The Primary Structure of the Flavoprotein D-Aspartate Oxidase from Beef Kidney*

(Received for publication, September 13, 1991)

Armando Negri, Fabrizio Ceciliani, Gabriella Tedeschi, Tatjana Simonic and Severino Ronchi‡

From the Istituto di Fisiologia Veterinaria e Biochimica, Università di Milano and the Centro Interuniversitario per lo Studio delle Macromolecole Informazionali, 20100 Milano, Italy

The complete primary structure of the peroxisomal flavoenzyme D-aspartate oxidase from beef kidney has been determined by analyses of the peptides obtained through fragmentation of the carboxymethylated protein with trypsin, CNBr, heptafluorobutyric acid/ CNBr and *Staphylococcus aureus* V8 protease. The protein consists of a single polypeptide of 338 residues, accounting for a M_r of 37,305 for the apoprotein. A form of the enzyme lacking Lys-338 and therefore ending with Pro-337 has been detected.

The N-terminal residue is blocked. Seven cysteines and no disulfide bridges are present. Residue 228 can be either Ile or Val. Thus, D-aspartate oxidase presents two types of heterogeneity in the polypeptide chain in addition to the one already described concerning the possible content of FAD or 6-hydroxyflavin adenine dinucleotide.

Comparison of the primary structure of D-aspartate oxidase with other known sequences reveals that Daspartate oxidase is homologous with D-amino acid oxidase (another flavo-oxidase) and does not present significant sequence similarities with any other protein, including flavoenzymes.

D-Aspartate oxidase (EC 1.4.3.1) is a flavoprotein that catalyzes the oxidative deamination of dicarboxylic D-amino acids by O_2 with the production of H_2O_2 , NH_3 , and the corresponding α -keto acids (1). Since the discovery of a Daspartate oxidase activity in tissue homogenates, it became clear that the enzyme responsible for this activity is distinct from the other well known flavoenzyme D-amino acid oxidase (EC 1.4.3.3), which catalyzes the same reaction but is active only toward neutral and basic amino acids.

D-Aspartate oxidase has been so far purified only from the hepatopancreas of *Octopus vulgaris*. (2) and beef kidney cortex (3). The spectrophotometric properties and kinetic mechanism of the beef enzyme have been investigated in detail (3, 4). These studies showed that the general properties of D-aspartate oxidase resemble those of D-amino acid oxidase.

The latter is still considered a model for the understanding of the mechanism of action of flavoprotein oxidases. For this reason, we extended the comparison between the two enzymes at the primary structure level by determining the amino acid sequence of beef kidney D-aspartate oxidase and comparing it with the one of D-amino acid oxidase from mammals (5-7), yeast (8), and bacteria (9). This information will be useful in order to understand the evolution of these molecules, in view of a possible explanation for their physiological role, which is still unclear.

This report presents the experimental work that allowed the determination of the amino acid sequence of D-aspartate oxidase.

EXPERIMENTAL PROCEDURES AND RESULTS^{1,2}

DISCUSSION

This paper reports for the first time the characterization of the primary structure of the flavoprotein D-aspartate oxidase from beef kidney. The enzyme contains 338 amino acid residues, accounting for a M_r of 37,305 for the apoprotein. This value agrees with the one determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and gel filtration, suggesting that also this peroxisomal enzyme is not a glycoprotein, despite the presence of N-glycosylation consensus sequences (Asn-X-Ser/Thr) at positions 181–183, 230–232, and 243–245.

The N-terminal residue is blocked. Seven cysteines and no disulfide bridges are present. The primary structure has been determined by sequencing the fragments obtained by digestion of the carboxymethylated enzyme with trypsin, CNBr, CNBr/HFBA, and *Staphylococcus aureus* protease. Experimental evidence for the amino acid sequence shown in Fig. 9 has been obtained for all but the first amino acid residue, which has been indicated as Met on the basis of the following observations: (a) the sequence of C2 shows an N-terminal extension of 4 amino acid residues with respect to T2, the first tryptic peptide sequenced; (b) the amino acid composition of the tryptic peptide with retention time of 25 min in Fig. 1, whose N terminus is blocked, agrees with the one expected for T1 (Table I); (c) the amino acid composition of the protein deduced from the sequence is in good agreement

^{*} This work was supported by grants from the Ministero della Università e della Ricerca Scientifica e Tecnologica and from the Consiglio Nazionale delle Ricerche (Target Project on "Biotechnology and Bioinstrumentation," contract 89.00212.70). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EMBL Data Bank with accession number(s) M91559.

[‡] To whom correspondence should be addressed: Istituto di Fisiologia Veterinaria e Biochimica, Università di Milano, Via Celoria 10, 20133 Milano, Italy. Tel.: 02-2664343; Fax: 02-2367788.

¹ Portions of this paper (including "Experimental Procedures" and "Results," Figs. 1–8, and Tables I and II) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press. ² The abbreviations used are: HFBA, heptafluorobutyric acid;

² The abbreviations used are: HFBA, heptafluorobutyric acid; PTH, phenylthiohydantoin; GnCl, guanidine hydrochloride; TFA, trifluoroacetic acid; DTNB, 5,5'-dithiobis(2-nitrobenzoic) acid; PVDF, polyvinylidene difluoride; TPCK, *N-p*-tosyl-L-phenylalanine chloromethyl ketone.

with the one determined by amino acid analysis (Table I); (d)no additional short peptides deriving from fragmentation with CNBr were isolated, suggesting that no other amino acid residues besides a single Met is present before C2. The identification of the C-terminal residue posed some problems due to the lack of information derived from the direct sequence of the entire protein through digestion with carboxypeptidase. In addition, since the last residue is Lys, no tryptic peptide lacking a basic residue at its C terminus was detected, except for T37a, which, however, appears to derive from a form of the enzyme that underwent partial C-terminal proteolysis either in the cell or during the purification procedure of the native enzyme. The assignment of T37 as the C-terminal tryptic peptide is based on the sequence of C7. This CNBr peptide, in fact, overlaps T35, T36, and T37 and lacks a methionine residue at its C terminus. In conclusion, these data show that the enzyme is isolated as a mixture of two forms differing for the C-terminal residue that can be either Pro-337 or Lys-338. This observation is important in view of the role proposed for the C-terminal tripeptide in many peroxisomal enzymes as a "peroxisomal targeting signal" (10). The data presented here for D-aspartate oxidase suggest that such a tripeptide (usually Ser/His(Lys)/Leu) is absent, indicating that some different structural feature fulfills the requirement for the targeting of this enzyme to peroxisomes. However, since Pro has been shown to be a possible substitute for Ser as the first residue of the tripeptide (10), it is also possible that a third residue following Lys-338 (Leu or a residue functionally equivalent for protein targeting) has been proteolytically lost in a maneer similar to what is observed for Lys-338.

Microheterogeneity has also been observed at position 228, which can either be Val or Ile. However, this substitution represents a conservative replacement, originated by a single point mutation in the gene, and is not expected to have a major effect on the general properties of the enzyme. In conclusion, this study reveals that D-aspartate oxidase shows molecular microheterogeneity in the polypeptide chain, in addition to the one already described concerning the possible content of FAD or 6-hydroxyflavin adenine dinucleotide in the isolated protein (3).

Fig. 10 shows the alignment of the amino acid sequence of D-aspartate oxidase from bovine kidney and that of several D-amino acid oxidases from various sources. The "consensus sequence," comprising 137 out of 387 residues (35%) was obtained, considering those amino acids that are identical or conservatively substituted in at least six out of the seven sequences shown. It is interesting to note that D-aspartate oxidase presents 125 of the 137 residues of the consensus sequence (91%), pig, human, and mouse D-amino acid oxidase presents 137 (100%), rabbit D-amino acid oxidase presents 134 (98%), Trigonopsis variabilis D-amino acid oxidase presents 107 (78%), and Fusarium solani D-amino acid oxidase presents 115 (84%). This observation suggests that indeed Daspartate oxidase retains most of the primary structural characteristics common to D-amino acid oxidases. Inspection of the consensus sequence reveals that it comprises residues belonging to all the different portions of the molecules: 39, 50, and 48 conserved amino acids are present between residues 1 and 120, 121 and 240, and 241 and 387 of the consensus sequence, respectively. It can be concluded that D-aspartate oxidase and D-amino acid oxidase derive from divergent evolution of a single gene.

When the comparative analysis is restricted to D-aspartate oxidase from beef kidney and D-amino acid oxidase from pig kidney (the usual and best studied source of this enzyme), the



FIG. 9. The complete amino acid sequence of beef kidney Daspartate oxidase. T, tryptic peptides; C, CNBr peptides; H, CNBr/ HFBA peptides; S, S. aureus V8 peptides.

two proteins appear to be homologous (41% identities and an additional 15% of conservative replacements). In addition, the observation that 75% of the glycines and tryptophans and 60% of the prolines are conserved suggests also that the tertiary structure of the two enzymes must be very similar.

The role of active site residues in pig kidney D-amino acid oxidase has been extensively investigated using both chemical modification and site-directed mutagenesis techniques (7, 11). According to the alignment shown in Fig. 10, Tyr-228 and His-307 (the numbers refer to pig kidney D-amino acid oxidase) are punctually conserved in all the enzymes, suggesting a common catalytic role for these residues that, however, has not yet been elucidated. None of the other "active site residues" (11) (Tyr-55, Met-110, Lys-204, His-217, and Tyr-224; marked with an *asterisk* in Fig. 10) are conserved in all the proteins considered, although in most cases, functionally equivalent residues can be found around the position of those amino acids in the sequence of the enzymes.

No other proteins besides those shown in Fig. 10 present extended primary structure similarity with D-aspartate oxidase, except for the N-terminal sequence, which is known to be common to most enzymes binding ADP-containing nucleotides (12). In particular, no significant similarities were detected between D-aspartate oxidase and the FMN- containing flavoproteins of the oxidase/dehydrogenase class that have been shown to share a common mechanism of catalysis with D-amino acid oxidase (13) (lactate oxidase, glycolate oxidase, and flavocytochrome b_2 (14–16)). This observation reinforces the proposal of Geigel *et al.* (14) that D-aspartate oxidase and D-amino acid oxidase (both FAD-containing enzymes) constitute a family of homologous proteins different from the one comprising the FMN-containing oxidases and

