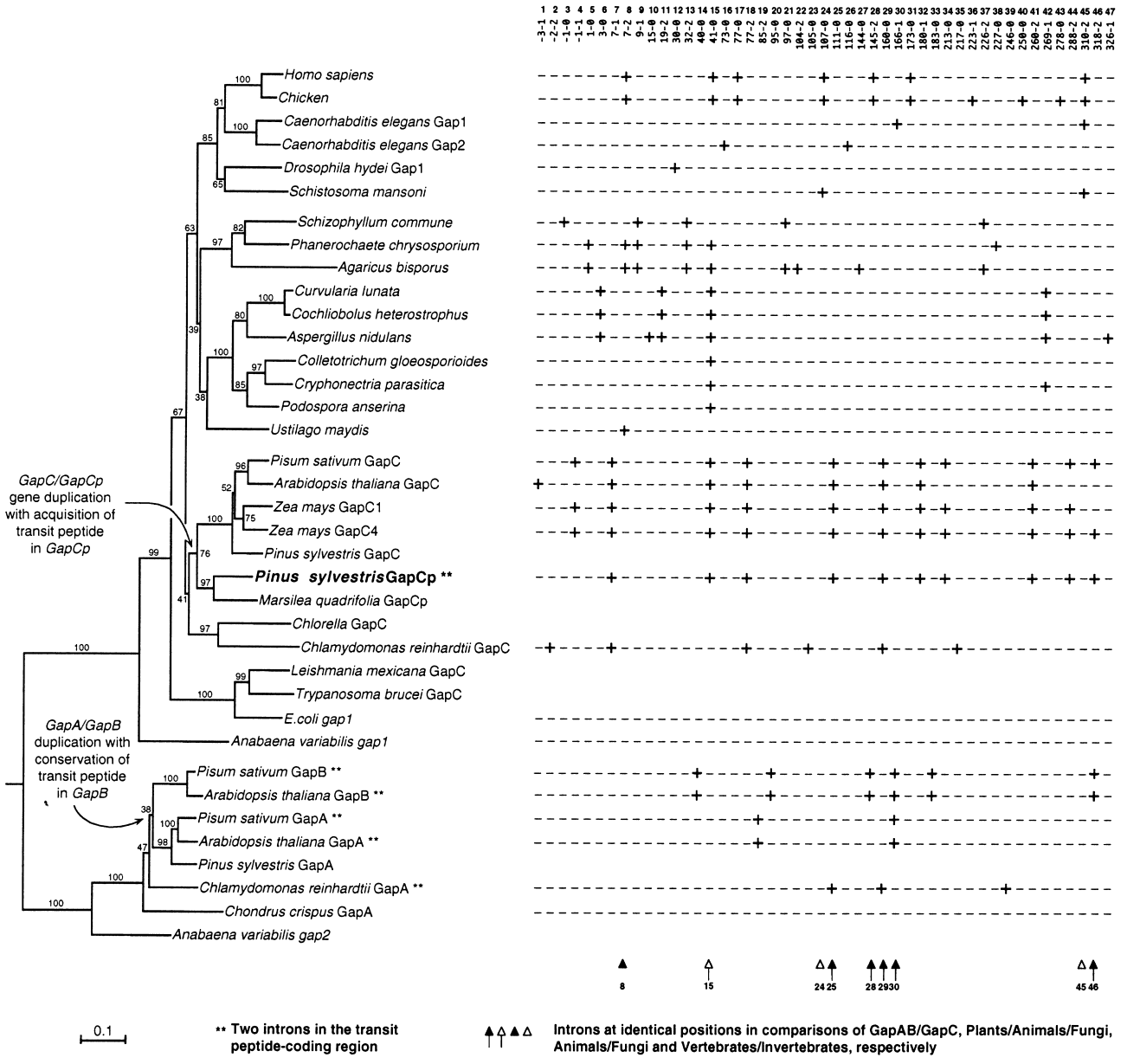



Fig. 3. Gene phylogeny and intron positions in GAPDH genes. The tree is rooted to the outgroup sequences from *E.coli* (*gap2*), *Rhodobacter* (*cbbG*), and *Anabaena variabilis* (*gap3*). Sequences were retrieved from GenBank (alignment available upon request). Positions in the intron matrix refer to those designated in Kersanach et al. (1994).

