cDNA cloning and expression of the flavoprotein D-aspartate oxidase from bovine kidney cortex

Tatjana SIMONIC*†§, Stefano DUGA*, Armando NEGRI*, Gabriella TEDESCHI*, Massimo MALCOVATI†‡, Maria L. TENCHINI†‡ and Severino RONCHI*†

*Istituto di Fisiologia Veterinaria e Biochimica and ‡Dipartimento di Biologia e Genetica per le Scienze Mediche, Università di Milano, Via Celoria 10, 20133 Milano, and †Centro Interuniversitario per lo Studio delle Macromolecole Informazionali, Milano, Italy.

The isolation and sequencing of the complete cDNA coding for a D-aspartate oxidase, as well as the overexpression of the recombinant active enzyme, are reported for the first time. This 2022 bp cDNA, beside the coding portion, comprises a 5′ untranslated tract and the whole 3′ region including the polyadenylation signal and the poly(A) tail. The encoded protein comprises 341 amino acids, with the last three residues (-Ser-Lys-Leu) representing a peroxisomal targeting signal 1 (PTS1), hitherto unknown for this protein. The overexpression of recombinant D-aspartate oxidase was achieved in a prokaryotic system, and a soluble and active enzyme was obtained which accounted for about 10 % of total bacterial protein. Comparisons with the known cDNAs for mammalian D-amino acid oxidase, another peroxisomal enzyme, are also made. The close structural and functional similarities shared by these enzymes at the protein level are not reflected at the nucleic acid level.

INTRODUCTION

D-Aspartate oxidase (EC 1.4.3.1) is a FAD-containing protein which catalyses the oxidative deamination of dicarboxylic D-amino acids by O_2, with the production of H_2O_2, NH_3 and the corresponding 2-oxo acid [1]. Although this enzyme activity has been demonstrated in several mammalian ([2] and references therein) and amphibian [3] tissues and in octopus [4], the protein has been purified to homogeneity only from the hepatopancreas of *Octopus vulgaris* [5] and bovine kidney cortex [6]. Furthermore, only in the latter case has the complete primary structure been determined [7]. The amino acid sequence of D-aspartate oxidase resembles, among known proteins, only that of the flavo-oxidase D-amino acid oxidase (EC 1.4.3.3). The latter catalyses the same reaction, but with a different substrate specificity, being active towards neutral and basic D-amino acids. Both are peroxisomal enzymes and exhibit similar general properties, as expected for proteins catalysing similar reactions [6–8]. On the basis of the common properties shared by these two enzymes, it has been suggested that they may be derived from the divergent evolution of a single gene [7].

Increasing information on the widespread distribution of D-amino acids in different biological materials raises renewed interest in the physiological roles of D-aspartate and D-amino acid oxidases, including whether they should still be considered only as 'detoxifying enzymes' [9] or whether they are involved in more complex metabolic pathways. Furthermore, the fact that concentrations of D-amino acids in both human and animal foodstuffs are often increased during processing calls for a careful examination of their nutritional and, possibly, toxicological properties, as well as their biological role. D-Aspartate is usually the most abundant D-amino acid in several organisms and foodstuffs, probably because aspartate has the highest racemization rate of all amino acids. However, little information is available in these research fields, because of the lack of a simple and routine method for the assay of dicarboxylic D-amino acids. In this regard D-aspartate oxidase could be a good candidate for constructing a biosensor for the assay of acidic D-amino acids, especially D-aspartate. Unfortunately, this enzyme is very poorly expressed in mammalian tissues, so that it is rather difficult and very time-consuming to obtain suitable amounts of purified protein. For this reason we decided to clone the coding sequence of this enzyme and to express the recombinant D-aspartate oxidase in *Escherichia coli*.

In the present paper we report the isolation and characterization of the complete cDNA encoding bovine kidney D-aspartate oxidase, together with the overexpression of the recombinant active flavoenzyme.

EXPERIMENTAL

Materials

All enzymes were from Promega, unless otherwise indicated. [γ-^{32}P]ATP (5000 Ci/mmol), [α-^{32}P]dCTP (3000 Ci/mmol), Hybond-N filters and Multiprime DNA Labelling System were from Amersham; plasmid and gel extraction kits were from Qiagen; TA Cloning Kit and P_L Expression System were from Invitrogen BV; Taq Dye Deoxy Terminator Cycle Sequencing Kit was from Applied Biosystems. Oligonucleotides were synthesized by Primm s.r.l., Milan, Italy.

PCR assays with genomic DNA

PCRs were carried out with 0.1–1.0 μl (30–300 ng) of genomic DNA extracted from bovine kidney [10]. The reactions were performed in a 480 Thermal Cycler (Perkin Elmer) in 50 μl mixtures containing 1 × reaction buffer [50 mM KCl/10 mM Tris/HCl (pH 9)/0.1 % Triton X-100], 1.5 mM $MgCl_2$, 200 μM

Abbreviations used: RT-PCR, reverse transcription-PCR; ss-cDNA, single-stranded cDNA; UTR, untranslated region.

§ To whom correspondence should be addressed.

The nucleotide sequence data reported have been submitted to the EMBL/GenBank/DDBJ Nucleotide Sequence Databases under accession no. X95310.