DASPO b	
DAAO D	1 K RVVVIGAGVIGLSTAL CIRE RVHSVLOP LOVEV VADDE
DAAO h	1 K RVVVIGAGVIGLETAL CIHE RYHSVLOP LHIKV YADRP
DAAO r	1 M RVVVIGAGVIGLETAL CIHE LYHSALOP LDNTI YADRF
DAAO B	1 M RVAVIGAGVIGLETAL CIHE RYH PTQP LHMKI YADRF
DAAO t	1 M AKIVVIGAGVAGLTTALQLLRKGHEVTIVSEFTPGD LSIGYTSPW
DAAO Y	1 M SNTIVVVGAGVIGLTSAL LLSKNKGNKITVVAKHNPGDYDVEYASPF
CONSENSUS	51T tva gl ped sis
	• • • • • • • • • • • • • • • • • • • •
DASPO b	39 TPETTSDVA AGHLIPPTYPDTPIQKQ KQWFKETPDHLFAIVNSAEAE
DAAO P	40 TPFTTTDVA AG LNQP YTSEPSNPQEANNNQQTPNYLLSHIGSPNAA
	40 TPLTTTUVA AG LNOP YLSDPHNPQEADWSQQTPDYLLSHVHSPNAE
DAAO	39 TETTSIVA AG LAOP VISDENDOVENCOVERVIJECTUSDAN
DAAO t	46 AGANWLTFYD GGKLADYDAVSYP ILRELARSSPEAGIRLINORSHVIK
DAAO f	49 AGANHSPMATEESSEWERRTWYEPKRLVE EVPEAGVHPQKSRIQRRNV
CONSENSUS	
CONDERDOD	
DASPO b	86 DAGVILVSGWOIFOSIPTEEVPYWADVVLGFRKWTKDELRKFPOHVF CH
DAAO p	86 NMGLTPVSGYNLF REAVPDPYWKDMVLGFRKLTPRELDMPPDYRY GW
DAAO h	86 NLGLFLISGYNLF HEAIPDPSWEDTVLGFRELTPRELDMFPDYGY GW
DAAO T	86 KNGLALISGYNLF RKAVPDPSWKDTVLGFRKLTLRELDNFPGYSY GW
	85 KNGLALISGYNLF RDEVPDPFNKNAVLGFRKLTPSEMDLFPDYGY GW
	97 ND LPALEGARSAICOK NPWFKNIVDSFEIIEDRSKIVHDDEAYLVE
	ST DIEGOGROG FFOREFOREFOREDIREDIREDIREDIREDIREDIREDIREDIREDIREDI
CONSENSUS	151 Teleg YLWL rlvk ke D
Digno b	
DANO D	135 AFTTLACEGPATLEWL QKRVKGNGGL ILTRRIE DLWE LHPSFD
DAAO h	133 FHISLILEGENTIONS TERLITERGYN FFLERVE SPEEVARGGAD
DAAO r	133 FNTSLILDGRSYLOWL TKRLTERGVK LFORKVE SFDEVAGGGVD
DAAO m	132 FNTSLLLEGKSYLPWL TERLTERGVK LIHRKVE SLEEVA RGVD
DAAO t	140 FRSVCINTG VYLNWLMSQCLSLGATVVKRRVNHIKDA NLLHSSGSRPD
DANO 2	141 EPTSVCINTAIYLPWLLGQCIK NGVIVKRAILNDISEAKKLSHAGKTPN
CONSENSUS	201 VIVNCTOVAL DLPRODITY py fit a
	•
DASPO b	178 IVVNCSGLGSRQL AGDSKIPPVRGOVLKV QAPWVKHFI RD SSG
DASPO B DAAO P	178 IVVNCSGLGSRQL AGDSKIPPVRGQVLKV QAPWVKHFI RD SSG 177 VIINCTGVWAGVL QPDPLLQPGRGQIIKV DAPWLKNFIITHDLERG
DASPO B DAAO P DAAO h	178 IVVNCSGLGSRQL AGDSKIPFVRGOVLKV QAPWVKHPI RD SSG 177 VIINCTGVMAGVL OPDPLLQPGRGOINKV DAPWIKHFIITHDLENG 177 VINCTGVMAGAL ORDPLLQPGRGOINKV DAPWIKHFIITHDPENG
DASPO b DAAO p DAAO h DAAO r DAAO m	176 IVVNCSGLGSRQL AGDSKIFFVRGQVLKV QAPWVKHFI RD SSG 177 VINCTGVMAQL QPDPLLQPGRGQIIKV DAPWLKHFITHDLERG 177 VIVNCTGVMAQL QPDPLLQPGRGQIIKV DAPWKHFITHDPERG 177 VIVNCTGVMAQL QDPPLLQPGRGQIIKV DAPWKHFITHDPERG
DASPO b DAAO p DAAO h DAAO r DAAO m DAAO t	178 UVWCSGLESKOL AGDSKIFFVRGOULKV QAPWYNHFI ED SSC 177 VIINCTSWNACH QPPPLLQFSRQIIKV DAPHIXHFITHDLERG 177 VIINCTSWNACH QRDPLLQFSRQIIKV DAPHIXHFITHDLERG 177 VINCTSWNACH QPPLLQFSRQIIVV DAPHIXHFITHDLERG 175 VIINCTSWNACH QPPLLQFSRQIIVV LAPHIXHFITHDLERG 185 VIINCTSWNACH DADALGFSRQIIVV LAPHIXHFITHDLERG 185 VIINCTSWNACH DADALGFSRQIIVV LAPHIXHFITHDLERG
DASPO b DAAO p DAAO h DAAO r DAAO r DAAO m DAAO t DAAO f	178 IVVNCSGLGSRQL AGDSKIFFVRGQVLKV QAPWVKHFI RD SSG 177 VINCTGVMAQUL QPDFILQFGRGQIIKV DAPWLKHFITHDLERG 177 VINCTGVMAQL QRDFLLQFGRGQIKV DAPWLKHFITHDPERG 175 VINCTGVMAQL QADSLQFGRGQIIVV DAPWLKHFITHDPEGG 175 VINCTGVMAQL QADSLQFGRGQIIVV DAPWLKHFITHDPEGG 188 VINCSGLFARFLGGVEDKNMFFIGQVVVVRNSLFPMAFFST PEKEM 190 IIVNGGLGFARFLGGVEDKNMFRGQVVVVRNSLFPMAFFST PEKEM
DASPO b DAAO p DAAO h DAAO r DAAO r DAAO m DAAO t DAAO f	176 IVVNCSGLGSRQL AGDSKIFFVRGQVLKV QAPWVKHFI RD SSG 177 VINCTGVMAQL QPDPLLQPGRGQIIXV DAPWLKHFITHDLERG 177 VIVNCTGVMAQL QPDPLLQPGRGQIIXV DAPWLKHFITHDPERG 177 VIVNCTGVMAQL QADSLQPGRGQIIXV DAPWLKHFITHDPERG 175 VIINCTGVMAQL QADSLQPGRGQIIQV EAPWIKHFITHDPESG 188 VIVNCSGLPARFLGGVEDKKNYPIRGQVVLVRNSLPFNASFSST PEKEN 190 IIVMATGLGSYKLGGVEDKTMAPARGQIVVV RNESFNLJTSGVEDG
DAASPO b DAAO p DAAO h DAAO r DAAO r DAAO m DAAO t DAAO f	178 IVVNCSGLGSRQL AGDSKIFFVRGOULW QAPWXNHFI RD SSG 177 VIINCTGVMAGL QPDPLLQPGRGQIIKV QAPWXNHFITHDLERG 177 VINCTGVMAGL QPDPLLQPGRGQIKV DAPWNKHFITHDLERG 177 VINCTGVMAGL QPDPLLQPGRGQIKV DAPWNKHFITHDERG 175 VIINCTGVMAGL QHDSLQPGRGQIKV DAPWNKHFITHDERG 186 VINCSGLFARFLGGVEDKNMFFIRGQVUVNHSLPFMASFSST PEKEN 196 UIVMAGLGSKKLGGVEDKNMFFIRGQVUVNHSLPFMASFSST PEKEN
DASPO b DAAO p DAAO h DAAO r DAAO r DAAO m DAAO t DAAO f CONSENSUS	178 IVVRCSGLGSRQL AGDSKIFFVRGQVLKV QAPWVKHFI RD SSG 177 VINCTGVMAAL QPDFLIQEGGQIKV DAPWLKHFITHDLERG 177 VINCTGVMAAL QPDFLIQEGGQIKV DAPWKHFITHDERG 175 VINCTGVMAAL QPDFLIQEGGQIKV DAPWKHFITHDERG 175 VINCTGVMAAL QDDALQEGGQIQV DAPWKHFITHDERG 175 VINCTGVMAAL QDADALQEGGQIQV DAPWKHFITHDESG 186 VINCSGLFARFLGGVEDKKNFFIRGQVUVNNSLFPMAFSST PEKEN 190 IIVMATGLGSVKLGGVEDKKNFFIRGQVUVNNSLFPMAFSST PEKEN 251 n YI1 t
DASPO b DAAO p DAAO r DAAO r DAAO r DAAO R DAAO t DAAO f CONSENSUS	178 VUNCSGLESKOL AGOSKIFFVRGOULW OAPWYNFI ED SSC 177 VUINCGOWNACH OPPELLOPGRGIIKW OAPWYNFI THDLERG 177 VUNCGOWNACH ORDPILLOPGRGIIKW DAPWINFITHDLERG 177 VUNCGOWNACH ORDPILLOPGRGIIKW DAPWINFITHDPERG 179 VUNCGOWNACH OADDILOPGRGIIVW DAPWINFITHDPERG 188 VUNCGOWNACH OADDILOPGRGIIVW DAPWINFITHDPERG 190 IIVNATGLOSYKLGGVEDKOMPIRGOVUNNSLFPMARSST PERG 190 IIVNATGLOSYKLGGVEDKOMPIRGOVUNNSLFPMARSST PERG 251 n YII t LOG Q gnW I 221 IVVINGLSWYT LGGTEDKOMPIN LAPDA EIGKKULARCCA
DASPO b DAAO p DAAO b DAAO r DAAO r DAAO t	174 IVVNCSGLGSKQL AGDSKIFFVRGQVLKV QAPWYKHFI RD SSG 177 VINCTGVMAGVL QPDPLLQPGKQIIKV DAPWIKHFITHDLERG 177 VINCTGVMAGAL QRDPLLQPGKQIKV DAPWKHFITHDLERG 177 VINCTGVMAGAL QRDPLLQPGKQIKV DAPWKHFITHDPEGG 175 VINCTGVMAGAL QADASLQPGKQIKV DAPWKHFITHDPEGG 175 VINCTGVMAGAL QADASLQPGKQIKV DAPWKHFITHDPEGG 186 VINCSGLPARFLGGVEDKTMIPIKGQVUVNNSKPTMAEFSST PEKEN 190 IIVNATGLGSYKLGGVEDKTMIPIKGQVUVNN RMESSPMLLTSGVEDG 251 n YI1 t 251 n YI1 t 251 n YI1 t 251 ITIYPGISHVT LGGTRGKOM N EISHKILSRCA 231 LYNYGISHVT LGGTRGKOM N SINHBCCR
DASPO b DAAO p DAAO r DAAO r DAAO r DAAO t DAAO t CONSENSUS DASPO b DAAO p DAAO h	178 VUNCSGLESKOL AGDSKIFFVRGOULW OAPWYNFI ED SSC 177 VIINGTGWNACL OPPELGPGRGUIKV OAPWYNFI ENDLERG 177 VINGTGWNACL ORDFLAGFGRGUIKV DAPHARMFI ENDLERG 177 VINGTGWNACL OPPELGPGRGUIKV DAPHARMFI ENDLERG 178 VINGTGWNACL OPPELGPGRGUIKV DAPWYNFI ENDLERG 188 VINGTGWNACL OADSLAGFGRGUIGV EAPWYNFI ENDESG 188 VINGTGWNACL OADSLAGFGRGUIGV EAPWYNFI ENDESG 190 IIVNATGLGSYKLGGVEKTWNFERGUVLWN ENPHARFSET PEREN 190 IIVNATGLGSYKLGGVEKTWNFERGUVLWN ENESPHLLTSGVEDG 251 n YII t LGG Q gnW I 251 LTYIYFGISHVT LGGTGRGKDW N LSPDA EISKHILSRCCA 231 IYNSFYIJFGGYVT LGGTGRGKDW N LSPDA EISKHILSRCCA 233 IYNSFYIJFGGYVT LGGTGRGKDW N ELNNI QOHNTIMEGCR
DASPO b DAAO p DAAO h DAAO h DAAO n DAAO t DAAO p DAAO p DAAO p DAAO p DAAO p DAAO r	178 IVVNCSGLGSNGL AGDSKIFFVRGOULW QAPWXNFII RD SSG 177 VIINCTGVMAGL QPDFLLQFGRGOINV QAPWXNFIITHDLERG 177 VINCTGVMAGL QPDFLLQFGRGOINV DAPWKNFIITHDLERG 177 VINCTGVMAGL QHDFLQFGRGOINV DAPWKNFIITHDERG 175 VINCTGVMAGL QHDSLQFGRGOINV DAPWKNFIITHDERG 186 VINCSGLFARFLGGVEDKNNFIRGQVUVNNSLPFMAFFST PEKEN 190 IIVMAGLGSIKLGGVEDKNMFPARGOIVV RNSSPHILTSGVEDG 251 n YII t LØG Q gnW I 221 LTYIPGISNVT LGGTRQKGM N LSPDA EISKELSRCCA 231 IVNSFIITEGGVVT LGGTRQVENW N EINNI QDHNTINEGCCR 231 IVNSFIITEGGVT LGGIFQLGNM S ELMNI QDHNTINEGCCR
DASPO b DAAO P DAAO h DAAO r DAAO r DAAO r DAAO t DAAO t DAAO f DASPO b DASPO b DASPO b DASPO h DAAO r DAAO r DAAO r	178 VUNCSGLESKOL AGDSKIFFVRGULKV OAPWYNFIF ED SSC 177 VINCTGVWAAL OPPFLAGPGRGUIKV OAPWINFIFIDLERG 177 VINCTGVWAAL ORDFLAGPGRGUIKV OAPWINFIFIDDERG 177 VINCTGVWAAL ORDFLAGPGRGUIKV OAPWINFIFIDDERG 178 VINCTGVWAAL OADSLAPGRGUIKV OAPWYNFIFIDDESG 188 VINCTGVWAAL QADASLAPGRGUIKV APWYNFIFIDDESG 190 IINNATGLGSVKLGGVEDKTWFIRGVUVNNSLEPTMAFSSF 190 IINNATGLGSVKLGGVEDKTWFIRGVUVNNSLEPTMAFSSF 251 n YII t L400 Q gnW 221 LTVIYPGISNVT LGGTRGKGDM N LSPDA 223 IYNSPYIIPGGUVVT LGGTGTGLMW S ELNNI QOHNTIWEGCCR 223 IYNSPYIIPGGVAVT LGGTGRGHW S ECNST 223 IYNSPYIIPGGAVT LGGTGRGHW S ECNST 223 IYNSPYIIPGGVAVT LGGTGRGHW S ECNST
DASPO b DAAO p DAAO r DAAO r DAAO r DAAO r DAAO t DAAO f DASPO b DASPO b DASPO b DAAO p DAAO r DAAO r DAAO r DAAO r	178 VUNCSGLESKOL AGDSKIFFVRGOULW OAPWCKHFIED SSG 177 VUNCGCWMACH GPDFLIGPGRGOINV OAPWCKHFITHDIEGG 177 VUNCGCWMACH GPDFLIGPGRGOINV OAPWCKHFITHDIEGG 177 VUNCGCWMACH GPDFLIGPGRGOINV OAPWCKHFITHDIEGG 178 VINCGCWACH OADSALGGERGOINV OAPWCKHFITHDIEGG 188 VINCGCUARTLAWLGGVEDKWHTINGGVUVWNSLFFNAFSST FEKEN 190 IIWANGLGSYKLGGVEDKWHTAGGUVVV RHSSFMLITSGVEDG 251 n YII t LØG Q gnW I 221 LFVINGISWYF LGGTROKGDW N LSPDA EIKKELSRCCA 223 IYNSPIIIPGIGAVT LGGTROKGDW N LSPDA EIKKELSRCCA 231 IYNSPIIIPGIGAVT LGGTROKGDW N ENNI GOHRTIMGCCCR 233 IYNSPIIIPGIGAVT LGGTROKGMW S ECHNT DOHRTIMGCCCS 231 IYNSPIIIPGKTVT LGGIGLGHW S ELMNI GOHRTIMGCCCS 231 IYNSPIIIPGKTVT LGGIGLGHW S ELMNI GOHRTIMGCCCS 231 IYNSPIIIPGKTVT LGGIGLGHW S ELMNI GOHRTIMGCCS 231 IYNSPIIIPGKTVT LGGIGLGHW S ELMNI GOHRTIMGCCS
