Margenau, H., & Murphy, G. M. (1956) The Mathematics of Physics and Chemistry, Van Nostrand, New York.

Mazurek, N., Schindler, H., Schurholz, Th., & Pecht, I. (1983) *Proc. Natl. Acad. Sci. U.S.A.* 81, 6841-6845.

Menon, A. K., Holowka, D., & Baird, B. (1984) J. Cell. Biol. 98, 577-583.

Menon, A. K., Holowka, D., Webb, W. W., & Baird, B. (1985) J. Cell Biol. 102, 534-540.

Menon, A. K., Holowka, D., Webb, W. W., & Baird, B. (1986a) J. Cell Biol. 102, 534-540.

Menon, A., Holowka, D., Webb, W. W., & Baird, B. (1986b) J. Cell Biol. 102, 541-550.

Pecht, I., Haselkorn, D., & Friedman, S. (1972) FEBS Lett. 24, 331-338.

Poljak, R. J. (1978) CRC Crit. Rev. Biochem. 45-84.

Rudolph, A. K., Burrows, P. D., & Wahl, M. R. (1981) Eur. J. Immunol. 11, 527-529.

Sagi-Eisenberg, R., & Pecht, I. (1984) EMBO J. 3, 497-500.

Schechter, I., & Berger, A. (1966) *Biochemistry 5*, 3362-3370. Schlessinger, J., Webb, W. W., Elson, E. L., & Metzger, H. (1976) *Nature (London) 264*, 550-552.

Schumaker, V. N., Green, G., & Wilder, R. D. (1973) Immunochemistry 10, 521-528.

Schumaker, V. N., Seegan, G. W., Smith, C. A., Ma, S. K., Rodwell, J. D., & Schumaker, M. F. (1980) Mol. Immunol. 17, 413-423.

Schwartz, L. B., & Austen, K. F. (1984) *Prog. Allergy 34*, 271-321.

Segal, D. M., Taurog, J. D., & Metzger, H. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 2993–2997.

Siraganian, R. P., Hook, W. A., & Levine, B. B. (1975) *Immunochemistry* 12, 149-157.

Warner, C., & Schumaker, V. (1970) *Biochemistry* 9, 451-459.

Wilder, R. L., Green, G., & Schumaker, V. N. (1975) Immunochemistry 12, 35-47.

Molecular Cloning and Sequence Analysis of cDNAs Encoding Porcine Kidney D-Amino Acid Oxidase[†]

Kiyoshi Fukui, Fusao Watanabe, Toshihiro Shibata, and Yoshihiro Miyake*

Department of Biochemistry, National Cardiovascular Center Research Institute, Fujishiro-dai, Suita, Osaka 565, Japan

Received November 17, 1986; Revised Manuscript Received January 22, 1987

ABSTRACT: Complementary DNAs encoding p-amino acid oxidase (EC 1.4.3.3, DAO), one of the principal and characteristic enzymes of the peroxisomes of porcine kidney, have been isolated from the porcine kidney cDNA library by hybridization with synthetic oligonucleotide probes corresponding to the partial amino acid sequences. Analysis of the nucleotide sequences of two clones revealed a complete 3211-nucleotide sequence with a 5'-terminal untranslated region of 198 nucleotides, 1041 nucleotides of an open reading frame that encoded 347 amino acids, and a 3'-terminal untranslated region of 1972 nucleotides. The deduced amino acid sequence was completely identical with the reported sequence of the mature enzyme [Ronchi, S., Minchiotti, L., Galliano, M., Curti, B., Swenson, R. P., Williams, C. H. J., & Massey, V. (1982) J. Biol. Chem. 257, 8824–8834]. These results indicate that the primary translation product does not contain a signal peptide at its amino-terminal region for its translocation into peroxisomes. RNA blot hybridization analysis suggests that porcine kidney D-amino acid oxidase is encoded by three mRNAs that differ in size: 3.3, 2.7, and 1.5 kilobases. Comparison of the sequences of the two cDNA clones revealed that multiple polyadenylation signal sequences (ATTAAA and AACAAA) are present in the 3'-untranslated region, making the different mRNA species. The efficiency of 3' processing of the RNA was quite different between the two signal sequences ATTAAA and AACAAA. Southern blot analysis showed the presence of a unique gene for D-amino acid oxidase in the porcine genome.

Flavoenzymes catalyze a variety of reactions by transferring one or two electrons between chemically diverse donor and acceptor molecules. D-Amino acid oxidase (EC 1.4.3.3, DAO)¹ is one of the representative flavoproteins with flavin adenine dinucleotide (FAD) as the prosthetic group that catalyzes the oxidative deamination of D-amino acids. Since the initial characterization (Krebs, 1935) and crystallization (Kubo et al., 1958; Massey et al., 1961; Yagi et al., 1962) of this enzyme, many investigations have been made to clarify the physicochemical properties and reaction mechanism of the enzyme. Systematic studies of DAO activity in various tissues

revealed its existence in liver, kidney proximal tubules, certain parts of brain (Dunn & Perkoff, 1963), and granules of neutrophilic leukocytes (Cline & Lehrer, 1969). However, D-amino acids have not been found in mammalian proteins, and they do not appear to be intermediates in normal metabolism. Several lines of evidence indicating the function of the enzyme were reported (Hamilton et al., 1979; Nakajima et al., 1981), but the biological significance of this enzyme still remains to be elucidated.

Intracellular localization of DAO is reported to be in the special organelle, i.e. peroxisome (de Duve & Baukhuin, 1966), and DAO is one of the principal and characteristic enzymes of the peroxisomes of porcine kidney. This organelle has a

[†]This study was supported in part by a Research Grant for Cardiovascular Diseases (59C-8) from the Ministry of Health and Welfare of Japan and Grant-in-Aids for Scientific Research (61570151) and for Special Project Research (61132005) from the Ministry of Education, Science and Culture of Japan.

¹ Abbreviations: DAO, D-amino acid oxidase; bp, base pair; kb, kilobase; FAD, flavin adenine dinucleotide; PAS, periodic acid-Schiff.

single membrane, metabolizing hydrogen peroxide and catalyzing the β -oxidation of fatty acids. Our previous study has demonstrated that DAO is synthesized on free ribosomes, and both the in vivo and in vitro synthesized DAO have the same size as that of the mature enzyme, indicating a posttranslational transfer of DAO into peroxisomes without any proteolytic modification (Fukui et al., 1986). Studies on the biosynthesis of other peroxisomal enzymes such as catalase, the enzymes catalyzing fatty acid β -oxidation, and urate oxidase agreed with our result (Goldman & Blobel, 1978; Miura et al., 1984), but not in the case of 3-ketoacyl-CoA thiolase. cDNA cloning analysis revealed that catalase and the enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase bifunctional enzyme have no terminal peptide extension as a signal for translocation into peroxisomes (Osumi et al., 1985; Furuta et al., 1986).

We have started molecular cloning of the cDNA for DAO with the following aims: (1) to facilitate the elucidation of the biological function of DAO; (2) to study the regulation of the biosynthesis of this enzyme; (3) to approach the translocation mechanism of DAO into peroxisomes at the molecular level in order to obtain some clues for the analysis of the biogenesis of this organelle. In this paper, we describe the molecular cloning and the complete nucleotide sequence of the cDNA encoding porcine kidney DAO. In addition, the expression in the tissue and the genomic organization of the DAO gene are also presented.

EXPERIMENTAL PROCEDURES

Materials. Reagents were obtained as follows: $[\gamma^{-32}P]ATP$ (sp act. 3000–4000 Ci/mmol) and $[\alpha^{-32}P]TTP$ and $[\alpha^{-32}P]$ dGTP (sp act. 3000 Ci/mmol) from New England Nuclear/Du Pont; $[\alpha^{-32}P]dCTP$ (sp act. 410 Ci/mmol) from Amersham, U.K.; avian myeloblastosis virus reverse transcriptase from Midwest Bio-Products; oligo(dT)-cellulose (type 7), oligo(dA)-cellulose (type 7), Escherichia coli ribonuclease H, E. coli DNA polymerase I, E. coli DNA ligase, the Klenow fragment of E. coli DNA polymerase I, terminal deoxynucleotidyl transferase, T4 polynucleotide kinase, and Okayama-Berg vector strain kit (pSV7186 and pSV1932) from Pharmacia, Sweden; restriction endonuclease from Toyobo Co., Japan; nitrocellulose filter from Advantec, Japan, and Schleicher & Schuell, West Germany; M13 mp18 and mp19 RF-DNA from Takara Shuzo Co., Japan; agarose type I from Sigma; guanidinium thiocyanate from Fluka, Switzerland; RNasin from Promega-Biotec; sequencing reagents from Takara Shuzo and Toyobo Co., Japan.

Oligonucleotide Synthesis. The two amino acid sequences corresponding to residues 106-111 (Tyr-Trp-Lys-Asp-Met-Val) and 66-71 (Asn-Trp-Asn-Gln-Gln-Thr) of DAO (Ronchi et al., 1982) were chosen for the synthesis of oligonucleotide probes (Figure 1). Mixtures of all possible oligonucleotides complementary to mRNA were synthesized by the phosphoramidite method (Matteucci & Caruthers, 1981) with an automated DNA synthesizer (Applied Biosystems, Model 380A). Full-length oligonucleotides were purified by high-performance liquid chromatography on a reverse-phase (C_{18}) column. The synthetic probes were phosphorylated at the 5' end by transfer from $[\gamma^{-32}P]ATP$ with T4 polynucleotide kinase. Unincorporated nucleotides were removed by chromatography over a DEAE-cellulose column.

RNA and DNA Preparation. High molecular weight DNA was obtained from porcine kidney as described (Yaoita & Honjo, 1980). Total RNA was extracted from a porcine kidney with guanidinium thiocyanate (Chirgwin et al., 1979), and poly(A+) RNA was isolated by subjecting the total RNA

to oligo(dT)-cellulose column chromatography (Aviv & Leder, 1972).

cDNA Library Construction and Isolation of cDNA for DAO. Vector primer and oligo(dG)-tailed linker for cDNA library construction were prepared from pSV7186 and pS-V1932, respectively, according to the original procedure described by Okayama and Berg (1982). A porcine kidney cDNA library was then constructed with the purified mRNA. E. coli HB101 was transformed and selected for ampicillin resistance. About 500 000 transformants derived from the cDNA library were screened by hybridization (Hanahan & Meselson, 1980) with the mixture of oligonucleotide probes as described in Figure 1. Hybridization and filter washing conditions were as described (Wallace et al., 1979). Twenty clones gave positive hybridization with both probes I and II, and two clones (pDAO-10, pDAO-13), which apparently carry the largest insert, were selected for further sequence analysis. Cloning procedures were carried out under P2 conditions according to the Japanese guidelines for recombinant DNA research.

DNA Sequence Determination. cDNA clones were digested with a variety of restriction enzymes to give convenient DNA fragments for subcloning into M13 phage vectors mp18 and mp19 (Yanisch-Perron et al., 1985). The nucleotide sequences were determined by the dideoxy chain-termination method of Sanger et al. (1977).

Northern and Southern Blot Analysis. For RNA blot analysis, RNA (5 µg of poly(A+) RNA per lane) was denatured with glyoxal and dimethyl sulfoxide, separated on a 1.5% agarose gel, and transferred to a nitrocellulose filter according to Thomas et al. (1980). For DNA blot analysis, high molecular weight DNA from porcine kidney was digested with restriction enzymes, electrophoresed on a 0.7% agarose gel (2) μg per lane), and transferred to a nitrocellulose filter according to Southern (1975). The filter was hybridized with DNA fragments derived from a cDNA insert that was labeled by nick translation with $[\alpha^{-32}P]dCTP$ to obtain a specific activity of 150-300 cpm/pg (Rigby et al., 1977). Hybridization and filter washing conditions were as described (Honjo et al., 1979). When the mixture of synthetic oligonucleotides was used as probes, the hybridization conditions were the same as those of library screening.

RESULTS

Isolation of cDNA Clones for DAO. The approach used to clone cDNA for DAO was to screen a library of cDNA clones by hybridization with synthetic oligonucleotide probes that had been designed on the basis of known amino acid sequence data. A cDNA library was constructed with the elaborated plasmid vector of Okayama and Berg (1982) by the use of poly(A+) RNA extracted from a porcine kidney cortex. Two stretches of amino acid sequence from the published data (Ronchi et al., 1982) were chosen for the oligonucleotide synthesis as shown in Figure 1. The cDNA library was screened by hybridization with two oligonucleotide probes containing all possible cDNA sequences for the selected amino acid sequences. Twenty hybridization-positive clones for both probes I and II were identified from approximately 500 000 transformants. Upon restriction enzyme analysis, these clones were found to share common restriction sites. Two clones (pDAO-10 and pDAO-13), which apparently harbored the largest cDNA insert, were selected for further analysis and subjected to nucleotide sequence determination.

Nucleotide Sequence Analysis. The restriction endonuclease map and outline of the strategy used to determine the whole nucleotide sequence of the cDNA insert in clone pDAO-10

3614 BIOCHEMISTRY FUKUI ET AL.

FIGURE 1: Synthetic oligonucleotides used as hybridization probes. Two portions of the amino acid sequence of DAO (I and II) were chosen for the synthesis of oligonucleotide probes: (a) amino acid sequence of DAO; (b) predicted mRNA sequences; (c) 17-mer oligonucleotide probes complementary to possible mRNAs.

as well as the partial sequence of the insert in pDAO-13 are shown in Figure 2. Comparison of restriction endonuclease sites reveals complete homology between the two clones except the limited region of the 5' end. Figure 3 shows the 3211-nucleotide sequence encoding the porcine DAO, determined with clones pDAO-10 and pDAO-13. This sequence is flanked by a poly(A) tract at the 3' end. Determination of the 5'-terminal sequence of pDAO-13 cDNA detected a 109-nucleotide deletion and a 5'-directed 54-nucleotide extention compared with that of pDAO-10, whereas the other nucleotide sequences analyzed were identical with each other.

As shown in Figure 3, pDAO-10 starts at nucleotide 55, and pDAO-13 lacks nucleotides from 81 to 189 (109 nucleotides). The first ATG codon is located at nucleotides 99-101. However, the TGA termination codon (nucleotides 114-116) appears in-frame 12 nucleotides downstream, and this reading frame codes only 5 amino acid residues. The second ATG codon exists at nucleotides 199-201 and appears downstream from the in-frame terminator TGA (nucleotides 82-84). This ATG is preceded by sequences that fulfill the Kozak (1981) criteria for initiation codon. The open reading frame is followed by the TGA termination codon at nucleotides 1240-1242. The 1041-nucleotide reading frame codes for the 347 amino acids, and the deduced amino acid sequence was in complete agreement with the reported sequence by Ronchi et al. (1982). One ambiguous amino acid residue in the reported sequence was Asx at position 192. This residue was assigned to be Asp on the basis of the nucleotide sequence determination. To confirm the N-terminal amino acid, DAO was purified from porcine kidney according to Curti et al. (1973). The N-terminal amino acid sequence (32 residues) of the purified enzyme preparation, determined by a gas-phase amino acid sequenator (Hewick et al., 1981), starts with methionine as does the deduced sequence and provides the same result (data not shown). Therefore, the ATG codon at nucleotides 199–201 is concluded to be the initiator of translation.

Taken altogether, the complete cDNA sequence is composed of a 5'-untranslated region of 198 nucleotides, a coding region of 1041 nucleotides for 347 amino acids, and a 3'-terminal untranslated region of 1972 nucleotides. These total 3211 nucleotides covered almost the full length of the mRNA sequence on the basis of the RNA blot analysis (see below). It should be also noted that the sequence difference between pDAO-10 and pDAO-13 exists only within the untranslated region, leaving the coding sequences identical with each other.

The 3'-untranslated regions contain 1972 and 1400 nucleotides in clones pDAO-10 and pDAO-13, respectively. The polyadenylation signal, ATTAAA (nucleotides 3178-3183), is present 29 residues upstream of the poly(A) tail in clone pDAO-10. In the case of clone pDAO-13, the sequence AA-CAAA (nucleotides 2628-2633) located 11 nucleotides before the site of poly(A) addition is presumed to be the polyadenylation signal.