Table 1 Oligonucleotides used as primers in PCR assays

Numbers refer to the nucleotide positions in the cDNA sequence shown in Figure 2. F and R indicate coding and anticoding sequences respectively. F9 carries at its 5′ end the trinucleotide sequence which completes the *Nde*I restriction site. 5A and AdT are 5′ and 3′ anchor oligonucleotides respectively. The sequence of FR is complementary to part of 5A and identical to part of AdT.

Primer	Sequence
F1	G^{934}CNATGCAYTGGGG^{947}
F2	C^{712}ARAARGGNGAYTGGAA^{728}
F3	T^{628}GGGTNAARCAYTTYAT^{644}
F4	A^{352}ARATGACNAARGAYGA^{368}
F5	A^{193}ARCAYTGGTTYAARGA^{209}
F6	A^{1}TGGAYACNGTNMGNAT^{17}
F7	G^{861}AAAGAGCTACTAGCTCAG^{879}
F8	C^{892}TGCCTGTGGTCCATC^{907}
F9	CAT A^{1}TGGATACAGTACGGATTGC^{20}
R1	G^{993}ACYTGVACACAYTCRTTSAC^{973}
R2	A^{974}CCAGCCTGGTGGCCTCCAGAGCAGTGCC^{946}
R3	G^{415}TGTCGTAAAAGCATGACC^{397}
R4	C^{403}ATGACCAAACACATGTTGG^{384}
R5	T^{368}CATCCTTGGTCATCTTTCG^{349}
R6	T^{1313}TAGATCCTTGGACCTGAG^{1295}
R7	A^{1246}GCTTAACGCCCTATGTC^{1229}
5A	GAATTCGTCGACAAGCTTAGTCA
AdT	CAGGAAACAGCTATGACTAAGCTTGTCGACG$(T)_{17}$
FR	TGACTAAGCTTGTCGACG

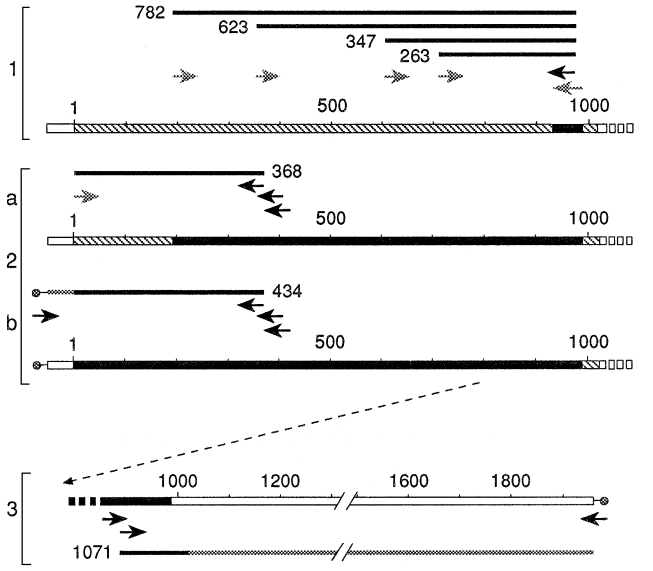


Figure 1 Schematic representation of the amplification strategy used to obtain D-aspartate oxidase cDNA

The phases of the strategy are denoted 1–3 (see the text). Bars indicate DNA segments: thick ones represent cDNA, and thin ones represent amplification products. Unknown sequences are indicated by empty bars, unknown sequences within the coding region are indicated by hatched bars, and sequenced segments are indicated by black bars when coding and by dotted bars when non-coding. Base numbering begins at base 1 of the translation start codon. Primers used in PCR assays are shown by arrows (not to scale): dotted, degenerate; black, unique. Dotted circles indicate anchor oligonucleotides.

dNTPs and 0.4 μM degenerate primers F1 and R1 (Table 1). *Taq* DNA polymerase (2.5 units) was added after 5 min of denaturation at 95 °C, immediately before starting 30 amplification cycles performed as follows: 95 °C for 1 min, 45 °C for 30 s and 72 °C for 30 s. Finally the samples were further extended for 5 min at 72 °C.

Reverse transcription-PCR (RT-PCR) assays

Total RNA was extracted from bovine kidney cortex by the method of Chomczynski and Sacchi [11] and used as a template (5 μg) for first-strand cDNA synthesis, using 30 units of avian myeloblastosis virus reverse transcriptase at 42 °C and, depending on the purpose of the experiment, different primers (Table 1). The complete strategy and the primer couples used for each experiment are reported in the Results section.

When PCRs were carried out using a degenerate primer coupled with a unique primer, the reactions were performed on one-tenth of the single-stranded cDNA (ss-cDNA) obtained. Annealing temperatures corresponded to the lowest T_m of the degenerate primer, as estimated by the simple formula $[2(A+T)+4(G+C)]$; the elongation times were modified on the basis of the expected product size. Semi-nested experiments were carried out on variable amounts of the reaction mixtures resulting from the first PCR step (0.01–0.5 aliquots) with the same degenerate primer and an inner unique primer for 25 amplification cycles. The annealing temperature and elongation time ranges were 48–52 °C and 0.5–1.5 min respectively. In some cases a second step of semi-nested amplification was required. All PCRs were concluded by a further 10 min of elongation at 72 °C.