DASPO b DAAO p DAAO h DAAO h DAAO n DAAO f CONSENSUS DAAO DAAO p DAAO p DAAO p DAAO p DAAO p DAAO n DAAO t DAAO t DAAO t	174 IVVNCSGLGSNQL AGDSKIFFVRGQVLKV QAPWYKHFI RD SSG 177 VINCTGVMAQLL QPDPLLQPGRGQITKV DAPWIKHFITHDLERG 177 VINCTGVMAQLL QPDPLLQPGRGQITKV DAPWIKHFITHDLERG 177 VINCTGVMAQLL QPDPLLQPGRGQITKV DAPWIKHFITHDERG 175 VINCTGVMAQLL QDADSLQPGRGQITKV DAPWIKHFITHDERG 175 VINCTGVMAQLL QDADSLQPGRGQITKV DAPWIKHFITHDERG 180 VINCSGLPARFLGGVEDKUMPARGQVUV RMESSPMLLTSGVEDG 251 n YI1 t L40 Q gnW I 251 IN YI1 LGGTRQKOM N ELNNI QOHNTIMESCCR 231 IXNSPVIIPGIGNVT LGGTRQKOM N ELNNI QOHNTIMESCCR 231 IXNSPVIIPGGVTV LGGIFQLANN S ELNNI QOHNTIMESCCR 231 IXNSPVIIPGGKTVT LGGIFQLANN S ELNNI DOHNTIMESCCR 231 IXNSPVIIPGGANS IIGGCSTPLANNSE EDDNES DOHNTIMESCCR </td
DASPO b DAAG p DAAG h DAAG r DAAG r DAAG r DAAG t COHSENSUS DAAG f DAAG p DAAG h DAAG r DAAG r DAAG t DAAG t DAAG t	174 VUNCSGLESKOL AGDSKIFFVRGQULW OAPWYNFIED SSG 177 VUNCGGWAGAL GRDPILGPGRGJIKV OAPWYNFIETHDLERG 177 VUNCGGWAGAL GRDPILGPGRGJIKV DAPWNNFIETHDDERG 177 VUNCGGWAGAL GPDFLLGPGRGJIKV DAPWNNFIETHDDESG 188 VUNCGGWAGAL GADSLGPGRGJIGV LAPWYNFIETHDESG 190 IIVNATGLSYKLGGVEKKMYRGQVVVRNSLEPMASF267 P2KS 190 IIVNATGLSYKLGGVEKKMYRGQVVVRNSLEPMASF267 P2KS 221 IVNSPYIFGGNYT LGGTGRKGDN N LSPDA EISKHILSRCCA 223 IVNSPYIFGGNYT LGGTGRKGDN N LSPDA EISKHILSRCCA 223 IVNSPYIFGGNYT LGGTGRKGDN N LSPDA EISKHILSRCCA 223 IVNSPYIFGGNYT LGGTGRKGN S ENNT DONNTMESCCR 231 IVNSPYIFGGNYT LGGIGGNAM SENST DONNTMESCCR 231 IVNSPYIFGSTYT LGGIGGNAM SENST DONNTMESCCR 231 IVNSPYIFGGNYT LGGIGGNAM SENST DONNTMESCCR 231 IVNSPYIFGSTYT LGGIGGNAM SENST DONNTMESCCR 231 IVNSPYIFGSTYT LGGIGGNAM SENST DONNTMESCCR 231 IVNSPYIFGSTYT LGGIGGNAM SENST DONNTMESCCR 231 IVNSPYIFGSTYT LGGIGGNAM SENST DONNTMESCR 231 IVNSPYIFGSTYT LGGIGGNAM SENST DONNTMESCR 231 IVNSPYIFGSTYT LGGIGGTGNAM SENST DONNTMESCR
DASPO b DAAO p DAAO r DAAO r DAAO r DAAO r DAAO r DAAO r DAAO t DAAO t DAAO t DAAO p DAAO p DAAO p DAAO r DAAO r DAAO t DAAO t DAAO t DAAO t DAAO f	17 VUNCSGLESKQL AGDSKIFFVRQUIAV QAPWVENFI RD SSG 17 VINCTGVMAQL QPDFLIQFGRQIIKV QAPWVENFI RDLERG 17 VINCTGVMAQL QPDFLIQFGRQIKV DAPWLKFIITDLERG 17 VINCTGVMAQL QPDFLIQFGRQIKV DAPWLKFIITDLERG 17 VINCTGVMAQL QDASLGFRKQIKV DAPWLKFIITDLERG 17 VINCTGVMAQL QDASLGFRKQIKV DAPWLKFIITDLERG 17 VINCTGVMAQL QDASLGFRKQIQVE DAPWLKFITDPBEG 18 VINCSGLFARFLGGVEDKTNFRGQVUVVNSLEPMAFSST PEKEM 19 IVMAGLSSYKLGGVEDKTMFARGQVUVVNSLEPMAFSST PEKEM 251 N YI LGGTRQKGDN N LSPDA EISKELLSGCC 221 LVSPYILFGGAVT LGGTFQVENW N EINN QOHNTIMEGCCR 221 221 LVSPYILFGGAVT LGGTFQUENW N ELNNI QOHNTIMEGCCR 221 IVMSPYILFGGAVT LGGIFQLENW S ELNNI QOHNTIMEGCCR 221 LYSPYILFGGAVT LGGIFQLENW S ELNNI QOHNTIMEGCCR 221 IYMSPYILFGGAVT LGGIFQLENW S ELNNI QOHNTIMEGCCR 221 IYMSPYILFGGAVT LGGIFQLENW S ELNNI QOHNTIMEGCCR 221 IYMSPYILFGRAVEL ELNKKRCCS 221 IYMSPYILFGGAVT LGGIFQLENW S ELNST DOHNTIMEGCCR 221 231 IYMSPYILFGGAVT LGGIFQLENW S ELNST DOHNTIMEGCCR 231 IYMSPYILFGARAVELGAT LGGTFDVGNH S ELNST DOHNTIMSCCE 24
DASPO b DAAO p DAAO h DAAO r DAAO r DAAO r DAAO r DAAO r DAAO r DAAO t DAAO t DAAO t DAAO t DAAO p DAAO p DAAO p DAAO n DAAO r DAAO r CONSENSUS	17 VINCSGLESKOL AGDSKIFFVRGULAV OAPWURFIF ED SSG 17 VINCGUWAAL OPPELOPSKOLIKU OAPWURFITIDLERG 17 VINCGUWAAL ORDFLAGESKOLIKU OAPWURFITIDDERG 17 VINCGUWAAL ORDFLAGESKOLIKU OAPWURFITIDDEGG 17 VINCGUWAAL OADSLOPSKOLIKU OAPWURFITIDDEGG 18 VINCGUMAAL OADSLOPSKOLIKU OAPWURFITIDDEGG 19 IINKATGUSLAVKAL GADALOPSKOLIKU OAPWURFITIDDEGG 19 IINKATGUSLAVKAL GADALOPSKOLIKU OAPWURFITIDDEGG 251 n YI1 t LOG Q gnW 221 LTYIYPGISKVT LGGTRCKADN N LSPDA EISKHILSKCCA 231 YNSPYIIPGGUAVT LGGTRCKADW N LSPDA EISKHILSKCCA 231 IYNSPYIIPGGUAVT LGGTRCHAM S ELNNI OOHNTIMSCCR 231 IYNSPYIIPGGUAVT LGGTRCHAM S ENST DOHNTIMKSCCK 231 IYNSPYIIPGGUAVT LGGIRCHAM S ENST DOHNTIMKSCCK 236 EDEMLXIMTRDOTS IIGGCROPHNNES EPOPS LTHRLSKALDEPE 301 Et 1k IV O GREE VELGET 1
DASPO b DAAO p DAAO r DAAO r DAAO r DAAO r DAAO r DAAO t DAAO t DAAO t DAAO p DAAO p DAAO p DAAO p DAAO r DAAO r DAAO r DAAO r DAAO t DAAO t DAAO h DAAO r DAAO r DAAO t DAAO r DAAO t DAAO t DAAO h DAAO r DAAO h DAAO t DAAO h DAAO r DAAO h DAAO t DAAO h DAAO h DAAO t DAAO h DAAO h	174 VUNCSCLESKOL AGDSKIFFVRCOULW OAPWUKNFITRD SSG 177 VUNCCTOWNACH ORDFILGPERGIIKV OAPWUKNFITRDERG 177 VUNCCTOWNACH ORDFILGPERGIIKV OAPWUKNFITRDERG 177 VUNCCTOWNACH ORDFILGPERGIIKV OAPWUKNFITRDERG 175 VINCCTOWNACH OADSLOTEGEN 180 IUWESGLAAFLGOVEDKUNTFIRGVULVENSLFPHARFST FXKEN 191 IUWESGLAAFLGOVEDKUNTFIRGVULVENSLFPHARFST FXKEN 191 IUWESGLAAFLGOVEDKUNTFIRGVULVENSLFPHARFST FXKEN 221 LTYIPGISWUT LGGTROKGDW N LEPDA EISKELLSRCCA 231 INSFIIIPGIGVUT LGGTROKGDW N LEPDA EISKELSRCCA 231 INSFIIIPGIGVUT LGGTROKGDW N LEPDA EISKELSRCCA 231 INSFIIIPGIGVUT LGGTROKGDW N EINNI OOHNTINKGCCS 231 INSFIIIPGIGVUT LGGTROKHM S ENNI OOHNTINKGCCS 231 INSFIIIPGIGVUT LGGIGULAHW S GLWSV ROHNTINKGCCS 231 INSFIIIPGKUT LGGIGULAHW S GLWSV ROHNTINKGCS 231 INSFIIIFGKUT LGGIGULAHWS S GLWSV ROHNTINKGCS 231 INSFIIIFGKUT LGGIGULAHWSK GUNSV ROHNTINKGCS 231 INSFIIIFGKUT LGGIGULAH S GUNSV ROHNTINKGCS 231 INSFIIIFGKUT LGGIGULAH S GUNSV ROHNTINKGCS 231 INSFIIIFGKUT LGGIGULAH S GUNSV ROHNTINKGCS 231 INSFIIIFGKUT
DASPO b DAAO p DAAO p DAAO n DAAO r DAAO r DAAO r DAAO r DAAO r DAAO f CONSENSUS DAAO p DAAO n DAAO r DAAO r DAAO r DAAO r DAAO h DAAO h DAAO h DAAO h DAAO f CONSENSUS CONSENSUS	174 VUNCSGLESKOL AGDSKIFFVRGOULW OAPWYNFI ED SSG 177 VUNCTGVWAAL GPDPILGPGRGOINK DAPHADFINIDLEAG 177 VUNCTGVWAAL GPDPILGPGRGOINK DAPHADFINIDLEAG 177 VUNCTGVWAAL GPDPILGPGRGOINK DAPHADFINIDDESG 188 VUNCTGVWAAL GADALGPGRGOINK DAPWYNFFIITDDESG 188 VUNCTGVWAAL GADALGPGRGOINK EAPWYNFFIITDDESG 190 IIVNATGLGSYKLGGVEKTMFRGQVUVW RNESFPHLTSGVECG 251 n YII t LGG Q gnW I 251 LTVIYPGISNVT LGGTRGKGDN N LSPDA EISKEILSRCCA 231 IYNSPYIPGGVYT LGGTRGKGDN N LSPDA EISKEILSRCCA 231 IYNSPYIPGGVYT LGGTRGKGDM S ELNNI QONNTIMEGCCR 231 IYNSPYIPGGVYT LGGTRGKGDM S ELNNI QONNTIMEGCCR 231 IYNSPYIPGGVYT LGGTRGKGDM S ELNNI QONNTIMEGCCR 231 IYNSPYIPGGVYT LGGTRGKAM S ENST DONNTIMKGCCS 231 GADALINHTRDGTS IIGGCCPMNNSS EPDPS LTHRILSRALDEF 301 E t k IV 0 G BP E VRL0 E 1 VIENTAR
DASPO b DAAG p DAAG p DAAG r DAAG R	174 VUNCSGLESBOL AGDSKIFFVRCQULW OAPWXRIFI RD SSG 177 VUNCGCWMACL GPDFLIGPCRQIIK OAPWXRIFI RDDERG 177 VUNCGCWMACL GPDFLIGPCRQIKV OAPWXRIFI RDDERG 177 VUNCGCWMACL GPDFLIGPCRQIKV DAPWXRIFI RDDERG 177 VUNCGCWACL GPDFLIGPCRQIVE DAPWXRIFI RDDERG 179 VINCGCWACL GPDFLIGPCRQIVE DAPWXRIFI RDDERG 189 VINCGCWACL GPDFLIGPCRQIVE DAPWXRIFI RDDERG 190 IVWATGLSYKLGOVEDKINFRGQVUV RNSSFFMAFSST FEXEN 201 IVWATGLSYKLGOVEDKINFRGQVUV RSSFFMAFSST FEXEN 201 IVWATGLSYKLGOVEDKINFRGQVVV RESSFFLIARCCA 201 IVMSPVIFGGVVV LGGTGRGKDN N LSPDA EISKILSRCCA 201 IVMSPVIFGGVVV LGGTGRGKDN SENST DDINTINKGCCR 201 IVMSPVIFGGVVV LGGTGRGKN SENST DDINTINKGCCR 203 IVMSPVIFGGVVV LGGTGRGKN SENST DDINTINKGCCR 203 IVMSPVIFGGVVV LGGTGRGKN SENST DDINTINKGCCR 204 UNTRODASGGTLGGTVDVGNNESGPDPN LNRINGRIVEVFEI 301 E t LK IV G BPR VELG B 1 VIENIGB G 251 LEFS LRGAVDLÆRVGLERVT GFPVF POVILERELAGDSRLPVNHVCHGG 266 LEFT LKINKIGETGFFVF POVILERELAGDSRLPVNHVCHGG
DASPO b DAAO p DAAO h DAAO h DAAO n DAAO t DAAO t DAAO t DAAO t DAAO p DAAO p DAAO p DAAO p DAAO p DAAO p DAAO t DAAO p DAAO t DAAO p DAAO p DAAO h DAAO p DAAO h DAAO h DAAO h DAAO h	179 VUNCTGUMAGUL AGDSKIFFVRGOULK QAPWVENFI RD SSG 177 VUNCTGUMAGUL GPDFLIGPGRGUIK DAPWLRHFITHDLERG 177 VUNCTGUMAGUL GPDFLIGPGRGUIK DAPWLRHFITHDLERG 177 VUNCTGUMAGUL QHDFLIGPGRGUIK DAPWLRHFITHDERG 175 VINCTGUMAGUL QHDSLIGFGRGUIK DAPWLRHFITHDERG 175 VINCTGUMAGUL QHDSLIGFGRGUIK DAPWLRHFITHDERG 175 VINCTGUMAGUL QHDSLIGFGRGUIK DAPWLRHFITHDERG 180 IUWAGUSHLGGVEDKTNFRGQUUV RHSSPHLTSGVEDG 251 n YII t LØG Q gnV I 251 n YII t LØG Q gnV I 251 n YII t LØG Q gnV I 251 LTYIPGISHVT LGGTGUGHN H EINH QOHNTINEGCCR 231 IVNSPIIDFGUAV LGGTGUGHN S ELANI QOHNTINEGCCR 231 GINSPIIDFGUAV LGGTGUAN S ELANI QOHNTINEGCCR 231 GINSPIIDFGUAV LGGTGUGHTGUAN S ELANI QOHNTINEGCCR 231 GINSPIIDFGUAV LGGTGUAN S ELANI QOHNTINEGCCR 231 GINSPIIDFGUAV LGGTGUGHN S ELANI QOHNTINEGCCR 232 GINSPIIDFGUAV LGGTGUGHTFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF
DASPO b DAAG p DAAG p DAAG r DAAG r DAAG r DAAG r DAAG r DAAG t DAAG t DAAG p DAAG p DAAG p DAAG n DAAG r DAAG r	179 VINCSCLESBOL AGDSKIFFVRCOULW OAPWYNFI ED SSC 177 VINCCTCWMACL GPDFLLGPCRCJIKV OAPWYNFI HTDPERG 177 VINCCTCWMACL GPDFLLGPCRCJIKV DAPWNNFI HTDPERG 177 VINCCTCWMACL GRDPLLGPCRCJIKV DAPWNNFI HTDPERG 179 VINCCTCWACL GADSLAPCRCJIVV DAPWNNFI HTDPERG 189 VINCCTCWACL GADSLAPCRCJIVV DAPWNNFI HTDPERG 190 HIVNTCLSYKLGVEDKTNFRQVIVNSLFYNNFIFTNEFI 201 HVNTCLSYKLGVEDKTNFRQVIVNSLFYNNFIFTNEFI 211 LTYIYPGISWYT LGGTRCKGDW N LSPDA EISKHILSRCCA 223 INNSPYIFPGCTVT LGGTGRCKDW N LSPDA EISKHILSRCCA 223 INNSPYIFPGCTVT LGGTGRCKDW N LSPDA EISKHILSRCCA 223 INNSPYIFPGCTVT LGGTGRCKDW N LSPDA EISKHILSRCCA 223 INNSPYIFPGCTVT LGGTGRCMW SENNT ODNTTIMSCCR 223 INNSPYIFPGCTVT LGGTGRCMW SENNT ODNTTIMSCCR 231 INNSPYIFPGCTVT LGGTGRCMW SENNT DONTTIMSCCR 231 INNSPYIFPGCTVT LGGIGGRCMW SENNT DONTTIMSCCR 231 INNSPYIFPGTTVT LGGIGGRCMW SENNT DONTTIMSCCR 231 INNSPYIFPGTTVT LGGIGGRCMW SENNT DONTTIMSCCR 231 INNSPYIFPGTTVT LGGIGGRCMWSCPDNIANTNOCKCG 232 LEBS LRGATOLERWGLRTT POVRLEELADSSRLPVNHYCHGG 241 LEBS LRGATOLERWGTRVTR POTRLEELGLARGSSTFUTINNYCHGG 244 LEFT LKMARTUGELTGFRVVT POINLEREDLARGSSTFUTINNYCHGG 245 LEDT LKMARTUGELTGFRVVT POINLEREDLARGSSTFUTINNYCHGG 244 LEFT LKMARTUGELTGFRVVTRY POINLEREDLARGSSTFUTINNYCHGG 245 LEDT LKMARTUGELTGFRVFTR POVELEELGLARGSSTFUTINNYCHGG 246 LEFT LKMARTUGELTGFRVFTR POYNLEREDLARGSSTFUTINNYCHGG 246 LEFT LKMARTUGELTGFRVFTR POYNLEREDLARGSSTFUTINNYCHGG 247 LKMARTUGELTGFRVFTR POYNLEREDLARGSSTFUTINNYCHGG 248 PELKARGCHUTGKGRGRGRGRGCHVCEKGRCRGCRGCRGCRGCHWYCHVCAL
DASPO b DAAC p DAAC p DAAC f DAAC T DAAC T DAAC T DAAC T DAAC T DAAC T DAAC P DAAC T DAAC T DAAC F DAAC T DAAC T DAAC T	177 VUINCTCVMACH GOPPILGPERGUINV OAPWIKHFI RD SSG 177 VUINCTCVMACH GPDFILGPERGUINV OAPWIKHFI THDLERG 177 VUINCTCVMACH GPDFILGPERGUINV OAPWIKHFI THDLERG 177 VUNCTCVMACH GPDFILGPERGUINV DAPWIKHFI THDLERG 175 VINCTCVMACH GADALGYEDKGUINV DAPWIKHFI THDLERG 175 VINCTCVMACH GADALGYEDKUTTRACULVNNSLPFMAFPST PXKN 180 IUWAGLSTHAFLGGVEDKUTTRACULVNNSLPFMAFPST PXKN 251 n YII t LØG Q gnW I 221 LIYIYPGISHVT LGGTRQKGDM N LSPDA EISKELLSRCCA 231 INSPIIIPGIGVT LGGTRQKGDM N LSPDA EISKELSRCCA 231 INSPIIIPGIGVT LGGTRQKGM N EINNI QOHNTIMEGCCR 231 INSPIIIPGIGVT LGGTRQKMM S ENNI QOHNTIMEGCCR 231 INSPIIIPGIGVT LGGTRQKMM S ENNI QOHNTIMEGCCR 231 INSPIIIPGGRVT LGGTRQKMM S ENNI QOHNTIMEGCCR 231 INSPIIIPGGRVT LGGTRQKMM S ENNI QOHNTIMEGCCR 231 GINAULHIMTRDGTS IIGGCUPUNNS EDDJ LITRILGALDERF 236 EDGLIHMTRDGENT LGGTRQKMM S ENNI QOHNTIMEGCCR 236 EDGLIHMTRDGENT LGGTRQKMM S ENNI QOHNTIMEGCCR 237 GADVMILAQRAAGGTILGCTIVUNNESGPDPNIAMRINGRIVEVEPI 301 E t k IV 0 G BP R VELG I VIENTG G
DASPO b DAAC p DAAC p DAAC r DAAC r DAAC r DAAC r DAAC f COMSENSUS DAAFO b DAAC r DAAC n DAAC r DAAC r	17 VUNCSGLESKOL AGOSKIFFVRGOULW OAPWUNFIF ED SS 17 VUNCGUWAAL GROPILGPGRGOINK DAPHUNFIFIDDERG 17 VUNCGUWAAL GROPILGPGRGOINK DAPHUNFIFIDDERG 18 VUNCGUWAAL GADALGERGOIGV APPUNFIFIDDERG 19 VINCGUWAAL GADALGERGOIGV APPUNFIFIDDERG 19 VINCGUWAAL GADALGERGOIGV APPUNFIFIDDERG 19 VINCGUMAAL GADALGERGOIGV APPUNFIFIDDERG 19 VINCGUMAAL GADALGERGOIGV APPUNFIFIDDERG 19 VINCGUMAAL GADALGERGUNGUW RNESFPLLTSGVEGG 251 n YII t LGG Q gnW I 21 LTVIYGISWYT LGGTRCKDW N LSPDA EISKHILSRCCA 23 IYNSPYIIPGGUVY LGGTRCKDW N LSPDA EISKHILSRCCA 23 IYNSPYIIPGGUAYT LGGTRCKDW N LSPDA EISKHILSRCCA 23 IYNSPYIIPGGUAYT LGGTRCKDW S ELNNI GOMNTIMSCCR 23 IYNSPYIIPGGUAYT LGGTRCHAW S ENST DOMNTIMSCCR 23 GADWYLMGRAAGGGILGGTYDVGNWS EPDPS LTHRLSRALDFF 23 GADWYLMGRAAGGGILGGTYDVGNWS EPDPS LTHRLSRALDFF 23 LEPS LRGAYDLREWGLRPT POYRLEKELAGDSRELPVVHYCHGG 26 LEFT LRGAHIGETGRPVR POYRLEKELAGDSRELPVVHYCHGG 26 LEFT LRGAHIGETGRPVR POYRLEBELAGDSRELPVVHYCHGG 26 LEFT LRGAHIVGETGRPVR POYRLEBELAGDSRELPVVHYCHGG 26 LEFT LRGAHIVGETGRPVR POYRLEBELAGDSRELPVVHYCHGG 26 LEFT LRGAHIVGETGRPVR POYRLEBELGGSMFTVHYCHGG 26 LEFT LRGAHIVGETGRPVR POYRLEBELGAGSSFTVHVHYCHGG 26 LEFT LRGAHIVGETGRPVR POYRLEBELGAGSSFTVHVHYCHGG 26 LEFT LRGAHIVGETGRPVR POYRLEBELFULAGSSFTVHYLHVCHGG 26 LEFT LRGAHIVGETGRPVR POYRLEBELFULAGSSFTVHVHYCHGG 26 LEFT LRGAHIVGETGRPVR POYRLEBELFULAGSSFTVHYLHVCHGG 26 LEFT LRGAHIVGETGRPVR POYRLEBELFULAGSSFTVHYLHVCHGG 26 LEFT LRGAHIVGETGRPVR POYRLEBELFULAGSSFTVHYLHVCHGG 26 LEFT LRGAHIVGETGRPVR POYRLEBELFULAGSSFTVHVHYCHGAG 26 LEFT LRGAHIVGETGRPVR POYRLEBELFULAGSSFTVHYLHVCHGAG 26 LEFT LRGAHIVGETGRPVR POYRLEBELFULAGSSFTVHYLHVCHGAG 26 LEFT LRGAHIVGETGRPVR POYRLEBELFULAGSSFTVHYLHVCHGAG 26 LEFT LRGAHIVGETGRPVR
DASPO b DAAO p DAAO r DAAO p DAAO p DAAO n DAAO r DAAO r	17 VUNCSCLESKOL AGDSKIFFVRCQUINV OAPWURKFIRD SSG 17 VUNCCGVWACHL GPDFLIGHCRQIINV OAPWURKFIITHDLERG 17 VUNCCGVWACHL GPDFLIGHCRQIINV OAPWURKFIITHDLERG 17 VUNCCGVWACHL GPDFLIGHCRQINV OAPWURKFIITHDPERG 19 VINCCGVWACHL GPDFLIGHCRQINV OAPWURKFIITHDPERG 19 VINCCGVWACHL GDDSLIGHCRQINV OAPWURKFIITHDPERG 19 VINCCGVWACHL GDDSLIGHCRQIVV RHSSFRST FXKM 20 INNTGLIGHTLGGVECKMURHARGQUVV RHSSFRST FXKM 21 IVNTGLIGHTLGGVECKMURHARGQUVV RHSSFRST FXKM 221 IVNFGLIGHT LGGTGVECKMUR SLIPHAG 221 IVNFGLIGHT LGGTGVECKMUR SLIPHAG 221 IVNFGLIGHT LGGTGVECHM S ELMIN ODHTINGCCR 221 IVNFFVILHQGVAVT LGGTGVECHM S ELMIN ODHTINGCCR 221 IVNFFVILHQGVAVT LGGTGVECHM S ELMIN ODHTINGCCR 231 IVNFFVILHQGVAVT LGGTGVECHM S ELMIN ODHTINGCCR 232 IVNFFVILHQGVAVT LGGTGVECHM S ELMIN ODHTINGCCR 233 INSFVILHQUAAY LGGTGVECHM S ELMIN ODHTINGCCR 234 INSFVILHQGVAVT LGGTGVELM S ELMST DUNTINGCCR 235 INSFVILHQUAAY LGGTGVECHM S ELMIN ODHTINGCCR 236 ELMININGTDGVECKURSFPT PCVELEELLADDSRELPVNHYCHGG 236 LEFF LMDAKICGUTGFFVF PCVELEELLADDSRESPFVIHYCHGG 236 LEFF LMDAKICGUTGFVFF PCVELEELLADDSRESPFVIHYCHGG 237 ANGKCVKGLSVIRHAVGHRPMFCGVFIEELL DDETWIVHYCHGG 239 ANGKCVKGLSVIRHAVGHRPMFCGVFIEELL DDETWIVHYCHGG
DASPO b DAAO h DAAO t DAAO t DAAO t DAAO t DAAO t DAAO t DAAO t DAAO h DAAO p DAAO h DAAO t DAAO t DAAO t DAAO t DAAO t DAAO t DAAO t DAAO t DAAO t DAAO t	1 UVURCSULSKUL AGOSKIFFURGOULW OAPWURFI ED SG 17 UVURCSUMAAL ORDFILGESGIIN DEPHINIPTINDEG 17 UVURCSUMAAL ORDFILGESGIIN DEPHINIPTINDEG 19 UVURSUMAAL OADSLOEGSGIIN DEPWINIPTINDEG 19 UVURSUMAAL OADSLOEGSGIIN DEPWINIPTINDEG 19 UVURSULSVALGUEDKUMPTINGVULWASHEPHINISTIDEG 10 UVURSULSVALGUEDKUMPTINGVULWASHEPHINISTIDEG 21 UVURSULSVALGUEDKUMPTINGVULWASHEPHINISTIDEG 22 UVURSULSVALGUEDKUMPTINGVULWASHEPHINISTIG 23 UVURSULSVALGUEDKUMPTINGVULWASHEPHINISTIG 23 UVURSULSVALGUEDKUMPTINGVULWASHEPHINISTIG 23 UVURSULSVALGUEDKUMPTINGVULWASHEPHINISTIG 23 UVURSULSVALGUEDKUMPTINGVULWASHEPHINISTIG 23 UVURSULSVALGUEDKUMPTINGVULWASHEPHINISTIG 23 UVURSULSVALGUEDKUMPTINGVULWASHEPHINISTIG 23 UVURSULSVALGUEDKUMPTINGVULWASHEPHINISTIG 23 UVURSULSVALGUEDKUMPTINGVULWASHEPHINISTIG 24 UVURSULSVALGUEDKUMPTINGVULWASHEPHINISTIG 24 UVURSULSVALGUEDKUMPTINGVULWASHEPHINISTIG 25 UVURSULSVALGUEDVENTVINISTIG 26 UPUT LUMARICUSTURGEPHINISSE SEDIS UTURTUNKIS 26 LEPS LEGAVIDHERVGUEPHI POVILEEBOLGEGSSKATUINISTIG 26 LEPS LEGAVIDHERVGUEPHI POVILEEBOLGUEGSSKATUINISTIG 26 LEPS LEGAVIDHERVGUEPHI POVILEEBOLGUEGSSKATUINISTIG 26 LEPT LUMARICUSTURGEPHIN POVILEEBOLGUEGSSKATUINISTIG 26 LEPT LUMARICUSTURGEPHIN POVILEEBOLGUEGSSKATUINISTIG 26 LEPT LUMARICUSTURGEPHIN POVILEEBOLGUESSSKATUINISTIG 26 LEPT LUMARICUSTURGEPHIN POVILEEBOLGUESSSKATUINISTIG 27 MORGUESSKATURGAN PONCOURDEPHINTINGEG 28 PELTROCHLISTIGATUREPHINTINGEG 29 PELTROCHLISTIGATURGEPH
