Molecular cloning and characterization of the thiolesterase glyoxalase II from *Arabidopsis thaliana*

Marianne RIDDERSTRÖM and Bengt MANNERVIK*

Department of Biochemistry, Uppsala University, Biomedical Center, Box 576, S-751 23 Uppsala, Sweden

cDNA encoding glyoxalase II from *Arabidopsis thaliana* has been cloned and sequenced. The isolated 894 bp segment included a sequence of 774 bp encoding a protein with a calculated molecular mass of 28 791 Da. The amino acid sequence deduced from the *A. thaliana* cDNA showed 54 % identity with that of the human enzyme. Searches in databanks identified seven additional DNA sequences from different species with high similarity to glyoxalase II. Certain limited regions, one rich in histidine

residues, shared 100 % identity. A 29 kDa protein with an isoelectric point of 6.2 was obtained by heterologous expression of the *A. thaliana* cDNA in *Escherichia coli*. Homogeneous enzyme was obtained by affinity purification and its catalytic parameters with thiolesters of glutathione were similar to those for human glyoxalase II. The structural and functional similarities between glyoxalase II from *A. thaliana* and from human tissues suggest a common evolutionary origin.

INTRODUCTION

The glyoxalase system catalyses the conversion of 2-oxoaldehydes into the corresponding 2-hydroxy acids [1–3]. The process involves two consecutive reactions mediated by glyoxalase I (EC 4.4.1.5; lactoylglutathione lyase) and glyoxalase II (EC 3.1.2.6; hydroxyacylglutathione hydrolase) respectively [4]. Glyoxalase I catalyses the formation of S-D-lactoylglutathione from the hemimercaptal adduct formed spontaneously from methylglyoxal and reduced glutathione. This thiolester product is hydrolysed by glyoxalase II to D-lactic acid and glutathione. The latter enzyme is present in both the cytoplasm and mitochondria [5]. The two distinct enzymes have been purified from a wide variety of species such as mammals, prokaryotes and plants [1–3]. They seem to be ubiquitous, in accord with their proposed role of detoxifying methylglyoxal, which is believed to be a physiological substrate. Methylglyoxal is produced mainly as a by-product of glycolysis [6], but it can also be formed from aminoacetone and hydroxyacetone derived from the catabolism of threonine and acetone [2]. The human enzymes have attracted interest in research targeted at diabetes and cancer therapy.

The glyoxalase system has been studied only to a limited extent in plants. A finding of possible significance is the apparent correlation between cell proliferation and glyoxalase activity [7]. Inhibitors of cell division have been observed to lower the activity of the enzymes.

Glyoxalase II purified from plants has been shown to have a molecular mass in the range 26 to 28 kDa [7–9]. The isoelectric point of the enzyme from different plant sources has been reported in the range between 4.7 and 6.2.

Human glyoxalase II has recently been cloned [10]. Searches in the EMBL databank for protein sequences with similarity to the human enzyme revealed a similarity of 71 % to the deduced amino acid sequence of two partial cDNA sequences in *Arabidopsis thaliana*. This striking similarity suggested that the plant sequences encode glyoxalase II. Elucidation of the function of expressed nucleic acid sequences is particularly important in *A. thaliana*, since this species has been chosen for complete genome analysis (as a representative of plants). In the present

investigation we have cloned the cDNA from *A. thaliana*. The full-length sequence encodes glyoxalase II with thiolesterase activity. Characterization of the recombinant enzyme reveals that the plant and the human enzyme are not only structurally but also functionally similar.

EXPERIMENTAL

Materials

The *A. thaliana* (variety L. colombia) cDNA library constructed from mRNA from 4.5-week-old plants (leaves and stem) was purchased from Clontech Laboratories (Palo Alto, CA, U.S.A.). Oligonucleotides were custom-synthesized by Operon Technologies (Alameda, CA, U.S.A.). Lambda Sorb and the vector pGEM-3Zf(+) were obtained from Promega (Madison, WI, U.S.A.). The expression vector pKK 223-3 was purchased from Pharmacia Biotech (Uppsala, Sweden) and modified as previously described [11]. Enzymes used for recombinant DNA work were bought from Boehringer-Mannheim (Mannheim, Germany). Sequenase version 2.1 and [α -³²P]dCTP were purchased from Amersham International (Amersham, Bucks., U.K.). Affi-Gel 10 was bought from Bio-Rad (Richmond, CA, U.S.A.). The thiolesters were enzymically synthesized as described previously [12].

Nested PCR and cloning

For the purification of *A. thaliana* cDNA, 200 μ l (3×10^{12} plaque-forming units) of λ phage lysate was used. The purification was carried out using the Lambda Sorb phage adsorbent according to the manufacturer's instructions. The purified DNA was dissolved in 1 ml of water.

The PCR was carried out in a 100 μ l reaction mixture consisting of 10 mM Tris/HCl, pH 8.3, 1.5 mM MgCl₂, 50 mM KCl, 0.2 mM each dNTP, 0.8 μ M 5' and 3' primers and 5 μ l of the *A. thaliana* cDNA library. The reaction mixture was overlaid with 100 μ l of mineral oil. Amplification involved 95 °C for 1 min, 55 °C for 2 min and 72 °C for 2 min for 30 cycles. Two consecutive

* To whom correspondence should be addressed.