The 5′ end of the cDNA was isolated by the method of single-strand ligation to ss-cDNA [12]. Reverse transcription was primed by oligonucleotide R5 and ss-cDNA was ligated, using 10 units of T4 RNA ligase (New England Biolabs), to the 5′-anchor oligonucleotide (5A), phosphorylated at the 5′ end and blocked at 3′ end by amino group addition. The ligation products were used as templates for PCR.

The coding sequence for protein expression (1246 bp) was obtained from semi-nested PCR, performed using oligonucleotide F9, carrying the complete *Nde*I site at the 5′ end, coupled with the R6 and then the R7 primer, on reverse-transcribed total RNA primed by oligonucleotide R6. Furthermore, the 1246 bp cDNA was annealed to the partially overlapping 1071 bp fragment containing the 3′ untranslated region (UTR) (Figure 1) and submitted to five PCR cycles (95 °C for 30 s; 72 °C for 2 min) followed, after addition of the F9 and FR primers, by 30 cycles at a lower annealing temperature (52 °C). The resulting 1962 bp cDNA comprised the whole coding region and the 3′ UTR.

Cloning of PCR products and sequence analysis

Products derived from genomic DNA amplification were ligated into the *Sma*I site of pUC9, and *E. coli* 71/18 competent cells were transformed with the recombinant plasmids. All the DNA fragments obtained from RT-PCR experiments were inserted into the pCRII vector by the TA cloning procedure and used to transform *E. coli* INVF′ competent cells. The recombinant plasmid containing the 1246 bp cDNA was digested by *Nde*I and *Eco*RV, and the insert DNA was ligated into the same restriction sites of the pLEX expression vector to give pED, which was used to transform *E. coli* GI724 competent cells.

Selection of recombinant plasmids was carried out by colony blot hybridization using suitable ^{32}P-labelled oligonucleotides as probes and/or direct evaluation of insert sizes after excision from extracted plasmid DNA.

Sequence analysis of the DNA fragments was performed fully on both strands, either using plasmids or directly using purified PCR products, by the Taq dye-deoxy terminator method, using an automated 370A DNA sequencer (Applied Biosystems).

Searches for sequence similarities were performed with the EMBL Data Base (release 43) using the program FASTA of the software PC-GENE.

Southern and Northern blot analysis

Southern blots of plasmids and PCR fragments were hybridized to [γ-^{32}P]ATP-labelled oligonucleotides. Northern blots of total (20 μg) and poly(A)$^+$ (1 μg) RNAs were hybridized, using the purified 1962 bp cDNA labelled with [α-^{32}P]dCTP, according to standard procedures [10]. Filter washings were carried out for 15 min each at 42 °C in 2 × SSC/0.1 % SDS and 1 × SSC/0.1 % SDS, followed by 10 min at 44 °C in 0.1 × SSC/0.1 % SDS.

Expression of recombinant D-aspartate oxidase and enzyme activity assays

E. coli GI724 cells, carrying the *cI* repressor gene engineered into the bacterial chromosome under the control of the *Trp* promoter, and transformed by the recombinant plasmid pED which contains the λ P_L promoter, were grown according to the manufacturer's instructions. Induction medium was inoculated with overnight precultures of GI724 cells harbouring pED (10–15 ml aliquot for 100 ml of medium) and grown at 30 °C with shaking up to $D_{550} = 0.5$–0.6. Expression of the recombinant protein was initiated by the addition of L-tryptophan at a final concentration of 100 μg/ml and by changing the temperature to 37 °C. Samples (25 ml) were withdrawn immediately before and at fixed times after the start of induction and centrifuged (2000 ***g***, 10 min, 4 °C). The bacterial pellet was washed once with 5 ml of 50 mM Tris/HCl, pH 8, centrifuged as before, resuspended in 1 ml of the same buffer and sonicated in an ultrasound generator (Labsonic 1510). Supernatants were assayed for protein content [13] and D-aspartate oxidase activity [6].

N-terminal sequencing of recombinant D-aspartate oxidase

Aliquots of supernatants containing enzyme activity were submitted to SDS/PAGE [14] and electroblotted on to a PVDF membrane (Immobilon; Millipore) as described by Matsudaira [15]. N-terminal sequence analysis of the expressed D-aspartate oxidase was performed using an automated Protein Sequencer (Applied Biosystems Model 477/A), following the manufacturer's instructions.

RESULTS

Isolation of a short genomic fragment coding for D-aspartate oxidase

On the basis of the D-aspartate oxidase amino acid sequence [7], a set of degenerate oligonucleotides was synthesized. The amplification experiments were carried out by PCR using genomic DNA. The primer couple F1/R1 (Table 1), out of several assayed, gave a positive result. In this case, a number of short DNA fragments, detectable after gel blot hybridization using the same 5′ primer as the ^{32}P-labelled probe (results not shown), were obtained by genomic DNA amplification. Only those displaying the expected size of approx. 60 bp were analysed further. For this purpose all the amplified fragments were cloned into the pUC9 *Sma*I site and plasmids containing DNA inserts of the expected size were selected for sequencing. Two DNA inserts, out of six sequenced, coded for a protein fragment identical to part of the D-aspartate oxidase sequence located near the C-terminus (positions 314–326 of the published amino acid sequence) (Figure 2).The nucleotide sequence of this fragment allowed us to design a specific and unique 40 bp DNA probe that was used to repeatedly screen both cDNA and genomic bovine libraries. Since no such attempts to isolate cDNA or genomic clones coding for this flavoprotein gave a positive result, an alternative strategy based on PCR applications was adopted.

