Purification and cloning of the two domain glyoxalase I from wheat bran

Katja S. Johansen a,*, Ib Svendsen b, Søren K. Rasmussen a

a Plant Biology and Biogeochemistry Department, Risø National Laboratory, PBK-301, DK-4000 Roskilde, Denmark
b Chemistry Department, Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Copenhagen, 2500 Valby, Denmark

Received 19 August 1999; received in revised form 25 October 1999; accepted 22 November 1999

Abstract

Investigation of proteins extracted from wheat bran lead to the isolation of a 37 kDa polypeptide extracted from a polyacrylamide gel. Extensive internal peptide sequence information of this protein identified it as a glyoxalase I. Glyoxalase I activity in crude wheat bran extract was measured to 1 U/mg protein (1 U = 1 µmol S-lactoyl glutathione formed/min). Degenerate primers were designed and used for PCR-RACE-based cloning of the corresponding composite cDNA sequence (AJ243528). The wheat bran glyoxalase I amino acid sequence is very similar to the translated sequence of a RNA transcript induced by desiccation of the resurrection grass Sporobulus stapfianus, suggesting a role for glyoxalase in de- or rehydration of plant tissue. The 37 kDa wheat enzyme belongs to a group of monomeric glyoxalases and is composed of two similar halves each representing the full-length human glyoxalase I enzyme. A survey of glyoxalase I sequences, including one (not previously reported) from Drosophila melanogaster, is presented and alignments of these sequences show that amino acid residues involved in co-ordinating zinc or interaction with the substrate are conserved. The alignments indicate a non-linear evolution of glyoxalase I enzymes. © 2000 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: Glyoxalase I; Triticum aestivum; Sequence alignment; Gene duplication; Drosophila melanogaster; Amino acid sequence

1. Introduction

The glyoxalase system consists of (1) glyoxalase I (S-D-lactoylglutathione methylglyoxal lyase (isomerising), EC 4.4.1.5) which is capable of forming S-(2-hydroxyacyl)glutathione from most 2-oxoaldehydes in the presence of glutathione, and of (2) glyoxalase II (S-2-hydroxyacylglutathione hydrolase, EC 3.1.2.6) which hydrolyse the thioester formed in the first reaction to the free 2-hydroxyacids and glutathione. Two-oxoaldehydes are electrophilic and cytotoxic compounds and the coupled reaction may therefore have arisen to detoxify these compounds. Research on the ubiquitous glyoxalase system in mammals was initiated more than 80 years ago, and at the time this system was believed to be of great importance and a major part in the pathway of triose catabolism. Now it is evident that the glyoxalase system is responsible for catalysing only a small percentage of the trioses generated in the citric acid cycle. It has been suggested that the glutathione thiolester S-D-lactoylglutathione, produced by the glyoxalase I reaction, may have specific cellular functions in cell proliferation, differentiation and other processes. The glyoxalase system is reviewed and discussed by Thornalley [1,2].

Glyoxalase I is believed to be a zinc metallo-protein, with the best characterised example being...
the human enzyme. The crystal structure of human glyoxalase I has recently been solved to a resolution of 2.2 Å [3]. The enzyme consists of two monomers, each built up from two equal domains, and it has been suggested that the active dimeric protein has evolved from monomeric proteins through 3D domain swapping [4]. Supporting this theory, the Pseudomonas putida glyoxalase I has now been shown to be active both as a monomer and as a dimer, albeit with a much lower $K_{cat}/K_m$ value for the monomer [5]. The dimeric form can be slowly converted to and from the monomer form as glutathione is removed from or added to the solution.

The residues involved in binding of the zinc ion were identified in the crystal structure of human glyoxalase I, but no structural homology to a glutathione binding domain like the one seen in glutathione-S-transferases could be found in the human sequence [3]. Instead three residues involved in polar interactions with glutathione were identified.

Interestingly, a non-zinc glyoxalase I from Escherichia coli, coded for by the gloA gene, has been characterised [6]. Zinc has no effect on the activity of the enzyme whereas nickel greatly enhances the activity and was shown to be bound in a 1:1 ratio with the dimer.