electrophoresis of the crude extract revealed that the 37 kDa band corresponding to GapCp constituted roughly 50% of the total protein in BL21/pETCp1 (Fig. 5). To obtain active GapCp for further study, protein from 20 ml of culture was prepared 2 h after induction and purified over Blue Sepharose. Recovery was 34%, the 0.5% background of endogenous *E. coli* GAPDH activity thus constitutes at most 1.5% of the activity in the Blue Sepharose eluate. The protein in the Blue Sepharose eluate was electrophoretically homogeneous and had a specific activity of 89 U/mg. Total yield was in the order of 60 mg per litre culture.

Using this preparation, the properties of the enzyme were investigated. The purified protein showed no detectable activity with NADPH (forward reaction) or $NADP^+$ (reverse reaction). For the forward reaction, $K_m^{app}{}_{(GA3P)}$ was determined as 344 μM over the range 5–500 μM at 250 μM $NAD^+$. $K_m^{app}{}_{(NAD)}$ was determined as 62 μM over the range 5–500 μM at 1 mM GA3P (Table 2).

## 4. Discussion

In a previous report on GapCp, we showed that the transit peptide encoded by the *Pinus* mRNA is func-

```
Pinus       GapCp MAMAASTMMR SSVTAGSFRE NHRPSVAVDA QIRSKITSTD GTVRLPFGSF

Arabidopsis GapC                                                MADKKIRI
Pisum       GapC                                                 MGAKIKI
Marsilea    GapCp                  <<<QTA ASRPL·SVKA TATEIPSKVH KSSSKNVVRI
Pinus       GapCp GSNFFGSTAE FGQGSALQCQ SARTLQPVRA TATEAPPRVQ RSSGKERTKI
                                    •    • •  • •  •••• •     •• •

Arabidopsis GapC  GINGFGRIGR LVARVVLQRD DVELVAVNDP FITTEYMTYM FKYDSVHGQW  50
Pisum       GapC  GINGFGRIGR LVARVALKRD DVELVAVNDP FITTDYMTYM FKYDSVHGQW
Marsilea    GapCp GINGFGRIGR LCLRVALARD DIEVVAVNDP FIDTKYMAYM FKYDSTHGIY
Pinus       GapCp GINGFGRIGR LVLRVALSKE DIEVVAINDP FIDAKYMAYM FKYDSTHGVY

Arabidopsis GapC  KHNELKIKDE KTLLFGEKPV TVFGIRNPED IPWAEAGADY VVESHGVFTD 100
Pisum       GapC  KNDELTVKDS NTLLFGQKPV TVFAHRNPEE IPWASTGADI IVESTGVFTD
Marsilea    GapCp K·DEIKAVDE STLQIGGQRV KVFGKRDPSE IPWGSAGADF VVESSGVFTT
Pinus       GapCp K·GSLKIVDD TTLEIDGHRI TVNSKRDPSE IPWGNYGAEY VVESSGAFTT
                           ⇑
Arabidopsis GapC  KDKAAAHLKG GAKKVVISEP SKDDPMFVVG VNEHEYKSDL DIVSNASCTT 150
Pisum       GapC  KDKAAAHLKG GAKKVIISAP SKDAPMFVVG VNENEYKPEF DIISNASCTT
Marsilea    GapCp IEKASSHLKG GAKKVIISAP SADAPMFVMG VNEETYKPDM QIVSNASCTT
Pinus       GapCp TEKASAHLKG GAKKVVISAP SADAPMFVVG VNEGSYKPEM SIVSNASCTT

Arabidopsis GapC  NCLAPLAKVI NDRFGIVEGL MTTVHSITAT QKTVDGPSMK DWRGGRAASF 200
Pisum       GapC  NCLAPLAKVI NDRFGIVEGL MTTVHSITAT QKTVDGPSSK DWRGGRAASF
Marsilea    GapCp NCLAPLAKVV DQEFGIVEGL MTTVHATTAT QKTVDGPSGK DWRGGRGAAQ
Pinus       GapCp NCLAPLAKVV NEEFGIAEAL MTTVHATTAT QKTVDGPSMK DWRGGRGAGQ

Arabidopsis GapC  NIIPSSTGAA KAVGKVLPAL NGKLTGMSFR VPTVDVSVVD LTVRLEKAAT 250
Pisum       GapC  NIIPSSTGAA KAVGKVLPAL NGKLTGMSFR VPTVDVSVVD LTVRLEKAAT
Marsilea    GapCp NIIPSSTGAA KAVGKVLPQL NGKLTGMAFR VPTPNVSVVD LTCRLEKGAS
Pinus       GapCp NIIPSSTGAA KAVGKVLPEL NGKLTGMAFR VPTPNVSVVD LTCRLEKPAS

Arabidopsis GapC  YEEIKKAIKE ESEGKLKGIL GYTEDDVVST DFVGDNRSSI FDAKAGIALS 300
Pisum       GapC  YDEIKAAIKE ESEGKLKGIL GYTEDDVVST DFIGDTRSSI FDAKAGIALN
Marsilea    GapCp YDDVKAAVKA ASEGSMKGIL GYTEDDVVST DFVGDSRSSI FDAKAGIALN
Pinus       GapCp YDDIKAAMKA ASEGSLKGIL GYTDEDVVSN DFVGDARSSI FDAKAGIALS

Arabidopsis GapC  DKFVKLVSWY DNEWGYSSRV VDLIVHMSKA *
Pisum       GapC  DKFVKLVSWY DNELGYSTRV VDLIVHIAKQ L*
Marsilea    GapCp KNFIKMVTWY DNEWGYSNRV VDLISHIAAV VAANKGN*
Pinus       GapCp STFVKLVSWY DNEWGYSNRV VDLISHM·AL VASRK*
```