There are two potential asparagine-linked glycosylation sites conforming to the consensus sequence of Asn-X-Ser/Thr (Marshall, 1974) at amino acid positions 134 and 180. However, our preliminary analysis indicated the absence of a carbohydrate moiety associated with the mature enzyme preparation as judged by the PAS staining (data not shown).

Codon choices for the DAO mRNA have been assigned (Table I), according to the amino acid sequence shown in Figure 3. There is a preferential use of certain codons for some amino acids by DAO mRNA. Compared with the mammalian codon usage pattern surveyed by Ikemura (1985), the UCU codon was used for Ser in preference to AGC, and for Ala, GCU and GCA were more often used than GCC. In addition, an apparent preference for G or C at the third

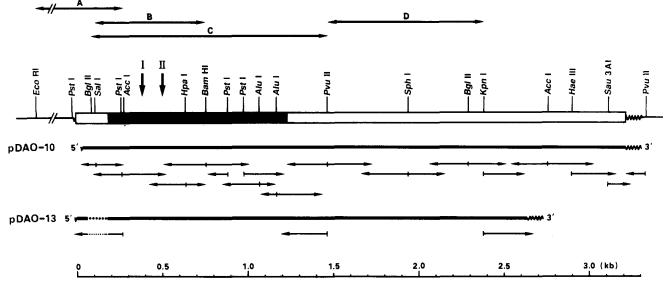


FIGURE 2: Restriction endonuclease map and sequencing strategy for the cDNA inserts. The map displays only the relevant restriction endonuclease sites. A solid line represents the vector DNA; a zigzag line represents the oligo(dG)/(dC) or poly(dA)/(dT) tail; a solid box represents the coding region for DAO; an open box represents the untranslated region. The cDNA inserts are shown by solid bars. The broken bar in pDAO-13 indicates the deleted region. The direction and extent of the sequence determination are indicated by horizontal arrows under each insert used. Vertical arrows above the restriction map show the positions of the oligonucleotide probes used for screening of the cDNA library. Fragments A-D show the probes used for Southern and Northern blot hybridization experiments (Figures 4 and 5).