The nucleotide sequence for *A. thaliana* glyoxalase II has been submitted to EMBL, GenBank and DDBJ Nucleotide Sequence Databases under accession no. Y08357.

```

                -21                -1
                GGC GAT CGA TCT AAG GAG ACG
1  M  K  I  F  H  V  P  C  L  Q  D  N  Y  S  45
   ATG AAG ATC TTC CAC GTT CCT TGT CTG CAA GAC AAC TAC TCT TAT
   .....
   -----T-----T-----A-----G-----C-----T-----G-----C-----
   L  I  I  D  E  S  T  G  D  A  A  V  V  D  90
   CTG ATA ATC GAT GAG AGC ACC GGA GAC GCG GCG GTT GTG GAT CCA
   -----
   V  D  P  E  K  V  I  A  S  A  E  K  H  Q  135
   GTT GAT CCG GAG AAG GTG ATC GCA TCG GCT GAG AAG CAC CAA GCT
   K  I  K  F  V  L  T  T  H  H  H  W  D  H  180
   AAG ATC AAG TTC GTA CTC ACT ACG CAT CAT CAC TGG GAT CAT GCC
   .....
   G  G  N  E  K  I  K  Q  L  V  P  D  I  K  225
   GGT GGA AAC GAG AAG ATT AAG CAG TTG GTT CCT GAT ATC AAA GTA
   Y  G  G  S  L  D  K  V  K  G  C  T  D  A  270
   TAT GGA GGT TCT CTG GAT AAG GTG AAG GGT TGC ACT GAT GCG GTT
   D  N  G  D  K  L  T  L  G  Q  D  I  N  I  315
   GAT AAT GGT GAC AAG CTG ACT TTG GGT CAG GAT ATT AAC ATA TTG
   A  L  H  T  P  C  H  T  K  G  H  I  S  Y  360
   GCT CTC CAC ACT CCT TGT CAC ACC AAG GGT CAC ATT AGT TAT TAT
   V  N  G  K  E  G  E  N  P  A  V  F  T  G  405
   GTG AAC GGA AAA GAA GGA GAA AAC CCA GCC GTG TTC ACC GGA GAT
   T  L  F  V  A  G  C  G  K  F  F  E  G  T  450
   ACA CTT TTT GTT GCT GGA TGT GGG AAG TTT TTT GAA GGG ACA GCT
   E  Q  M  Y  Q  S  L  C  V  T  L  A  A  L  495
   GAG CAG ATG TAT CAG TCG TTG TGT GTG ACT CTG GCT GCA TTA CCT
   K  P  T  Q  V  Y  C  G  H  E  Y  T  V  K  540
   AAA CCG ACA CAG GTT TAC TGC GGC CAT GAG TAC ACC GTT AAG AAC
   .....
   L  E  F  A  L  T  V  E  P  N  N  G  K  I  585
   TTG GAA TTT GCT CTA ACT GTG GAA CCA AAC AAC GGG AAG ATA CAA
   Q  K  L  A  W  A  R  Q  Q  R  Q  A  D  L  630
   CAG AAG TTA GCA TGG GCC CGT CAA CAG CGC CAA GCA GAT CTT CCC
   T  I  P  S  T  L  E  E  E  L  E  T  N  P  675
   ACA ATC CCT TCA ACA CTG GAG GAA GAG CTC GAA ACA AAT CCA TTT
   M  R  V  D  K  P  E  I  Q  E  K  L  G  C  720
   ATG CGT GTT GAT AAG CCA GAG ATA CAG GAG AAA CTC GGT TGC AAG
   S  P  I  D  T  M  R  E  V  R  N  K  K  D  765
   TCG CCA ATT GAT ACA ATG AGG GAA GTT CGG AAC AAG AAG GAT CAG
   .....
   W  R  G  END
   TGG AGG GGA TGA tga taa atg gag agc ttc ttg cct atg ttt tgt
   -----TAA TAA
   gta ctt gtg ttg aat gat cct gta aaa ggt gta tga acc agt aat
   gtt gat att ctg tgt gaa
                873

```

Figure 1 cDNA sequence encoding glyoxalase II from *A. thaliana*

The deduced amino acid sequence is shown above. Dotted lines correspond to regions covered by primers (denoted A, B and C in order from the 5' towards the 3' end) used for PCR and sequencing. Dashed lines indicate primers used for optimizing the expression in *E. coli*. The altered bases in the 5' region of the selected expression clone are marked.

nested PCRs were performed to isolate the partial glyoxalase II cDNA sequences in two overlapping parts. The 5' part of the *A. thaliana* glyoxalase II cDNA was isolated using the λ gt11 vector-

specific primers V1 and V2 (5'-CG GAATTC GAG CTC ACA CCA GAC CAA CTG GTA ATG-3' and 5'-CTC GAATTC ACC AAC TGG TAA TGG TAG CG respectively; the restriction sites for *Eco*RI are underlined) in combination with the primers directed upstream of C (Figure 1) (5'-TG AAGCTTGTCTGAC GGT GTA CTC ATG GCC GCA GTA-3'; the restriction endonuclease sites for *Hind*III and *Sal*I are underlined). The 3' part of the cDNA sequence was isolated using the primers A and B directed downstream (5'-TTC CCG GAATTC ATG AAG ATC TTC CAC GTT CCT TGT-3' and 5'-ATC GAATTC CTC ACT ACG CAT CAT CAC TGG-3' respectively; the restriction endonuclease sites for *Eco*RI are underlined).