Fig. 4. Comparison of deduced GapC (*Arabidopsis* and *Pisum*) and GapCp (*Pinus* and *Marsilea*) sequences. Transit peptide regions are indicated in italics, conserved residues in the transit peptides are indicated by ●, the typical GapCp indel is indicated by ⇑.

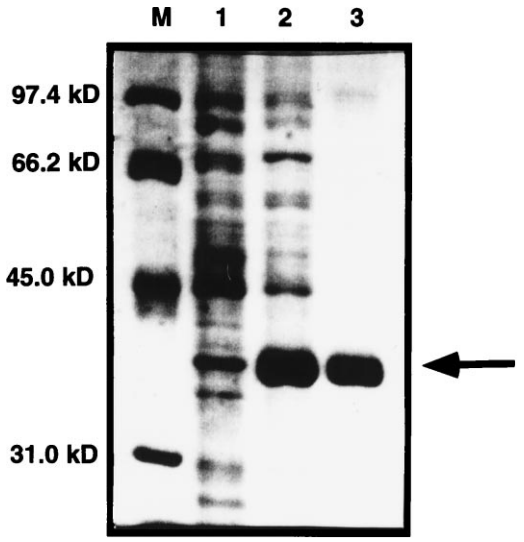



Fig. 5. Expression of GapCp in *E. coli*. Sizes of molecular weight standards are given. Lane 1, crude extract of *E. coli* BL21/pET3a control (lacking insert). Lane 2, crude extract of *E. coli* BL21/pETCp1. Lane 3, GapCp from BL21/pETCp1 after one-step chromatography on Blue Sepharose. The arrow indicates the 37 kDa band expected for GapCp.