PstI CTGC/GGGGGGGGG

1																				_				GGGG	
GGAG	AACA	GGAC	CCTC	CCTC	CAGG	AGAG	AGGA	CCCT	GTGC	CCAC	AGCA	GCTG	CTGC	TTTT	CTCC	AGAA	TTTC	GGCA	CGCC	AGGT	GÄGT	GGGT	TGGG	GGCA	98
XYGG	CGGA	GCAC	ATC	GXG	GTCC	TCCC	AGTG	TCGA	CAGA	AAAA	GTTC	ACGT	GAGA	GTCT	GACG	CAGG	TGAC	AACC	TTGA	AGGG	ACCA	CAGG	CTGG	CACG	198
ATG Met																									273
1	9	141	*41	5	116	GI,	VI a	GIY	10	116	GIY	Leu	261	15	N1a	Leu	Cys	116	20	GIU	AIG	ıyı	urs	25	
CTC	CMC	CAC	000	O.B.O	C.m	CEC		cmc	m. c																• • •
GTC Val	Leu	Gln	Pro	Leu	Asp	Val	Lvs	Val	TAC	GCA Ala	ASD	CGC	TTC	ACC Thr	Pro	TTC	ACC	ACC Thr	ACT Thr	GAT	GTA Val	GCT	GCT	GGC	348
26				30			-,-		35					40			• • • •	•	45	p	•••		714	50	
СТС	TCC	CAG	ccc	TAC	NCC.	mcm.	CAC	coo	100		001	010	CAC	coc		mcc			~~~						422
Leu	Trp	Gln	Pro	Tyr	Thr	Ser	Glu	Pro	Ser	Asn	Pro	Gln	Glu	Ala	Asn	Tro	Asn	Gln	Gln	Thr	Phe	Asn	TAT	Leu	423
51	_			55					60					65				I	70				- • -	75	
CTG	AGT	CAC	ATC	GGT	тст	ccc	AAT	GCT	GCA	AAC	ATG	GGT	СТС	ACC	CCA	GTC	TCA	GGC	TAC	AAC	СТС	ጥ ጥር	ССТ	GAA	498
Leu	Ser	His	Ile	Gly	Ser	Pro	Asn	Ala	Ala	Asn	Met	Gly	Leu	Thr	Pro	Val	Ser	Gly	Tyr	Asn	Leu	Phe	Arg	Glu	470
76				80					85					90					95					100	
GCT	GTT	CCG	GAC	ССТ	TAC	TGG	AAA	GAC	ATG	GTC	CTG	GGA	TTC	CGA	AAG	CTG	ACT	ccc	AGA	GAA	CTG	GAC	ATG	ттт	573
Ala	Val	Pro	Asp	Pro	Tyr	Trp	Lys	Asp	Met	<u>Val</u>	Leu	Gly	Phe	Arg	Lys	Leu	Thr	Pro	Arg	Glu	Leu	Asp	Met	Phe	
101				105			1	I	110					115					120					125	
CCT	GAT	TAT	AGA	TAT	GGC	TGG	TTC	AAC	ACA	TCC	CTG	ATT	CTG	GAG	GGA	AGG	AAG	TAC	СТА	CAG	TGG	CTG	ACA	GAA	648
Pro	Asp	Tyr	Arg	Tyr	Gly	Trp	Phe	Asn	Thr	Ser	Leu	Ile	Leu	Glu	Gly	Arg	Lys	Tyr	Leu	Gln	Trp	Leu	Thr	Glu	•
126				130					135					140					145					150	
AGG	TTA	ACT	GAG	AGG	GGA	GTG	AAA	TTC	TTC	CTG	AGG	AAG	GTG	GAA	тст	TTT	GAG	GAG	GTG	GCA	AGA	GGT	GGC	GCT	723
Arg	Leu	Thr	Glu	Arg	Gly	Val	Lys	Phe	Phe	Leu	Arg	Lys	Val	Glu	Ser	Phe	Glu	Glu	Val	Ala	Arg	Gly	Gly	Ala	
151				155					160					165					170					175	
GAT	GTG	ATT	ATC	AAC	TGC	ACT	GGG	GTG	TGG	GCT	GGG	GTG	CTG	CAA	CCG	GAT	ccc	CTG	CTG	CAG	CCA	GGC	CGG	GGG	798
Asp	Val	Ile	Ile		Сув	Thr	Gly	Val		Ala	Gly	Va 1	Leu		Pro	Asp	Pro	Leu		Gln	Pro	Gly	Arg		
176				180					185					190					195					200	
CAG	ATC	ATT	AAG	GTG	GAT	GCC	CCT	TGG	CTG	AAG	AAC	TTC	ATT	ATC	ACC	CAT	GAC	CTA	GAG	AGA	GGC	ATC	TAC	AAC	873
Gln	Ile	Ile	Lys	Val	Asp	Ala	Pro	Trp	Leu	Lys	Asn	Phe	Ile	Ile	Thr	His	Asp	Leu	Glu	Arg	Gly	Ile	Tyr	Asn	
201				205					210					215					220					225	
TCT	CCA	TAC	ATC	ATT	CCA	GGG	CTG	CAG	GCA	GTG	ACA	CTT	GGA	GGC	ACC	TTC	CAG	GTG	GGG	AAC	TGG	AAT	GAG	ATA	948
Ser	Pro	Tyr	Ile		Pro	Gly	Leu	Gln		Val	Thr	Leu	Gly		Thr	Phe	Gln	Val		Asn	Trp	Asn	Glu		
				230					235					240					245					250	
226																									
226	AAT	ATC	CAG		CAC	AAC	ACC	ATC		GAA	GGC	TGC	TGC	AGA	CTG	GAG	ccc	ACA	CTG	AAG	GAT	GCA	AAA	ATT	1023
226 AAT Asn				GAC Asp					TGG Trp	GAA Glu				Arg					Leu					Ile	1023
226 AAT				GAC					TGG																1023
AAT Asn 251 GTT	Asn GGT	Ile GAA	Gln TAT	GAC Asp 255	His GGC	Asn TTC	Thr	Ile	TGG Trp 260	Glu CGC	Gly	Cys CAG	Cys GTT	Arg 265 CGG	Leu	Glu GAA	Pro	Thr GAA	Leu 270 CAG	Lys CTT	Asp	Ala TTT	Lys GGA	Ile 275 TCT	1023
AAT Asn 251 GTT Val	Asn GGT	Ile GAA	Gln TAT	GAC Asp 255 ACT Thr	His GGC	Asn TTC	Thr	Ile	TGG Trp 260 GTA Val	Glu	Gly	Cys CAG	Cys GTT	Arg 265 CGG Arg	Leu	Glu GAA	Pro	Thr GAA	Leu 270 CAG Gln	Lys CTT	Asp	Ala TTT	Lys GGA	Ile 275 TCT Ser	
AAT Asn 251 GTT	Asn GGT	Ile GAA	Gln TAT	GAC Asp 255	His GGC	Asn TTC	Thr	Ile	TGG Trp 260	Glu CGC	Gly	Cys CAG	Cys GTT	Arg 265 CGG	Leu	Glu GAA	Pro	Thr GAA	Leu 270 CAG	Lys CTT	Asp	Ala TTT	Lys GGA	Ile 275 TCT	
226 AAT Asn 251 GTT Val 276 TCA	Asn GGT Gly AAC	GAA Glu ACA	Gln TAT Tyr GAG	GAC Asp 255 ACT Thr 280 GTC	GGC Gly ATC	Asn TTC Phe CAC	Thr CGG Arg	CCA Pro	TGG Trp 260 GTA Val 285	CGC Arg	CCC Pro	Cys CAG Gln GGC	Cys GTT Val	Arg 265 CGG Arg 290	CTA Leu CTC	GAA Glu ACC	AGA Arg	GAA Glu CAC	Leu 270 CAG Gln 295 TGG	CTT Leu GGC	CGC Arg	TTT Phe	GGA Gly CTA	TCT Ser 300	
226 AAT ASN 251 GTT Val 276 TCA Ser	Asn GGT Gly AAC	GAA Glu ACA	Gln TAT Tyr GAG	GAC Asp 255 ACT Thr 280 GTC Val	GGC Gly ATC	Asn TTC Phe CAC	Thr CGG Arg	CCA Pro	TGG Trp 260 GTA Val 285 GGC Gly	Glu CGC Arg	CCC Pro	Cys CAG Gln GGC	Cys GTT Val	Arg 265 CGG Arg 290 GGG Gly	CTA Leu CTC	GAA Glu ACC	AGA Arg	GAA Glu CAC	Leu 270 CAG Gln 295 TGG Trp	CTT Leu GGC	CGC Arg	TTT Phe	GGA Gly CTA	TCT Ser 300 GAG Glu	1098
226 AAT Asn 251 GTT Val 276 TCA	Asn GGT Gly AAC	GAA Glu ACA	Gln TAT Tyr GAG	GAC Asp 255 ACT Thr 280 GTC	GGC Gly ATC	Asn TTC Phe CAC	Thr CGG Arg	CCA Pro	TGG Trp 260 GTA Val 285	Glu CGC Arg	CCC Pro	Cys CAG Gln GGC	Cys GTT Val	Arg 265 CGG Arg 290	CTA Leu CTC	GAA Glu ACC	AGA Arg	GAA Glu CAC	Leu 270 CAG Gln 295 TGG	CTT Leu GGC	CGC Arg	TTT Phe	GGA Gly CTA	TCT Ser 300	1098
AAT Asn 251 GTT Val 276 TCA Ser 301 GTG	Asn GGT Gly AAC Asn GCC	GAA Glu ACA Thr	Gln TAT Tyr GAG Glu CTC	GAC Asp 255 ACT Thr 280 GTC Val 305	GGC Gly ATC Ile	TTC Phe CAC His	Thr CGG Arg AAC Asn GTC	CCA Pro TAT Tyr	TGG Trp 260 GTA Val 285 GGC Gly 310	CGC Arg	Gly CCC Pro GGA Gly	Cys CAG Gln GGC Gly	Cys GTT Val TAT Tyr	Arg 265 CGG Arg 290 GGG Gly 315	CTA Leu CTC Leu	Glu GAA Glu ACC Thr	AGA Arg ATC Ile	Thr GAA Glu CAC His	Leu 270 CAG Gln 295 TGG Trp 320	CTT Leu GGC Gly	CGC Arg TGT Cys	TTT Phe GCC Ala	GGA Gly CTA Leu	TCT Ser 300 GAG Glu 325	1098
AAT Asn 251 GTT Val 276 TCA Ser 301 GTG Val	Asn GGT Gly AAC Asn GCC	GAA Glu ACA Thr	Gln TAT Tyr GAG Glu CTC	GAC Asp 255 ACT Thr 280 GTC Val 305 TTT Phe	GGC Gly ATC Ile	TTC Phe CAC His	Thr CGG Arg AAC Asn GTC	CCA Pro TAT Tyr	TGG Trp 260 GTA Val 285 GGC Gly 310 GAA Glu	Glu CGC Arg CAT His	Gly CCC Pro GGA Gly	Cys CAG Gln GGC Gly	Cys GTT Val TAT Tyr	Arg 265 CGG Arg 290 GGG Gly 315 CTC Leu	CTA Leu CTC Leu ACA Thr	Glu GAA Glu ACC Thr	AGA Arg ATC Ile	Thr GAA Glu CAC His	Leu 270 CAG Gln 295 TGG Trp 320 TCC Ser	CTT Leu GGC Gly	CGC Arg TGT Cys	TTT Phe GCC Ala	GGA Gly CTA Leu	TCT Ser 300 GAG Glu 325	1098
AAT Asn 251 GTT Val 276 TCA Ser 301 GTG	Asn GGT Gly AAC Asn GCC	GAA Glu ACA Thr	Gln TAT Tyr GAG Glu CTC	GAC Asp 255 ACT Thr 280 GTC Val 305	GGC Gly ATC Ile	TTC Phe CAC His	Thr CGG Arg AAC Asn GTC	CCA Pro TAT Tyr	TGG Trp 260 GTA Val 285 GGC Gly 310	CGC Arg	Gly CCC Pro GGA Gly	Cys CAG Gln GGC Gly	Cys GTT Val TAT Tyr	Arg 265 CGG Arg 290 GGG Gly 315	CTA Leu CTC Leu ACA Thr	Glu GAA Glu ACC Thr	AGA Arg ATC Ile	Thr GAA Glu CAC His	Leu 270 CAG Gln 295 TGG Trp 320	CTT Leu GGC Gly	CGC Arg TGT Cys	TTT Phe GCC Ala	GGA Gly CTA Leu	TCT Ser 300 GAG Glu 325	1098
AAT ASN 251 GTT Val 276 TCA Ser 301 GTG Val 326	GGT Gly AAC Asn GCC Ala	GAA Glu ACA Thr AAG Lys	TAT Tyr GAG Glu CTC Leu	GAC Asp 255 ACT Thr 280 GTC Val 305 TTT Phe 330	GGC Gly ATC Ile GGG Gly	TTC Phe CAC His AAA Lys	CGG Arg AAC Asn GTC Val	CCA Pro TAT Tyr CTG Leu	TGG Trp 260 GTA Val 285 GGC Gly 310 GAA Glu 335	Glu CGC Arg CAT His GAA Glu	Gly CCC Pro GGA Gly AGG Arg	Cys CAG Gln GGC Gly AAT Asn	GTT Val TAT Tyr	Arg 265 CGG Arg 290 GGG Gly 315 CTC Leu 340	CTA Leu CTC Leu ACA Thr	Glu GAA Glu ACC Thr ATG Met	AGA Arg ATC Ile CCA Pro	GAA Glu CAC His CCA Pro	Leu 270 CAG Gln 295 TGG Trp 320 TCC Ser 345	CTT Leu GGC Gly CAC	Asp CGC Arg TGT Cys CTC Leu	TTT Phe GCC Ala TGA END	GGA Gly CTA Leu AGA	TCT Ser 300 GAG Glu 325	1098 1173 1249
AAT ASN 251 GTT Val 276 TCA Ser 301 GTG Val 326	GGT Gly AAC Asn GCC Ala	GAA Glu ACA Thr AAG Lys	TAT Tyr GAG Glu CTC Leu	GAC Asp 255 ACT Thr 280 GTC Val 305 TTT Phe 330	GGC Gly ATC Ile GGG Gly	TTC Phe CAC His AAA Lys	CCCT	CCA Pro TAT Tyr CTG Leu	TGG Trp 260 GTA Val 285 GGC Gly 310 GAA Glu 335	Glu CGC Arg CAT His GAA Glu	Gly CCC Pro GGA Gly AGG Arg	Cys CAG GIn GGC Gly AAT Asn	GTT Val TAT Tyr TTG Leu	Arg 265 CGG Arg 290 GGG Gly 315 CTC Leu 340	CTA Leu CTC Leu ACA Thr	Glu GAA Glu ACC Thr ATG Met	AGA Arg ATC Ile CCA Pro	GAA Glu CAC His CCA Pro	Leu 270 CAG Gln 295 TGG Trp 320 TCC Ser 345	CTCC	CGC Arg TGT Cys CTC Leu	TTT Phe GCC Ala TGA END	GGA Gly CTA Leu AGA	TCT Ser 300 GAG Glu 325 CATC	1098 1173 1249
AAT Asn 251 GTT Val 276 TCA Ser 301 GTG Val 326 AGTG	GGT Gly AAC Asn GCC Ala GACC	GAA Glu ACA Thr AAG Lys ACCG GAAG	Gln TAT Tyr GAG Glu CTC Leu CCTC	GAC Asp 255 ACT Thr 280 GTC Val 305 TTT Phe 330	GGC Gly ATC Ile GGG Gly ACAA	TTC Phe CAC His AAA Lys	CGG Arg AAC Asn GTC Val	CCA Pro TAT Tyr CTG Leu GCTC	TGG Trp 260 GTA Val 285 GGC Gly 310 GAA Glu 335 CCCT	CGC Arg CAT His GAA Glu CAGC	Gly CCC Pro GGA Gly AGG Arg	Cys CAG GIn GGC Gly AAT Asn TACT	GTT Val TAT TYr TTG Leu CAAT	Arg 265 CGG Arg 290 GGG Gly 315 CTC Leu 340 GTGC	CTA Leu CTC Leu ACA Thr TCCA	Glu GAA Glu ACC Thr ATG Met	AGA Arg ATC Ile CCA Pro	GAA Glu CAC His CCA Pro	CAG Gln 295 TGG Trp 320 TCC Ser 345	CTT Leu GGC Gly CAC His	CGC Arg TGT Cys CTC Leu TCCA	TTT Phe GCC Ala TGA END CCCC AGTC	GGA Gly CTA Leu AGA TGGC	TCT Ser 300 GAG Glu 325 CATC	1098 1173 1249 1349 1449
AAT Asn 251 GTT Val 276 TCA Ser 301 GTG Val 326 AGTG TCTG	Asn GGT G1y AAC Asn GCC Ala	GAA Glu ACA Thr AAG Lys ACCG GAAG CCTG	Gln TAT Tyr GAG Glu CTC Leu CCTC CACG	GACC ASP 255 ACT Thr 280 GTC Val 305 TTT Phe 330 CCCC AGGTTACA	GGC Gly ATC Ile GGG Gly ACAA GAGA	Asn TTC Phe CAC His AAA Lys AACTG	CGG Arg AAC Asn GTC Val	CCA Pro TAT Tyr CTG Leu GCTC TGAG	TGG Trp 260 GTA Val 285 GGC Gly 310 GAA Glu 335	CAGC GTTC ATCC	Gly CCC Pro GGA Gly AGG Arg CAAC	Cys CAG GIn GGC Gly AAT Asn TACT AGGC	Cys GTT Val TAT Tyr TTG Leu CAAT AAGT	Arg 265 CGG Arg 290 GGG Gly 315 CTC Leu 340 GTGC CCAA GAAA	CTA Leu CTC Leu ACA Thr TCCA CCCT GAAG	Glu GAA Glu ACC Thr ATG Met TTGA ACCT GAAA	AGA Arg ATC Ile CCA Pro TTCC GGGG GCTC	CAC His CCA Pro	Leu 270 CAG Gln 295 TGG Trp 320 TCC Ser 345 CCCC	CTCC ATGG	Asp CGC Arg TGT Cys CTC Leu TCCA	TTT Phe GCC Ala TGA END CCCC AGTC	GGA Gly CTA Leu AGA TGGC TTTC AGAA	TCT Ser 300 GAG Glu 325 CATC	1098 1173 1249 1349 1449 1549
AAT ASN 251 GTT Val 276 TCA Ser 301 GTG Val 326 AGTG TCTG TCGT	Asn GGT Gly AAC Asn GCC Ala SACC GCAA CGGGG	GAAAGLYB ACCGGAAGCCTGAATGG	Gln TAT Tyr GAG Glu CTC Leu CCTC CACG ACAT	GAC Asp 255 ACT Thr 280 GTC Val 305 TTT Phe 330 CCCC AGGT	GGC Gly ATC Ile GGG Gly ACAA GAGA AAGA	TTC Phe CAC His AAA Lys AACT GGAA ACAG	Thr CGG Arg AAC Asn GTC Val CCCTT ACCA CTGG GAGA	CCA Pro TAT Tyr CTG Leu GCTC TGAG AGGC	TGG Trp 260 GTA Val 285 GGC Gly 310 GAA Glu 335 CCCT ATCA	Glu CGC Arg CAT His GAA Glu CAGC GTTC ATCC	Gly CCC Pro GGA Gly AGG Arg CAAC CTGG CACA	Cys CAG GIn GGC Gly AAT ASN TACT AGGC AGTC	Cys GTT Val TAT Tyr TTG Leu CAAT AAGT CGCA	Arg 265 CGG Arg 290 GGG Gly 315 CTC Leu 340 GTGC CCAA	CTA Leu CTC Leu ACA Thr TCCA GAAG	GAAA GIU ACC Thr ATG Met TTGA ACCT	AGA Arg ATC Ile CCA Pro	Thr GAA Glu CAC His CCA Pro CGTT CCCC AGGA CTAA	Leu 270 CAG G1n 295 TGG Trp 320 TCC Ser 345 CCCCCTCTG.	CTT Leu GGC Gly CAC His	Asp CGC Arg TGT Cys CTC Leu TCCA CTGA TGGT ACTG	Ala TTT Phe GCC Ala TGA END CCCC AGTC GAGC	Lys GGA Gly CTA Leu AGA TGGC TTTC AGAA CATT	TCT Ser 300 GAG Glu 325 CATC	1098 1173 1249 1349 1449 1549 1649
AAT ASD 251 GTT Val 276 TCA Ser 301 GTG Val 326 AGTG TCTG GGGA CAGT	Asn GGT G1y AAC Asn GCC Ala GACC GCAA GCGGG TGGG TGGG	GAAA Glu ACA Thr AAG Lys ACCG GAAG CCTG	Gln TAT Tyr GAG Glu CTC Leu CCTC CACG ACAT AAACC	GAC Asp 255 ACT Thr 280 GTC Val 305 TTT Phe 330 CCCC AGGT TACA AAGT TATC	GGC Gly ATC Ile GGG Gly ACAA GAGA AAGA AACC TGGC	Asn TTC Phe CAC His AAA Lys AACT GGAAA ACAG TCAT	Thr CGG Arg AAC Asn GTC Val CCCT ACCA CTGG GAGA	CCA Pro TAT Tyr CTG Leu GCTC TGAG AGGC AATC	TGG Trpp 260 GTA Val 285 GGC Gly 310 GAA Glu 335 CCCT ATCA TCTT ACGC TGTC	CAT HIS GAA Glu	Gly CCC Pro GGA Gly AGG Arg CAAC CTGG CACA ACAT AGCA	CYS CAG GIn GGC Gly AAT AS TACT AGGC AGTC TANA CCAC	Cys GTT Val TAT Tyr TTG Leu CAAT AAGT CGCA AGTT AGAA	Arg 265 CGG Arg 290 GGG Gly 315 CTC Leu 340 GTGC CCAA GAAA CTGA	CTA Leu CTC Leu ACA Thr TCCA CCCT GAAG	GAAAACCTGAAAACCATGGAAAACCATGGAAAACCATGGGAAAACCATGGGAAACCATGGGAAACCATGGAAACCATGGAAACCATGGAAACCATGGAAACCATGGAAACCATGGAAACCATGGAACCATGAACCATGAACCATGAACCATGAACCATGAACCATGAACCATGAACCATGAACCATGAACCATGAACCATGAACCATGAACCATGAACCATGAACAACATGAACAACAACAACAACAACAACAACAACAACAACAACAAC	AGA Arg ATC Ile CCA Pro TTCC GGGG GCTC ACAG	Thr GAA Glu CAC His CCA Pro CGTT CCCC AGGA CTAA TTGG	Leu 270 CAG Gln 295 TGG Trp 320 TCC Ser 345 CCCCC TCTG. AATT	Lys CTT Leu GGC Gly CAC His CTCC ATGG CCAG TCGA GATT	Asp CGC Arg TGT Cys CTC Leu TCCA CTGA TGGT ACTG	Ala TTTT Phe GCC Ala TGA END CCCC AGTC GAGC ATGT GGCA	GGA GGA GGA AGA TGGC TTTC AGAA CATT GGAA	TCT Ser 300 GAG Glu 325 CATC TCCTT CACCA AATTG TGGTC GTTAT	1098 1173 1249 1349 1449 1549 1649 1749