DASPO b DAAG p DAAG p DAAG r DAAG r DAAG r DAAG r DAAG r DAAG r DAAG r DAAG r DAAG r DAAG p DAAG p DAAG p DAAG r DAAG R D	17 VUNCSCLESEL AGOSKIFFVRCOULTV OAPWYLFI ED SSG 17 VUNCTSUMAAL GEDPILGEGEGIIKV DAPWIGHTITHDERG 17 VUNCTSUMAAL GEDPILGEGEGIIKV DAPWIGHTITHDERG 19 VINCTSUMAAL GADALGEGEGIIV DAPWIGHTITHDERG 19 VINCTSUMAAL GADALGEGEGIIV DAPWIGHTITHDERG 19 VINCTSUMAAL GADALGEGEGIIV DAPWIGHTITHDERG 19 VINCTSUMAAL GADALGEGEGIIV DAPWIGHTITHDERG 20 VINCTSUMAAL GADALGEGEGIIV DAPWIGHTITHDERG 21 VINSSIFARTUSCEEKHTIGUUVUNNIFFANTIST PERD 221 LTVINGTUNATUSCEEKHTIGUUVUNNIFFANTIST PERD 221 LTVINGTUNATUSCEEKHTIGUUVUNNIFFANTIST PERD 221 LTVINGTUNATUSCEEKHTIGUUVUNNIFFANTIST PERD 221 VINSFYILFGGUVU LGGTERGEGUN LESPALETSKEEKCA 223 VINSFYILFGGUVU LGGTERGEGUN SENST DOINTINGSCCR 223 VINSFYILFGGUVU LGGTERGENN SENST DOINTINGSCCR 231 VINSFYILFGGUVU LGGTERGENN SENST DOINTINGSCCR 232 INSFYILFGGUVU LGGTERGENN SENST DOINTINGSCCR 233 CADUNYLMGRAGGGILGGTUVUNNESOPPNIARINGRIVEVRPEI 301 E L & V O B F VLO I VINTGSYNENG 234 ELES LRGAVDLERWGLERTF PEVRLEELLAGDSREIPVNHYCHGG 245 LEFS LRGAVDLERGEGERFF PEVRLESGSMFEVINNYCHGG 246 LEFT LXMARIGEGERFFFF PEVRLEELLAGDSREIPVNHYCHGG 246 LEFT LXMARIGEGERFFFF PEVRLEELLAGDSREIPVNHYCHGG 246 LEFT LXMARIGEGERFFFF PEVRLEELLAGDSREIPVNHYCHGG 247 ANGKYCLSVIRHARVGERFFF PEVRLEELLAGDSREIPVNHYCHGG 248 LEDALVINGTGFFFFF PEVRLEELLAGDSREIPVNHYCHGG 249 ANGKYCLSVIRHAVGERFFFF PEVRLEELLAGDSREIPVNHYCHGG 240 LEFT LXMARIVGELGEFFFFF PEVRLEELLAGSSFFFFUNKYCHGG 241 EDALVINGEGENFFFFFF PEVRLEELLAGDSREIPVNHYCHGG 242 ANGKYCLSVIRHAVGERFFFFF PEVRLEELL DDEFNIVNCHGG 243 PENDERGENFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF
DASPO b DAAO P DAAO T DAAO T DAAO T DAAO T DAAO T DAAO T DAAO T DAAO T DAAO P DAAO P DAAO T DAAO T	17 VUVNCGCVMAGAL GOPPILGPCRQUINV QAPWIKHIN RD SSG 177 VUVNCGCVMAGAL QPDFLIGPCRQUINV QAPWIKHIN RD SSG 177 VUVNCGCVMAGAL QPDFLIGPCRQUINV DAPWIKHINTINDERG 177 VUNCGCVMAGAL QPDFLIGPCRQUINV DAPWIKHINTINDERG 175 VINCGCVMAGL 175 VINCGCVMAGL 176 VINCGCVMAGL 176 VINCGVMAGL 177 VINCGCVMAGL 178 VINCGVMAGL 178 VINCGVMAGL 180 IVWAGLSVIKLOVENTPIRGVUVNNSLPMAGFST PXKN 180 IVWAGLSVIKLOVENTPIRGVUVNNSLPMAGFST PXKN 180 IVWAGLSVIKLOVENTPIRGVUVNNSLPMAGFST PXKN 181 VINGGUSULGVENTMARAGUIVV RNSSFNLISGVEGG 251 n XII t LØG Q gnV I 251 LIYIPGISNVT LGGTRQKGDM N LEPDA EISKELLSRCCA 231 INSPIIDFLGVT LGGTRQKGDM N LEPDA EISKELLSRCCA 233 INSPIIDFLGVT LGGTRQKGM N EGNST DDINTINKGCCS 231 INSPIIDFLGVT LGGTRQKMM S ENNI QOMNTINGCCS 231 INSPIIDFLGVT LGGTRQKMM S ENNI QOMNTINGCCS 232 GINAGUINTRYDGTST LGGIGLGHM S GLMSV RDINTINGCCS 233 GINAGUINTRYDGTST LGGIGLGHM S GLMSV RDINTINGCCS 234 INSPIIDFLGVT LGGIGLGHM S GLMSV RDINTINGCCS 235 INSPIIDFLGVT LGGIGLGHM S GLMSV RDINTINGCCS 236 EDENLINTRYDGTGTT LGGIGLGHM S GLMSV RDINTINGCCS 237 GADWYLAQRAAGGTILGGTTVVCHMESQPDPNIAMRINGRIVEVRPEI 301 E t k V O G BP K VELG T VIENTG 241 LEPS LRGAYDLRERVGLRPTR PCVRLENELLAQDSRLFVVIHVGHGG 244 LEPT LKMARIGERGTGRVR PQIRLENELLADDSRLFVVIHVGHGG 245 LEPS LRGAYDLRERVGLRPTR PCVRLENELLADDSRLFVVIHVGHGG 246 LEPT LKMARIGERGRENGERVR POJRLENGELAGESSAVEVIHVGHGG 247 ANGCKRELSVIHAVGRENGRENGERVRLEK KIG VUPVINYGHGG 351 G GAAE 1 1 309 GGIAMMMTALEATRL VNECYGRRAGESSAVELEK KIG VUPVINYGHGG 351 G GAAE 1 1 309 GGIAMMMTALEATRL VNECYGRRAGESSAVELEK KIG VUPVINYGHGG 351 G GAAE 1 1
DASPO b DAAG p DAAG p DAAG r DAAG r DAAG r DAAG r DAAG r DAAG r DAAG t DAAG p DAAG p DAAG p DAAG n DAAG r DAAG R D	17 VUNCSCLESKOL AGDSKIFFVRCOULW OAPWUNFIF ED SS 17 VUNCTSUMAAL GRDFLAGESCIIK DAFWINFIFTED SS 17 VUNCTSUMAAL GRDFLAGESCIIK DAFWINFIFTEDBEG 17 VUNCTSUMAAL GRDFLAGESCIIK DAFWINFIFTEDBEG 18 VUNCTSUMAAL GADASLGESCIIK DAFWINFIFTEDBEG 19 IIVNTSUSGLEAKTAGUEENGEN FLEGVUNNSLEFMAAFLEDBEG 19 IIVNTSUSGLEAKTAGUEENGEN FLEGVUNNSLEFMAAFLEDBEG 251 n YII t LOG Q gnf I 221 LTVIPGISWUT LGGTERGREDN N LSPDA EIEKEILSRCCA 233 INNSPYIFTEGUT LGGTERGREDN N LSPDA EIEKEILSRCCA 233 INNSPYIFTEGUT LGGTERGREDN N LSPDA EIEKEILSRCCA 233 INNSPYIFTEGUT LGGTERGREDN N EINNI ODENTIMESCCR 233 INNSPYIFTEGUT LGGTERGREDN S EINNS DENTIMESCCR 233 INNSPYIFTEGUT LGGTERGREDN S ENDS DENTIMESCCR 234 INNSPYIFTEGUT LGGTERGREDN S LENNI ODENTIMESCCR 235 INNSPYIFTEGUT LGGTERGREDN S LENNI ODENTIMESCCR 236 EDEALYIMTEDGTS IIGGCTERNNES EPDPS LTHERLISRCCA 236 EDEALYIMTEDGTS IIGGCTERNNES EPDPS LTHERLISRCCA 236 EDEALYIMTEDTS IIGGCTERNNESCPOPNIARTHORIVEVERFEI 301 E t LK IV • G BP R VEL B 1 VIENTER 236 LEPS LREAVOLAERWULKTERTP FOYRLEELLAGDSRLPVNHVCHGG 246 LEPT LKNAKIGELGTERVER PURIEEDLARGESHTEVINNCHGG 246 LEPT LKNAKIGELGTERVER PURIEEDLARGESHTEVINNCHGG 246 LEPT LKNAKIGELGTERVER PURIEELLAGDSRLPVNHVCHGG 247 ANGEGVELSVIRGENGRERTR POYRLEELLAGDSRLPVNHVCHGG 248 ANGEGVELSVIRGENGRERTR POYRLEELGLARGESHTEVINNCHGG 249 ANGEGVELSVIRGENGRERTR POYRLEELSLAGSSTEVINVCHGG 240 LEPT LKNAKIVGELGERFER PURIEELLAGDSRLPVNHVCHGG 241 EDALVIVELGERFERFER PURIEELLAGDSRLPVNHVCHGG 243 PARTICELTGREVER PURIEELSLAGSSTEVINVCHGG 244 LEPT LKNAKIVGELGERFER PURIEELLAGDSRLPSTEVINVCHGG 245 ANGEGVELSVIRHAVGERFER PURIEELLAGDSRLPSTEVINVCHGG 246 LEPT LKNAKIVGELGERFER PURIEELSLAGSSTEVINVCHGG 247 ANGEGVELSVIRHAVGERFER PURIEELLAGDSREPSTEVINVCHGG 248 PLANKERVELKGERFER PURIEELLAGDSREPSTEVINVCHGG 249 ANGEGVELSVIRHAVGERFER PURIEELLAGDSREPSTEVINVCHGAG 240 ANGEVELSVIRHAVGERFER PURIEELLAGTERFER PORL
DASPO b DAAO P DAAO P DAAO T DAAO T DAAO T DAAO T DAAO T DAAO T DAAO T DAAO T DAAO P DAAO P DAAO T DAAO T	177 VUINCGCWAAXL QOPPLLQPCRQJIKV QAPWIKHIIRD SSG 177 VUINCGCWAAXL QOPPLLQPCRQJIKV QAPWIKHIIRDLERG 177 VUINCGCWAAXL QOPPLLQPCRQJIKV DAPWIKHIIRDLERG 177 VUNCGCWAAXL QOPPLLQPCRQJIKV DAPWIKHIIRDPERG 179 VINCGCWAAXL QOPPLLQPCRQJIKV DAPWIKHIIRDPERG 179 VINCGCWAAXL QOPPLLQPCRQJIV DAPWIKHIIRDPERG 180 IVWCGLAVLAGUVEKNUPARGUIVV RHSSPKLISPCA 251 n YII t LOG Q gnV I 221 LIYINGISWVT LGGTQVCNW N EINII QOHNTINGCCC 233 INSPIIIPGLQVT LGGTQVCNW N EINII QOHNTINGCCC 233 INSPIIIPGLQVT LGGTQUCNW S ENNI QOHNTINGCCC 233 INSPIIIPGLAVT LGGTQUCNW S ELMII QOHNTINGCCC 234 INSPIIIPGLAVT LGGTQUCNW S ELMII QOHNTINGCCC 235 INSPIIIPGRAVT LGGTQUCNW S ELMII QOHNTINGCCC 236 INSPIIIPGRAVT LGGTQUCNW S ELMII QOHNTINGCCC 237 INSPIIIPGRAVT LGGTQUCNW S ELMII QOHNTINGCCC 238 INSPIIIPGRAVT LGGTQUCNW S ELMII QOHNTINGCCC 239 INSPIIIPGRAVT LGGTQUCNW S ELMII QOHNTINGCCC 230 INSPIIIPGRAVT LGGTQUCNW S ELMII QOHNTINGCCC 231 INSPIIIPGRAVT LGGTQUCNW S ELMII QOHNTINGCCC 232 INSPIIIPGRAVT LGGTQUCNW S ELMII QOHNTINGCCC 233 INSPIIIPGRAVT LGGTQUCNW S ELMII QOHNTINGCCC 234 ENDIINTRYDTT LGGTQUCNW S ELMII QOHNTINGCCC 235 INSPIIIPGRAVT LGGTQUCNW S ELMII QOHNTINGCCC 236 EDALIHTRYDTPT IGGCTUPYNNSS EDDDJ LITRILGAULDPR 237 GADVMILAQUAAGGTILGGTTUVGNMESGPOPNIAARINGNIVENPIG 236 ELEFT LGAARINGTPTPY POYLLEEDLARGSSMTVINNYCHGG 237 ANGGVRELAVINAVGRAWKRCAREGEPKVELKING VUNYCHGG 239 ANGGVRELAVINAVGRAWKRCAREGEPKVELKING OFFFFFFININYCHGG 230 GIANNGVALAVGRAWKRCAREGEPKVELKING VUNYCHGG 231 G GALL 1 1 300 GIANNGTALEYRL VNEY QULLERALARGSSALVINNYCHGG 232 ANGGVRELAVINAVGRAWKRCAREGEPKVELKING OFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF
DASPO b DAAC p DAAC p DAAC p DAAC r DAAC r DAAC r DAAC r DAAC t DAAC f CONSENSUS DAAC p DAAC h DAAC r DAAC r	17 VUVNCSGLESKOL AGOSKIFFVRGOULW OAPWURFIE DD SG 17 VUINCGUWAAL GROPILAGESGIIK DAPHADENIITIDLEGG 17 VUVNCGUWAAL GROPILAGESGIIK DAPHADENIITIDDEGG 13 VUNCGUWAAL GROPILAGESGIIK DAPHUTEFILTDDEGG 13 VUNCGUWAAL GADALGESGIIK DAPWURFFILTDDEGG 13 VINCGUWAAL GADALGESGIIK DAPWURFFILTDDEGG 13 VINCGUWAAL GADALGESGIIK DAPWURFFILTDDEGG 13 VINCGUMAAL GADALGESGIIK DAPWURFFILTDDEGG 14 VINCGUMAAL GADALGESGIIK DAPWURFFILTDDEGG 15 VINCGUMAAL GADALGESGIIK DAPWURFFILTDEGG 15 VINCGUMAAL GADALGESGIIK DAPWURFFILTDEGG 25 VINCGUPARLAGGUENGNFILGGUVUNN KEPTALSTA 26 VINCGUPARLAGUENGNFILGGUVUNN KEPTALSTA 27 VINSTULSGUVUN LGGTEGKADN N LSPDA EISKHILSKCCA 23 INNSPYILFGUAAT LGGTEGKADN N LSPDA EISKHILSKCCA 23 INNSPYILGGUAAT LGGTEGKADN N LSPDA EISKHILSKCCA 23 INNSPYILGGUAAT LGGTEGKADN S ELNNI GOUNTIMSCCK 23 INNSPYILGGUAAT LGGTEGKADN S ELNNI GOUNTIMSCCK 23 GADWULMGRAAGGGILGGTUPUNNS EEDPS LTHRILSKALDEF 23 GADWULMGRAAGGGILGGTUPUNNSS EPDPS LTHRILSKALDEF 24 LEDS LRGAVDLREWULRTR FORLEKELAGDSKRIPVINYCHGG 26 LEFT LADALVGUENGTRUK PEVELEN VINCHGGSSMTUINVCHGG 26 LEFT LADALVGUENGTRUK POIRLEBGLRGSSMTUNVCHGG 26 LEFT LADALVGUENGTRUK POIRLEST SENDEF 37 O GALMINGTALEARTEL VNCCV QUENTAPK 37 ANGLOVGLSVINAVGUENGTRUK POIRLEBEL DDEFWINVCHGG 38 PEITAGEDUENGENGUENGENGENE POFIL 31 VULTIMGCALEARALFUGT LLEEK KISRUP PEHL 32 VULTIMGCALEARALFUGT LLEEK KISRUP PEHL 32 VULTIMGCALEARALFUGT LLEE