tional, that the precursor is imported into chloroplasts, and that the primary structure of the protein contains several sequence characteristics typical of $NAD^+$-dependent GAPDH enzymes (Meyer-Gauen et al., 1994). Here we have shown that the *Pinus* GapCp protein expressed in *E. coli* is completely specific for $NAD^+$ and NADH, showing no detectable activity with $NADP^+$ or NADPH. That GapCp is a strictly $NAD^+$-specific chloroplast GAPDH enzyme is in contrast to all chloroplast GAPDH enzymes purified to date (GapAB), which are $NADP^+$-dependent (Cerff, 1995). The designation $NADP^+$-dependent for chloroplast (Calvin cycle) GAPDH is sometimes confusing, because GapAB enzymes from various sources show nearly equal $v_{max}$ with both $NAD^+$ and $NADP^+$, but the $K_m^{app}{}_{(NAD)}$ of GapAB is 10-fold higher than its $K_m^{app}{}_{(NADP)}$, so that the $NAD^+$-dependent activity is physiologically irrelevant during photosynthesis (Cerff, 1978b). The properties found for GapCp purified from the overexpressing *E. coli* strain are compared with published values for $NAD^+$-specific GAPDH from the higher plant cytosol (GapC) in Table 2.

That a GapCp homologue was found in leaves of the fern *Marsilea* suggests that $NAD^+$-dependent chloroplast GAPDH may be more widespread than previously thought. Notably, $NAD^+$-specific GAPDH enzymes have been found in specialized chloroplasts of two angiosperms, but have not been purified. In the first example, Neuhaus et al. (1993) found that purified non-photosynthetic plastids from cauliflower heads possessed 10-fold greater $NAD^+$-dependent GAPDH activity than $NADP^+$-dependent, suggesting that this angiosperm may possess a plastid-specific $NAD^+$-GAPDH. More recently, Backhausen et al. (1998) found that isolated chromoplasts from ripe red peppers possess an $NAD^+$-specific GAPDH activity that conceivably could be attributable to a GapCp-like enzyme. Also, Neuhaus and Schulte (1996) found that plastids of *Mesembryanthemum crystallinum* reduce oxaloacetate at the expense of triose phosphate in the dark, suggesting the

Table 1
GAPDH activity in *E. coli* expressing *Pinus* GapCp

| Time after induction (h) | BL21/pET3a (U/ml, NADH) | BL21/pETCp1 (U/ml, NADH) | BL21/pETCp1 (U/ml, NADPH) |
|---|---|---|---|
| 0 | 0.10 | 4.2 | 0 |
| 1 | 0.10 | 5.7 | 0 |
| 2 | 0.07 | 12.3 | 0 |
| 3 | 0.09 | 11.6 | 0 |
| 4 | 0.10 | 8.1 | 0 |
| 5 | 0.05 | 3.5 | 0 |

Units per ml crude extract, assayed for the forward reaction.

Table 2
Comparison of GapCp from *Pinus* chloroplasts with other plant GAPDH enzymes

| Parameter | *Pinus* GapCp chloroplast | Pea[a] cytosol | Pea[b] cytosol | Spinach[c] cytosol | Mustard[d] chloroplast |
|---|---|---|---|---|---|
| $K_m^{app}{}_{(GA3P)}$ (μM) | 344 | 240 | 80 | 400 | 300 |
| $K_m^{app}{}_{(NAD)}$ (μM) | 62 | 97 | 78 | 250 | 300 |
| $K_m^{app}{}_{(NADP)}$ (μM) | n.u. | n.u. | n.u. | n.u. | 23 |
| Spec. Act. (U/mg) | 89 | 217 | 133 | 62 | 120 |

n.u., not utilized as substrate.
[a]Data from Duggelby and Dennis (1974).
[b]Data from McGowan and Gibbs (1974).
[c]Data from Speranza and Gozzer (1978).
[d]Data from Cerff (1978b).

presence of an $NAD^+$-dependent, plastid GAPDH activity in this angiosperm as well, although it is not known whether this $NAD^+$-dependent activity is attributable to the bispecificity of $A_2B_2$ chloroplast GAPDH (Baalmann et al., 1994, 1996; Backhausen et al., 1998) or to a distinct $NAD^+$-dependent, GapCp-like isoenzyme.

