Molecular cloning and nucleotide sequences of the complementary DNAs to chicken skeletal muscle myosin two alkali light chain mRNAs

Yo-ichi Nabeshima<sup>\*</sup>, Yoshiaki Fujii-Kuriyama<sup>+</sup>, Masami Muramatsu<sup>+</sup> and Kikuo Ogata<sup>\*</sup>

\*Department of Biochemistry, Niigata University School of Medicine, Niigata 951, and \*Department of Biochemistry, Cancer Institute, Japanese Foundation for Cancer Research, Kami-Ikebukuro, Toshimaku, Tokyo 170, Japan

Received 28 July 1982; Revised and Accepted 9 September 1982

#### ABSTRACT

We report here the molecular cloning and sequence analysis of DNAs complementary to mRNAs for myosin alkali light chain of chicken embryo and adult leg skeletal muscle. pSMA2-1 contained an 818 base-pair insert that includes the entire coding region and 5' and 3' untranslated regions of A2 mRNA. pSMA1-1 contained a 848 base-pair insert that included the 3' untranslated region and almost all of the coding region except for the N-terminal 13 amino acid residues of the A1 light chain. The 741 nucleotide sequences of A1 and A2 mRNAs corresponding to C-terminal 141 amino acid residues and 3' untranslated regions were identical. The 5' terminal nucleotide sequences corresponding to N-terminal 35 amino acid residues of A1 chain were quite different from the sequences corresponding to N-terminal 8 amino acid residues and of the 5' untranslated region of A2 mRNA. These findings are discussed in relation to the structures of the genes for A1 and A2 mRNA.

## INTRODUCTION

Chick embryo muscle cells have been extensively used as a model system for the study of the control of gene expression during cell differentiation.

Myosin is a major protein component in the contractile apparatus in the cell, consisting of two large subunits of 200 k daltons each and three to four smaller ones ranging from 15 to 30 k daltons (1). The fusion of mononucleated myoblasts into multinucleated fibers is associated with the synthesis of myosin peptides and other specific proteins involved in muscle contraction (2-7). Furthermore, multiple isozymic forms of the myosin heavy and light chains have also been identified (8-15). To study myosin gene expression, studies of DNA cloning of these proteins have been reported for chick embryonic cardiac myosin light chain (16), rat skeletal myosin heavy and light chains (17-19) and chick skeletal myosin heavy chain (20). Although, the primary structures of  $A_1$  and  $A_2$  light chains and DTNB light chains of rabbit skeletal muscle myosin have been known for some time (21), those of chicken skeletal muscle myosin were reported relatively recently (22, 23). As for chicken myosin light chains, the primary structures of heart and gizzard muscle myosin light chains have also been determined (23).

From a comparison of the primary structures of  $A_1$  and  $A_2$  light chains of chicken and rabbit skeletal muscle, Matsuda et al. (22) discussed a possible structure for these genes.

Here we describe the construction of a bacterial plasmid containing sequences of mRNAs for chicken skeletal muscle myosin light chains  $A_1$  and  $A_2$ , and their nucleotide sequences. It was found that, while the 5' terminal region of the  $A_1$  and  $A_2$  mRNA is quite different, the 3' side of the mRNA including most of the coding region is identical.

## MATERIALS AND METHODS

## Preparation of mRNA

The microsome fraction was prepared from the leg skeletal muscles of 15 day chick embryos as described by Kennedy and Heywood (24) and  $poly(A)^+$  mRNA was prepared as described previously (25).

For the preparation of poly(A)<sup>+</sup> RNA from 45 day chicken leg skeletal muscle, the method described by Duguid et al. (26) was used. Briefly, leg skeletal muscle was homogenized with 10 volumes of a buffer (0.1 M Tris-HCl pH 9.0, 13 mM DTT, 1% SDS) and 10 volumes of phenol-chloroform-isoamylalcohol (50:50:1) by a Waring blendor for 30 sec. After 5 times extraction, described above with the phenol-chloroform solution, a quarter volume of 10 M LiCl was added to the aqueous fraction containing total nucleic acid and the fraction was kept in ice for 1 h. Precipitated RNA was purified by two-cycles of oligo(dT) cellulose column chromatography.