AAT ASD 251 GTT Val 276 TCA Ser 301 GTG Val 326 AGTG TCTG GGGA CAGT	Asn GGT G1y AAC Asn GCC Ala GACC GCAA GCGGG TGGG TGGG	GAAA Glu ACA Thr AAG Lys ACCG GAAG CCTG	Gln TAT Tyr GAG Glu CTC Leu CCTC CACG ACAT AAACC	GAC Asp 255 ACT Thr 280 GTC Val 305 TTT Phe 330 CCCC AGGT TACA AAGT TATC	GGC Gly ATC Ile GGG Gly ACAA GAGA AAGA AACC TGGC	Asn TTC Phe CAC His AAA Lys AACT GGAAA ACAG TCAT	Thr CGG Arg AAC Asn GTC Val CCCT ACCA CTGG GAGA	CCA Pro TAT Tyr CTG Leu GCTC TGAG AGGC AATC	TGG Trpp 260 GTA Val 285 GGC Gly 310 GAA Glu 335 CCCT ATCA TCTT ACGC TGTC	CAT HIS GAA Glu	Gly CCC Pro GGA Gly AGG Arg CAAC CTGG CACA ACAT AGCA	CYS CAG GIn GGC Gly AAT AS TACT AGGC AGTC TANA CCAC	Cys GTT Val TAT Tyr TTG Leu CAAT AAGT CGCA AGTT AGAA	Arg 265 CGG Arg 290 GGG Gly 315 CTC Leu 340 GTGC CCAA GAAA CTGA	CTA Leu CTC Leu ACA Thr TCCA CCCT GAAG	GAAAACCTGAAAACCATGGAAAACCATGGAAAACCATGGGAAAACCATGGGAAACCATGGGAAACCATGGAAACCATGGAAACCATGGAAACCATGGAAACCATGGAAACCATGGAAACCATGGAACCATGAACCATGAACCATGAACCATGAACCATGAACCATGAACCATGAACCATGAACCATGAACCATGAACCATGAACCATGAACCATGAACCATGAACAACATGAACAACAACAACAACAACAACAACAACAACAACAACAAC	AGA Arg ATC Ile CCA Pro TTCC GGGG GCTC ACAG	Thr GAA Glu CAC His CCA Pro CGTT CCCC AGGA CTAA TTGG	Leu 270 CAG Gln 295 TGG Trp 320 TCC Ser 345 CCCCC TCTG. AATT	Lys CTT Leu GGC Gly CAC His CTCC ATGG CCAG TCGA GATT	Asp CGC Arg TGT Cys CTC Leu TCCA CTGA TGGT ACTG	Ala TTTT Phe GCC Ala TGA END CCCC AGTC GAGC ATGT GGCA	GGA GGA GGA AGA TGGC TTTC AGAA CATT GGAA	TCT Ser 300 GAG Glu 325 CATC	1098 1173 1249 1349 1449 1549 1649 1749
AAT ASN 251 GTT Val 276 TCA Ser 301 GTG Val 326 AGTG TCTG GGGA CAGT	ASN GGT Gly AAC ASN GCC ASN GCCAAN GCGGGGCTGGG	GAAAG CCTG ATGG CCTA GCTG	Gln TATT Tyr GAG Glu CTC Leu CCTC CACG ACAT GGTT AACC GAGG	GACC Asp 255 ACT Thr 280 GTC Val 305 TTT Phe 330 CCCC AGGT TACA AAGT TATC GTGG	GGC Gly ATC Ile GGG Gly ACAA GAGA AAGA AAGC ATGGC	TTC Phe CAC His AAA Lys AACTT GGAAA ACAG TCATT TATG GACAC	CGG Arg AAC Asn GTC Val CCCTTACCA CCTGG GGAGA GGAAA GGAAA	CCA Pro TAT Tyr CTG Leu GCTC TGAG AGGC AATC CTTTT AGGC	TGG Trp 260 GTA Val 285 GGC Gly 310 GAA Glu 335 CCCT ATCA TCTT ACGC TGTC	CACT CAGC GTTC ATCC CAGC CAGC CAGC CAGC	GCCC Pro GGA GGY AGG Arg CAAC CTGG CACA ACAT AGCA AGCA GAAC	Cys CAG Gln GGC Gly AAT AS TACT AGGC AGTC TAGC TAGA	Cys GTT Val TAT Tyr TTG Leu CAAT AAGT CGCA AGTT AGAA AACG	Arg 265 CGG Arg 290 GGG Gly 315 CTC Leu 340 GTGC CCAA GAAA CTGA CACA	CTA Leu CTC Leu ACA Thr TCCA GAAG AAAG CTTT CACC	GAAA GTTTGA ATGGAAA TCCC ATGGTCTTTGA	AGA Arg ATC Ile CCA Pro TTCC GGGG GCTC ACAG AACA CCAC	Thr GAA Glu CAC His CCA Pro CGTT CCCC AGGA TTGG ATCC	Leu 270 CAG G1n 295 TGG Trp 320 TCC Ser 345 CCCCC TCTG. AATT GACAAACAAACGA	Lys CTT Leu GGC Gly CAC His CTCC ATGG CCAG GATT TTAT	Asp CGC Arg TGT Cys CTC Leu TCCA CTGA TGGT ACTG CCAT	Ala TTTT Phe GCC Ala TGA END CCCC AGTC GAGC ATGT GGCA GGCA	Lys GGA GGY CTA Leu AGA TGGC TTTC AGAA CATT GGAA TCAG	TCT Ser 300 GAG Glu 325 CATC TCCTT CACCA AATTG TGGTC GTTAT	1098 1173 1249 1349 1449 1549 1649 1749 1849
AAT ASD 251 GTT Val 276 TCA Ser 301 GTG Val 326 AGTG TCTG GGGA CAGT AGGT CACT	ASN GGT GIY AAC ASN GCC AIA GCCAAIA CGGGG CTT CGGA	GAAAG Lys ACCG GAAG CCTG ATGG CCTA GCTGCTGCCTGCCTGCCTGCCTGCCTGCCTGCCTGCCTG	Gln TAT Tyr GAG Glu CTC Leu CCTC CACG ACAT GGTT AACC GAGG CCGT	GACC Aspp 255 ACT Thr 280 GTC Val 305 TTT Phee 330 CCCC AGGT TACA AAGT TATC GTGG	GGC Gly ATC Ile GGG Gly ACAA GAGA AAGA AACC AGGGA AACC AGGA AACC	TTC Phe CAC His AAA Lys AACTGGAAACTGGAATTATGGACAGTTCATGGACAGTTCATGGACAGTTCATGTA	Thr CGG Arg AAC Asn GTC Val CCCTTGG GGAGA GAAA GAAA GAAA GAAAA GAAAA GAAAA GAAAA GAAAA GAAAA GAAAA	TAT Tyr CTG Leu GCTC TGAG AGGC AATC CTTTT AGGC TTGA	TGG Trpp 260 GTA Val 285 GGC Gly 310 GAA Glu 335 CCCT ATCA TCTT ACGC TGTC AAGGT GGAG	Glu CGC Arg CAT His GAA Glu CAGC GTTC ATCC TAAC CCAG CACT GGCC	GIY CCC Pro GGA GIY AGG Arg CAAC CTGG CACA ACAT AGCA GGAAC CACA CACA	Cys CAG GIn GGC Gly AAT ASI TACT AGGC AGTC TAGA CCAC	Cys GTT Val TAT Tyr TTG Leu CAAT AAGT CGCA AGTT AGAA AACG AGAT	Arg 265 CGG Arg 290 GGS Gly 315 CTC Leu 340 CCAA CCAA CTGA CACA GGGA CACA GGGA CACA GGGA CACA GGGA CACA CA	CTA Leu CTC Leu ACA Thr TCCA CCCT GAAG AAAG CTTT CACC	GAAA ACC Thr ATGA Met TTGA ACCT GAAA TCCC ATGG TGTC	AGA Arg ATC Ile CCA Pro TTCC GGGG GCTC ACAG AACA CCAC	Thr GAA Glu CAC His CCA Pro CGTT CCCC AGGA TTGG ATCC TCAT	Leu 270 CAG G1n 295 TGG Trp 320 TCC Ser 345 CCCCC TCTG. AATT GACA' AACA' AAGGA	Lys CTT Leu GGC Gly CAC His CTCC ATGG CCAG GATT TTAT AATT	Asp CGC Arg TGT Cys CTC Leu TCCA TCGA TGGT ACTG CCAT CCAT CAAT	Ala TTTT Phe GCC Ala TGA END CCCC GAGC ATGT GGCA TGCA TCAC	Lys GGA GGA GGA GGA GGA GGA GGA GGA GGA GG	TCT Ser 300 GAG Glu 325 CATC TCCTT CACCA AATTG TGGTC GTTAT CTTCC	1098 1173 1249 1349 1449 1549 1649 1749 1849 1949
AAT AS N 251 CTT Val 276 CAG	ASN GGTTGIY AAC ASN GCC Ala GCAA GGGGA TGGA TGGAA TGGGA TGGA TG	Ile GAA Glu ACA Thr AAG Lys ACCG GAAG CCTG ATGG CCTA GCTG TTGG TTGG	Gln TAT Tyr GAG Glu CTC Leu CCTC CACG ACAT GGTT AACC GAGG CCGT GTCT	GAC Asp 255 ACT Thr 280 GTC Val 305 TTT TAC AGGT TACA AAGT TATC GTGG GGGC CCCT	GGC Gly ATC Ile GGG Gly ACAA GAGA AAGA AACC AGGGA ACCA ACCA CACC	ASA TTC Phe CAC His AAA Lys AACT GGAA ACAG TCAT TATG GACA TTATG	Thr CGG Arg AAC Asn GTC Val CCGT ACCA CTGG GAGA GAAA	CCA Pro TAT Tyr CTG Leu GCTC TGAG AGGC AATC CTTT AGGC TTGA	TGG Trpp 260 GTA Val 285 GGC Gly 330 GAA TCA TCTT ACGC TGTC AAGT GGAG CCCT	Glu CGC Arg CAT His GAA Glu CAGC GTTC ATCC TAAC CCCAG CACT GGCC CTGA	Gly CCC Pro GGA Gly AGG Arg CAAC CTGG CACA ACAT AGCA GGAAC CACA GATG	Cys CAG G1n GGC G1y AAT ASD TACT AGC AGTC TAXA CCAC TAXA GGAG GGAG	Cys GTT Val TAT Tyr TTG Leu CAAT AAGT CGCA AGTT AGAA AACG AGAT ACAG	Arg 265 CGG Arg 290 GGG Gly 315 CTC Leu 340 GTGC CCAA GAAA CTGA CACA GGGA GATC ACTT	CTA Leu CTC Leu ACA Thr TCCA CCCT GAAG ANAG CTTT CACC CAAA	GAAA Glu ACC Thr ATGA ACCT GAAA TCCC ATGG TGTC GGGA ATGT	AGA Arg ATC Ile CCA Pro TTCC GGGG GCTC ACAG AACA CCAC AACT	Thr GAA Glu CAC His CCA Pro CGTT CCCC AGGA TTGG ATCC TCAT	Leu 270 CAG G1n 295 TGG Trp 320 TCC Ser 345 CCCCC TCTG. AATT GACAAACAAACAAACAAACAAACAAACAAACAAACAAAC	Lys CTT Leu GGC Gly CAC His CTCC ATGG CCAG GATT TTAT TATTGAG	Asp CGC Arg TGT Cys CTC Leu TCCA CTGA TGGT ACTG CCAT CAAT	Ala TTTT Phe GCC Ala TGA END CCCC GAGTC GAGC ATGT GGCA TTCAC CCCA	Lys GGA GGA GGAAGGAAGGAAGGAAGGAAGGAAGGAAGG	TCT Ser 300 GAG G1u 325 CATC TCCTT CACCA AATTG TGGTC GTTAT CTTCC CATGC	1098 1173 1249 1349 1449 1549 1649 1749 1849 1949 2049
AAT AS N 251 CTT Val 276 CAG	ASN GGT G1y AAC ASN GCC A1a GCC A1a GCC A1C GCAA GGGG GGGG GGGG GGGG GGGG	Ile GAA Glu ACA Thr AAG LyB ACCG GAAG CCTG ATGG CTTGG CTTGG CTTCG CTTCG	GIn TAT Tyr GAG Glu CTC Leu CCTC CACG ACAT GGTT AACC GAGG CCGT GTCT ATCC	GAC Asp 255 ACT Thr 280 GTC V305 TTT TAT TATC GTGG GGGC CCCT CTCA	GGC Gly ATC Ile GGG Gly ACAA GAGA AAGC AGGGA ACCA ACCA CCCTGAT	TTC Phe CAC His AAA Lys AACTTAGGAAA ACAGGACATTATGGACATTTCT CTGC	Thr CGG Arg AAC Asn GTC Val CCCTTACCA GGAGA GGAAA GAAG GAAG GAAG G	CCA Pro TAT Tyr CTG Leu GCTC TGAG AGGC TTTT AGGC TTGAG GCCA GCCA	TGG Trp 260 GTA Val 285 GGC Gly 310 GAA TCTT ACGC TGTC AAGT GGAG CCCT AA CCCT CTAA	Glu CGC Arg CAT His GAA Glu CAGC GTTC ATCC CAGC CCAGC CCAGC GGCC CTGA	Gly CCC Pro GGA Gly AGG Arg CAAC CTGG CACA ACAT AGCA GAAC CACA GAAG GAAC	Cys CAG G1n GGC G1y AAT AGC AGTC TAXA CCAC TAGG GGAG GCAA TTTC	CYS GTT Val TAT TYI TTG Leu CAAT AAGT CGCA AGTT AAAA AACG AAAAAAAAAA	Arg 265 CGG Arg 290 GGG Gly 315 CTC Leu 340 GTGC CCAA GAAA CTGA GAGA GACA GACA GACA	CTA Leu CTC Leu ACA Thr TCCA AAAG CCCT GAAG CTTC CAAAA	GAAA ACCT Thr ATGA ACCT GAAA TCCC ATGG TGTC GGGA ATGT TTTT	AGA Arg ATC Ile CCA Pro TTCC GGGG GCTC ACAG AACA TTTG	Thr GAA Glu CAC His CCA Pro CGTT CCCC AGGA TTGG ATCC TCAT GAAA AAAG	Leu 270 CAG G1n 295 TGG Trp 320 TCC Ser 345 CCCCC TCTGG AATT GACAA AAGAA AGGA AGGGA AGGGA ACCG	Lys CTT Leu GGC Gly CAC His CTCC ATGG CCAG GATT TTAT TTAT TTGAG GAGC	Asp CGC Arg TGT Cys CTC Leu TCCA TGGT ACTG CCAT CAAT GCAT CAAT GCAT CAAT AAAG GCAT	Ala TTT Phe GCC Ala TGA END CCCC AGTC GAGC ATGT GCCA TCAC CCCA GGGA	Lys GGA GGA GGAAG CCAGGCCAGGA GGAAG CCAGGCAGG	TCTT Ser 300 GAG Glu 325 CATC TCCTT CACCA AATTG TGGTC GTTAT CTTCC CATGC CATGC CAGGCG GGGCG	1098 1173 1249 1349 1449 1549 1649 1749 1849 1949 2049 2149
AAT AS IN 251 CTC AS CAST CAST CAST CAST CAST CAST CA	ASIN GGT G1y AAC ASIN GCC A1a GCCAAN GGGG TGGAN TGGAN TGGGGT TGGAN TGGGT TGGAN TGGGT TAGGT TAGGT	Ile GAA Glu ACA Thr AAG Lys ACCG GAAG CCTG CCTA GCTG TTGG CTTCC CACT	GIn TAT TYr GAG Glu CTC Leu CCTC CACG ACAT AACC GAGG CCGT GTCT ATCC	GAC Asp 255 ACT Thr 280 GTC Val 305 TTT Phee 330 CCCC AGGT TACA AGT TATC GTGG GGGC CCCT CTCA AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA	GGC Gly ATC Ile GGG Gly ACAA AAACC TGGC AGGA ACCA ACCA TGGT TGGT TGGT TGTT	TTC Phe CAC His AAA Lys AACT TATG GAAA TTCTT CTGC TTGAT	Thr CGG Arg AAC Asn GTC Val CCCTTACCA GGAAA GAAAG GAACC AGCT CTCA TGAC	TAT TYY CTG Leu GCTC TGAG AGGC TTTTA AGGC TTTGA GGCA ATC CTTTTA AGGC TTTGA GCCA CTTTTA CTGC TTTGA GCCA CTTTTA CTGCA CTTTTA CTGCA CTTTTA CTGCA CTTTTA CTGCA CTTTTA CTGCA CTGTT CTGTT CTGCA CTGTT CTGTT CTGCA CTGCA CTGTT CTGCA CTGCA CTGTT CTGCA CTGTT CTGCA CTGTT CTGCA	TGG TTPD 260 GTA Val 285 GGC Gly 310 GAA Glu 335 CCCT ATCA TCTT ACGC TGTC AAGT CCCT AAGT CCCT AAGT CCCT CTAA	Glu CGC Arg CAT His GAA Glu CAGC GTTC ATCC CCAG CCAGT GGCC CTGA GGCT TCTA	Gly CCC Pro GGA Gly AGG Arg CAAC CTGG CACA ACAT AGCA GAAC CACA GAAC CACA GATG GAAT TGTC	CAGGGIN GGC GIY AATTACT AGGC AGTC TAXA CCAC GGAG GCAA TTTC CCAAG	CYS GTT Val TAT TYr TTG Leu CAAT AAGTT AAGAA AACG AACAG AACTT ACAG ATTG	Arg 265 CGG Arg 290 GGG Gly 315 CTC Leu 340 GTGC CCAA ACTGA CTGA CAGA ACTGA CAGA ACTGA CACTGA	CTA Leu CTC Leu ACA Thr TCCA AAAG CCTT CACC CAAAA CCAGGGGGGGCCTTT	Glu GAA Glu ACC Thr ATG Met TTGA ACCT GAAA TCCC GGGA ATGT TTTT CTCT	AGA ATG Ile CCA Pro TTCC GGGG GCTC ACAG ACAG ACAC TTTG TGTG GTCA	Thr GAA Glu CAC His CCA Pro CGTT CCCC AGGA ATCC TCAT GAAA AAAG CTTC	Leu 270 CAG G1n 295 TGG Trp 320 TCC Ser 345 CCCCC TCTG AATT GACA' AACA' AACA' AAGA GCCT ACCG CCCCC CCCC CCCC CCCC CCCC CCCC CCCC	CTCC GGCY CAC His CTCC ATGG CCAG GATT TTAT AATT TGAG GATG GAT	CGC Arg TGT Cys CTC Leu TCCA TCGAT ACTG CCAT CAAT CAAT GAAT GGCT	Ala TTT Phe GCC Ala TGA END CCCC AGTC GAGC ATGT TCAC GCCA GCCA CCCCA CCCCA CCCCA CCCCA	Lys GGA G1y CTA Leu AGA TGGC TTTC AGAA CATT GGAA TCAG CAAG CA	TCTT Ser 300 GAG Glu 325 CATC TCCTT CACCA AATTG TGGTC GTTAT CTTCC CATGC CAGGCG AGGCA	1098 1173 1249 1349 1449 1549 1649 1749 1849 1949 2049 2149 2249
AGTO TCA GGGA AGTO TCAC TCAC TCAC TCAC TCAC TCAC TCAC TC	ASIN GGT G1y AAC ASIN GCC A1a GCCAAA GGGGG GGGA GGGGG GGGG GGGG G	Ile GAA Glu ACA Thr AAG Lys ACCG GAAG CCTG GCTG CCTA GCTGC CTTCC CACT TGAT	GIn TAT TYr GAG Glu CTC Leu CCTC CACG ACAT AACC GAGG CCGT GTCT ATCC GTAA	GAC Asp 255 ACT Thr 280 GTC Val 305 TTT Phee 330 CCCC AGGT TACA AGT CCCC CTCA AGAG GGTC GTCG GGGC CTCA AGAG GTTG	GGC GTy ATC ITE GGG GTy ACAA GAGA AACC TGGC AGGA ACCA TGGC TGAT GTCT ACCC	TTC Phe CAC His AAA Lys AACT TATG GAAA TTCT CTGC TGAT CAAG	Thr CGG Arg AAC Asn GTC Val CCCTTGG GAGA GAAA GAAA GAAA GAAC CTCA GCTCA GGGA	TAT TYF CTG Leu GCTC CTTTAGGGCAAATC CTTTGAGGCAATCAGCCTTTGAGGCACTTTGAGGCATTGAGGCATTGAGGCATTGAGGCATTGAGGAGAGGAGAGAGA	TGG Trpp 260 GTA Val 285 GGC Gly 310 GAA Glu 335 CCCT ATCA TCTT ACGC TGTC AAGT GGAG CCCT CTAA AGT GCCA AGTG	Glu CGC Arg CAT His GAA Glu CAGC GTTC ATCC CCAG CACT GGCC CTGA GGCT TCTA AAAAG	GLY CCC Pro GGA GLY AGG ALG CTGG CACA ACAT AGCA GAAC CACA GATG GAAC CACA AGAT GATG GAAC CACA AGAT GATG CACA AGAT AGCA AGCA	CAGGGIN GGC GIY AAT AGGC AGTC TAGG GGAG GCAG TTTC CAAGAAAAAAAAAA	CYS GTT Val TAT TYr TTG Leu CAAT AAGT AGAA AACG AGAT ACAG ATTG TCAG TCA	Arg 265 CGG Arg 290 GGG Gly 315 CTC Leu 340 CCAA ACTGA CACA GGGA CACA GGGA CTGA CTGA CACA GGGA GATC GCTTC GGGAG	CTA Leu CTC Leu ACA Thr TCCA GAAG CTTT CACC CAAA CCAG GAGG GAGG	GAAA Glu ACC Thr ATG Met TTGA ACCT GAAA TCCC ATGG GGGA ATGT TTTT TCTCT	AGA ATC Ile CCA Pro TTCC GGGG GCTC ACAG AACA CCAC TTTTG GTCG GTC	Thr GAA Glu CAC His CCA Pro CGTT CCCC AGGA TTGG GAAA AAAG CTTC TTAT	Leu 270 CAG G1n 295 TGG Trp 320 TCC Ser 345 CCCCC TCTG. AATT GACA' AACA' AACA' AACA' AACA' CCCT CCCT C	CTT Leu GGC Gly CAC His CTCC ATGG GATT TTAT TGAG GATG GATG GATG	CGC Arg TGT Cys CTC Leu TCCA TGGT ACTGA TGGT ACTG CCAT TCAAT GCT TGAG TGAG	Ala TTT Phe GCC Ala TGA END CCCC AGTC GAGC ATGT GCA CCCA CCCA CCCA	GGA AGA CCAGG GAAGA TATA	TCT Ser 300 GAG Glu 325 CATC TCCTT CACCA AATTG TGGTC CATGC CATG	1098 1173 1249 1349 1449 1549 1649 1749 1849 2049 2149 2249 2349