Intron positions in GAPDH genes have figured prominently in the 'introns-early' vs 'introns-late' debate (see de Souza et al., 1997 for arguments from both sides), since introns have been found to occur at identical positions in GAPDH genes that arose through duplication early in evolution. Such data have been interpreted as evidence favouring conservation of introns that may have been involved in the assembly of primordial GAPDH genes (Kersanach et al., 1994), and alternatively, as evidence favouring evolutionarily recent parallel insertion of introns at identical positions in GAPDH genes in independent lineages (Stoltzfus et al., 1994). The intron–exon organization of the *GapCp* gene from *Pinus* reveals strict conservation of introns with *GapC* homologues from plants, indicating that these introns were in place prior to the gene duplication that gave rise to *GapCp* from a *GapC* antecedant (Fig. 3). However, no additional identity of *GapCp* intron positions with genes encoding higher plant GapA and GapB was found, so these data do not contribute in a decisive manner to the ongoing 'introns-early' (Long et al., 1995; de Souza et al., 1996; Gilbert et al., 1997) vs 'introns-late' (Stoltzfus et al., 1994; Stoltzfus et al., 1997) debate.

Genes for both GapAB and GapC of higher plants were acquired by the nucleus from the eubacterial antecedants of chloroplasts and mitochondria, respectively (Martin et al., 1993; Henze et al., 1995). The $NAD^+$-dependent GapCp of *Pinus* is chloroplast-specific, but arose via duplication of a nuclear gene of mitochondrial origin which encoded a cytosolic enzyme. The transit peptide-coding region was acquired concomitant with or subsequent to the gene duplication event, and resulted in rerouting of the mature subunit to a novel cellular compartment. It is conceivable that the transit peptide was acquired by exon-shuffling, as was recently shown for higher plant cytochrome *c* (Long et al., 1996), but in contrast to that study, BLAST searches with the GapCp transit peptide did not reveal significant similarity to any previously characterized protein. Thus, GapCp is another example in a growing list of enzymes of carbohydrate metabolism in plants that have been rerouted to a new compartment subsequent to transfer of a eubacterial gene to the nucleus via endosymbiotic gene transfer (reviewed by Martin and Schnarrenberger, 1997). Further work is needed to clarify the phylogenetic distribution and function of GapCp in higher plants and to determine whether the $NAD^+$-dependent GAPDH activities recently detected in specialized angiosperm chloroplasts are attributable to homologues of GapCp, or whether yet another novel variant of evolutionary rerouting underlies the compartmentation of those enzymes.

## Acknowledgement

We thank Christiane Lindemann for excellent technical assistance, Dorothea Tholl for valuable discussions and Katrin Henze for critical reading of the manuscript. This work was funded by a grant from the Deutsche Forschungsgemeinschaft (Ma 1426/3-2).

## References

Baalmann, E., Backhausen, J.E., Kitzmann, C., Scheibe, R., 1994. Regulation of NADP-dependent glyceraldehyde-3-phosphate dehydrogenase activity in spinach chloroplasts. Bot. Acta 107, 313–320.

Baalmann, E., Scheibe, R., Cerff, R., Martin, W., Functional studies of chloroplast glyceraldehyde-3-phosphate dehydrogenase subunits A and B expressed in *Escherichia coli*: formation of highly active $A_4$ and $B_4$ homotetramers and evidence that aggregation of the $B_4$ complex is mediated by the B subunit carboxyterminus. 1996. Plant Mol. Biol. 32, 505–514.

Backhausen, J.E., Vetter, S., Baalmann, E., Kitzmann, C., Scheibe, R., 1998. $NAD^+$-dependent malate dehydrogenase and glyceraldehyde-3-phosphate dehydrogenase isoenzymes play an important role in dark metabolism of various plastid types. Planta, in press.

Bradford, M.M., 1976. A rapid and sensitive method for the quantification of microgram quantities of protein using the principle of protein-dye binding. Anal. Biochem. 72, 248–254.

Brinkmann, H., Cerff, R., Salomon, M., Soll, J., 1989. Cloning and sequence analysis of cDNAs encoding the cytosolic precursors of subunits GapA and GapB of chloroplast glyceraldehyde-3-phosphate dehydrogenase from pea and spinach. Plant Mol. Biol. 13, 81–94.