 $Poly(A)^+$  mRNA was then sedimented at 198,000 X g for 15 h on a 5-20% linear sucrose density gradient containing 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.1% SDS in a Beckman SW 40 Rotor. Fractions (0.4 ml) were collected and assayed for translation of myosin alkali light chain with a nuclease-treated rabbit reticulocyte lysate cell-free system (27). The translation products were analysed on 15% SDS acrylamide slab gel electrophoresis (34) followed by fluorography. Fractions containing mRNAs for myosin light chains were pooled and sedimented on a second sucrose gradient. As thus purified these are referred to as the enriched skeletal myosin light chain mRNA. For colony-filter hybridization, the enriched myosin light chain mRNA was further purified by 98% formamide acrylamide gel electrophoresis (32) and is referred to as the  $A_1$  myosin light chain mRNA fraction (manuscript in preparation).

# <u>Preparation of double-stranded cDNA, insertion into pBR322, and</u> transformation of $\chi$ 1776

Complementary DNA molecule (cDNA) was synthesized from the enriched myosin light chain mRNA by using avian myeloblastosis virus reverse transcriptase (generously supplied by Dr. J.W. Beard, Life Sciences, St. Petersburg, FL), in a 400 µl reaction mixture containing 50 mM Tris-HCl pH 8.3, 10 mM MgCl<sub>2</sub>, 50 mM KCl, 10 mM DTT, 2.4  $\mu$ g oligo(dT)<sub>10</sub> (Collaborative Research), 500  $\mu$ M each dCTP, dGTP, dTTP, dATP (dCTP: <sup>32</sup>P radioactivity at the specific activity of 2.95 X  $10^5$  cpm/nmol) and 20 µg the enriched myosin light chain mRNA. The second strand of cDNA was synthesized by using a Escherichia coli DNA polymerase I (Klenow fragment, Bethesda research laboratories). After Sl nuclease treatment, the double-stranded cDNA was tailed with  $oligo(dC)_{20-30}$  at the 3'-OH terminus by using calf thymus terminal transferase in the presence of CoCl<sub>2</sub> (28). The vector pBR322 was cleaved with restriction endonuclease PstI and was tailed with  $oligo(dG)_{20-30}$  as described above. Oligo(dC)-tailed double-stranded cDNA was anealed with oligo(dG)-tailed plasmid DNA and then used to transform the EK $_2$  E. coli strain  $\chi 1776$ .

Bacterial colonies containing the recombinant plasmid were screened as described by Grunstein and Hogness (29) with highly labelled cDNA that was synthesized from the A<sub>1</sub> myosin light chain mRNA.

# Identification of the plasmid by the positive hybridization translation method

Five  $\mu g$  of plasmid DNA was cleaved by restriction endonuclease Eco RI and coupled to DBM-paper after heat denaturation according to the procedure of Hoeijmakers et al. (30).

Total poly(A)<sup>+</sup> RNA which was prepared from 45 day chicken skeletal muscle was hybridized to the coupled DNA and unbound RNA was washed off as described Hoeijmakeres et al. (30). The hybridized RNA was eluted from DBM-paper by heating for 1 min at 100°C in 200 µl 10 mM Tris-HCl (pH 7.4), 1 mM EDTA and 0.1% SDS and precipitated with 5 µg of carrier rat liver tRNA. The eluted RNA was assayed for the synthesis of myosin alkali light chain  $A_1$ and  $A_2$  by the reticulocyte lysate system (27). The products were analyzed by 15% SDS acrylamide slab gel electrophoresis (34) followed by fluorography.

## Restriction endonuclease mapping of plasmid DNA

Conditions for restriction endonuclease cleavage of plasmid DNA were essentially as indicated by the suppliers. Fragments were electrophoresed on 1% agarose gels containing ethidium bromide and visualized by UV irradiation.