The amplified DNA was digested with restriction enzymes at sites introduced via the PCR primers and ligated into the *Eco*RI and *Sal*I sites of the vector pGEM-3Zf(+) for transformation of *Escherichia coli* JM109. The cloned DNA was sequenced on both strands [13].

Sequence analyses, alignments and homology searches were carried out using the GCG computer program package for sequence analyses (Genetics Computer Group, 575 Science Drive, Madison, WI 53711, U.S.A.).

Heterologous expression and purification of glyoxalase II

For optimal expression in *E. coli* [14], a 5'-oligonucleotide primer was designed by randomizing silent positions of the nucleotide triplets encoding the first 14 amino acid residues after the initiator methionine: AGIISTART, 5'-TCTA GAATTC ATG AAR ATY TTY CAY GTD CCD TGY CTB CAR GAY AAY TAY TCB TAY CTG ATA ATC GAT G-3' (the restriction site for *Eco*RI is underlined). In the 3' end of the coding region, two stop codons (TAA) in tandem were introduced via the primer AGIISTOP, 5'-TTAC GTCGAC TTA TTA TCC CCT CCA CTG ATC CTT-3' (the restriction site for *Sal*I is underlined).

A 10 μ l aliquot of the *A. thaliana* cDNA library was used in a first PCR as described above using the primers AGII, 5'-TT CCCGGGAATTC ATG AAG ATC TTC CAC GTT CCT TGT-3' (restriction sites for *Sma*I and *Eco*RI are underlined) and the vector-specific primer V1 (described above). The product from the first PCR was used in a second PCR using the primers for optimal expression, AGIISTART and AGIISTOP, as described except for the annealing temperature with five cycles at 40 °C for 2 min followed by 25 cycles at 50 °C for 2 min. The fragment obtained was digested with *Eco*RI and *Sal*I and ligated to the expression vector pKK-D [11]. Screening for clones expressing glyoxalase II was via activity measurements using S-D-lactoylglutathione. The cDNA of the clone selected for expression of the enzyme was sequenced for verification.

Recombinant *A. thaliana* glyoxalase II was purified from a 3-litre culture as described previously [10] using an affinity gel consisting of glutathione coupled to Affi-Gel 10 [15]. The recombinant enzyme was eluted from the column with 0.2 M NaCl in 10 mM Mops, pH 7.1. Protein concentration was determined by the method of Bradford [16]. Native PhastGel 8–25 (w/v) (Pharmacia Biotech) was used for determination of the subunit composition of the recombinant plant enzyme. The relative molecular mass of the enzyme was estimated using SDS/12.5 % (w/v) PAGE [17]. The protein bands were visualized by silver staining [18]. Isoelectric focusing was performed in the interval pH 3–9 on precast gels PhastGel IEF 3–9 from Pharmacia Biotech using the PhastSystem. N-Terminal sequence determination of the protein was carried out after electroblotting to poly(vinylidene difluoride) membranes [19].

Kinetic measurements

The kinetic determinations were performed at 37 °C in a reaction volume of 1 ml of 100 mM Mops, pH 7.1. The concentrations of *S*-D-lactoylglutathione and *S*-D-mandeloylglutathione were varied in the intervals 10–2000 μ M and 1–60 μ M respectively. The formation of glutathione was monitored at 412 nm ($\epsilon = 13.6 \text{ mM}^{-1} \cdot \text{cm}^{-1}$) in the presence of 200 μ M 5,5'-dithiobis-(2-nitrobenzoate) [20]. The amount of enzyme in the assay system was 5 ng. The kinetic data were analysed using the SIMFIT computer program [21].

RESULTS

Cloning of the *A. thaliana* glyoxalase II cDNA

Primers for PCR amplification and isolation of the glyoxalase II cDNA were designed on the basis of the similarity between the DNA sequence encoding human glyoxalase II [10] and two partial cDNA sequences from *A. thaliana* in the GenBank/EMBL databank (under accession numbers Z29948 and T22987). The two subsequently isolated partial cDNA sequences combined contained a sequence of 894 bp (Figure 1). DNA sequencing gave the structure of the 3' coding and non-coding region that was missing in the cDNA sequences identified in the databank. The coding region consisted of 774 bp. Analysis of the codon frequencies in *A. thaliana* glyoxalase II revealed a preference for A or T in the third position of the coding triplets in the plant