Isolation of cDNA fragments covering the entire D-aspartate oxidase mRNA

All experiments were carried out on ss-cDNA that had been reverse-transcribed from total RNA extracted from bovine kidney cortex. The experimental strategy can be divided into three phases on the basis of the use of first-strand cDNAs that were differently primed, i.e. using oligo(dT) or random primers, a D-aspartate oxidase specific primer and finally an anchor oligonucleotide. The general procedure is summarized in Figure 1. Table 1 reports all the oligonucleotides used to isolate the cDNA fragments described below. All the DNA fragments isolated were inserted into the pCRII vector by the TA cloning procedure.

(1) The starting point for the first phase of this investigation was the 60 bp genomic fragment corresponding to nucleotides 934–993 of the cDNA (Figure 1), which allowed the synthesis of a unique 3′ primer (R2). Four degenerate 5′ primers (F2–F5) were synthesized, on the basis of the D-aspartate oxidase amino acid sequence, in regions with the lowest codon ambiguities. Reverse transcription was performed using oligo(dT) or random primers. Four parallel amplifications were performed using the four upstream primers (one for each experiment) coupled with the same 3′ primer (R1). An aliquot of each PCR reaction was used as the template for a second amplification (semi-nested), in which the 3′ primer R1 was replaced by the inner primer R2. Only this second step of amplification allowed the detection of the expected fragments after separation by gel electrophoresis and ethidium bromide staining, as well as by Southern blot hybridization using the internal oligonucleotide F1 as the ^{32}P-labelled probe (results not shown).

The resulting four DNA fragments, showing different degrees of extension to the 5′ end of the cDNA and overlapping at their 3′ ends (Figure 1, part 1), were cloned and sequenced. These fragments were 263, 347, 623 and 782 bp in length. Sequence analysis of the cloned fragments (pRD263–pRD782) demonstrated that all segments were from the D-aspartate oxidase cDNA, with the longest one corresponding to the protein region ranging from amino acid 65 to amino acid 325 of the published primary structure [7].

(2) The cDNA sequence information derived from the first phase of this work enabled us to approach a second phase, which led to the isolation both of the complete 5′ coding region (phase a) and of the 5′ UTR (phase b). For these purposes, reverse transcription was performed using the specific 3′ primer R3.

(a) For the isolation of the complete 5′ coding sequence, one more PCR step was required, owing to the extremely high degree of degeneracy of the 5′ primer F6. Therefore three unique downstream primers, located close to the 5′ end of the sequenced part of the cDNA, were synthesized. The most external of these was used both for reverse transcription and for the first step of amplification (primer couple F6/R3), which was followed by two steps of semi-nested PCR using the inner 3′ primers R4 and R5 respectively. The resulting 368 bp fragment (Figure 1, part 2a) encoded the first 123 amino acids of the flavoprotein, as