FIG. 10. Comparison of the primary structure of D-aspartate oxidase with that of D-amino acid oxidase from various sources. DASPO b, bovine D-aspartate oxidase; DAAO p, pig Damino acid oxidase (5); DAAO h, human D-amino acid oxidase (6); DAAO r, rabbit D-amino acid oxidase (7); DAAO m, mouse D-amino acid oxidase (7); DAAO t, T. variabilis D-amino acid oxidase (8); DAAO f, F. solani D-amino acid oxidase (9). The consensus sequence was determined, considering the residues that were identical in all seven sequences (bold upper case), identical or conservatively replaced in all seven sequences (upper case), and identical or conservatively replaced in at least six out of seven sequences (lower case). Asterisks mark the active site residues of mammalian D-amino acid oxidase proposed on the basis of chemical modification and site-directed mutagenesis studies (7, 11).

that the two families evolved toward a common chemical mechanism of catalysis through convergent evolution.

No extended sequence similarities were found between Daspartate oxidase, L-aspartate oxidase from Escherichia coli, and L-amino acid oxidase from Neurospora crassa. The latter two enzymes are FAD-containing oxidases constituted by 540 and 566 amino acids, respectively (17, 18). The different masses and the absence of sequence similarities between the oxidases active on D- or L-amino acids, which result in the completely different substrate specificity exerted by the two groups of enzymes, probably derive from the need to mantain a strict control on the entrance of D- or L-amino acids in different metabolic pathways. The growing number of information concerning the presence and the possible biological role of D-amino acids will permit a conclusion whether Daspartate oxidase and p-amino acid oxidase are still to be considered "detoxifying enzymes" (19) or are involved in more complex metabolic pathways.