Table 1 Codon frequencies in *A. thaliana* glyoxalase II cDNA

Amino acid	Codon	Frequency	Amino Acid	Codon	Frequency
Gly	GGG	3	Trp	TGG	3
Gly	GGA	8	End	TGA	1
Gly	GGT	7	Cys	TGT	4
Gly	GGC	1	Cys	TGC	3
Glu	GAG	10	End	TAG	0
Glu	GAA	8	End	TAA	0
Asp	GAT	14	Tyr	TAT	5
Asp	GAC	3	Tyr	TAC	3
Val	GTG	7	Leu	TTG	5
Val	GTA	2	Leu	TTA	2
Val	GTT	10	Phe	TTT	5
Val	GTC	0	Phe	TTC	3
Ala	GCG	3	Ser	TCG	3
Ala	GCA	4	Ser	TCA	1
Ala	GCT	7	Ser	TCT	2
Ala	GCC	3	Ser	TCC	0
Arg	AGG	2	Arg	CGG	1
Arg	AGA	0	Arg	CGA	0
Ser	AGT	1	Arg	CGT	2
Ser	AGC	1	Arg	CGC	1
Lys	AAG	19	Gln	CAG	9
Lys	AAA	4	Gln	CAA	5
Asn	AAT	2	His	CAT	4
Asn	AAC	9	His	CAC	6
Met	ATG	4	Leu	CTG	6
Ile	ATA	4	Leu	CTA	1
Ile	ATT	4	Leu	CTT	2
Ile	ATC	6	Leu	CTC	4
Thr	ACG	1	Pro	CCG	2
Thr	ACA	7	Pro	CCA	6
Thr	ACT	6	Pro	CCT	5
Thr	ACC	4	Pro	CCC	1

AGII	1	MKIFHVPLCLDQNSYLIIDESTGDAAVDPVDPEKVIASAEKHQAKIKFV	50
		: : : : : : : : : : : : : : : : : : : :	
hGII	1	MKVEVLPAITDNYMYLVIDDETKAAIVDPVQPKVDAARKHGKLTTV	50
		: : : : : : : : : : : : : : : : : : : :	
AGII	51	LTTHHHWDHAGGNEKIKQLVPDIKVGSLDKVKGCTDAVDNGDKLTLGQ	100
		: : : : : : : : : : : : : : : : : : : :	
hGII	51	LTTHHHWDHAGGNEKLKLESLGKLVYGGD-DRIGALTKHITHLSTLQVG.	98
		: : : : : : : : : : : : : : : : : : : :	
AGII	101	DINILALHTPCHTKGHSIYYVNGKEG-ENPAVFTGDTLTVAGCGKFFEGT	149
		: : : : : : : : : : : : : : : : : : : :	
hGII	99	SLNVKCLATPCHTSGHICYFVSKPGSEPPAVFTGDTLTVAGCGKFFEGT	148
		: : : : : : : : : : : : : : : : : : : :	
AGII	150	AEQMYQSLCVTLAALPKPTQVYCGHEYTVKNLEFALTVEPNNGKIQQKLA	199
		: : : : : : : : : : : : : : : : : : : :	
hGII	149	ADEMKALLLEVLGRIPDPTRVYCGHEYITNNLKFARHVEPGNAAREKLA	198
		: : : : : : : : : : : : : : : : : : : :	
AGII	200	WARQQRQADLPTIPSTLEELETNPFMRVDKPIEQELGCKSPIDTMREV	249
		: : : : : : : : : : : : : : : : : : : :	
hGII	199	WAKEKYSIGEPTVPSTLAEETYNPFMRVREKTVQGHAGETDPVTMRAV	248
		: : : : : : : : : : : : : : : : : : : :	
AGII	250	RNKKDQWRG*... 258	
		: : : : : : : : : : : : : : : : : : : :	
hGII	249	RREKDDQFKMPRD* 260	

Figure 2 Alignment of the deduced amino acid sequences of *A. thaliana* and human glyoxalase II

AGII corresponds to *A. thaliana* and hGII to human glyoxalase II. Vertical bars indicate identical residues, hyphens gaps in the alignment and dots similar residues.

sequence (Table 1). The 5' non-coding region contained 21 bp and the 3' region 96 bp. The isolated cDNA sequences lacked a poly(A) tail, and attempts to isolate the cDNA using an oligo(dT) primer were not successful. The calculated molecular mass of the protein of 258 amino acid residues was 28791 Da.

Alignment of the human glyoxalase II cDNA sequence with that from *A. thaliana* demonstrated 60% identity. The amino acid sequences deduced from the cDNA sequences shared 54% identity and 71% similarity (Figure 2). Certain regions such as those composed of residues 50–65 and residues 127–150 (except for a Phe-146/Tyr substitution) were 100% identical. Searches in the EMBL databank using the cDNA and conserved amino acid regions (50–65 and 127–150) identified several additional matches. The deduced amino acid sequences from seven DNA sequences are aligned in Figure 3. The alignment includes sequences with similarity to that of glyoxalase II that have not previously been identified as coding for any known protein from the platyhelminth *Sch. mansoni* (EMBL GenBank accession number U30264), the prokaryotes *Rh. blastica* (EMBL SWISSPROTBank accession number P05446), *Bu. aphidicola* (EMBL SWISSPROTBank accession number Q08889), *Ps. aeruginosa* (EMBL SWISSPROTBank accession number P20581), *B. subtilis* (EMBL SWISSPROTBank accession number P37502), *E. coli* (EMBL GenBank accession number D83536) and the yeast *S. cerevisiae* (EMBL SWISSPROTBank accession number X87331). Certain positions of the aligned sequences have identical residues (Figure 3, e.g. 76 (Thr-53; the numbering in parentheses refers to the *A. thaliana* sequence), 77 (His-54), 79 (His-56), 81 (Asp-58), 82 (His-59), 167 (His-112), 238 (His-174), 286 (Glu-219), 290 (Asn-223) and 294 (Arg-227).