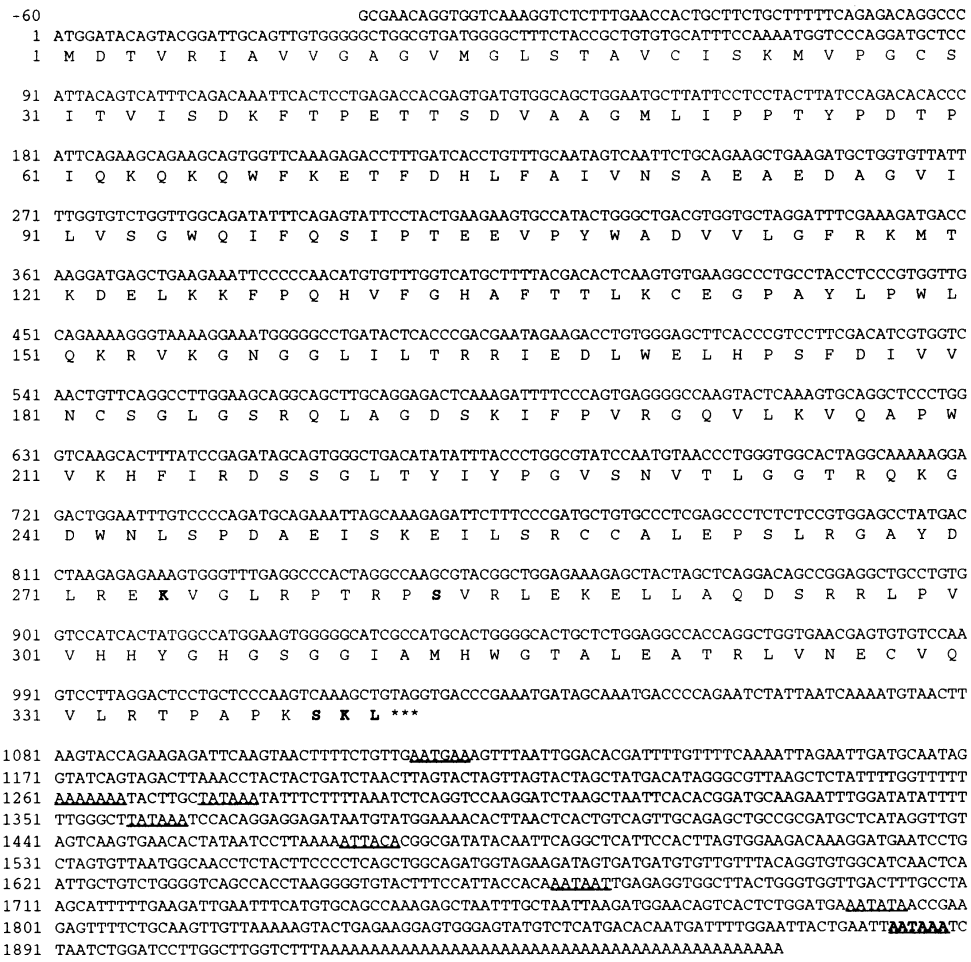


Figure 2 Nucleotide sequence and deduced amino acid sequence of bovine kidney D-aspartate oxidase

The first base of the translation start codon is designated +1, and the stop codon is indicated by asterisks. Potential polyadenylation signals are underlined, and that located 28 nt upstream of the poly(A) tail is emboldened. The emboldened amino acids K^{274} and S^{283} show the differences compared with the elucidated protein primary structure [7]. The C-terminal tripeptide (emboldened) represents the peroxisomal targeting signal.

demonstrated by sequence analysis of the cloned DNA insert (pRD368).

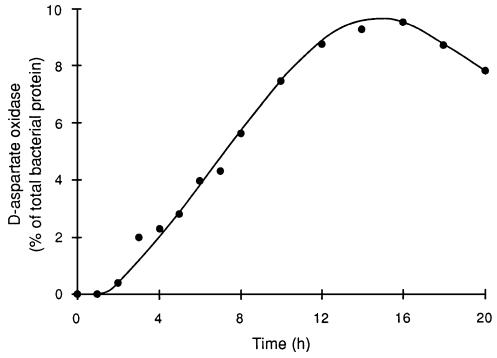
(b) An anchor oligonucleotide (5A) was ligated to the first-strand cDNA [12] in order to generate a suitable template for amplification of the 5′ UTR. Three subsequent PCR steps, the last two being semi-nested, were performed using an oligonucleotide complementary to the anchor sequence as 5′ primer (FR), coupled with the same three 3′ primers reported above (R3–R5). A faint band of about 400 bp appeared as the major product of the final amplification step. Sequence analysis of the resulting recombinant plasmid pRD428 showed that this DNA fragment, of 428 bp in size, contained 60 bp belonging to the 5′ UTR and overlapped exactly in its 3′ region with the 368 bp fragment previously described (Figure 1, part 2b).

(3) In order to isolate the 3′ part of the coding region that was still lacking, as well as the 3′ UTR, reverse transcription was primed by an oligo(dT) extended by an anchor sequence to the 5′ end (AdT). Subsequently, two PCR steps (the second semi-nested) were performed, using in turn two 5′ primers (F7 and F8) located within the already sequenced portion of the cDNA, coupled with a single 3′ primer (FR) annealing to the anchor sequence. By this procedure a DNA fragment spanning about 1 kb was obtained. Sequence analysis demonstrated that the terminal coding region and the whole 3′ UTR of the cDNA were contained in this 1071 bp DNA fragment (Figure 1, part 3) included in plasmid pRD1071.

Features of D-aspartate oxidase cDNA

The entire cDNA sequence was obtained from four main overlapping DNA fragments (Figure 1) derived from semi-nested PCR experiments. In order to avoid errors caused by *Taq* polymerase activity in the final compilation of the cDNA sequence, each region was checked by sequencing two or more fragments obtained from separate experiments.

The nucleotide sequence of bovine kidney D-aspartate oxidase cDNA and the deduced amino acid sequence are shown in Figure 2. The 2022 bp long cDNA exhibits a 5′ UTR of 60 nt, followed by the coding sequence with a single open reading frame of 1023 nt, interrupted by the stop codon TAG (nt 1024–1026). The 3′ UTR spans 936 nt, with the final 46 nt belonging to the poly(A) tail, which is preceded by a polyadenylation signal (AATAAA) (nt 1883–1888) located 28 nt upstream. The encoded protein comprises 341 amino acids, i.e.



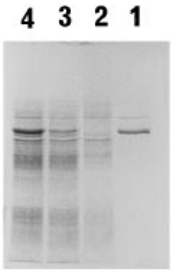


Figure 3 Expression of recombinant D-aspartate oxidase in *E. coli*

Samples of liquid culture of pED-transformed bacterial GI724 cells were taken immediately before and at fixed times after induction. Samples were sonicated and submitted to D-aspartate oxidase activity assays and SDS/PAGE. Upper panel: the amount of D-aspartate oxidase, calculated as a percentage of the total bacterial protein, is plotted as function of induction time. Each point is the mean value of at least two independent experiments. Lower panel: SDS/PAGE (12% gel) of bacterial protein after different induction times. Lane 1, bovine native D-aspartate oxidase; lanes 2–4, bacterial protein 0, 8 and 16 h respectively after induction.

three more residues at the C-terminus in comparison with the published D-aspartate oxidase sequence [7]. Moreover, two amino acid substitutions were found, i.e. Lys^{274} and Ser^{283} instead of arginine and glycine respectively.