AGTO TCA TCA TCA TCA TCA TCA TCA TCA TCA TCTC TCTC TCTC TCTC TCTC TCTC TCTC TCTC	ASIN GGT GCA ASIN GGCC ASIN GGGG ATGC GCATT	Ile GAA Glu ACA Thr AAGGLys ACCG GAAGCCTG CCTA GCTGC CTTCG CACT TGAT CCTAG	Gln TATTYT GAG Glu CCTC CACG ACAT GGTT AACC GAGG CCGT GTCT ATCC GTAA	GAC Asp 255 ACT Thr 280 GTC Val 305 TTT THE 330 CCCC AGGT TACA AGGT CCCT AGAG GGGC CTCA AGAG GTTG ATTT	GGC GTY ATC ITE GGG GTY ACAA GAGA AAGA AACC AGGA ACCA ACCA GGTT ACCC GGTT	TTC Phe CAC His Lys AAA Lys AACTTATGGAA. ACAGTTATGTATCTTGCTGATCTATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGA	Thr CGG Arg AAC ABn GTC Val CCCTTGG GAAA GAAAG CTCA ACCC AGCT CTCA CTGAC GGAA	CCA Pro TAT Tyr CTG Leu GCTC TGAG AGGC AATC CTTTA AGGC ACTTTA GGCA GCCA CTTTA AGCA ACTG ACTG	TGG Trp 260 GTA Val 285 GGC Gly 310 GAA Glu 335 CCCT ATCA TCTT ACGC TGTC AAGT GGAG CCCT AAGT GGAG GCCA AGT GGCA AGT GGCA AGT GGCA	Glu CAGC GAAGC GATC CAGC GTTC ATCC TAAC CCAG GCC CTGA GGCT TCTA AAAG	Gly CCC Pro GGA Gly AGG Arg CAAC CTGG CACA ACAT AGCA GAAC CACA GAAG GAAT GAAC CACA GATG GAAT TGTC AGCC TTGTC TTGT TT	Cys CAG Gln GGC Gly AAT ASD TACT TACT TACT TACA CCAC TAGA GCAA TTTC CCAC GGAG GCAA TTTC CCAC GCAA GCAA	Cys GTT Val TAT Tyr TTG Leu CAAT AGA AAGT AGAA AACG TCAG TCAG TCAG	Arg 265 CGG Arg 290 GGG GGly 315 CTC Leu 340 GTGC CCAA GAAA CTGA CACA GGAA CACA GGTGC GGAG GGTGC GGAG	CTC Leu CTC Leu ACA Thr TCCA CCCT GAAG CTTT CCACA CCAG GAGG CTTT TTCCA	GAAA ACCT ATGAAA TCCC ATGGGGAAA TCTTTTTTTTTT	AGA ATC Ile CCA Pro TTCC GGGG GCTC ACAG TTTTG GTGG GTCA CCACT TTTG GTCA CCACT ACCT AC	Thr GAA Glu CAC His CCA Pro CGTT CCCC AGGA TTGG ATCC TCAT GAAA AAAG CTTA TAAT	Leu 270 CAG GInn 295 TGG Trp 320 TCC Ser 345 CCCCC TCTG. AATT GACA' AACA' AACA' AACA' ACGC CCCT ACCA CCCTA CCCTA CCCTA CCCTA CCCTA	CTT Leu GGC Gly CAC His CTCC ATGG GATT TTAT TTAT TGAG GATG GATG	CTC Leu TCCA TCGTTAAAAG TGGTTAAAAAG TGGTTGAAAAAG TGGGTTGAAAAAG TGGGTTGAAAAG TGGGTTGAAAAAG TGGGTTGAAAAAG TGGGTTGAAAAAG TGGGTTGAAAAAG TGGGTTGAAAAAG TGGGTTGAAAAAG TGGGTTGAAAAAG TGGGTTGAAAAAAG TGGGTTGAAAAAAG TGGGTTGAAAAAAAA	Ala TTTT Phe GCC Ala TGA END CCCC AGTC GAGC ATGT GGCA TCAC CCCA GCGA CTTC CCCA AGTC CCCA AGTC CCCA	GGA CAAGA CA	TCT Ser 300 GAG Glu 325 CATC TCCTT CACCA AATTG TGGTC CATGC CATG	1098 1173 1249 1349 1449 1549 1749 1849 1949 2049 2149 2249 2349 2449
AAT AS IN 251 CTC TCC TCC TCC TCC TCC TCC TCC TCC TC	ASIN GGT GIY AAC ASIN GGC AIA CGAMATGGA CGAMTGGA CGATTCA CATTTTGTTGTTTGTTTGTTTGTTTGTTTGTTTGTTTGTT	Ile GAA Glu ACA Thr AAGG Lys ACCG GAAG CCTG ATGG CTTGG CTTCG CTTCT CTAG CTTGG CTTCC CACT	Gln TAT Tyr GAG Glu CCTC CACG ACAT AACC GAGG CCGT TATCC GTAA TTTG CTAT TATT	GAC Asp 255 ACT Thr 280 GTC Val 305 TTT THP 330 CCCC AGGT TACA AGT GTGG GGGC CCT CTCA GGGG GTTG CCCC AGAG GTTG CCCC AGAG AGAG	GGC GTY ATC ITE GGG GTY ACAA AGAA AACC AGGA ACCA CCCC TGAT GTCT ACCC GGTT	TTC Phe CAC His Lys AAA Lys AACTTATGGAA. ACAGTTATGTATTCTCTGC TGAT CAGAGGATAAGGCATATAG	Thr CGG Arg AAC Asn GTC Val CCCTT ACCA GGAA GAAC GTCA GGAA TCTA GGGAA GTCTA	CCA Pro TAT Tyr CTG Leu GCTC TGAG AGGC AATC CTTTA GGCA GCCT TTGA GCA ACTG ACTG	TGG Trpp 260 GTA Val 285 GGC G1y 310 GAA G1u 335 CCCT ATCA TCTT ACGC AAGT GGAG CCCT CTAA GCCA AGTG GGCA AGTG GGTA AGTG	Glu CGC Arg CAT His GAA Glu CAGC GTTC ATCC TAAC CCAG GCC TTCT AAAG GCC CCTGA GGCT CCTGA GGCC GGCG GGCG	Gly GGA GGA GGA AGG CAAC CTGG CACA AGCA GAAC CACA GAAC GAAC CACA GAAC	CYS CAG GIN GGC GIY AAT AGC AGTC TACT AGGC AGTC CCAC TAGC GGAG GCAA TTTC CCAC GAAG AAAC GAGG	Cys GTT Val TAT Tyr TTG Leu CAAT AGAA AACG AGAT ACAG ATTG TCAG TCAG	Arg 265 CGG Arg 290 GGG Gly 315 CTC Leu 340 GTGC CCAA CACA GGAA GATC GCTT GCTT GCTT	CTC Leu CTC Leu ACA Thr TCCA CCCT GAAG CTTT CACC CAAA CCAG GAGG CTTT TTCC CACAT CCACT CACAT CCACT CACAT CCACT CCACAT C	Glu GAA Glu ACC Thr ATGA Met TTGA ACCT GAAA TCCC GGGA ATGT TTTT TGAC GGACT AACA	Pro AGA Arg ATC Ile CCA Pro TTCC GGGG GCTC ACAG AACT TTTG GTCG GTCA CCAAT	Thr GAA Glu CAC His CCA Pro CGTT CCCC AGGA ATCC TCAT GAAA AAAG CTTC TAAT TAAT	Leu 270 CAG G1n 295 TGG G1r 320 TCC Ser 345 CCCCC CCTCGAATT AACAAAGAAGGAGGCCTAACCAACCATAACCGGGAACCCTTGGGGAACCCTTGGGGAACCTTGGGGAACCTTGGGGAACCTTGGGGAACCTTGGGGAACCTTGGGGAACCCTTGGGGAACCTTGGGGAACCTTGGGGAACCTTGGGGAACCCTTGGGGAACCCTTGGGGAACCCTTGGGGAACCCTTGGGGAACCCTTGGGGAACCCTTGGGGAACCCTTGGGGAACCCTTGGGGAACCCTTGGGGAACCCTTGGGGAACCCTTGGGGAACCCTTGGGGAACCCCTTGGAACCCCTTGGAACCCCTTGGAACCCCTTGGAACCCCTTGGAACCCCTTGGAACCCCTTGGAACCCCTTGGAACCCCTTGGAACCCCTTGGAACCCCTTGGAACCCCTTGGAACCCCTTGGAACCCTTGGAACCCCTTGGAACCCCTTGGAACCCTTTGGAACCCTTTGGAACCCTTTGGAACCCTTTGGAACCCTTTGGAACCCTTTGGAACCCTTTGGAACCCTTTGAACCCTTTGGAACCCTTTGAACCCTTTGAACCCTTTGAACCCTTTGAACCCTTTGAACCCTTTGAACCCTTTGAACCCTTTGAACCCTTTGAACCCTTTTGAACCCTTTTGAACCCTTTTGAACCCTTTTGAACCCTTTTGAACCCTTTTGAACCCTTTTGAACCCTTTTGAACCCTTTTTGAACCCTTTTTTTT	CTTCC CTCCATGATTTTATTTGAGGATGATTATATATATATATAT	CGC Arg TGT Cys CTC Leu TCCA TCGA TGGT ACTG ACTG ACTG ACTG ACTG	Ala TTTT Phe GCC Ala TGAD CCCC AGTC GAGC ATGT GGCA TCAC CCCA GGGA CCTC AGGT CCCA GGGA CCCCA GGGA	Lys GGA G1y CTA Leu AGA TGGC TTTC AGAA CATT GGAA CCAG GAAG CCAG ATAA TTATT CACC	TCTT Ser 300 GAG G1u 325 CATC TCCTT CACCA AATTG TGGTC CATGC CATGATAT CCCTA CGGAGA	1098 1173 1249 1349 1449 1549 1649 1749 1849 2049 2149 2249 2349 2449 2549
AGTO TCA	ASIN GGC ASIN GGC ASIN GGG ASIN GGGG ATGG ATGG GGC ATT FAGG GGG ATT FAGG GGGG ATT FATT GGGGG GCA TATT GGGGGG ATGG AGGG GCA TATT GGGGGGG ATGG AGGGGGGGGGG	Ile GAA Glu ACA Thr AAGGLys ACCG GAAG CCTG CCTA GCTG CTTCG CACT TGAT CCTAG CCCAA	GIn TATTY GAG GIU CCTC CACG ACAT GGTT AACC GTCA ATTCC GTAA TTTG CTATT AAAT	GAC Asp 255 ACT Thr 280 GTC Val 305 TTT TAT 330 CCCC AGGT TACA AAGT TATC GTGG GGCC CCTC AGAG GTTG GTT	GGC GTY ATC ITE GGG GTY ACAA GAGA AAGA AACC TGGC TGAT GTCT ACCC GGTT TTTC CGAT	TTC Phe CAC His Lys AAAA Lys AACTTATGGAAACAGTTCTTCTCTGCTATTCTGATCAAGGCATAAAAATAGTCCA	Thr CGG Arg AAC Asn GTC Val CCCTT ACCA GGAA ACCC GGGAA TCTA GGGAA TCTA GTGAC GTCA TTGAC	TILE CCA Pro TAT Tyr CTG Leu GCTC TGAG AGGC AATC CTTTA AGGC ATTGA GCA ACTG ACTG	TGG Trpp 260 GTA Val 285 GGC Gly 310 GAA Glu 335 CCCT ATCA TCTT ACGC CTAA GGAG GCCA AGTG GCCA AGTG GCCA AGTG GCTA AGTT TTTC	Glu CAGC GAAG GAA Glu CAGC GTTC ATCC TAAC CCAG GGCC CTGA AAAG GCCAG GGCG AACC	Gly GGA GGA AGG CCTGG CACA AGCA GATG GAAC CACA GATG GAAC CACA GATG GAAC AGCA	Cys CAGGGIN GGC GIY AATT AGGC AGTC TAAA CCAC TAGG GGAG GCAA TTTC CAAG GAAG AAAC GAGG CACA	Cys GTT Val TAT Tyr TTG Leu CAAT AGAA AACG ATTG AGAA ATTG TCAT TCAT	Arg 265 CGG Arg 290 GGG GGly 315 CTC Leu 340 GTGC CCAA CTGA CTGA CACA GGGA GATC GCTTC GCTTC GCTC GGGAG GCCCAA	CTA Leu CTC Leu CTC Leu ACA Thr TCCA CCCT GAAG CTTT CCAGGGGCTTT TTCC ACAT GGGTC TCCAGGTC TCCCACAT CCCAGGTC CCACAT CCCAGGTC CCCACAT CCCAGGTC CCCACAT CCCAGGTC CCCACAT CCCAGGTC CCCACAT CCCAGGTC CCCACAT	GAAA ACCT ATGA ACCT GAAA TCCC ATGG TTTTT TCTCT TGAC GACT AACA	Pro AGA Arg ATC Ile CCA Pro TTCC GGGG GCTC ACAG AACA TTTG GGTCA CCAAT GGGG TTGGG	Thr GAA Glu CAC His CCA Pro CGTT CCCC AGGA TTGG ATCC TCAT GAAA AAAG CTTAA TTAAT TAAT	Leu 270 CAG G1n TGG G1n TTCC Ser 345 TCC CCCC TCTG. AATT GACA AAGA ACGA CCCT ACTA CCCT CCCT	CTT Leu GGC Gly CAC His CTCC ATGG CCAG GATT TATT TGAG GAGC GATG ATAA ATAG GGGG ACAA	Asp CGC Arg TGT TGT Cys CTC Leu TCCA CTGAT CCAT ACTG CCAT CAAT GGCT CCACT CCACT CCACA AGG	Ala TTTT Phe GCC Ala TGA END CCCC AGTC GAGC ATGT GCCA GCCA CCCA GCCA AGTC CCCA AGT CCCC GGTA ATGA	Lys GGA G1y CTA Leu AGA TGGC TTTC AGAA CATT GGAA CCAG ATAA TATT CCACC TTAT AGGA	TCT Ser 300 GAG Glu 325 CATC TCCTT CACCA AATTG TGGTC CATGC	1098 1173 1249 1349 1449 1549 1649 1749 1849 2049 2149 2249 2349 2449 2549 2649
AGTO TCA	ASIN GGC ASIN GGC ASIN GGG ASIN GGGG ATGG ATGG GGC ATT FAGG GGG ATT FAGG GGGG ATT FATT GGGGG GCA TATT GGGGGG ATGG AGGG GCA TATT GGGGGGG ATGG AGGGGGGGGGG	Ile GAA Glu ACA Thr AAGGLys ACCG GAAG CCTG CCTA GCTG CTTCG CACT TGAT CCTAG CCCAA	GIn TATTY GAG GIU CCTC CACG ACAT GGTT AACC GTCA ATTCC GTAA TTTG CTATT AAAT	GAC Asp 255 ACT Thr 280 GTC Val 305 TTT TAT 330 CCCC AGGT TACA AAGT TATC GTGG GGCC CCTC AGAG GTTG GTT	GGC GTY ATC ITE GGG GTY ACAA GAGA AAGA AACC TGGC TGAT GTCT ACCC GGTT TTTC CGAT	TTC Phe CAC His Lys AAA Lys AACT TATG GAAA TTCT CTGC TGAT CAAG GCAT ATAG TCCA	Thr CGG Arg AAC Asn GTC Val CCCTT ACCA GGAA ACCC GGGAA TCTA GGGAA TCTA GTGAC GTCA TTGAC	TILE CCA Pro TAT Tyr CTG Leu GCTC TGAG AGGC AATC CTTTA AGGC AATC CTTTA AGCA ACTG ACTG	TGG Trpp 260 GTA Val 285 GGC Gly 310 GAA Glu 335 CCCT ATCA TCTT ACGC CTAA GGAG GCCA AGTG GCCA AGTG GCCA AGTG GCTA AGTT TTTC	Glu CAGC GAAG GAA Glu CAGC GTTC ATCC TAAC CCAG GGCC CTGA AAAG GCCAG GGCG AACC	Gly GGA GGA AGG CCTGG CACA AGCA GATG GAAC CACA GATG GAAC CACA GATG GAAC AGCA	Cys CAGGGIN GGC GIY AATT AGGC AGTC TAAA CCAC TAGG GGAG GCAA TTTC CAAG GAAG AAAC GAGG CACA	Cys GTT Val TAT Tyr TTG Leu CAAT AGAA AACG ATTG AGAA ATTG TCAT TCAT	Arg 265 CGG Arg 290 GGG GGly 315 CTC Leu 340 GTGC CCAA CTGA CTGA CACA GGGA GATC GCTTC GCTTC GCTC GGGAG GCCCAA	CTA Leu CTC Leu CTC Leu ACA Thr TCCA CCCT GAAG CTTT CCAGGGGCTTT TTCC ACAT GGGTC TCCAGGTC TCCCACAT CCCAGGTC CCACAT CCCAGGTC CCCACAT CCCAGGTC CCCACAT CCCAGGTC CCCACAT CCCAGGTC CCCACAT CCCAGGTC CCCACAT	GAAA ACCT ATGA ACCT GAAA TCCC ATGG TTTTT TCTCT TGAC GACT AACA	Pro AGA Arg ATC Ile CCA Pro TTCC GGGG GCTC ACAG AACA TTTG GGTCA CCAAT GGGG TTGGG	Thr GAA Glu CAC His CCA Pro CGTT CCCC AGGA TTGG ATCC TCAT GAAA AAAG CTTAA TTAAT TAAT	Leu 270 CAG G1n TGG G1n TTCC Ser 345 TCC CCCC TCTG. AATT GACA AAGA ACGA CCCT ACTA CCCT CCCT	CTT Leu GGC Gly CAC His CTCC ATGG CCAG GATT TATT TGAG GAGC GATG ATAA ATAG GGGG ACAA	Asp CGC Arg TGT TGT Cys CTC Leu TCCA CTGAT CCAT ACTG CCAT CAAT GGCT CCACT CCACT CCACA AGG	Ala TTTT Phe GCC Ala TGA END CCCC AGTC GAGC ATGT GGCA TCAC CCCA GGGA CTTC CCCA AGT CCCC GGTA ATGA	Lys GGA G1y CTA Leu AGA TGGC TTTC AGAA CATT GGAA CCAG ATAA TATT CCACC TTAT AGGA	TCTT Ser 300 GAG G1u 325 CATC TCCTT CACCA AATTG TGGTC CATGC CATGATAT CCCTA CGGAGA	1098 1173 1249 1349 1449 1549 1649 1749 1849 2049 2149 2249 2349 2449 2549 2649
226 AAT ASD 251 GTT Val 276 TCA Ser 301 GTG Val 326 AGTG TCTG GGGA CAGT TTCA TTCA TTCA TTCA	ASIN GGC ASIN GGC ASIN GGG ASIN GGGG ATGC GGATTAGGGGCATTTG	Ile GAA Glu ACA Thr AAGG Lys ACCG GAAG CCTG CCTA GCTG CTTCG CACT TGAT CCTAG CCTCA CCTCA CCTCA	GIn TATTY GAG GIU CCTC CACG ACAT GGTT AACC GTCA TTTG TTTTG TTTTT AAAT TCCA	GAC Asp 255 ACT Thr 280 GTC Val 305 TTT 330 CCCC AGGT TACA AAGT TATC GTGG GGCC CTCA AGAG GTTG TTT TTC TTC TTC TTC TTC TTC TTC T	GGC GTY ATC ITE GGG GTY ACAA GAGA AAGA AACC TGGC AGGAT TTTC GGTT TTTC CGAT	TTC Phe CAC His Lys AAA Lys AACT TATG GAAA TTCT CTGC TGAT CAAG GCAT ATAG GCAT GCA GGCA G	Thr CGG Arg AAC Asn GTC Val CCCTT ACCA GGAA GGAA TCTA GGAA TCTA GTGAC GGAA TCTA GTGAC ACCC ACCC	CCA Pro TAT Tyr CTG Leu GCTC TGAG AGGC AATC CTTTA GGCA GCAT GTAG ACTG ACT	TGG Trpp 260 GTA Val 285 GGC Gly 310 GAA Glu 335 CCCT ATCA TCTT ACGC CTAA GGAG GCCA AGTG GCCA AGTG GCCA AGTG AGTG AGTA AGTG AGTA AGTG AGAA	Glu CAGC GATG CATT His GAA Glu CAGC GTTC CAGC CCAG GCC TAAC CCAG GGCC CTGA AAAG GCCAG GGCG AACCA ACCA	Gly GGA GGA AGG CCTGG CACA AGCA GATG GAAC CACA GATG GAAC GATG GAAC GATG GAAC GATG GAAC GACA GATG GAAC GACA GATG GAAC GACA GATG GAAC GACA GATG GACA GCC GAGA ACCA GCC GAGA GCC GAGA GCC GCC GAGA GCC GCC GCC GAGA GCC GCC	Cys CAGGIN GGC GIY AAT AGGC AGTC TACAC TACAC GGAG GCAA TTTC CAAG AAAA CCAC CAAG AAAA CCAC CAAG CAAG AAAA CCAC CAAG CACA TTTC CAAG CACA TTTC CAAG CACA TTTC	Cys GTT Val TAT Tyr TTG Leu CAAT AGA AACG AACG TCAAT CAAT ACAG ACC TCAAT ACC CTCA	Arg 265 CGG Arg 290 GGG GGly 315 CTC Leu 340 GTGC CCAA CTGA ACTGA CACA GGGA GATC GGGA GCCCAA GCCCAA CCCCAA	CTA Leu CTC Leu CTC Leu ACA Thr TCCA CCCT GAAG CTTT CCAGGGGGCTTT CCAGGTCTTCCCACAT CCAGGTCTTTCCCTACAT CCCTTTCCCTACAT CCCTTTCCCTACAT CCCTTTCCCTACAT CCCTTTCCCTACAT CCCTTTCCCTACAT CCCTTTCCCTACAT CCCTTTCCCTACAT CCCTTTCCCTACAT CCCTTTCCCTTTCCCTACAT CCCTTTCCCTTTCCCTACAT CCCTTTCCCTTTCCCTTTCCCTTTCCCTTTCCCTTTCCCTTTCCCTTTCCCTTTCTTTCCTTTCTTTCCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTT	GAAA ACCT GAAAA TCCC ATGG GGGA ATGT TTTT TGAC GGGCA ATGT TTTT TGAC GACT AACA AGGT	AGA ATC Ile CCA Pro TTCC GGGG GCTC ACAG TTTGG GTCA ACAT TTTGG GTCA GTCA	Thr GAA Glu CAC His CCA Pro CGTT CCCC AGGA TTGG ATCC TCAT GAAA AAAG AAGA	Leu 2770 CAG G1n 295 TGG G1n 320 TCC Serr 345 CCCCC TCTG. AATT GACA' AACA' AACA' ACGG CCCT ACTA CCTT GGGGA CCTT TGGGA CCATT	CTT Leu GGC Gly CAC His CTCC ATGG CCAG GATT TATT TGAG GAGC GATG ATAA ATAG GGGG ACAA	CGC Arg TGT CYS CTC Leu TCCA CTGAT CCAT CCTT C	Ala TTTT Phe GCC Ala TGA END CCCC AGTC GAGC ATGT GGCA CCCA GGCA CCCA GGGT AGGT CCCC GGTA ATGA	Lys GGA G1y CTA Leu AGA TGGC TTTC AGAA CATT GGAA CCAG ATAA TCAC TTAT AGGA AGAC	TCT Ser 300 GAG Glu 325 CATC TCCTT CACCA AATTG TGGTC CATGC	1098 1173 1249 1349 1449 1549 1649 1749 1849 2049 2149 2249 2349 2449 2549 2649 2749