Expression and purification

Primers designed to optimize the expression in *E. coli* were introduced into the cDNA encoding glyoxalase II via PCR. The resulting cDNA library of variants with silent mutations was cloned into the expression vector pKK-D [11]. Screening for clones expressing glyoxalase II was performed using assays for enzyme activity with *S*-D-lactoylglutathione. Four independent clones expressing glyoxalase II were identified. One was chosen for large-scale purification of protein and DNA sequencing. The 5' part of the cDNA sequence derived from the PCR primer, with

1		50	
<i>A. thaliana</i>		MKIFHPVCLQ	D..NYSYLII DESTGDAAVV
<i>H. sapiens</i>		MKVEVLPAIT	D..NYMYLVI DDETKEAIV
<i>S. mansoni</i>			MSSEQLO
<i>R. blastica</i>		MA	LNLLTVPCLK D..NFAFLH DAASGRALV
<i>B. aphidic.</i>		MILTSISVLK	D..NYVWILY NDMCS.CIII
<i>S. cerevis.</i>		MK	FLQQIRNMH VKPIKMRWLT GGVNYSYLLS TDRRRNSWLI
<i>P. aerugin.</i>		MLRLSAPGQ	LDDDLCLLGD VQVPV..FLL RLGEASWALV EGGISRDA..
<i>B. subtilis</i>		MNLRLQSEHV	FQCEFFELDVP LRIPVHTWFI KDGDDVY..IV DTGIERFA..
<i>E. coli</i>			MNLNSIPAFD D..NYIIVLNL DEA.GRCLIV
51		100	
<i>A. thaliana</i>	DPVDPKQVIA	SAEKHQAK.I	KFVLTHHHH DHAGGNEKI.KQLVP
<i>H. sapiens</i>	DPVDPKQVVD	AARKHGVK.L	TTVLTTHHHH DHAGGNEKL.VKLES
<i>S. mansoni</i>	DPVEPKILS	AVTQRGLR.I	ESLITTHHHH DHAGGNLDLV TK.CRKQGLE
<i>R. blastica</i>	DAPEAAPVLS	ALAAQGNR.L	TDILLTHHHH DHQAVPEI.HAAT
<i>B. aphidic.</i>	DPGVSEDIK	KIEKKNNK.L	IAILLTHHHH DHVGVGVEII RR.Y.....P
<i>S. cerevis.</i>	DAPEPLEVSP	KLSAEKKSI	DAIVNTHHHY DHSGGNLALY SILCOENSGH
<i>P. aerugin.</i>	E.LVWADLCR	WVADP..SQV	HYMLITHHHY DHCGLLPYLC PRLPNVQVLA
<i>B. subtilis</i>	DAQIRALA.	IGNP..KAI	...LLTHGHS DHIGGASKWL ERF.DIP IFA
<i>E. coli</i>	DFGDAEPVLN	ATAANNW.QP	EAIPLTHHHH DHVGVGKELV EKFPQIVVYG
101		150	
<i>A. thaliana</i>	DIKVYGGSLD	..KVGCT..	DAVDNGDK
<i>H. sapiens</i>	GLKVYGGD.D	..RIGALT..	HKITHLST
<i>S. mansoni</i>	GLKVYGGD.D	..RIGGLT..	DVTSHGYK
<i>R. blastica</i>	GARVHGAAAD	AHRLPLD..	HALREGDR
<i>B. aphidic.</i>	NVTVFPGP..E	ETKTRNVN..	K.IKVQGV
<i>S. cerevis.</i>	DIKIIGGSKS	SPGVTEVP..	DNLQQYHL
<i>P. aerugin.</i>	SERTQAWKS	ESAVRVVERL	NRLQLRAEQR LPEACADWL PVRAVADGEW
<i>B. subtilis</i>	HQK.....		ELKYINGEEP YPNKNEVENT GVAHIVQPLT
<i>E. coli</i>	PQET.Q...D	..K..GTT..	QVVKDGET
151		200	
<i>A. thaliana</i>	LTLGQDINIL	ALHTPCHTRG	HISYYVNGKE G.ENPAVF.. .TGDTLFVAG
<i>H. sapiens</i>	LQVG.SLNVK	CLATPCHTRG	HICYFVSKPG GSEPPAVF.. .TGDTLFVAG
<i>S. mansoni</i>	XRLVATSTFI	AWLHHRHTTG	HICYLVTEN STKEGAVF.. .TGDTLFVAG
<i>R. blastica</i>	VAIGAESAIV	.IDVPGHTRG	HIAFHPG.. .SALAF.. .TGDSLMAAG
<i>B. aphidic.</i>	IKL.LKSEIH	VFLTPGHTLG	HVSYYL.. .KPYIF.. .CGDTLFSGG
<i>S. cerevis.</i>	GNL...RVT	CIRTPCHTKD	SICYIKDLE TGEQ.CIF.. .TGDTLFVAG
<i>P. aerugin.</i>	LELGPRHRLV	VIEAHGHSDD	HVVFY.....DVRRLRLF.. .CGDALGEFD
<i>B. subtilis</i>	EQTLAHLPLK	YYLTPGHSPG	HVVY.....HKIDRTLL.. .TGDLFITSK
<i>E. coli</i>	AFVLGHEFS	VIATPGHTLG	HICYF.SKP.YLF.. .CGDTLFSGG
201		250	
<i>A. thaliana</i>	CGKFFEGTAE	QMYQSLCVTL	AALPKPT... .QVYCGHE. YT.VKNLEFA
<i>H. sapiens</i>	CGKFFEGTAD	EMCKALLEVL	GRLPDDT... .RVYCGHE. YT.INNLKFA
<i>S. mansoni</i>	CGKFFEGTAE	QMFKALIEVL	SKLPTT... .KVCYCGHE. YT.VKNLEFA
<i>R. blastica</i>	CGRLFEPTA	EMWASLSK.L	AALPET... .LICSGHD. YL.DGNLRF
<i>B. aphidic.</i>	CGRVFKNKFF	DMYQSI.NFI	KSLPKKT... .ILCCSHE. YT.LSNLWFA
<i>S. cerevis.</i>	CGKFFEGTGR	DMDMALNQIM	LRAVGETNNW KVKIYYPGHEY YT.KGNVSFI
<i>P. aerugin.</i>	EAEGVWRPLV	FDDMEAYLES	LERLQRLPTL.LQLIPCHGG LL.RGR..LA
<i>B. subtilis</i>	EDLHPPIRRF	SV..DINENI	DSGSIIDQIK PTLICSSHGE EI.LYNEEL..
<i>E. coli</i>	CGRLFEPTAS	QMYQSLKK.L	SALPDDT... .LVCCAEH. YT.LSNMKFA
251		300	
<i>A. thaliana</i>	LTVEPNNGKI	QOKLAWARQQ	RQADL.PTIP STLEEELETN PFMRVDKPEI
<i>H. sapiens</i>	RHVEPGNAAI	REKLAWAKER	YSIGE.PTVP STLAEEFYTN PFMRVREKTV
<i>S. mansoni</i>	LTVEPKNEAL	KHRLAIVKRL	RASNQ.ASPV GTIGEEELATN PLMRVSEPDV
<i>R. blastica</i>	LALPGNPAL	ISRQGRLESEM	RREGR.LPMP WTLAEELATN PFLRAPLQPM
<i>B. aphidic.</i>	MSILPFDKKI	KKYKKIKKH	ISQNK.TSLP VSLETEKKIN IFLRTNEKTI
<i>S. cerevis.</i>	RAKIYSDIGQ	NKEFDALQEQ	CKSNECTTGH FTLRDELGIN PFMRLLDDRAV
<i>P. aerugin.</i>	..ADGAESAY	TECLRLCRLR	LWRQSMGESL DELSEELH... .RAWGGQS
<i>B. subtilis</i>	Y	KNYVRYRD..	
<i>E. coli</i>	LSILPHDLIS	NDYRKVKEL	RAKNQ.ITLP VILKNERQIN VFLRTEDIDL
301		332	
<i>A. thaliana</i>	QKLGKCSPI	DTM....REV	RNKKDQWRG..
<i>H. sapiens</i>	QKHAGTDPV	TTM....RAV	RREKDPKMP RD
<i>S. mansoni</i>	LAHAKTTDPI	KAM....RTI	REEKDRF...
<i>R. blastica</i>	KAAVGLPQAS	DTV..VFAEI	RARKDLF...
<i>B. aphidic.</i>	KKAMGLKDDT	SSFE.VFILL	RKEKDF...
<i>S. cerevis.</i>	RLAVGDTAGT	YPRSVVMQEL	RKLKNAM...
<i>P. aerugin.</i>	VDPLPGLHL	GSMRMLLEIL	SRQALPLD..
<i>B. subtilis</i>			
<i>E. coli</i>	INVINEETLL	QOPEERFAWL	RSKKDRF...