Restriction endonuclease Eco RI, BglII, PstI, AluI, HaeIII, KpnI, PvuII, HinfI and HpaII were purchased from Takara Shuzo Co. Ltd. and Sau3A from Bethesda Research Laboratories Inc. Sequence determination of plasmid DNA

After restriction enzyme mapping of the cDNA clones, appropriate restriction fragments were labelled at their 5' ends with  $T_4$  polynucleotide kinase (Takara Shuzo) and  $[\gamma^{-32}P]ATP$  (Amersham, 5,000 Ci/mmol) after removal of the 5' phosphate groups with bacterial alkaline phosphatase (Worthington). The 3' end of the PstI site was labelled with the  $T_4$  DNA polymerase (Takara Shuzo) and  $[\alpha^{-32}P]dCTP$  (Amersham, 2,500 Ci/mmol).

DNA fragments were separated on a thin polyacrylamide gel (8 and 20%) after chemical cleavage as described by Maxam and Gilbert (31).

### RESULTS

## Identification of recombinant plasmids containing cDNA sequence for mRNA of chick embryo myosin alkali light chain

Fifteen  $\mu g$  of poly(A)<sup>+</sup> RNA fraction enriched with the myosin alkali light chain A<sub>1</sub> mRNA that was prepared from 15 day chicken embryo leg skeletal muscle, was used to prepare the double stranded cDNA for the subsequent bacterial transformation.

About 5 X 10<sup>5</sup> tetracycline resistant  $\chi$ 1776 colonies were produced per 1 µg of double-stranded cDNA and 5 µg of oligo(dG) tailed pBR322. Five hundred transformed colonies were replica plated and screened, by colony-filter hybridization, with a  $^{32}$ p labelled cDNA made from A<sub>1</sub> myosin light chain mRNA fraction which was purified by 98% formamide gel electrophoresis (32).

Recombinant plasmids were prepared from eight positive colonies and was assayed by positive hybridization-translation assay methods. Poly(A)<sup>+</sup> RNA prepared from 45 day chicken skeletal muscle was used for this assay, for the following reasons. This poly(A)<sup>+</sup> RNA fraction contained nearly equimolar amounts of mRNAs for  $A_1$  and  $A_2$  chains, while poly(A)<sup>+</sup> RNA prepared from 15 day chicken embryo skeletal muscle lacked mRNA for the  $A_2$  chain (unpublished data). Furthermore, since the Cterminal 141 amino acid residues of  $A_1$  and  $A_2$  chains were identical (22), it was assumed that cDNA inserts of the myosin alkali light chain hybridized to both  $A_1$  and  $A_2$  mRNAs.

Of the colonies tested, one plasmid, pSMA-1, showed a fluorographic pattern of hybridization to both  $A_1$  and  $A_2$  mRNAs (data not shown), and was chosen for further characterization. Partial restriction map of pSMA-1 and identification of chicken skeletal myosin alkali light chain sequence

The size of the inserted cDNA sequence was determined by digestion of pSMA-1 with PstI. A cleavage map of several restriction enzymes in pSMA-1 DNA was constructed by analyses of their single or double enzyme digests (Fig. 1). DNA of pSMA-1 was digested with BglII, and <sup>32</sup>P-end-labelled fragments were purified and sequenced as described above.

The nucleotide sequence determined, corresponded to the 170-190 amino acid residues of a chicken myosin alkali light chain and 3' untranslated regions. Plasmid pSMA-1 was thus confirmed as a clone for a chicken myosin alkali light chain. <u>Identification of recombinant plasmid containing chicken myosin</u> alkali light chain A1 or A2 sequences

Poly(A)+ RNA was prepared from 45 day chick skeletal muscle and myosin alkali light chain mRNA was partially purified by sucrose density gradient centrifugation (data not shown) which separated A1 and A2 mRNA. Poly(A)<sup>+</sup> RNA enriched with  $A_1$  or  $A_2$  chain was used to prepare the double stranded cDNA, which was used for subsequent bacterial transformation. One thousand transformed colonies were replicaplated and hybridized with <sup>32</sup>P labeled insert DNA of plasmid pSMA-1. The largest plasmid containing  $A_1$  or  $A_2$  sequence was selected from strongly hybridized colonies. Plasmid pSMA1-1 and plasmid pSMA2-1 was identified as a clone for  $A_1$  chain and  $A_2$ chain, respectively. Restriction maps of the insert DNAs of pSMA1-1 and pSMA2-1 are shown in Fig. 1.