AGTO TCA	ASIN GGC ASIN GGGG ATGG GCATTGGGGGGTTTTGGGGGTTTTGGGGGTTTTGGGGGTTTTGGGG	Ile GAA Glu ACA Thr AAGG Lys ACCG GAAG CCTG GCTT TGGT TTGGT TTGGT CTAG CCTAG CCTCT GGAG CCTCT GGAG CCTCT GGAG	GIn TATTY GAG GIU CCTC Leu CCTC CACG ACAT AACC GTCA ATTG GTTAT TTTG AAAT TTCA	GAC Asp 255 ACT Thr 280 GTC Val 305 TTT TH 330 CCCC AGGT TACA AGT TATC GTGG GGCC CTCA AGAG GTTG CCCA TTT TT T	GGC GTY ATC GGG GTY ACAA AGAA AACC TGGC AGGAT TTTC GGTT TTTC ACCT ACCT	TTC Phe CAC His Lys AAAT Lys AACT TATG GAAA TTCT CAAG GCAT TATG GCAT TTCT CAAG GCAT TTCA TTCA	Thr CGG Arg AAC Asn GTC Val CCCTT ACCA GGAA GGAA TCTA GGGAA TCTA GTGAC GGAA TCTA GTGAC GGAA TTGAC ACCT ACCT	CCA Pro TAT Tyr CTG Leu GCTC TGAG AGGC AATC CTTTA AGGC AATC CTGT GTAG ACTG TCAC GCAA TCAC GGCA TGGC	TGG Trp 260 GTA Val 285 GGC Gly 3310 GAA Glu 335 CCCT ATCA TCTT ACGC TGTC AGGA GCCA AGTG GCCA AGTG GCCA AGTG AGTG	Glu CAGC GATG GAA Glu CAGC GTTC CAGC CCAGC CCAGC CCTGA AAAG GCCA GGCG AACCA ACCA	Gly GGA GGA AGG CAAC CTGG CACA AGCA GATG GAAC CACA GATG GAAC CACA GATG GAAC TTGC GAGA ACCA GATG GAGA CTTAC	Cys CAGGIN GGC GIY AATT AGGC AGTC TATA CCAC CGAGG GCAA TTTC GGAGG AAGA AAAC GAGG CACA CCAAC CCAAC	Cys GTT Val TAT Tyr TTG Leu CAAT AGAA AACG TCAG ATTG TCTT TCTT CAAT ACTC CCTA	Arg 265 CGG Arg 290 GGG GGly 315 CTC Leu 340 GTGC CCAA CACA GGAA CCCCAA GGGAG GCCCAA GGGAG GCCCAA CCCAA ACCACA CACA ACCAA ACCAAA ACCAA ACCAAA ACCAA ACCAA ACCAA ACCAA ACCAA ACCAA ACCAA ACCAA ACCAA ACCAAA ACCAAA ACCAAA ACCAAA ACCAAA ACCAAA ACCAAA ACCAAA ACCAAA ACCAAAA ACCAAA ACCAAA ACCAAA ACCAAA ACCAAA ACCAAA ACCAAA ACCAAA ACCAAAA ACCAAA ACCAAAA ACCAAAA ACCAAAA ACCAAAA ACCAAAA ACCAAAA ACCAAAAA ACCAAAAA ACCAAAAA ACCAAAAAA	CTA Leu CTC Leu CTC Leu ACA Thr TCCA CCCT GAAG CTTT CCAAAA CCAG GAGG CTTT TCCAAAA CCAG GAGG CTTT CCCAACAT CCCACAC CCAACAT CCCACAC CCAACAT CCCACAC CCCACAC CCCAACAT CCCCCACAC CCCACAC CCCACACAC CCCACAC CCCACAC CCCACAC CCCACAC CCCACACAC CCCACACAC CCCACACAC CCCACACAC CCCACACAC CCCACACACACAC CCCACACACACACACACACACACACACACACACACACACA	GAAA ACCT GAAA TCCC ATGG GGGA ATGT TTTT TGAC GACT AACA AGGT TTTT CTCT TGAC CTCT CTCT	AGA ATC Ile CCA Pro TTCC GGGG GCTC ACAG TTTG GGGG GTCA ACAT TTTG GGGGGTCA AGAT AGAT	Thr GAA Glu CAC His CCA Pro CGTT CCCC AGGA TTGG GAAA AAAG CTAA AAAG AATG GCA CTAA CGCA CTAA CGCA CTAA CGCA CTAA CGCA CTAA CGCA CG	Leu 2770 CAG G1n 295 TGG G1r 320 TCC Ser 345 CCCCC TCTG. AATT GACA' AACA' AACA' ACGG CCCT ACTA CCCT TCTG. CCCCT TCTG. CCCT TCTG.	CTT Leu GGC G1y CAC His CTCC ATGG CCAG GATT TATT TGAG GAGC GATG GAGG ATAA ATAG GGGG ACAA GGGC ACCA	CGC Arg TGT Cys CTC Leu TCCA TCTGA TGGTT ACTG GCAT CAAT GGCT CAAT AAAG GTT CAAT CAA	Ala TTTT Phe GCC Ala TGA END CCCC AGTC GAGC ATGT GGCA CCCA GGCA CCCA GGGA CTCC GGGT AAGT CCCC GGTA AAGG CCCG AAGGA AAGGA	Lys GGA GIY CTA AGA TGGC TTTC AGAA CATT GGAA TCAG GAAG CCAGA TATA AGGA AGAC TTATA AGGA AGAC TTATA AGGA AGAC TTATA AGGA AGAC TTATA TGGA AGAC TTATA AGGA AGAC TTATA TGGA AGAC TTATA AGGA AGAC TTATA TGGA AGAC TGGA TGGA	TCT Ser 300 GAG G1u 325 CATC TCCTT CACCA AATTG GTTAT CTTCC CATGC ACCCC GGGCG AGGCA ATAAT CCCTA GGAGA ATAAT CCCTA GGAGA ATCG TTCCT TTCTC TTCTCA	1098 1173 1249 1349 1449 1549 1649 1749 1849 2049 2149 2249 2349 2449 2549 2649 2749 2849
AGTO TCA	ASIN GGC ASIN GGGG ATTG FAGGGG ATTTG GGGG ATTTTG GGGG ATTTTG GGGG ATTTTG GGGG ATTTTG GGGG ATTTTG GGGG	Ile GAA Glu ACA Thr AAGGLYB ACCGGAAGCCTAGGCCTAGGCTTGGCCTAGGTCTCCCAA CCTCGGAGCCTCTCCCAA	Gln TATTY GAG Glu CCTC CACG CACGT AACC GGTT ATTG ATTG CTAT TATT AAAT TCCA	GAC Aspp 255 ACT Thr 280 GTC Val 3305 TTT 330 CCCC AGGT TACA AAGT TATC CCCA AGAG GTCG GTC	GGC GTY ATC GGG GTY ACAA AGAAAACC TGGC AGGA ACCA TGGTT TTTC CGAT TTTC ACAT ACAT	TTC Phe CAC His AAA Lys AACT GGAA ACAG TCAT TTCT CTGC TGAT ATAG GCAT TCTA ATAG GCAT TTATAG ACAC ACAC	Thr CGG Arg AAC Asn GTC Val CCCTT ACCA GTAGA GAAA GAAA GAAA GGAA TCTA GGGAA TCTA TGCC AACT TTTT AGTG	CCA Pro TAT Tyr CTG Leu GCTC TGAG AGGC AATC CTTTA AGGCA ACTG GCAG ACTG GTAG ACTG TGAG ACTG CTGT TGAG ACTG CTGT TGAG ACTG CTGT CTGT	TGG Trp 260 GTA Val 285 GGC Gly 3310 GAA Glu 335 CCCT ATCA TCTT ACGC TGTC AGGA GCCA AGTG GCA AGTG GCA AGTG AGTA ACGC AAAA ACAAA	Glu CAGC GAAGC GATC CAGC GATC CAGC GATC CCAGC CCAGC GCC CTGA AAAAG GGCG ACCA GCCA ACCA A	GIY CCCC Pro GGA GIY AGG ATG CAAC CTGG CACA AGCA GATG GAAC CACA GATG GAAC CACA GATG GAAC TTTTC GAGA CCTT GAGA CCTT GAGA CCTT GAGA CCTT GAGA CCTT C	Cys CAG GIn GGC Gly AAT ASS TACT AGGC AGTC TANA CCAC GGAG GCAA ANAC GAGG CACA TCTC CACC CAAC TTGGG	CYS GTT Val TAT TYI TTG Leu CAAT AAGT CGCA AGAT ACAG ATTG TCTT TCTT	Arg 265 CGG Arg 290 GGG GG1y 315 CTC Leu 340 GTGC CCAA CTGA CACA GGAA CCCCAA GGAA CCCCAA GGAA CCCCAA GGAA CCCCAA GGAA CCAGAA CCAGAA CCAGGAA CCAGAA CCAGGAA CCAGGAA CCAGGAA CCAGGAA CCAGGAA CCAGGAA CCAGGAA CCAGAA CCAGGAA CCAGGAA CCAGAA CCAGGAA CCAGGAA CCAGGAA CCAGGAA CCAGAA CCAGGAA CCAGAA CCA	CTA Leu CTC Leu CTC Leu ACA Thr TCCA CCCT GAAG CCTT CCCAGGCTTT CCCAGGTCTTT CCCAGGTCTT CCCGTT CCCCTT CCCTT CCCCTT CCCTT CCCCTT CCCTT CCCCTT CCCTT CCCCTT CCCTT	GAAA ACCT GAAA TCCC ATGG GGGA ATGT TTTT TGAC TCTCT TGAC TCTCT TGAC TCTCT	AGA ATC Ile CCA Pro TTCC GGGG GCTC ACAG TTTG GGGG GTCA ACAT TTTG GGGGGGGGGG	Thr GAA Glu CAC His CCA Pro CGTT CCCC AGGA ATTGG ATCAT GAAA AAAG CTAA AAAG AAAG	Leu 2770 CAG GInn 1295 TGG Trp 320 TCC Ser 345 CCCCC CCCC CCCC CCCC CCC CCC CCC CCC C	CTT Leu GGC Gly CAC His CTCC ATGG CCAG GATT TATT TGAG GATG ATAA ATAG GGGG ACAA GGGCA ACCA TGAC	CGC Arg TGT Cys CTC Leu TCCA CTGA CTGGT ACTGGT ACTGGCT CAAT GGCT CACT CGAA AAGG AGTT CTTT CT	Ala TTTT Phe GCC Ala TGA END CCCC AGTC GAGC ATGT GGCA CCCA GGGA CTTC CCCC GGTA ATGA CCTGA ACCTGA ACC	Lys GGA GIY CTA AGA TGGC TTTC AGAA TCAGG CCAGA TATA TCAGC TTATA AGGA AGAC TTATA TGGGA TGGGT TGGT TGGGT TGGT TGGGT TGGGT TGGGT TGGGT TGGGT TGGGT TGGGT TGGGT TGGT TGGGT TGGGT TGGT TGGGT TGGT	TIE 275 TCT Ser 300 GAG Glu 325 CATC TCCTT CACCA AATTG TGTTAT CTTCC CATCC CACCC ACCC	1098 1173 1249 1349 1449 1549 1649 1749 1849 2049 2149 2249 2349 2449 2549 2649 2749 2649 2749 2949
AGTO TCA	ASIN GGT GCC ASIN GCC ASIN GCCAN GCC	Ile GAA Glu ACA Thr AAGG Lys ACCG GAAG CCTG GCTG CCTA CCTA CCTA CCTA	Gln TATTY GAG Glu CCTC CACG CACGT AACC GGTT ATTG ATTG TATT AAAT TCCA TAGA TTTG	GAC Aspp 255 ACT Thr 280 GTC Val 3305 TTT TAT 330 CCCC AGGT TATA GGGGC CCT AAAG TTTT CCCA AGAG GTTG GGGC CCTTA AGAG GTTG CCCA AGAG TTTT CCCA AGAG TTTT CCCA AGAG CCCA AGAG CCCA AGAG CCCA AGAG CCCA AGAG CCCA AGAG CCCA AGAG CCCA AGAG CCCA AGAG CCCA AGAG CCCA AGAG CCCA AGAG CCCA AGAG CCCA AGAG CCCA AGAG AG	GGC GTY ATC GGG GTY ACAA AGAAAACC AGGGAACCA ACCA ACCA TGTTT TTTC CGAT TTTC CAGA TATG	TTC Phe CAC His AAA Lys AACT TATG GAAA CAGTTATG GACATTCT CAAG GCAT TATG GCATTCT ATAG GCATTAG G	Thr CGG Arg AAC Asn GTC Val CCCTT ACCA GTAGA GAAAA GAAAA GAAAA GTGAC GAAAA TCTA GTGAC GGAA TTTT TTTT	TIE CCA Pro TAT Tyr CTG Leu GCTC TGAG AGGC AATC CTTTA AGGC AATC CTGT GTAG ACTG TCAC GCA TTGA CTGT TTGA	TGG Trp 260 GTA Val 285 GGC Gly 3310 GAA Glu 335 CCCT ATCA TCTT ACGC TGTC AGGT GGAG GCCA AGTG CCTAA AGTG CCAA AGTG CCAAA CCACA	Glu CAGC GATG CATC GAACC GATC CAGC GATC CAGC GATC CAGC CAG	GIY CCCC Pro GGA GIY AGG ATG CAAC CTGG CACA AGCA GATG GAAC CACA GATG GAAC CACA GATG CACA CA	Cys CAG GIn GGC GIy AAT ASS TACT AGGC CAAG GCAA ATTC GGAG GCAA AAAC CAAC AATTT TACT TAC	CYS GTT Val TAT TYI TTG Leu CAAT AAGT CGCA AGAT ACAG ATTG TCTT TCTT	Arg 265 CGG Arg 290 GGG GG1y 315 CTC Leu 340 GTGC CCAA CTGA CACA GGAA CCCCAA GGAA CCCCAA GGAA CCCCAA GGAA CCGGAA CCCCAA CCGGAA CCGAA CCGA	CTA Leu CTC Leu CTC Leu ACA Thr TCCA CCCT GAAG AAAG CCAG GAGT TCCA ACAT CCCAAAA CCAG GAGT TTCCC ACAAA CCAG GAGT TTCCC ACAAA	GAAA ACCT GAAA TCCC ATGG GGAA ATGT TTTT TGAC TCTCT TGAC GACT TCTCT GACT GA	AGA ATC Ile CCA Pro TTCC GGGG GCTC ACAG TTTG GGGG GTCA ACAT TTTG GGGGG GGCC AGAG GGGCC AGGGGCC AGGGCC AGGGCC AGGGGCC AGGGCC AGGGGCC AGGGGCC AGGGGCC AGGGGCC AGGGCC AGGCC AGCC AGGCC AGCC AGCC AGCC AGCC AGCC AGCC AGCC AGCC	Thr GAA Glu CAC His CCA Pro CGTT CCCC AGGA ATTGG ATCA TAAT TAAT T	Leu 2770 CAG GIn TGG GIN 320 TCC Ser 345 CCCCC TCTGG. AATT GACAAAGAG CCCT ACTA ACGA ACGA CCCT TCTGG. TTGG CCCT ACTA ACGA TTGG CCCT TTGG	CTT Leu GGC Gly CAC His CTCC ATGG CCAG ATGG GATT TATT TGAG GATG ATAA ATAG GGGG ACAA GGGC ACCA ACCA	ASP CGC Arg TGT TGT CYS CTC Leu TCCA CCTGA CCTGA TGGCT CCAT CAAT GGCT CCAAT GGCT CCAAT AAGG ACCC CCCA	Ala TTTT Phe GCC Ala TGA END CCCC AGTC GAGC ATGT GGCA CCCC AGTC CCCC AGTC CCCC AGTC CCCC CCC	Lys GGA GIY CTA GGA AGA CATT AGGA AGA CATT AGGA AGA CATT AGGA AGA CATT ATT CACC TTAT AGGA AGA CATA GGA AGA CAGGA CAGA AGA CAGA C	TIE 275 TCT Ser 300 GAG Glu 325 CATC TCCTT CACCA AATTG TGTTAT CTTCC CATCC ACCC AC	1098 1173 1249 1349 1449 1549 1649 1749 1849 2049 2149 2249 2349 2449 2549 2649 2749 2849 2949 3049
AGTO TCA TCA TCA TCA TCA TCA TCA TCA TCA TCTC TCT TCTC TCT TCTC TCT TCTC TCT TCTC TCT TCTC TCT TCTC T	ASIN GGT GCC ASIN GCC ASIN GCCAA GCAA GCGAA GCGAA GCGCA TGCG GCAT TGCG GCAT TGCT GCTG GCCA TGTT GCT GCCA TGTT GCT GCCA TGT GCCA T	Ile GAA Glu ACA Thr AAGG Lys ACCG GAAG CCTA GCTG CCTA CCTA CCTA CCTA	GIn TATTY GAG GIU CCTC CACGG ACAT AACC GGTT ATCC ATTG TTATT TCCA TTATT TCCA TTAGA TTAGA TTAGA TTAGA	GAC Aspp 255 ACT Thr 280 GTC Val 3305 CCCC AGGT TACA AGG GGGC CCCT AGAG GTTG GGGC CCCT AGAG GTTG CCCA AGAG TTTT CCCA AGAG TTTT CCCA AGAG TTTT CCCA AGAG TTTT CCCA AGAG TTTT CCCA AGAG TTTT CCCA AGAG TTTT CCCA AGAG TTTT CCCA AGAG TTTT CCCA AGAG TTTT CCCA AGAG TTTT CCCA AGAG TTTT CCCA AGAG TTTT CCCA AGAG TTTT CCCA AGAG TTTT CCCA AGAG TTTT CCCA AGAG TTTT CCCA AGAG TTTT CCCA AGAG TTTTC CCCA AGAG TTTTT CCCA AGAG TTTTT CCCA AGAG TTTTT CCCA AGAG TTTTT CCCA AGAG TTTTTT CCCA AGAG TTTTT CCCA AGAG TTTTT CCCA AGAG TTTTTT CCCA AGAG AGAG	GGC GTY ATC GGG GTY ACAA GAGA AAGA AACC TGGC TGAT TTTC CGAT TTTC CAGAT TCACC TGAT TCACC TGAT TTTCACC	TTC Phe CAC His AAA Lys AACT GGAA ACAG TTATG GACAT TTCT CAAG GCAT ATAG GCA TTAG GCA TTAG GCA ACAC GCTT ATAG ACAC ACAC	Thr CGG Arg AAC Asn GTC Val CCCTT ACCA CTGG GAGA GTGAC CTCA CTGAC GGAA TGAC ACTT CCTT CCTT	TIE CCA Pro TAT Tyr CTG Leu GCTC GCTC TGAG AGGC AATC CTTTA AGGC AATC CTGT GTAG ACGC ACGC	TGG Trpp 260 GTA Val 285 GGC GGIY 3310 GAA GGL ATCA TCTT ACGC TGTC AGGG GGTA AGTG GCC AGGT GGTA AGTG CCT AGGG AGTG AGT	Glu CAGC GATG CATC GAACC GATC CAGC GATC CAGC GATC CAGC CAG	Gly GGA AGG ATG CACA CACA AGCA GACA TGCT GAGA CACA TGCT TGCT TGCT TGCT	Cys CAG GIn GGC Gly AAT ASS TACT AGGC AGTC CCAC GGAG GCAA AAAC CCAC CC	CYS GTT Val TAT TYI TTG Leu CAAT AAGT CGCA AGTT AGAA ACGT TCAG TCAG	Arg 265 CGG Arg 290 GGG GG1y 315 CTC Leu 340 GTGC CCAA CTGA CACA GGAA CCCCAA CCCAA CCCCAA CCCAA CCCAA CCCCAA CCCAA C	CTA Leu CTC Leu CTC Leu ACA Thr TCCA CCCT GAAG CCTT CCCAAAA CCAG GCTT TTCCA ACAT CCCAAAA CCAG GCTT TTCCCAACAT CAGGT CCCAAAAC CCAGGT CCCAAAAC CCAGGT CCCAAAAC CCAGGT CCCAAACAC CCAGGT CCCAAACAC CCAGGT CCCAAACAC CCAGGT CCCAAACAC CCAGGT CAAACAC CCAGGT CAAACACACACACACACACACACACACACACACACACA	GAAA ACCT GAAA TCCC ATGG GAAA TCCC GACT TTTT TCTCT TCTCT TCTCT TCTCT GACT TCTCG GCCA TCTCG GCCA TCTCG GCCA TCTCG GCCA	AGA ATC Ile CCA Pro TTCC GGGG GCTC ACAG TTTG GGGG GTCA ACAT TTTG GGGGG GGCC AGAG GGGCC AGGGCC AGGCC AGCC AGGCC AGCC AGGCC AGCC AGCC AGCC AGGCC AG	Thr GAA Glu CAC His CCA Pro CGTT CCCC AGGA ATTGG ATCA TAAT TAAT T	Leu 2770 CAG GIn TGG GIN 320 TCC Ser 345 CCCCC TCTGG. AATT GACAAAGAG CCCT ACTA ACGA ACGA CCCT TCTGG. TTGG CCCT ACTA ACGA TTGG CCCT TTGG	CTT Leu GGC Gly CAC His CTCC ATGG CCAG ATGG GATT TATT TGAG GATG ATAA ATAG GGGG ACAA GGGC ACCA ACCA	ASP CGC Arg TGT TGT CYS CTC Leu TCCA CCTGA CCTGA TGGCT CCAT CAAT GGCT CCAAT GGCT CCAAT AAGG ACCC CCCA	Ala TTTT Phe GCC Ala TGA END CCCC AGTC GAGC ATGT GGCA CCCC AGTC CCCC AGTC CCCC AGTC CCCC CCC	Lys GGA GIY CTA GGA AGA CATT AGGA AGA CATT AGGA AGA CATT AGGA AGA CATT ATT CACC TTAT AGGA AGA CATA GGA AGA CAGGA CAGA AGA CAGA C	TIE 275 TCT Ser 300 GAG Glu 325 CATC TCCTT CACCA AATTG TGTTAT CTTCC CATCC CACCC ACCC	1098 1173 1249 1349 1449 1549 1649 1749 1849 2049 2149 2249 2349 2449 2549 2649 2749 2849 2949 3049