Figure 3 Alignment of unidentified deduced amino acid sequences with high similarity to glyoxalase II

The residues identical with those of *A. thaliana* glyoxalase II are in bold. Residues with chemical similarity are in italics. *H. sapiens*, *Homo sapiens*; *S. mansoni*, *Schistosoma mansoni*; *R. blastica*, *Rhodospseudomonas blastica*; *B. aphidic.*, *Buchnera aphidicola*; *S. cerevis.*, *Saccharomyces cerevisiae*; *P. aerugin.*, *Pseudomonas aeruginosa*; *B. subtilis*, *Bacillus subtilis*.

the modifications introduced, is shown in Figure 1. Silent mutations involving replacements in wobble positions were present in eight codons of the coding sequence. No mutations other than those introduced via the synthetic oligonucleotide primer were present in the expression clone selected.

Large-scale purification (3-litre culture) of recombinant glyoxalase II yielded 15 mg of protein after affinity purification. The purified enzyme gave a single protein band of 30 kDa on native PhastGel 8–25 (results not shown), demonstrating that

Figure 4 Isoelectric focusing of recombinant *A. thaliana* glyoxalase II

The arrow indicates the position of the affinity-purified recombinant glyoxalase II on PhastGel IEF 3–9 after silver staining. The pI markers are Pharmacia Biotech broad pI kit (pH range 3–10). The lanes contained, from left to right, pI markers, glyoxalase II and glyoxalase II.