REFERENCES

- 1. Still, J. L., Buell, M. V., Knox, W. E., and Green, D. E. (1949) J. Biol. Chem. 179, 831-837
- 2. D'Aniello, A., and Rocca, E. (1972) Comp. Biochem. Physiol. B Comp. Biochem. 41B, 625-633
- 3. Negri, A., Massey, V., and Williams, C. H., Jr. (1987) J. Biol. Chem. 262, 10026-10034
- 4. Negri, A., Massey, V., Williams, C. H., Jr., and Schopfer, L. M. (1988) J. Biol. Chem. 263, 13557-13563
- 5. Ronchi, S., Minchiotti, L., Galliano, M., Curti, B., Swenson, R. P., Williams, C. H., Jr., and Massey, V. (1982) J. Biol. Chem. 257, 8824-8834
- 6. Momoi, K., Fukui, K., Watanabe, Y., and Miyake, Y. (1988) FEBS Lett. 238, 180-184
- 7. Miyake, Y., Fukui, K., Momoi, K., Watanabe, F., Tada, M., Miyano, M., and Takahashi, S. (1991) in Flavins and Flavoproteins 1990 (Curti, C., Ronchi, S., and Zanetti, G., eds) pp. 135-142, Walter de Gruyter & Co., Berlin
- 8. Komatsu, K., Matsuda, A., and Sugiura, K. (1988) Chem. Abstr. 108, 49619s
- 9. Isogai, T., Ono, H., Ishitani, Y., Kojo, H., Ueda, Y., and Kohsaka, M. (1990) J. Biochem. 108, 1063-1069
- 10. Gould, S. J., Keller, G. A., Hosken, N., Wilkinson, J., and Subramani, S. (1989) J. Cell Biol. 108, 1657-1664
- 11. Curti, B., Ronchi, S., and Pilone Simonetta, M. (1991) in Chemistry and Biochemistry of Flavoenzymes (Muller, F., ed) pp. 69-94, CRC Press Inc., Boca Raton, FL
- 12. Wierenga, R. K., Terpstra, P., and Hol, W. G. J. (1986) J. Mol. Biol. 187, 101-107
- 13. Ghisla, S., and Massey, V. (1989) Eur. J. Biochem. 181, 1-17 14. Geigel, D., Williams, C. H., Jr., and Massey, V. (1990) J. Biol.
- Chem. 265, 6626-6632 15. Cederlund, E., Linqvist, Y., Soderlund, G., Branden, C.-I., and
- Jornvall, H. (1988) Eur. J. Biochem. 173, 523-530 16. Lederer, F., Cortial, S., Becam, A. M., Haumont, P. Y., and Perez,
- L. (1985) Eur. J. Biochem. 139, 59-65 17. Flachman, R., Kunz, N., Seifert, J., Gutlich, M., Wientjes, F.-J., Laufer, A., and Gassen, H. G. (1988) Eur. J. Biochem. 175,
- 221 22818. Niedermann, D. M., and Lerch, K. (1990) J. Biol. Chem. 265, 17246-17251
- 19. Jakoby, W. B., and Ziegler, D. M. (1990) J. Biol. Chem. 265, 20715-20718
- 20. Fujiwara, M., Ishida, Y., Nimura, N., Toyama, A., and Kinoshita, T. (1987) Anal. Biochem. 166, 72-78
- 21. Ellman, G. L. (1959) Arch. Biochem. Biophys. 82, 70-77
- 22. Swenson, R. P., Williams, C. H., Jr., Massey, V., Ronchi, S., Minchiotti, L., Galliano, M., and Curti, B. (1982) J. Biol. Chem. 257, 8817-8823
- 23. Fontana, A., and Gross, E. (1986) in Practical Protein Chemistry: A Handbook (Dabre, A., ed) pp. 67-118, John Wiley & Sons, New York
- 24. Matsudaira, P. (1987) J. Biol. Chem. 262, 10035-10038

Continued on next page.

Primary Structure of D-Aspartate Oxidase

Supplemental Material to

THE PRIMARY STRUCTURE OF THE FLAVOPROTEIN D-ASPARTATE OXIDASE

FROM BEEF KIDNEY

Armando Negri, Fabrizio Ceciliani, Gabriella Tedeschi, Tatjana Simonic and Severino Ronchi

EXPERIMENTAL PROCEDURE

Materiais. Beef kidney D-aspartate oxidase was purified as described in (3). TFA, 6 N HCl, HFBA, CNBr and GnCl were from Pierce Chemical Company. TPCK-trypsin, *Staphylococcus aureus* protease and carboxypeptidase B were obtained from Boehringer Marnheim. Iodo/2¹⁴⁷/2 (acetic acid was from Amersham. All the other reagents used were of the highest purity commercially available.

Amino acid analysis. Gas-phase hydrolysis was carried out in 6 N HCl containing 1 % (v/v) phenol at 105 °C for 24 h. Amino acid analyses were carried out by a post-column o-phtalhaldeyde derivatization procedure (20), using a Jasco (Japan Spectroscopy) amino acid analyzer equipped with a 880-PU pump and a 820-FP detector. Cysteine content was determined either as carboxymethylosetine after reduction and carboxymethylation of the protein or by titration with DTNB under denaturing non riducing conditions in 50 mM KPi, 6 M GnCl, pH 7.4 (21).

<u>Carboxymethylation</u>. Reduction and carboxymethylation of the protein were performed as described in (22). The material to be subjected to tryptic digestion was derivatized using iodo[2-¹⁴C]acetic acid 1 μ Ci / μ mol.

Enzymatic cleavages. Tryptic digestion was carried out in 0.1 % NH4HCO3, pH 8.0 at 37°C for 4 h using a protease : carboxymethylated D-aspartate oxidase ratio of 1: 100. A second aliquot of protease was then added and the digestion was allowed to proceed for an additional hour under the same conditions. The reaction was stopped by lowering the pH below 20 with TFA followed by lopphylization. The material was then resuspended in 500 $_{-2}$ l of 0.1 % TFA and centrifuged for 8 min using an Eppendorf 5414 centrifuge. The pellet was washed with 100 $_{-1}$ l of 0.1 % TFA and centrifuged for 8 min using an Eppendorf 5414 centrifuge. The pellet was washed with 100 $_{-1}$ l of 0.1 % TFA and centrifuged is above. The supernatants from the two centrifugations were pooled and the peptides (0.1 % TFA insoluble periodics) chromatographed on a RP-HPLC column as described below. The pellet (0.1 % TFA insoluble material) was dissolved in 400 $_{-1}$ l of 30 % CH5COOH, centrifuged and separated on a second RP-HPLC chromatography. No insoluble material was present after solubilization with 30 % CH5COOH.

Digestion with S. aureus protease was carried out for 24 h at 37 °C using a protease : carboxymethylated D-aspartate oxidase ratio of 1 : 25 (w /w), in 50 mM ammonium acetate, pH 4.0.

Carboxypeptidase B digestion was conducted both on the carboxymethylated enzyme (using a protease : substrate ratio of 1:25 (ww) in 0.2 M N-ethylmorpholine acetate, 1 M urea, pH 8.5. Aliquots of the incubation mixture were withdrawn at different times, the proteins precipitated with 10 % trichloroacetic acid and centrifuged, and the supernatant analyzed for amino acid content.

Chemical cleavages. Cleavage at methionyl residues was accomplished by dissolving 0.5 mg of carboxymethylated protein in 70 % formic acid and adding a 100-fold molar excess of CNBr over methionyl residues. After 24 h at room temperature in the dark the reaction was stopped by a 10-fold dilution with water followed by liophylization.

Further cleavage of the CNBr peptides at tryptophanyl residue was carried out by dissolving the peptides in 2 ml of 70 % formic acid: HFBA 1: 1 and adding a 10,000-fold molar excess of solid CNBr over tryptophanyl residues (23). After 24 h at room temperature the mixture was dried with a stream of nitrogen in hood, dissolved in 10 ml of water and liophylized.

<u>Reverse-phase HPLC chromatography</u>. Purification of peptides was carried out by RP-HPLC using a JASCO (Japan Spectroscopy) HPLC equipped with a n Uvidec-100 V detector set at 220 nm and an Aquapore C-8 RP-300 (4.6 x 25 nm) column (Applied Biosystems). The chromatographic conditions used are indicated in the Figure legenda. Aliquots from each fraction were withdrawn and analyzed for amino acid content. [¹⁴C] incorporation and amino acid sequence.

[¹⁴C] incorporation measurement [¹⁴C] incorporation was measured by dissolving aliquots of [¹⁴C]-carboxymethylated enzyme, HPLC- purified peptides or PTH-amino acid derivatives originated during the Edman degradation and recovered from the sequencer in Instagel (Packard) and counting in a Packard Minaxi Tricarb liquid scintilitation counter.

<u>Sequence analysis</u>. Sequence analyses were carried out on an Applied Biosystems Model 477-A pulse liquidphase sequencer equipped with a Model 120-A PTH-analyzer according to the manufecturer's instructions. Identification of PTH-[$^{+}$ [$^{+}$ C] incorporation measurements. Blotting of the protein on PVDF membrane prior to sequencing was performed according to (24).

Computer analysis. A search for similarities between the primary structure of D-aspartate oxidase with the one of other proteins was performed using the Swiss-Prot Protein Database.

RESULTS

The N-terminal residue of D-aspartate oxidase is blocked, as indicated by the fact that no sequence informations were obtained by submitting either the native, carboxymethylated or PVDF- membrane blotted protein to Edman degradation. Thus, the determination of the armino acid sequence of D-aspartate oxidase was accomplished through the characterization of all the peptides obtained by tryptic digestion of the carboxymethylated protein. Alignment of the tryptic peptides was subsequentally achieved by submitting the carboxymethylated protein to different enzymatic or chemical fragmentations.

lated protein to dimercial enzymatic or chemise insegmentations. Torptic digest. Tryptic digestion of 3 mg D-aspartate oxidase gave two sets of peptides: (a) soluble in 0.1% TFA and (b) insoluble in 0.1% TFA. The two sets were separated by centrifugation as indicated under "Experimental Procedure". Peptides contained in set (a) were separated directly by RP-HPLC (Fig. 1). The peptides co-eluted in the fractions with retention time of 32, 35 and 43 min were further purptied. The separate directly by RP-HPLC (Fig. 1). The peptides co-eluted in the separated by RP-HPLC under different conditions (Figs. 2, 3 and 4). The 0.1% TFA insoluble material was first solubilized in 30% CH5COVH and then separated by RP-HPLC (Fig. 1). The rotions from each chromatography were analyzed for amino acid composition, 1⁺C1-incorporation and amino acid sequence. Tables I and II and Figure 9 report the results of such analyses for all the tryptic peptides characterized. Since the digestion conditions (Figs. 2) and the separate dimension of a significant number of partially digested peptides were obtained. This, however, allowed immediate alignment of many short peptides, as shown in Figure 9. No amino acid sequence was obtained by submitting the peptide with retention time 25 min of Figure 1 to Edman degradation. This peptide was assigned as T1 since its N-terminus is blocked, as expected from the analysis of the emile protein, and its amino acid composition (Table I) corresponds to the one expected for this portion of the molecule on the basis of the result aothared through the analysis of the CNBT fragments (see below).

In conclusion, the characterization of the tryptic peptides allowed to estabilish portions of the sequence accounting for 333 out of the total 338 amino acid residues which constitute the enzyme. Complete alignment of the tryptic peptides and identification of the sequence of the 5 amino acid residues still missing (positions 1-5) were achieved by the characterization of the peptides obtained by fragmentation of the carboxymethylated protein with CNBr, CNBr/HFBA and *S. aureus* V8 protease.

CNBr and CNBr/HEBA digests RP-HPLC separation of the petides obtained by fragmentation of the carboxymethylated protein with CNBr is shown in Figure 6. The number of peaks largely exceeded the number of pepides expected from the methionine content as determined by amino acid analysis of the protein (Table 1). Characterization of all the HPLC fractions revealed that some peaks constained the same peptide, sometimes in various mixtures with others, or partially fragmented portions. However, C2, C3, C4 and C7 were obtained as pure peptides and their sequence was numbipoutly determined (Table III and Fig. 9). The large fragments C3 and C6 instead were isolated only as a mixture. Sequence analysis performed on aliquots of the fractions eluting between 80 and 86 min and containing C5 and C6 resulted in more than one amino acid released for each degradation cycle. In this case, tentative assignment of the amino acid sequence for each fragment was deduced by comparison with the known primary sequence of the corresponding typic peptides previously characterized. The remaining material from the HPLC fractions containing C5 and C6 was pooled and further fragmented at tryptophanyl residues as described under "Experimental Conditions". RP-HPLC separation of the peptides to bained is shown in Figure 7. Again, only the data obtained from the analysis of the pure peptides have been used to determine the alignment described in Figure 9. Staphylococcus aureur V8 digest. Experimental evidence for the alignment of the tryptic peptides comprising the C-terminal portion of the molecule (T27-T37) was obtained by the characterization of a few peptides originated by digesting the protein with *Saureur* V8 protease (Fig. 8). Both native and carboxymethylated D-aspartate oxidase resulted to be very resistant to digestion. If fact, only the C-terminal portion of the molecule (starting after Glu253) appeared to be sensitive to the action of the protease.