Glyoxalase-I activity has been studied in several higher plant species, and in some cases the enzyme has been further characterised. In tomato (Lycopersicon esculentum) [7] an 848 bp cDNA clone was identified by differential screening for salt-induced genes, and the glyoxalase activity confirmed by expression in yeast. Using a similar approach in the resurrection grass Sporobolus staphianus [8] a 1.2 kb cDNA clone was found in desiccating plants. In addition a cDNA clone encoding a 186-residue long glyoxalase I has been isolated from epicotyls of Cicer arietinum grown under osmotic stress conditions [9]. Recently, the glyoxalase I cDNA from Brassica juncea was cloned and shown to confer resistance towards stress when expressed in E. coli and tobacco [10]. It has been suggested that the increased expression of glyoxalase I is linked to a high demand for ATP generation and enhanced glycolysis in stressed plants [7,9,10].

Glyoxalase I protein has also been purified from B. juncea seedlings (27 kDa protein) [11], from Glycine max cell suspensions (in which the enzyme is a dimer of 26 and 29 kDa polypeptides) [12], and from Aloe vera (44 kDa protein) [13], among others.

Blue light was shown to promote cell proliferation and glyoxalase I activity in Amaranthus paniculatus cells [14]. Incubation of dark-grown callus in the presence of the calcium ionophore A23187 overcame the requirement for light, suggesting a role for calcium in glyoxalase I activation and cell proliferation. Likewise auxin (NAA) induces glyoxalase I activity and cell division in tobacco protoplasts [15] and soybean suspension cultures [12].

Here we present for the first time the cDNA sequence and extensive protein sequence data for a wheat bran glyoxalase I. The duplicate nature of the sequence and the relationship to other known glyoxalase I sequences is discussed.

2. Materials and methods

2.1. Plant material

Commercially available wheat bran (Ringsted Dampmølle A/S) was used as starting material.

2.2. Protein extraction

Wheat bran was extracted for 30 min at 4°C in 1:5 w/v water containing 5 mM β-mercaptoethanol and 10% (g/g wheat bran) PVPP. The extract was passed through a mesh and then centrifuged for 20 min at 8000 × g to precipitate the starch. Ammonium sulphate was added to 30% saturation and the resulting supernatant saturated to 60% ammonium sulphate. The pellet was solubilised and dialysed against 50 mM sodium acetate (pH 4.6) overnight. The sample (20 ml, 80 mg protein) was loaded on a 6 ml Resource S column (Pharmacia-LKB) equilibrated with the same buffer and eluted with a linear gradient of acetate buffer containing 1.0 M sodium chloride.

Proteins in the fractions eluting at 250–400 mM sodium chloride (6.5 ml, 16 mg protein) were separated on a NuPage (NOVEX) 4–12% SDS-PAGE gel and the major band at 37 kDa cut out. The gel was run with Mes buffer at 200 mV (as suggested by the manufacture) and with M12 (NOVEX) as the molecular weight standards.
2.3. Amino acid sequencing

Cleavage of the protein in the gel was done according to Kellner [16]. EndoLysC protease was used in the ratio 1:4 (with the target protein) in 1% ammonium bicarbonate at room temperature overnight. A control gel piece not containing protein was treated in the same way. The resulting peptides were extracted with 70% TFA and separated by HPLC on a Vydac C18 using an acetonitrile gradient in 0.1% TFA from 5–60% over 1 h. The eluate was monitored at 216 nm, the relevant peaks collected by hand and dried in a Savant rotor vaporator. Before sequencing, the samples were redissolved in 30% acetic acid. Amino acid sequencing was performed on a model 470A sequenator connected to a model 120A phenylthiohydantion analyser (both Applied Biosystems) according to the manufacturers instructions.

2.4. cDNA cloning and sequencing

Total RNA was extracted from wheat seedlings [17] and mRNA isolated with the help of Dynalbeads mRNA DIRECT kit (Dynal, Norway) according to the manufacturers protocols. First strand DNA synthesis was performed using random hexamers. PCR was then performed using degenerate primers (#GSP3 5'-GGC CTA CAA CTA CGG NGT NGA CTA CG, #GSP2 5'-CGA CAT CAT NGC GAT NGT GTA CTT G). Specific primers (GSP01, GSP02, GSP03, GSP04, GSP05, GSP08) were designed and employed in 5’ and 3’ race using the 5’ RACE System kit from GIBCO BRL (Life Technology).

PCR products were extracted from a 1% agarose gel and blunt end cloned in the Sma I site of pUC 18 (Pharmacia).