Table 2 Comparison of kinetic parameters of recombinant *A. thaliana* and human glyoxalase II

The values for the human enzyme are from ref. [10].

Parameter	<i>A. thaliana</i>	Human
<i>S</i> -D-Lactoylglutathione		
K_m (μ M)	98 $\pm$ 7	187
k_{cat} (s^{-1})	466 $\pm$ 9	780
k_{cat}/K_m ($M^{-1} \cdot s^{-1}$)	4.7 $\times 10^6$	4.17 $\times 10^6$
<i>S</i> -D-Mandeloylglutathione		
K_m (μ M)	9 $\pm$ 1	29
k_{cat} (s^{-1})	74 $\pm$ 5	201
k_{cat}/K_m ($M^{-1} \cdot s^{-1}$)	8.8 $\times 10^6$	6.9 $\times 10^6$

glyoxalase II occurs as a monomer. The apparent molecular mass of the recombinant enzyme as determined by SDS/PAGE (not shown) was estimated to be 29 kDa, in agreement with the value calculated from the amino acid composition deduced from the cDNA. N-Terminal amino acid sequence analysis of the first four residues of the recombinant protein gave the sequence M-K-I-F, identical with the amino acid sequence deduced from the cDNA. The isoelectric point was estimated to be 6.2 on PhastGel IEF 3–9 (Figure 4).

Determination of kinetic constants

Steady-state kinetic studies were carried out with two substrates, *S*-D-lactoylglutathione and *S*-D-mandeloylglutathione. The experimental data were consistent with Michaelis–Menten kinetics and the estimated parameters are shown in Table 2. The K_m values were determined as 98 μ M and 9 μ M for *S*-D-lactoylglutathione and *S*-D-mandeloylglutathione respectively. The k_{cat} values were determined as 466 s^{-1} for *S*-D-lactoylglutathione and 74 s^{-1} for *S*-D-mandeloylglutathione. The calcu-

lated k_{cat}/K_m values were in excess of $10^6 \text{ s}^{-1} \cdot \text{M}^{-1}$ for both substrates. All kinetic values are of the same order of magnitude as those determined for human glyoxalase II (Table 2).

DISCUSSION

Cloning of the cDNA sequence encoding glyoxalase II from *A. thaliana* reported in this paper provides information on the primary structure that is useful for comparative purposes. Human glyoxalase II has already been cloned, and a comparison of the available sequences identifies conserved regions. Further, the *A. thaliana* cDNA can serve as a useful tool in studies of the expression of the enzyme under different conditions in the various tissues of the plant.

The 774 bp *A. thaliana* cDNA coding sequence was 6 bp shorter (2 amino acids) than the coding sequence of human glyoxalase II. Alignment of the two amino acid sequences deduced from the cDNA sequences (Figure 2) showed 54% identity, indicating that they are homologous proteins derived from the same ancestor. Several conserved regions were recognized. The *A. thaliana* segments 50–65 and 127–150 are almost 100% identical with the corresponding sequences in the human enzyme, and the first region includes four conserved histidine residues. Searches in the sequence databanks based on the regions conserved between two species (residues 50–65 and 127–150) revealed several unidentified DNA and protein sequences with high similarity to *A. thaliana* and human glyoxalase II (Figure 3). Presumably these sequences encode glyoxalase II. Glutathione thioesterase activity has been found in most species. Glyoxalase II has been purified from *S. cerevisiae* [22], an organism that was noted here to contain a related sequence (see Figure 3).

Glyoxalase II is assumed to have an imidazole group of significance for its catalytic activity [23]. A nucleophilic residue is supposed to be responsible for the attack on the carbonyl carbon of the thiolester substrate, forming an acylated enzyme with concomitant release of glutathione [23]. In the following hydrolysis of the acyl-enzyme intermediate the free hydroxy acid is produced. Interestingly, certain positions in the aligned amino acid sequences are conserved (Figure 3). The region 76–82 contains the motif T-H-X-H-X-D-H, which is found in all sequences. However, several histidine residues (eight) are conserved between the human and plant glyoxalase II; His-54, His-56, His-59, His-112 and His-174 are conserved in all the sequences (Figure 3). These residues are obviously of significance in these related proteins. Information on the three-dimensional structure may be necessary to elucidate their possible catalytic role further, but is not yet available. Nevertheless, the regions of the primary structures with high similarity are probably of importance for the catalytic mechanism, binding of the substrate or the structure of the enzyme.

The molecular mass of the enzyme was calculated as 28 791 Da, similar to that of the human enzyme (28 861 Da) (Table 3).

Glyoxalase II from *A. thaliana* with an isoelectric point of 6.2 is more acidic than the human enzyme (pI 8.5). The acidic plant enzyme has been purified from spinach leaves [8] and *Aloe vera* [9], for which the isoelectric points were determined to be in the ranges 5.3–6.2 and 4.7–5.0 respectively. The recombinant *A. thaliana* glyoxalase II obtained by heterologous expression was shown to retain the N-terminal methionine residue, which is also the case for the human recombinant enzyme. Retention of methionine is in accord with the presence of lysine in the second position [24].