Special problems. Apart from the failure in sequencing the N- terminal portion of the entire protein , the unambigous assignment of a few other portions of the molecule required special attention. In particular, no C-terminal sequence was obtained by submitting the carboxymethylated protein to digestion with carboxypeptidae. Thus, the C-terminus of the protein has been identified as Ly338 from the amino acid sequence of T37 and C87. However the sequence of tryptic peptide T37a is equal to the one of T37 accepts for the absence of the terminal Lys residue. Determination of the amino acid sequence of T37 act T37 act onfirms that these peptides differ only for one Lys residue (Table I). In addition, S18 also ends with Pro337. Taken together these results suggest that partial digestion of the C-terminal residue might occur, either in the cell or during the purification of the native protein, originating two forms of the enzyme ending with Lys338 or with Pro337.

Another type of molecular microheterogeinity has been detected, deriving from genetic variants of the enzyme. In fact, either Val or Ile were found at position 228 (peptide T24 and T24a, respectively). Finally, overlapping of the tryptic peptides was in some cases complicated by a few peculiar "tryptic" cleavages, such as between His134 and Ala135, Leu160 and Ile 161 and Asn181 and Cys182 (Fig. 9).

Table II and Figure 9 summarize the data used to determine the complete primary structure of beef kidney D-asparate oxidase. The enzyme consists of 338 amino acid residues, accounting for a Mr, of 37305 for the apoprotein, in accordance with the value reported in previous studies (3). Table I shows the good correlation between the amino acid composition of the entire protein determined by amino acid analysis with the one deduced from the sequence. The data relative to the Cys content suggest that the enzyme contains 7 Cys and no disulfide bridges, since all the Cys can be titrated by DTNB under denaturing non-reducing conditions.



Fig 1. RP-HPLC of the 0.1 % TFA-soluble peptides obtained by tryptic digestion of [14 C]-carboxymethylathed D-aspartate oxidase. Conditions: column C-8 Aquapore RP-300 (250x 7 mm); buffer A: 0.1 % TFA; buffer B: CH3CN + 0.075 % TFA; llow rate 2 ml/min. Gradient: from 0 to 60% B in 85 min. Fraction numbers refer to the tryptic peptides found in each peak and aligned as shown in Fig. 9. An asterisk marks peptides containing [14 C]-carboxymethylosteine determined by radioactivity counting both on an aliquot of each fraction and of the PTH-derivative recovered from the Edman degradation. The peptides containing threations with retention time of 32, 35 and 43 min were identified after re-purification as described in Fig. 3. 4 and 5, respectively.



Fig. 2. RP-HPLC purification of peptides contained in fraction with retention time of 32 min in Fig.1. Conditions: column C-8 Aquapore RP- 300 (250 x 4.6 mm), flow rate 0.7 ml/min; buffer A: 5 mM NaPa, pH 6.0; buffer B: CH3CN. Gradient: from 0 to 60 % B in 85 min.



Fig. 3. RP-HPLC purification of peptides contained in fraction with retention time of 35 min in Fig. 1. Conditions as in Fig. 3.



Eig. 4. RP-HPLC purification of peptides contained in fraction with retention time of 43 min in Fig. 1. Conditions as in Fig. 3.



Fig. 5. RP-HPLC of the 0.1 % TFA-insoluble peptides obtained by tryptic digestion of [¹⁴C]-carboxymethylathed D-aspartate oxidase. Conditions as in Fig. 1.



Eig. 6. RP-HPLC of peptides obtained by CNBr fragmentation of S- carboxymethylated D-aspartate oxidase. Conditions: column C-8 Aquapore RP-300 (250 x 4.6 mm); flow rate 0.7 ml/min; buffer A: 0.1 % TFA; buffer B: CH₃CN + 0.075 % TFA. Gradient: from 0 to 60% B in 85 min.



Eig_7. RP-HPLC of peptides obtained by CNBr/HFBA fragmentation of peptides CS and C6 contained in fractions with retention time ranging from 80 to 85 min in Fig. 6. Conditions as in Fig. 6.





λm	ino ac genera	id tea	a co	mp Y	os tr	iti ypt	ion ic	oi di	t ge	he sti	en on	ti o	re f	s-	ro Ca	te rb	in loxy	ar /æq	d c ithy	of t /lat	he j ed i	pej en:	tides yme	
AA		_	(ho	le	_		1	_	<u>T2</u>			т	3				T4			T5			6	
Cys	•	7.0	2(7)				, c	. 6	(1)		<u>.</u>	8 (1)										
Thr	20	3.(0.9)(2)(2	2)	ł	0.7	11)) {	. 8	(1)		1. D.	1($1) \\ 1)$		4.	3(2	2)						
Glx	31	0.1	5(3	1)				1	.,	(2)			ь (2)		1.	3(2	2)	٥.	9 (1)	1.0	(1)	
Gly	29	9.0	2(3	2)				3	. 2	(3)		0.	9 (1) 1}		3. 1.	9(1	5) L)						
Val	27	7.2	2(3	0)	4	5.6	(1)) 3	.3	(3) (4)		1.	8 (2)		1.	9(2	2) L)						
Ile	18	3.0	(2	0)	1		(1)	1		(1) (2)		1.1	5 (1) 2)	-	2.	0(1	2)						
Tyr		5.6		7)				1		(1)						1.	0(1							
Lya	19	. 4	(1	9)				1	. 0	(1)		L.(0 (1)		1.	0(1	5	1.	1 (1)	1.0	(1)	
Arg Trp	ь 20	5.0	12	0) B)	3	1.0	(1))																
Pos	1	1-3	38	Ċ		1-	5	6	-2	4	2	5-:	37			38	-63		64	-65	Ì	56-	-69	
(*)						8		1	5		:	20				5	1		2	3		22		
٨٨	T6+7		Т	3+9	9		T1(<u>)+1</u>	1		T1	2			т	13			T15		T1	;+1	6+17+1	8
Сув	2 9 (4)														0.9	• (1)					q	.6(1)	
Thr	1.4(2)		1.0)(1	L)		1.4	"(1	,	2	. 1	(2))									0	.2(4)	
Glx	6.9(8)						1.1	l (1)	1	.1	(1)	,		2.3	LC.	2)					1	.7(2)	
Gly	2.9(3)									1	.2		Ś		1.0	• (1)					4	.6(5)	
Val Met	5.3(6)		0.1	571						ō	.7	(1)	5		.		1,		1.0	(1)		Z	.4(3)	
Ile Leu	3.2(4) 3.1(3)		•••		• /		1.0		,	1	. 0				1.3		21					3	.5(3)	
Tyr Phe	0.4(1) 4.4(5)								,	2	. 8	(3)		j	0.1	i (ī)					1	.0(1)	
Lys His	1.0(1) 0.7(1)		2.()(2	?)		2.0)(2)	0 1	.9	1) 2)		;	1.0)(1)		0.9	(1)		1	.2(1)	
Arg Trp	0.9(1) n.d(3)													,	n. d	1	1)					2	.6(3) .d(1)	
Pos	66-117	1	18-	-12	21	1	22-	12	6	12	7-:	40	5	14	41-	-1!	52		154	-155	5	15	4-188	
(\$)	15		9	•			11	l			9				87	'			1	o			6	
								Ta	b 1e	• 1	(or	١t	in	1eć	I)								
<u> </u>	T17+18	_		18			1	19			T	0			1	2	1			T22			<u>T23</u>	
Сув Авх	0.7(1) 2.8(3)		2.	6 (9 (1) 3)		1.	ο(1)															
Thr Ser	1.9(3)		2.	0 (3)		ο.	6(1)															
Fro	2.1(2) 1.0(1)		2.	1(2) 1)		1.	2(1)	0	. 8 (1)			1.	0	(1)		1. 1.	4(1) 1(1)				
Ala	2.4(2)		2.	. 0 ((Z)		о. о.	9(1)	•					1.	.3:	(1)		1.	1(1)				
Met	1.7(2)		,	. / ((2) (2)					0	. • •	1)			0.	. 8	(1)		1.	9(2)			• (•)	
Leu	3.2(3)		3	1(3)		1.	D (1)			- ,			1.	1	(1)					Ű	(1)	
Phe Lvs	1.1(1)		٥.	9 (1)		1.	1(11	1	. 0 (1)			1.	0	(1)		1.	0(1)		1	.0(1)	
His Arg	1.0(1) 2.0(2)		1.	0($^{1)}_{1)}$			`	•	1	. 0 (1)								• •		0	.6(1) .0(1)	
Trp	n.d(1)		n.	d (1)														n.,	d(1)				
Pos (*)	165-18 23	8	166	5-1 51	.88		189	3	95	19	96- 77	20	00	1	201	6	205		206 5	-212 3		21	3-216 67	
٨٨	T24a			1	24			T	25			1	2	6			T2	7		T28		Т	29	
Суя Авх	2.1(2				. ()							٦.	0	0					1	. 8 (2	:)	1.	2(1)	
Thr	3.2(3	í.		.3	(3	ij l						1.	, 7	(2)		1.	. 1 (11	٥	. 973		•	• (• /	
Glx	0.7(1	5). 8				ο.	8 ()	1)		1.	2	(1)		ī.	. 3 (īj	1 2	.2(1	ij -			
Gly Ala	4.1(4)	4	1.4	(4	j.						0.	8	(1)					1	. 2 (1	0	o. o.	9(1) 9(1)	
Val Met	0.8(1)	1	9	(2)														•				
Ile Leu	1.9(2)	2	2.0) (1) (2	.) !)						0. 1.	9 0	(1) (1))	0	.6(.D(1) 1)	2	. 4 (2	!)	1.	2(1)	
Tyr Phe	2.0(2	:)	1	1.5	6(2	:)																0.	5(1)	
Lys His								1.	0(:	1)		1.	. 0	(1))									
Arg Trp	1.0(1)	3)(1	.)						л.	đ	(1))	1.	. 0 (1)	1	.0(1)	1.	0(1)	
Pos	217-2	37	21	17-	23	7	23	8-	239	•	24	0-	2	52	2	5	3-2	57	2	58-2	65	26	7-272	
(•)				23				2.5				-											205	
			-					Ta	514		د د	or	nt.	1.111	ad.	•)		•		-	. ~			
<u>88</u>	T30	•	T:	11			<u>T32</u>	<u> </u>	-	TJ	.	•••	T	34.	-35	>		<u>T3</u>	6 	T	17		<u>T378</u>	
Asx					,				1	1.0	(1)		1				1	.2	(1)					
Ser Glv	1.0/11		J.8	. (1	.,	,	. 0 '	1 1	ç	2.0	(1) (2)		01	. 8	(1)		,	, я	(21	0.5				
Pro Glv	,		1.4	(2)	-		- 1			,		15	. 0	(1) (5)					2.1	(2)		1.9(2)	
Ala Val			1.e	5(2	;)				¢	1.7	(1)		3 1	. 4	(3) (2)		3	. 3	()	1.1	(1)		1.1(1)	
Met Ile													0	.7 .8	(1) (1)									
Leu Tyr			1.2	!(1	.)	1	.1(1)	1	2.0	(2)		2 0	. 0	(2) (1)		1	. 8	(2)					
Phe Lys						0	.9(1)												1.1	(1)		0.1(0)	
AIG Tro	1.0(1)		3.0) (3)				;		(1)		31,00	. 6	(4) (2) (1)		1	. 0	(1)					
POR	273-27	4	275	i-2	85		86-	28	8 :	198-	-29	6	n 2'	.u.	-33	3	32	4-	333	314	-11	8	334-11	7
(1)	45			8		-	42		-	10	00			30	ົ	,		86			6	-	11	÷

Table I

Footnotes to Table I. "Cys content of the entire protein was determined both by amino acid analysis as S-carboxymethylcysteine and by DTNB titration. "Datarmined by second derivative spectracopy (from Ser. 3). Egg indicates the posicion of the peptide in the primary sequence. (8) Indicates the recovery yield of each peptide.





Footnotes to Table II. Values are given in pmoles. T. C. H and S indicates peptide obtained by fragmentation with trypsin, CNBr, CNBr/HFBA and S-aureus V8 protease, respectively; hal, indicates homoserine and homoserine lactone; d, indicates a PTH-derivative of tyrosine eluting after PTH-leucine under the experimental condition used for PTH-amino acid separation; % indicates that peptides C5 and C6 were sequenced as a mixture.