Northern blot analysis of total and poly(A)$^+$ bovine kidney RNA after 4 days of exposure revealed one very faint band, of approximate size 2 kb (results not shown).

Overexpression of recombinant D-aspartate oxidase

Before starting expression experiments, the insert DNA of the pED plasmid was checked by sequence analysis for the presence of undesired nucleotide substitutions, as well as to define precisely the 5′ end boundaries.

Samples from the liquid culture of pED-transformed bacterial GI724 cells were obtained immediately before and at fixed times after the induction of expression. All enzyme activity was detected in the supernatant following cell sonication, suggesting that the expression system resulted in a soluble and active enzyme. No basic bacterial D-aspartate oxidase activity was detectable before induction of heterologous protein expression. Enzyme activity was undetectable 1 h after induction, became detectable after 2 h and increased up to the maximum value at approx. 16 h after induction (Figure 3). The enzyme specific activity at that time was calculated to be 0.95 unit/mg of protein. One unit of D-aspartate oxidase is defined as the amount of enzyme producing 1 μmol of H_2O_2/min at 25 °C under the conditions previously described [6]. The fraction of the expressed flavoprotein, out of the total protein content, can be calculated by assuming that recombinant D-aspartate oxidase exhibits the same specific activity as the bovine kidney cortex purified enzyme, i.e. 10 units/mg of protein [6]. Figure 3 (upper panel) shows that at around 16 h after induction, D-aspartate oxidase represented approx. 10% of the total bacterial protein. A lower yield of recombinant enzyme was obtained when the cellular density at the time of induction exceeded D_{550} 0.5–0.6.

The native and overexpressed D-aspartate oxidases, compared by SDS/PAGE analysis, showed the same electrophoretic behaviour, and the expression level appeared satisfactory (Figure 3, lower panel). Finally, the N-terminal identity of the two proteins was demonstrated by sequencing the first 20 amino acids of the recombinant product.

DISCUSSION

The isolation and sequencing of the complete cDNA coding for a D-aspartate oxidase, as well as the overexpression of the recombinant active enzyme, are reported here for the first time.

Several attempts to isolate the coding sequence for this flavoprotein from an RNA source, by using classical experimental approaches (cDNA library screening, one-step PCR), were unsuccessful, most probably because of the low amount of D-aspartate oxidase mRNA in vertebrate tissues, as confirmed by Northern blot analysis, where only a very faint band was detectable. Only the characterization of the short 60 bp genomic DNA fragment (Figure 2; nt 934–993) and the use of a sensitive technique such as semi-nested PCR allowed us to obtain, step by step, the sequence of the complete D-aspartate oxidase cDNA (Figure 2).

The features of this 2022 nt cDNA, described in the Results section, allow some speculation. It is known that sequences flanking the ATG initiation codon constitute a ribosomal binding site and play a role in modulating the translational efficiency [16]. The sequence ranging from positions -5 to $+4$ (GGCCCATGG, $+1$ being the first nucleotide of the start codon) present in the D-aspartate oxidase cDNA does not completely agree with the consensus sequence CCA/GCCATGG, identified for strongly expressed eukaryotic mRNAs [16,17]. In spite of a favourable downstream context, i.e. GA at positions $+4$ and $+5$, the rather unusual triplet found just before the start site (CCC) exhibits C at position -3, which could account for the greatly decreased translation level [16–18]. The importance of this suboptimal sequence is supported by the absence of any downstream hairpin, able to improve the translation start [19]. Another element of note is the stop translation site, comprising the stop codon TAG (nt 1024–1026), which is the least frequently used [18,20]. In this respect, it has been reported that the contexts surrounding the stop codons are able to modulate termination efficiency [21,22]. The 3′ nucleotide following the stop codon, required to direct binding of release factor to the ribosome, may be involved, together with the stop codon, in signalling the translation stop [20]. A purine base, primarily A followed by G, was suggested as the most effective 3′ flanking base for release factor binding as well as for protein release [20,22]. The termination activity of the tetranucleotide TAGG (nt 1024–1027) found in D-aspartate oxidase cDNA could, according to Brown et al. [20] and McCaughan et al. [22], be less efficient than the more frequent signals that are present in highly expressed mammalian genes, for instance TAAA and TGAA, which are probably best fitted to the active site of the protein release factor, thus allowing translation to proceed quickly. The above features are consistent with the low amounts of D-aspartate oxidase observed in vertebrate tissues.

A and T bases constitute about two-thirds of the 3′ UTR of the D-aspartate oxidase cDNA, and are equally represented. A number of recent reports have emphasized the importance of 3′ UTRs in the regulation of the fate of eukaryotic mRNAs, i.e. intracellular localization, translation and degradation [23–25]. Some A/T-containing sequences, beside the polyadenylation hexanucleotides, have been extensively studied, such as the AU-rich elements and cytoplasmic polyadenylation elements, which are involved in promoting degradation of mRNAs and poly(A) elongation respectively. None of these elements can be recognized within the 3′ UTR of D-aspartate oxidase cDNA, in spite of the presence of the eight-base sequence TTTTAAAT (positions 1287–1294), which in some RNAs seems to function as a cytoplasmic polyadenylation element [26]. In bovine D-aspartate oxidase mRNA, this eight-base sequence is located about 600 nt upstream of the polyadenylation signal AATAAA (nt 1883–1888), and therefore is not in a potentially effective context, at least in this case. In fact, seven other potential polyadenylation signals (Figure 2) are recognizable [27,28] and differently sized 3′ UTRs could occur, as already demonstrated for D-amino acid oxidase [29]. The 3′ UTR of D-aspartate oxidase cDNA, which is considerably longer than in many other cDNAs, might contain as yet unknown novel regulatory elements.