FIGURE 3: Nucleotide and corresponding amino acid sequences of porcine kidney D-amino acid oxidase. Nucleotides are numbered in the 5' to 3' direction, beginning with the first nucleotide of the cDNA insert preceded by the oligo(dG)/(dC) linker. The nucleotide sequence was deduced by combining the sequences of the cDNA inserts of clones pDAO-10 and pDAO-13. In the 5'-untranslated region, the nucleotide residues 1–54 are absent in pDAO-10, and pDAO-13 has a deletion between 81 and 189. The deduced amino acid residues are indicated below the nucleotide triplets. Four possible polyadenylation signals, ATTAAA and AACAAA, are boxed. The sequences that correspond to the two synthetic oligonucleotide probes (I and II) are underlined, respectively. An upstream in-frame stop codon at nucleotides 82–84 is indicated by closed circles. A small open reading frame in the 5'-untranslated region that could encode only five amino acids is shown by open circles at the ATG and TGA codons. In the 3'-untranslated region, pDAO-13 is included up to the position indicated by the arrowhead.

3616 BIOCHEMISTRY FUKUI ET AL.

Table I: Codon Usas	e in Porcine Kidn	ev DAO mRNA and Amino A	Acid Composition Predicte	d from the Nucleotide Sequence ^a
---------------------	-------------------	-------------------------	---------------------------	---

amino acid	residues	codon	no.	amino acid	residues	codon	no.
Leu	36	UUA	1	Ser	13	UCU	6
		UUG	1			UCC	3 2
		CUU	2			UCA	2
		CUC	9			UCG	0
		CUA	4			AGU	1
		CUG	19			AGC	1
Arg	21	CGU	19 2	Thr	22	ACU	6
		CGC	4			ACC	10
		CGA	1			ACA	6
		CGG	3			ACG	0
		AGA	6	Ile	20	AUU	8
		AGG	5			AUC	11
Pro	22	CCU	3			AUA	1
		CCC	8	Asn	19	AAU	5
		CCA	8			AAC	14
		CCG	3	Phe	15	UUU	4
Gln	14	CAA	2			UUC	11
		CAG	12	Tyr	14	UAU	6
Lys	12	AAA	4			UAC	8
		AAG	8	Glu	22	GAA	10
Ala	17	GCU	6			GAG	12
		GCC	4	Cys	5	UGU	1
		GCA	6			UGC	4
		GCG	1	His	9	CAU	2
Val	26	GUU	3			CAC	7
		GUC	7	Asp	13	GAU	7
		GUA	2			GAC	6
		GUG	14	Met	5	AUG	5
Gly	32	GGU	4	Trp	10	UGG	10
8		GGC	13				
		GGA	7				
		GGG	8				

^aThe amino acid of the porcine kidney DAO was predicted solely from the nucleotide sequence of the mRNA. Numbers next to codons indicate the numbers of amino acids using particular codons.

position was observed; 101 codons (29%) terminate in G, 120 (35%) in C, 66 (19%) in U, and 60 (17%) in A.