Cerff, R., 1978a. Glyceraldehyde-3-phosphate dehydrogenase (NADP) from *Sinapis alba*: NAD(P) induced conformation changes of the enzyme. Eur. J. Biochem. 82, 45–53.

Cerff, R., 1978b. Glyceraldehyde-3-phosphate dehydrogenase (NADP) from *Sinapis alba*: steady state kinetics. Phytochemistry 17, 2061–2067.

Cerff, R., 1982. Separation and purification of NAD- and NADP-linked glyceraldehyde-3-phosphate dehydrogenases from higher plants. In: Edelmann, M., Hallick, R.B., Chua, N.-H., (Eds.), Methods in Chloroplast Molecular Biology. Elsevier Biomedical Press, Amsterdam, pp. 683–694.

Cerff, R., 1995. The chimeric nature of nuclear genomes and the antiquity of introns as demonstrated by the GAPDH gene system. In: Go, M., Schimmel, P., (Eds.), Tracing Biological Evolution in Protein and Gene Structures. Elsevier, Amsterdam, pp. 205–227.

Cerff, R., Kloppstech, K., 1982. Structural diversity and differential light control of mRNAs coding for angiosperm glyceraldehyde-3-phosphate dehydrogenases. Proc. Natl. Acad. Sci. USA 79, 7624–7628.

Cerff, R., Quail, P.H., Glyceraldehyde 3-phosphate dehydrogenases and glyoxylate reductase. II. Far red light-dependent development of glyceraldehyde 3-phosphate dehydrogenase isoenzyme activities in *Sinapis alba* cotyledons. 1974. Plant Physiol. 54, 100–104.

Corbier, C., Clermont, S., Billart, P., Skarzynski, T., Branlant, C., Wonacott, A., Branlant, G., 1990. Probing the coenzyme specificity of glyceraldehyde-3-phosphate dehydrogenases by site-directed mutagenesis. Biochemistry 29, 7101–7106.

de Souza, S.J., Long, M., Schoenbach, L., Roy, S.W., Gilbert, W., 1996. Intron positions correlate with module boundaries in ancient proteins. Proc. Natl. Acad. Sci. USA 93, 14632–14636.

de Souza, S.J., Fischer, W., Logsdon, J.M., Long, M., Martin, W., Stoltzfus, A., 1997. The origin and evolution of introns: A debate. *HMS Beagle: A BioMedNet Publication* (http://hmsbeagle.com/1997/01/cutedge/day1.htm), Vol. 1, Issue 1.

Duggelby, R.G., Dennis, D.T., 1974. Nicotinamide adenine dinucleotide-specific glyceraldehyde-3-phosphate dehydrogenase from *Pisum sativum*. J. Biol. Chem. 249, 162–166.

Felsenstein, J., 1993. PHYLIP (Phylogeny inference package) manual, version 3.5c. Distributed by the author. University of Washington, Seattle: Department of Genetics.

Genetics Computer Group, 1994. Program manual for Version 8, 575 Science Drive, Madison, Wisconsin, USA, 53711.

Gilbert, W., de Souza, S., Long, M., 1997. Origin of genes. Proc. Natl. Acad. Sci. USA 94, 7698–7703.

Heber, U., Pon, N.G., Heber, M., 1963. Localization of carboxydismutase and triosephosphate dehydrogenases in chloroplasts. Plant Physiol. 38, 355–360.

Henze, K., Schnarrenberger, C., Kellermann, J., Martin, W., 1994. Chloroplast and cytosolic triosephosphate isomerase from spinach: purification, microsequencing and cDNA sequence of the chloroplast enzyme. Plant Mol. Biol. 26, 1961–1973.

Henze, K., Badr, A., Wettern, M., Cerff, R., Martin, W., 1995. A nuclear gene of eubacterial origin in *Euglena* reflects cryptic endosymbioses during protist evolution. Proc. Natl. Acad. Sci. USA 92, 9122–9126.

Kersanach, R., Brinkmann, H., Liaud, M.-F., Zhang, D.-X., Martin, W., Cerff, R., 1994. Five identical intron positions in ancient duplicated genes of eubacterial origin. Nature 367, 387–389.