Searches for sequence similarities with other nucleic acid molecules, performed using the EMBL Data Base (CDEM43), did not show close similarity of D-aspartate oxidase cDNA with any known sequences, except for a slight similarity to the known mammalian D-amino acid oxidase cDNAs [28–31]. Nevertheless, in this latter case the similarity, ranging from less than 57% for human to about 53% for mouse mRNA, was mainly limited to one-third of the coding sequence, located within the 3′ region. This finding is also supported by the fact that no other genomic fragment, besides the corresponding gene, cross-hybridizing with the D-amino acid oxidase cDNA probe was detected by Southern blot analysis [32]. The lack of close similarity at the nucleotide level between D-aspartate and D-amino acid oxidases is also true for the start and stop translation sites, where different features are present. Except for that of the rabbit, which exhibits a wholly unfavourable context, the translation start sites of D-amino acid oxidase cDNAs show a purine base at the −3 position and poor 3′ features [16–22], i.e. a situation opposite to that for D-aspartate oxidase cDNA. The terminal tetranucleotides of D-amino acid oxidase cDNAs are always TGAA/G, which have been reported [16,18,20–22] to be much more active than TAGG, found in D-aspartate oxidase cDNA.

In contrast, a search performed using the four EST43 Data Bases revealed three newly submitted human cDNA fragments (L. Hillier, N. Clark, T. Dubuque, K. Elliston, M. Hawkins, M. Holman, M. Hultman, T. Kucaba, M. Le, G. Lennon, M. Marra, J. Parsons, L. Rifkin, T. Rohlfing, M. Soares, F. Tan, E. Trevaskis, R. Waterston, A. Williamson, P. Wohldmann and R. Wilson, accession nos. R19042 and R20147; G. G. Lennon, accession no. T26466) that exhibited more than 80% similarity with the bovine D-aspartate oxidase cDNA. These DNA fragments overlap but are differently extended at their ends; they appear to comprise the 5′ UTR and 300–400 nt of a coding sequence, more than 200 nt of which is very closely related to bovine kidney D-aspartate oxidase. Although the bovine and human 5′ UTRs are different, it is noteworthy that the nucleotide sequences (nt −22 to −1) immediately preceding the translation start codon ATG are almost identical (only one base is different: T in the human instead of C at position −2). This substitution concerns the triplet immediately upstream of the translation start site, which becomes CTC, a sequence even less frequent than CCC [18], previously discussed for the bovine cDNA. However, the properties are similar for both triplets, because of the presence of C at position −3 [16]. Curiously, rabbit D-amino acid oxidase exhibits the lowest enzyme expression and is the only case, among known mammalian cDNAs for this enzyme, where C is found at position −3 instead of a purine base. After deletion of most of the long 5′ UTR and transcription, a role for this nucleotide was suggested in determining low expression levels in *in vitro* translation experiments. In addition, no enzyme synthesis was observed for the native form of this rabbit mRNA, so that it was suggested that protein synthesis could be translationally restricted [28]. The failure to detect [^{35}S]methionine-labelled protein after *in vitro* translation of bovine kidney cortex poly(A)$^+$ RNA followed by specific immunoprecipitation using purified anti-(bovine kidney D-aspartate oxidase) antibodies (results not shown) is consistent with the result observed with the rabbit D-amino acid oxidase.

The amino acid sequence deduced from bovine D-aspartate oxidase cDNA substantially confirms the primary structure previously reported [7] except for two residues, namely Arg^{274} and Gly^{283}, substituted by lysine and serine respectively. Both of these substitutions involve a single base change: the first is conservative and corresponds to a position occupied by different residues in D-amino acid oxidases, and the second lies at a position occupied by glutamine in mammalian D-amino acid oxidases, so that this glycine should not have a special structural function. Furthermore, the tertiary structure of D-amino acid oxidase, recently resolved, does not indicate any critical role for residues at these two positions [33]. The possibility cannot be excluded that these residues represent additional cases of D-aspartate oxidase polymorphism. Indeed, various cases of molecular microheterogeneity, for instance the presence at times of valine or isoleucine at position 228, were noted during determination of the primary structure of the protein [7].

A more interesting observation concerns the presence of 341 amino acids in the deduced sequence instead of the 338 previously reported for this protein. These three additional residues are located at the C-terminus of the protein. It is likely that the failure to detect this terminal tripeptide by protein sequencing was due to proteolysis occurring during purification of the protein, as already observed for Lys^{338} [7]. The observed sequence (-Ser-Lys-Leu) represents a peroxisomal targeting signal 1 (PTS1) [34,35]. Since D-amino acid oxidases usually have the tripeptide -Ser-His-Leu at the end of the sequence [28–31], both D-aspartate oxidase and D-amino acid oxidases exhibit this additional common mechanism of peroxisomal targeting.