Presence of Three DAO mRNAs in Porcine Kidney. To determine the size of the DAO mRNA, Northern blot hybridization experiments were carried out with the cDNA insert of pDAO-10 and a synthetic oligonucleotide as hybridization probes (Figure 4). The oligonucleotide or cDNA fragment covering the protein-coding region of DAO detected three bands, the mobilities of which corresponded to sizes of approximately 3.3, 2.7, and 1.5 kb (Figure 4, lanes a and b). D-Amino acid oxidase with 347 amino acid residues requires at least 1041 bases for coding sequence. Therefore, these three species of mRNA are large enough to share a common sequence for the DAO coding region. The nucleotide sequences deduced from the two clones pDAO-10 and pDAO-13 differ in the length of the 3'-untranslated region (Figure 3). The size difference of the 3.3- and 2.7-kb bands, estimated from their mobilities on the electrophoretic gel, was consistent with the difference in the length of the 3'-untranslated region between the clones pDAO-10 and pDAO-13. These results indicate that cDNAs of pDAO-10 and pDAO-13 correspond to the largest and the middle-size mRNA, respectively. The cDNA corresponding to 1.5-kb mRNA has not been identified. However, when the cDNA fragment derived from the 3'-untranslated region (PvuII-KpnI) was used for hybridization, the band corresponding to 1.5 kb became weak (Figure 4, lane c). It is suggested, therefore, that a large portion of the 3'-untranslated region should be absent in the 1.5-kb mRNA.

The hybridization signals of the larger and the smaller mRNA were significantly more intense than that of the middle-size mRNA. This suggests that the efficiencies of their expression are considerably different.

Genomic Organization of DAO Gene. Southern blot analysis of genomic DNA was used to examine the copy number of

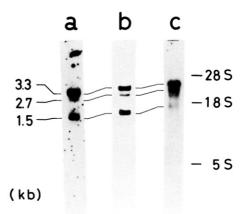


FIGURE 4: Northern hybridization analysis of porcine D-amino acid oxidase mRNAs. Five micrograms of poly(A+) RNA per lane was fractionated by electrophoresis on a 0.7% agarose gel and transferred to a nitrocellulose filter. Blots were hybridized with nick-translated cDNA fragments (Figure 2) or 5'-end labeled oligonucleotide probe (Figure 3), followed by washing and autoradiography. Hybridization probes were as follows: (lane a) synthetic oligonucleotide (17-mer) (I in Figure 2); (lane b) Bg/II-PvuII fragment (1367 bp) (C in Figure 2); (lane c) PvuII-KpnI fragment (913 bp) (D in Figure 2). The size markers used were porcine rRNAs.

the DAO gene in the porcine genome. When high molecular weight DNA was digested with *Hind*III and hybridized with the 0.65-kb fragment (SalI-BamHI; Figure 2, probe B) corresponding to the N-terminal half of the coding sequence or the 1.4-kb fragment (BglII-PvuII; Figure 2, probe C) covering the whole protein-coding region, a sharp single 1.8-kb band was observed for both probes (Figure 5, lanes a and b). No other cross-hybridizing band was detected under stringent conditions. These results suggest that the DAO gene exists as a single copy in the porcine genome.

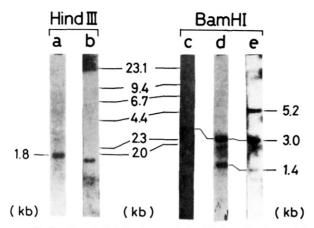


FIGURE 5: Southern hybridization analysis. High molecular weight DNA from porcine kidney (2 μ g/gel lane) was digested with *Hind*III (lanes a and b) and BamHI (lanes c-e). The following hybridization probes were used: (lanes a and d) SalI-BamHI fragment (641 bp) (B in figure 2); (lanes b and e) Bg/II-PvuII fragment (1367 bp) (C in Figure 2); (lane c) EcoRI-AccI fragment (557 bp) (A in Figure

However, the Southern hybridization pattern seen with BamHI digestion was rather complex (Figure 5, lanes c and d), because of the presence of a recognition site for this enzyme within the cDNA. We therefore examined BamHI-digested DNA with three different probes (Figure 2, probes A-C). Probe A corresponding to the 5'-untranslated region as well as the extremity of the N-terminal coding region recognized a single 3.0-kb band (Figure 5, lane c). This confirms that we are dealing with a single gene. The whole coding region (probe C) detected three bands: 5.2, 3.0, and 1.4 kb (Figure 5, lane d). The N-terminal half of the coding region (probe B) recognized two bands of 3.0 and 1.4 kb. These patterns could be most readily explained by the presence of multiple introns in the DAO gene, and the arrangement of the three BamHI fragments in the porcine genome is presumed to be in the following order: 3.0, 1.4, and 5.2 kb from 5' to 3'. DISCUSSION

Biogenesis of DAO. The assignment of the initiation site of translation was based on the fact that ATG at nucleotides 199-201 was preceded by the in-frame stop codon TGA at 82-84 and followed by 1041 nucleotides of an open reading frame that encoded 347 amino acids. In addition, this ATG is preceded by sequences that fulfill the Kozak (1981) criteria for initiation codon. The predicted amino acid sequence starting with methionine was confirmed by the N-terminal amino acid sequence determination of the mature enzyme for 32 residues. The reported sequence for DAO was also in complete agreement with the deduced sequence shown in Figure 3, and the ambiguous residue Asx was concluded to be Asp. The molecular weight of the deduced sequence was calculated to be 39 335, which is also in a good agreement with the value determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (Fukui et al., 1986). These results demonstrate directly the absence of a cleavable presequence at the amino terminal of the primary translation product of DAO mRNA, indicating the posttranslational translocation of DAO into peroxisomes without any proteolytic modification. It would seem reasonable to speculate the presence of a recognizable signal for translocation to peroxisomes within the polypeptide. Plots of the relative hydrophobicity at each amino acid position in DAO were made according to the Kyte-Doolittle method (Kyte & Doolittle, 1982) and revealed a highly hydrophobic region, which encompasses the first 17 amino acids of the protein (data not shown). This sequence

might function as a signal for DAO to associate with peroxisomal membrane, but a more thorough analysis and survey of other peroxisomal proteins are necessary. Two potential asparagine-linked glycosylation sites were observed within the coding sequence, although the mature enzyme showed the absence of a carbohydrate moiety. These results also support the idea that DAO is synthesized on free ribosomes and transported through the cytosol into peroxisomes.

Nucleotide Sequence of cDNA Clones. The 5'-untranslated region of pDAO-13 lacks 109 nucleotides compared with pDAO-10. This could be caused by the secondary structure of the relatively long 5'-untranslated region of DAO mRNA or simply by the cloning artifact for sequencing. However, this result is of interest in relation to mRNA processing, since the site of deletion coincided with the consensus sequences of donor and acceptor for splicing (Breathnach & Chambon, 1981). In addition, the branch site sequence (Wallace & Edmonds, 1983) was found at nucleotides 157-161. We have to wait for the genomic sequence analysis to clarify these possibilities. The analysis of the 5' upstream region of the DAO gene will also provide information for the mechanism of regulation governing the biosynthesis of DAO.

Northern and Southern Hybridization. Our Southern analysis revealed that DAO was coded by a single gene in the porcine genome. This conclusion is consistent with the genetic studies by Konno and Yasumura (1984), showing that both the brain and kidney DAO were coded by the same gene in the mouse. On the other hand, the analysis of the expression of this gene showed the presence of three DAO mRNAs in the pig kidney. Therefore, we postulated that the multiple DAO mRNAs arise from a single gene by utilization of four possible polyadenylation signals, which are found in the 3'untranslated region as indicated by boxes in Figure 3. Among these signals, it was shown from this study that the third and the last signals were utilized by 2.7- and 3.3-kb mRNAs. The utilization of the first ATTAAA signal by 1.5-kb mRNA is inferred from its size. However, the mRNA that utilizes the second ATTAAA signal was not found in the Northern blot. These results are not unusual, since the generation of multiple mRNAs from a single gene has been reported for several mRNAs, such as IL-2 receptor (Leonard et al., 1984; Nikaido et al., 1984) and prekiningen mRNAs (Kageyama et al., 1984). But the biological significance of this phenomenon remains unclear. Moreover, Northern hybridization signals of the larger and the smaller mRNAs were more intense than that of the medium mRNA (Figure 4, lane b), suggesting that the sequence of ATTAAA is used more effectively than that of AACAAA as a polyadenylation signal for the generation of the DAO mRNAs. A sequence of AACAAA was reported to be recognized as a polyadenylation signal in human coagulation factor XI cDNA (Fujikawa et al., 1986). Point mutation studies by Wickens and Stefenson (1984) showed that the AACAAA signal yielded a 3' end of RNA at a very much reduced level. This study clearly demonstrates the difference in efficiencies for 3' processing of the RNA between the polyadenylation signals in the physiological system. In this respect it would be of great interest to examine the tissuespecific alternative expression of the different DAO mRNAs, since the same gene can produce two or more mRNAs in different cell types or at different stages of differentiation. This study opens the way to a detailed analysis of the regulation of the biosynthesis of this enzyme and biogenesis of peroxisomes. Furthermore, molecular genetic studies with the DNA probe would lead us to the elucidation of the biological function of DAO.

3618 BIOCHEMISTRY FUKUI ET AL.

ACKNOWLEDGMENTS

We thank Drs. Retsu Miura, Kyoko Momoi, and Saori Takahashi for stimulating discussions during the course of this work.

Registry No. DNA (rat pancreas cationic trypsinogen messenger RNA complementary), 108007-28-9; trypsinogen, 9002-08-8; trypsin, 9002-07-7; pretrypsinogen, 81989-96-0; trypsinogen (rat pancreas cationic precursor reduced), 108007-29-0; trypsinogen (rat pancreas cationic reduced), 108007-30-3.

REFERENCES

- Aviv, H., & Leder, P. (1972) Proc. Natl. Acad. Sci. U.S.A. 69, 1408-1412.
- Breathnach, R., & Chambon, P. (1981) *Annu. Rev. Biochem.* 50, 349-383.
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., & Rutter, W. J. (1979) Biochemistry 18, 5294-5299.
- Cline, M. J., & Lehrer, R. I. (1969) Proc. Natl. Acad. Sci. U.S.A. 62, 756-763.
- Curti, B., Ronchi, S., Brazoli, U., Ferri, G., & Williams, C. H. J. (1973) Biochim. Biophys. Acta 327, 266-273.
- de Duve, C., & Baukhuin, P. (1966) Physiol. Rev. 46, 323-357.
- Dunn, J. T., & Perkoff, G. T. (1963) Biochim. Biophys. Acta 73, 327-331.
- Fujikawa, K., Chung, D. W., Hendrickson, L. E., & Davie, E. W. (1986) *Biochemistry* 25, 2417-2424.
- Fukui, K., Momoi, K., Watanabe, F., & Miyake, Y. (1986) Biochem. Biophys. Res. Commun. 141, 1222-1228.
- Furuta, S., Hayashi, H., Hijikata, M., Miyazawa, S., Osumi, T., & Hashimoto, T. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83, 313-317.
- Goldman, B. M., & Blobel, G. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 5066-5070.
- Hamilton, G. A., Buckthal, D. J., Mortensen, R. M., & Zerby, K. W. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 2625–2629.
- Hanahan, D., & Meselson, M. (1980) Gene 10, 63-67.
- Hewick, R. M., Hunkapiller, M. W., Hood, L. E., & Dreyer, W. J. (1981) J. Biol. Chem. 256, 7990-7997.
- Honjo, T., Obata, M., Yamawaki-Kataoka, Y., Kataoka, T., Kawakami, T., Takahashi, N., & Mano, Y. (1979) Cell (Cambridge, Mass.) 18, 559-568.
- Ikemura, T. (1985) Mol. Biol. Evol. 2, 13-34.
- Kageyama, R., Ohkubo, H., & Nakanishi, S. (1984) *Biochemistry* 23, 3603-3609.
- Konno, R., & Yasumura, Y. (1984) J. Neurochem. 42, 584-586.

- Kozak, M. (1981) Nucleic Acids Res. 9, 5233-5252.
- Krebs, H. A. (1935) Biochem. J. 29, 1620-1644.
- Kubo, H., Yamano, T., Iwatsubo, M., Watari, H., Soyama,
 T., Shiraishi, J., Sawada, S., Kawashima, N., Mitani, S.,
 & Ito, K. (1958) Bull Soc. Chim. Biol. 40, 431-447.
- Kyte, J., & Doolittle, R. F. (1982) J. Mol. Biol. 157, 105-132.
 Leonard, W. J., Depper, J. M., Crabtree, G. R., Rudikoff, S., Pumphrey, J., Robb, R. J., Kronke, M., Svetlik, P. B., Peffer, N. J., Waldmann, T. A., & Green, W. C. (1984) Nature (London) 311, 626-631.
- Marshall, R. D. (1974) Biochem. Soc. Symp. 40, 17-26.
 Massey, V., Palmer, G., & Bennet, R. (1961) Biochim. Biophys. Acta 48, 1-9.
- Matteucci, M. D., & Caruthers, M. H. (1981) J. Am. Chem. Soc. 103, 3185-3191.
- Miura, S., Mori, M., Takiguchi, M., Tatibana, M. Furuta, S., Miyazawa, S., & Hashimoto, T. (1984) *J. Biol. Chem.* 259, 6397-6402.
- Nakajima, H., Ohta, M., Yamano, T., & Miyake, Y. (1981) Biomed. Res. 2, 154-165.
- Nikaido, T., Shimizu, A., Ishida, N., Sabe, H., Teshigawara, K., Maeda, M., Uchiyama, T., Yodoi, J., & Honjo, T. (1984) Nature (London) 311, 631-635.
- Okayama, H., & Berg, P. (1982) Mol. Cell. Biol. 2, 161-170. Osumi, T., Ishii, N., Hijikata, K., Ozasa, H., Furuta, S., Miyazawa, S., Kondo, K., Inoue, K., Kagamiyama, H., & Hashimoto, T. (1985) J. Biol. Chem. 260, 8905-8910.
- Rigby, P. W. J., Dieckmann, M., Rhodes, C., & Berg, P. (1977) J. Mol. Biol. 113, 237-251.
- Ronchi, S., Minchiotti, L., Galliano, M., Curti, B., Swenson, R. P., Williams, C. H. J., & Massey, V. (1982) J. Biol. Chem. 257, 8824-8834.
- Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463-5467.
- Southern, E. M. (1975) J. Mol. Biol. 98, 503-517.
- Thomas, P. S. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 5201-5205.
- Wallace, J. C., & Edmonds, M. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 950-954.
- Wallace, R. B., Schaffer, J., Murphy, R. F., Bonner, J., Hirose, T., & Itakura, K. (1979) Nucleic Acids Res. 6, 3543-3557.
- Wickens, M., & Stephenson, P. (1984) Science (Washington, D.C.) 226, 1045-1051.
- Yagi, K., & Ozawa, T. (1962) Biochim. Biophys. Acta 56, 420-426.
- Yanisch-Perron, C., Vieira, J., & Messing, J. (1985) Gene 33, 103-119.
- Yaoita, Y., & Honjo, T. (1980) Biomed. Res. 1, 164-175.