Liaud, M.-F., Valentine, C., Martin, W., Bouget, F.-Y., Kloareg, B., Cerff, R., 1994. The evolutionary origin of red algae as deduced from the nuclear genes encoding cytosolic and chloroplast glyceraldehyde-3-phosphate dehydrogenases from *Chondrus crispus* . J. Mol. Evol. 38, 319–327.

Liaud, M.-F., Brandt, U., Scherzinger, M., Cerff, R., 1997. Evolutionary origin of cryptomonad microalgae: two novel chloroplast/cytosol-specific GAPDH genes as potential markers of ancestral symbiont and host cell components. J. Mol. Evol. 44, S38–S43.

Long, M., Rosenberg, C., Gilbert, W., 1995. Intron phase correlations and the evolution of the intron/exon structure of genes. Proc. Natl. Acad. Sci. USA 92, 12495–12499.

Long, M., de Souza, S.J., Rosenberg, C., Gilbert, W., 1996. Exon shuffling and the origin of the mitochondrial targeting function in plant cytochrome c1 precursor. Proc. Natl. Acad. Sci. USA 93, 7727–7731.

Martin, W., Brinkmann, H., Savona, C., Cerff, R., 1993. Evidence for a chimaeric nature of nuclear genomes: eubacterial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenase genes. Proc. Natl. Acad. Sci. USA 90, 8692–8696.

Martin, W., Schnarrenberger, C., 1997. The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: a case study of functional redundancy in ancient pathways through endosymbiosis. Curr. Genet. 32, 1–18.

McGowan, R.E., Gibbs, M., 1974. Comparative enzymology of the glyceraldehyde-3-phosphate dehydrogenases from *Pisum sativum*. Plant Physiol. 54, 312–319.

Meyer-Gauen, G., Schnarrenberger, C., Cerff, R., Martin, W., 1994. Molecular characterisation of a novel, nuclear-encoded, $NAD^+$-dependent glyceraldehyde-3-phosphate dehydrogenase in plastids of the gymnosperm *Pinus sylvestris* L.. Plant Mol. Biol. 26, 1155–1166.

Neuhaus, H.E., Batz, O., Thom, E., Scheibe, R., 1993. Purification of highly intact plastids from various heterotrophic plant tissues: analysis of enzymic equipment and precursor dependency for starch biosynthesis. Biochem. J. 296, 395–401.

Neuhaus, H.E., Schulte, N., 1996. Starch degradation in chloroplasts isolated from C-3 or CAM (crassulacean acid metabolism)-induced *Mesembryanthemum crystallinum* L.. Biochem. J. 318, 945–953.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Sambrook, J., Fritsch, E., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, New York.

Scagliarini, S., Trost, P., Pupillo, P., Valenti, V., 1993. Light activation and molecular mass changes of NAD(P)-glyceraldehyde 3-phosphate dehydrogenase of spinach and maize leaves. Planta 190, 313–319.

Schwarz-Sommer, Z.S., Gierl, A., Klosgen, R.B., Wienand, U., Peterson, P.A., Saedler, H., 1984. The Spm (En) transposable element controls the excision of a 2 kb DNA insert at the *wx-m8* allele of *Zea mays*. EMBO J. 3, 1021–1028.

Speranza, M.L., Gozzer, C., 1978. Purification and properties of $NAD^+$-dependent glyceraldehyde-3-phosphate dehydrogenase from spinach leaves. Biochim. Biophys. Acta. 522, 32–42.

Stoltzfus, A., Spencer, D.F., Zuker, M., Logsdon, J.M., Doolittle, W.F., 1994. Testing the exon theory of genes: the evidence from protein structure. Science 265, 202–207.

Stoltzfus, A., Logsdon, J.M., Palmer, J., Doolittle, W.F., 1997. Intron 'sliding' and the diversity of intron positions . Proc. Natl. Acad. Sci. USA 94, 10739–10744.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., Dubendorff, J.W., 1990. Use of $T_7$ RNA polymerase to direct expression of cloned genes. Meth. Enzymol. 185, 60–89.