In conclusion, the data presented suggest that the similarities shown by these two flavo-oxidases at the protein level are not reflected at the nucleic acid level. Characterization of the gene structure could resolve the open question of their genetic relationship. Expression of recombinant D-aspartate oxidase was achieved in a prokaryotic system, giving a soluble and active enzyme. This recombinant product showed the correct N-terminal amino acid sequence and the same molecular mass as the native enzyme. The experimental conditions reported allowed us to achieve overexpression of this flavoprotein (to represent about 10% of total bacterial protein), thus making available amounts of protein several hundred times higher than from the best mammalian sources.

This work was supported by grants from the Consiglio Nazionale delle Ricerche (CNR Target Project on Biotechnology and Bioinstrumentation; contract no. 93.01079.PF38.115.14604) and from the Ministero dell'Università e della Ricerca Scientifica e Tecnologica (MURST 40 and 60%). Dr. A. Mattevi is gratefully acknowledged for providing experimental data on the crystal structure of D-amino acid oxidase prior to publication.

REFERENCES

1 Still, J. L., Buell, M. V., Knox, W. E. and Green, D. E. (1949) J. Biol. Chem. **179**, 831–837
2 Hamilton, G. A. (1985) Adv. Enzymol. **57**, 85–178
3 Kera, Y., Nagasaki, H., Iwashima, A. and Yamada, R. (1992) Comp. Biochem. Physiol. B **103**, 345–348
4 Rocca, E. and Ghiretti, F. (1958) Arch. Biochem. Biophys. **77**, 336–349
5 D'Aniello, A. and Rocca, E. (1972) Comp. Biochem. Physiol. B **41**, 625–633
6 Negri, A., Massey, V. and Williams, Jr., C. H. (1987) J. Biol. Chem. **262**, 10026–10034
7 Negri, A., Ceciliani, F., Tedeschi, G., Simonic, T. and Ronchi, S. (1992) J. Biol. Chem. **267**, 11865–11871
8 Negri, A., Massey, V., Williams, Jr., C. H. and Schopfer, L. M. (1988) J. Biol. Chem. **263**, 13557–13563
9 Jakoby, W. B. and Ziegler, D. M. (1990) J. Biol. Chem. **265**, 20715–20718
10 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., pp. 6.53–6.54, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
11 Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. **162**, 156–159
12 Dumas Milne Edwards, J. B., Delort, J. and Mallet, J. (1991) Nucleic Acids Res. **19**, 5227–5232
13 Stoscheck, C. M. (1990) Methods Enzymol. **182**, 50–68
14 Laemmli, U.K. (1970) Nature (London) **227**, 680–685
15 Matsudaira, P. (1987) J. Biol. Chem. **262**, 10035–10038
16 Kozak, M. (1986) Cell **44**, 283–292
17 Grünert, S. and Jackson, R. J. (1994) EMBO J. **13**, 3618–3630
18 Cavener, D. R. and Ray, S. C. (1991) Nucleic Acids Res. **19**, 3185–3192
19 Kozak, M. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 8301–8305
20 Brown, C. M., Stockwell, P. A., Trotman, C. N. A. and Tate, W. P. (1990) Nucleic Acids Res. **18**, 6339–6345
21 Arkov, A. L., Korolev, S. V. and Kisselev, L. L. (1995) Nucleic Acids Res. **23**, 4712–4716
22 McCaughan, K. K., Brown, C. M., Dalphin, M. E., Berry, M. J. and Tate, W. P. (1995) Proc. Natl. Acad. Sci. U.S.A. **92**, 5431–5435
23 St Johnston, D. (1995) Cell **81**, 161–170
24 Curtis, D., Lehmann, R. and Zamore, P. D. (1995) Cell **81**, 171–178
25 Beelman, C. A. and Parker, R. (1995) Cell **81**, 179–183
26 Paris, J. and Richter, J. D. (1990) Mol. Cell. Biol. **10**, 5634–5645
27 Guo, Z. and Sherman, F. (1995) Mol. Cell. Biol. **11**, 5983–5990
28 Momoi, K., Fukui, K., Tada, M. and Miyake, Y. (1990) J. Biochem. (Tokyo) **108**, 406–413
29 Fukui, K., Watanabe, F., Shibata, T. and Miyake, Y. (1987) Biochemistry **26**, 3612–3618
30 Momoi, K., Fukui, K., Watanabe, F. and Miyake, Y. (1988) FEBS Lett. **238**, 180–184
31 Tada, M., Fukui, K., Momoi, K. and Miyake, Y. (1990) Gene **90**, 293–297
32 Fukui, K. and Miyake, Y. (1992) J. Biol. Chem. **267**, 18631–18638
33 Mattevi, A., Vanoni, M. A., Todone, F., Rizzi, M., Teplyakov, A., Coda, A., Bolognesi, M. and Curti, B. (1996) Proc. Natl. Acad. Sci. U.S.A. **93**, 7496–7501
34 Purdue, P. E. and Lazarow, P. B. (1994) J. Biol. Chem. **269**, 30065–30068
35 Rachubinski, R. A. and Subramani, S. (1995) Cell **83**, 525–528

Received 2 July 1996/2 October 1996; accepted 30 October 1996