Automated DNA sequencing was done with an Applied Biosystems 377 Prism. The ABI BigDye terminator cycle sequencing ready reaction kit was used to make the sequencing reactions on a Perkin Elmer thermocycler. Sequence traces were proof-read using Sequencher 3.0 (GeneCode). The cDNA sequences were translated by the MacVector 6.5 (Oxford Molecular) software.

2.5. Database search and sequence analysis

Glyoxalase I sequences were found using either the Entrez search engine or the sequence retrieval system at EMBL. The FASTA3 program at EBI identified more sequences with similarity to the known glyoxalases. The sequences were imported into the MacVector 6.5 (Oxford Molecular) program and aligned using the default parameter for ClustalW formatted alignments. The N-terminal peptide sequence was tested with SignalP v1.1 software at Center for Biological Sequence Analysis (CBS) (www.cbs.dtu.dk) for the presence of secretory signal peptide motifs.

2.6. Glyoxalase assay

The glyoxalase activity was measured in crude wheat bran extract, and the assay was done according to Racker [19]. The assay mixture contained 100 mM sodium phosphate buffer pH 7.5, 3.5 mM methylglyoxal, 1.7 mM reduced glutathione, and 16.0 mM magnesium sulphate in a final volume of 1 ml. The mixture was transferred to quartz cuvettes and incubated at 25°C. Then 0.02 ml wheat bran extract (or boiled extract as control) was added, and the formation of ester monitored by measuring the increase in absorbance at 240 nm in a Shimadzu spectrophotometer for 20 min.

3. Results and discussion

3.1. The sequence of wheat bran glyoxalase I

Following ammonium sulphate fractionation of the crude wheat bran extract, ion exchange chromatography was carried out. The protein containing fractions were analysed by SDS-PAGE and a major protein band visible after staining with coomassie blue, corresponding to a molecular weight of 37 kDa, was cut out (Fig. 1). The protein was cleaved in the gel by EndoLysC protease and successfully subjected to amino acid sequencing. From the extensive internal amino acid sequence data obtained the 37 kDa peptide sequence was identified as a glyoxalase I due to its high degree of similarity to known glyoxalase I sequences.

Primers were designed based on the peptide sequence and were utilised to isolate the corresponding cDNA sequence (in several fragments).
via coupled reverse transcription PCR. Fig. 2 shows the alignment of the translated wheat sequence, and the peptide sequences obtained from the 37 kDa protein, together with the monocot sequences from \textit{S. stapfianus} and rice (\textit{Oryza sativa}). There are a few positions at which the translated cDNA sequence and the peptide data do not agree. When these residues are compared with the equivalent residues found in homologous proteins (Fig. 4) it seems plausible that the cDNA sequence represents a gene other than that encoding the 37 kDa protein. Residue 89 is Ser in the wheat translation, \textit{Citrus paradise}, \textit{Brassica oleracea} and \textit{Arabidopsis thaliana} but Lys in the wheat 37 kDa peptide, \textit{S. stapfianus}, \textit{O. sativa} and \textit{A. thaliana}. Because the protein was isolated from bran and the cDNA sequence from seedlings it is very likely that the cDNA represents another gene.

The 40 N-terminal amino acids of the translated wheat bran sequence were tested and found not to contain any signal peptide motif. The wheat sequence presented here does not include a start Met.

---

**Fig. 1.** Coomassie stained 4–12% NuPage gel of wheat bran proteins. Lane 1, 60 µg protein of pooled fractions from ion exchange chromatography of wheat bran proteins; lane 2, M12 molecular weight markers. The arrow indicates the position of the major protein band that was cut out of the gel and subjected to peptide sequencing.

**Fig. 2.** Alignment of the wheat, rice and \textit{S. stapfianus} sequences with the amino acid sequences from the purified wheat glyoxalase I peptide. Identical residues are in bold on a dark background; similar residues are on a grey background. Areas of similarity are boxed and the consensus sequence is written as the last line. The abbreviations correspond to the accession numbers as listed in Table 1.
Several glyoxalase I sequences from both pro- and eukaryotic organisms can be found in EMBL/Genbank (Table 1). Glyoxalases can be grouped in two according to their size and Figs. 3 and 4 show alignments of the long (280–295 amino acids, Fig. 3) and short (up to 190 amino acids, Fig. 4) glyoxalase sequences, respectively. Most of the published sequences encode proteins of the short type and, these enzymes may all function as homodimers.