Steady-state kinetic constants were determined for *A. thaliana* glyoxalase II at 37 °C for comparison with the previously determined parameters for the human enzyme. The K_m and k_{cat} values for both *S*-D-lactoylglutathione and *S*-D-mandeloylglutathione were of the same magnitude as those of the human enzyme. The properties of the *A. thaliana* and the human glyoxalase II are similar (Tables 2 and 3). The similarities of the sequences and other characteristics argue for a common evolutionary origin of the two enzymes. The difference in the isoelectric points probably does not invalidate this reasoning, since pI values are usually determined by superficial charged residues without functional importance.

Bivalent metal ions, such as Mg^{2+} and Zn^{2+} , added to the growth medium, did not have any influence on the specific activity of recombinant glyoxalase II with *S*-D-lactoylglutathione (results not shown). In the case of the metalloenzyme glyoxalase I, Zn^{2+} was found to be essential for production of any significant amounts of the recombinant enzyme [25]. However, there is no evidence that glyoxalase II is dependent on metal ions for activity.

We thank Dr. Ulf Hellman at the Ludwig Institute for Cancer Research, Uppsala Biomedical Center, for the N-terminal amino acid sequence determination, Dr. David Eaker and Mrs. Marie Sundquist of our department for freeze-drying the thiolester substrates, Mrs. Birgit Olin in our laboratory for helpful advice and Mr. Thierry Brulavoin for his contribution to the preliminary experiments. This work was supported by the Swedish Natural Science Research Council and the Swedish Council for Forestry and Agriculture Research.

REFERENCES

- Vander Jagt, D. L. (1989) in Coenzymes and Cofactors (Dolphin, D., Poulson, R. and Avramovic, O., eds.), vol. 3A, pp. 597–641, John Wiley and Sons, New York
- Thornalley, P. J. (1993) Mol. Aspects Med. **14**, 287–371
- Mannervik, B. (1980) in Enzymatic Basis of Detoxication (Jakoby, W. B., ed.), vol. 2, pp. 263–273, Academic Press, New York
- Racker, E. (1951) J. Biol. Chem. **190**, 685–696
- Talesa, V., Uotila, L., Koivusalo, M., Principato, G., Giovannini, E. and Rosi, G. (1989) Biochim. Biophys. Acta **993**, 7–11
- Richard, J. P. (1991) Biochemistry **30**, 4581–4585
- Deswal, R., Chakravarty, T. N. and Sopory, S. K. (1993) Biochem. Soc. Trans. **21**, 527–530
- Talesa, V., Rosi, G., Contenti, S., Mangiabene, C., Lupattelli, M., Norton, S. J., Giovannini, E. and Principato, G. B. (1990) Biochem. Int. **22**, 1115–1120
- Norton, S. J., Talesa, V., Yuan, W.-J. and Principato, G. B. (1990) Biochem. Int. **22**, 411–418
- Ridderström, M., Saccucci, F., Hellman, U., Bergman, T., Principato, G. and Mannervik, B. (1996) J. Biol. Chem. **271**, 319–323
- Björnstedt, R., Widersten, M., Board, P. G. and Mannervik, B. (1992) Biochem. J. **282**, 505–510
- Uotila, L. (1981) Methods Enzymol. **77**, 424–430
- Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
- Andersson, S. G. E. and Kurland, C. G. (1990) Microbiol. Rev. **54**, 198–210
- Norton, S. J., Principato, G. B., Talesa, V., Lupattelli, M. and Rosi, G. (1989) Enzyme **42**, 189–196
- Bradford, M. M. (1976) Anal. Biochem. **72**, 248–254
- Laemmli, U. K. (1970) Nature (London) **227**, 680–685

Table 3 Properties of recombinant *A. thaliana* and human glyoxalase II

Property	<i>A. thaliana</i>	Human
Subunit composition	Monomer	Monomer
No. of amino acid residues	258	260
Molecular mass	28 791	28 861
Isoelectric point	6.2	8.5

- 18 Blum, H., Beier, H. and Gross, H. J. (1987) *Electrophoresis* **8**, 93–99
- 19 Hellman, U., Wernstedt, C., Góñez, J. and Heldin, C.-H. (1995) *Anal. Biochem.* **224**, 451–455
- 20 Ellman, G. L. (1959) *Arch. Biochem. Biophys.* **82**, 70–77
- 21 Bardsley, W. G., McGinlay, P. B. and Roig, M. G. (1989) *J. Theor. Biol.* **139**, 85–102
- 22 Murata, K., Inoue, Y., Watanabe, K., Fukuda, Y., Saikusa, T., Shimosaka, M. and Kimura, A. (1986) *Agric. Biol. Chem.* **50**, 135–142
- 23 Ball, J. C. and Vander Jagt, D. L. (1981) *Biochemistry* **20**, 899–905
- 24 Dalbøge, H., Bayne, S. and Pedersen, J. (1990) *FEBS Lett.* **266**, 1–3
- 25 Ridderström, M. and Mannervik, B. (1996) *Biochem. J.* **314**, 463–467

Received 30 September 1996; accepted 18 October 1996