It should be noted that EMBL accession X68819, which has been denoted soybean glyoxalase I cDNA, shares very little identity with the rest of the glyoxalase I sequences. Most importantly the well-conserved metal binding residues (see Section 3.4) are not evident in this sequence. The soybean glyoxalase expression was initially found to be induced by auxin [12], and Kalia et al. [15] has illustrated the similarity of this sequence to other auxin-regulated gene products. The related group of enzymes, the glutathione S-transferases (GST), are known to be induced by auxin, and in fact the GST from G. max (AF048978) [20] has a very high degree of identity to the soybean X68819 sequence (alignment not shown). Therefore, the sequence X68819 is not included in the alignments shown here.

The Drosophila glyoxalase I sequence was found by alignment of the human glyoxalase peptide to the translated sequence of nucleotide 46873-48936 of chromosome 2R (2R, region 43B2–43C2, is the accession number listed in Table 1). This sequence has, to our knowledge, not previously been identified as an open reading frame and is therefore the first example from insects, and only the second of animal origin glyoxalase I to be reported. The Drosophila peptide shown in the alignments presented here has been generated manually by removing four introns from the nucleotide sequence. The introns were identified by alignment to the human polypeptide and it was not possible to identify the N-terminal Drosophila sequence.

The alignment in Fig. 3 further suggests that the V. parahaemolyticus sequence is truncated by at least six amino acids in the N-terminus. It starts at the fully conserved Met (Met 36 in the human and Met 7 in the E. coli sequence) and is thus lacking the essential His-or Gln-34 that is involved with the binding of zinc.

### 3.3. Gene duplication

The sequences from S. stapfianus, wheat, B. oleracea, A. thaliana, C. paradisi and O. sativa encodes 280–290 amino acid proteins (Fig. 4) from almost twice as long transcript as the sequences shown in Fig. 3. The two halves, residues (according to Fig. 4) 1–150 and 151–295, respectively, each represents a subunit of the human glyoxalase I enzyme. This is also the case for the Schizosaccharomyces pombe and Saccharomyces cerevisiae yeast enzymes (not shown). The gene duplication has most likely led to a functional glyoxalase I monomer.

The alignment in Fig. 4 shows large areas of identity between all long plant glyoxalases. The yeast sequences are also of the long type but are not included in the figure because each of the two halves of these glyoxalase I sequences are, when cut in two halves, more similar to the relatively short plant and animal enzyme sequences (see also...
Fig. 5). The similarity is obvious because these sequences (together with the P. putida sequence) have some areas, residue 82–96 and 105–124, that are not present in the rest of the microbial or the long plant sequences. In that respect, the wheat half-sequences have more in common with the microbial (short) sequences and interestingly Deswal and Sopory [21] has noticed that B. juncea and the microbial glyoxalases are sharing an alkaline pH optimum.

Fig. 5 shows the tree-diagram relating all glyoxalases. Wheat and the other long glyoxalase I sequences are clustered in two branches: one for monomer I and a second for monomer II. The yeast sequences are grouped together and are distant from the rest of the long sequences and closer to the short sequences. The two animal sequences are closely related, and are unexpectedly placed closer to B. juncea than B. juncea is to B. oleracea in the dendrogram.

Unequal crossovers can explain the duplicate nature of the glyoxalase I, but the rather strange phylogenetic relationship is intriguing.

3.4. Residues co-ordinating zinc

Fig. 3 shows that the residues predicted from the crystal structure of the human enzyme to be co-ordinating zinc ions are conserved among all glyoxalase I sequences. The four residues involved are Gln-34, Glu-100 from one domain and His-128 and Glu-176 from the other domain (numbers...
Fig. 4. Sequence alignment of the relatively long glyoxalase I sequences. Identical residues are in bold on a dark background; similar residues are on a grey background. Areas of similarity are boxed and the consensus sequence is written as the last line. The abbreviations correspond to the accession numbers as listed in Table 1. The entry AC002130 is named *A. thaliana* 1 and the entry AC002131 *A. thaliana* 2. AC002130 has a long N-terminal extension, and the first 62 amino acids has been deleted in *A. thaliana* 1b in order to optimise the alignment.

according to the human sequence [3]. Glu-100, His-128 and Glu-176 are present in all the glyoxalase sequences, but in many of the sequences the Gln-34 is changed to a His via a single base pair change. Both residues are known to be able to co-ordinate zinc and most likely this change do not disrupt the zinc binding site. All microbial sequences and the long plant sequences have His as the first zinc co-ordinating residue instead of Gln which is seen in the rest of the short glyoxalase I. This is in agreement with His in *B. juncea* glyoxalase I being involved in the catalysis [21]. As discussed by Cameron et al. [3] the Glu-100 and Glu-176 might be involved in both co-ordination of zinc and directly in the catalytic reaction as active base.

In the long glyoxalase I sequences the corresponding zinc binding residues are His-29, Glu-80 His-98 and Glu-147 (numbers according to the alignment in Fig. 4). In the second half of the enzyme another putative zinc binding site can be found and the residues involved are Gln-160, Glu-211 and Gln-229 but the last residue seems to be substituted with a valine (Val-281), a substitution likely to disrupt the binding of zinc [3]. Based on the alignment in Fig. 4 it is possible that the Glu-291 could be involved instead of Val-281 because this
3.5. Residues interacting with glutathione

Three residues in the human glyoxalase I crystal structure were found to be making polar interactions with the $\gamma$-glutamyl residue of glutathione. Of these, only two are conserved among all the sequences (Arg-38 in the short corresponds to Arg-33 in the long and Asn-104 in the short corresponds to Asn-84 in the long glyoxalase I sequences, respectively) whereas the last (Arg-124) is only found in some of the short sequences and not at all in the long glyoxalase sequences. Two aromatic rings (Phe-68 and 167) were found to be in the plane of the peptide bond of glutathione, but only the equivalent of the last of these (Phe-167) is conserved. Phe-68 is present in some of the short but a conservative substitution has rendered a Tyr at that position in the \textit{E. coli}, \textit{V. parahaemolyticus}, \textit{S. typhimurium}, \textit{L. influenzae} (Tyr-63) and all of the long sequences (Tyr-194). The significance of these observations in relation to the reaction mechanism is uncertain.

3.6. Different spatial distribution of introns in the gene sequences

The human and the \textit{Drosophila} gene sequences each contains four introns, but not at the same positions, relative to the protein sequence, although at the second and the fourth intron the position is only shifted one amino acid residue from each other. The glyoxalase I gene from \textit{A. thaliana} contains a total of seven introns with the first located at the very beginning of the \textit{A. thaliana} 1b sequence. Neither of those seven splice sites coincides with the afore-mentioned introns from the animal genes. The introns have most likely been introduced after duplication of the ancestral glyoxalase I sequence.

3.7. The function of glyoxalase I in wheat bran

Glyoxalase I activity was measured in crude extract of wheat bran to 1 U/mg protein. This activity is comparable to the levels found in other plant extracts, 0.2–0.4 U/mg in \textit{N. tabacum} protoplasts [15], 0.5 U/mg in un-stressed to 1.8 U/mg in salt stressed \textit{B. juncea} seedlings [10]. Wheat bran is not an active tissue and is not likely to have a high demand for ATP. It has been speculated that increased glycolysis to generate ATP leads to increased levels of methylglyoxal and induction of glyoxalase I activity in stressed plants [7,9,10]. Induction of transcription of the glyoxalase I gene during desiccation of resurrection grass [22], and the presence of the protein in dry wheat bran might therefore suggest new roles for glyoxalase I, either in protecting cells during server water deficiency or during re-hydration when the tissue becomes metabolically active.

A number of enzymes having the potential to be active upon hydration are found in resting cereal grains. Among these are $\beta$-amylase in barley [23] and wheat [24], which is secreted to the endosperm in order to utilise the carbon reserves and stored energy in the starch grains. It seems more likely that glyoxalase I is functioning in the living cells of the grain where it destroys methylglyoxal, a toxic by-product from the citric acid cycle.
4. Conclusion

A wheat bran glyoxalase I peptide and cDNA sequence have been identified, adding to the growing number of glyoxalase I sequences being published from all species and plants in particular. Two distinct types of glyoxalases exist and plant enzymes of both types are found. One type is functioning as a homo-dimer the other type has through a gene-duplication evolved as a functional monomer of almost the size of the dimer glyoxalase I enzyme. The three known sequences from monocots are all belonging to the long monomeric type. The long plant glyoxalase I enzymes have clearly not evolved directly from the ancestral short plant sequence and the evolutionary relationship between all glyoxalase I sequences is intriguing.

Acknowledgements

The Danish Research Academy and the first framework (1996–2001) under the Cereal Network in Denmark in part, supported this work. The authors would like to thank Dr Tomas H. Roberts for proofreading the manuscript.

References