Fig. 1. Sequencing strategy and the restriction map of pSMA1-1, pSMA2-1 and pSMA-1. The open bar indicates the coding region of A1 or A2 mRNA, and the solid bar indicates the untranslated region of the corresponding myosin light chain mRNA. The restriction sites predicted by the DNA sequence were confirmed by digestion with restriction enzymes followed by sizing with 1% agarose gels. Restriction fragments were labeled at the sites of the solid circles and DNA sequencing proceeded in the direction of the arrow heads for the distance indicated by the length of the arrow. The 5'- and 3'-end regions of the cDNA were sequenced after labelling at the PstI site (open circle) by the T4 DNA polymerase in the presence of  $\alpha$ -32P dCTP and 1 µM dCTP. The scale is indicated by the calibration in base pairs and shown at the bottom of the figure.

Sequence determination of pSMA1-1 and pSMA2-1

The nucleotide sequences of chicken skeletal muscle myosin  $A_1$  and  $A_2$  mRNAs were determined (Fig. 2). The strategy of

PSNN2-1 C CTC TCA GCT ANT CCC TCC 10 AA GAA CCA GCC ATC GAC CTC AAG AGC ATC AAG ATC GAG TTC TTC AGG GA CAG GAC AT Lys Glu Pro Ala 11e Aap Leu Lys Ser 11e Lys 11e Glu Phe Ser Lys Glu Gin Aap GGC CCG TCC GCT CTA CTT TTC CAA CTC TCA ATC ATG TCC TTC TCA CCT GAC GAA ATC AAT Ser Phe Ser Pro Asp Glu IIe Aan GAC TTC AAG GAG GCC TTC CTC CTC CTT GAC AGG ACT GGT GAT GCC AAG ATT ACC CTG AGC AAp Phe Lys Glu Ala Phe Leu Leu Phe Asp Arg Thr GIY Ang ALa Lys 11e Thr Leu Ser CAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATT ACC ACA AAG CAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AAG CAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AAG CAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AAG CAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AAG CAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AAG CAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AATG AAT	₽SM2	1-1		1 A	CCA Pro	GCA Ala	CCC	GCA Ala	CCC	GCG Ala	CCA Pro	GCA Ala	CCT Pro	GCA Ala	CCA Pro	GCG Ala	CCC	GCT Ala	AAA Lvs	CCC
AMA GAA CCA GCC ATC GAC CTC AMA AGC ATC AMA ATC GAG TTC TTC AMA GAG CAG CAG CAT Lys Glu Pro Ala 11e Asp Leu Lys Ser 11e Lys 11e Glu Phe Ser Lys Glu Gln Amp GGC CCG TCC GCT CTA CTT TTC CAA CTC TCA ATC ATG TCC TTC TCA CT AGC GAA ATC AAT Ser Phe Ser Pro Asp Glu 11e Asn Land GMA GCC TTC CTC CTC TTT GAC AGG ACT GGT GAT GCC AMA ATT ACC CTG AGC GAC TTC AMA GAG GCC TTC CTC CTC TTT GAC AGG ACT GGT GAT GCC AMA ATT ACC CTG AGG CGC CTG GGT GAC ATC GTT CGG GCA CTG GGA CAG ACC CCC ACA AC GCT GAG ATT ACC CTG AGG GGT GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AC GCT GAG ATC AAC AMA CGG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AC GCT GAG ATC AAC AMA CGG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AC GCT GAG ATC AAC AMA CGG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AAG GIn Val Gly Asp 11e Val Arg Ala Leu Gly Gln Asn Pro Thr Asn Ala Glu 11e Asn Lys 230 ATC CTG GGC AMA CCC AGC AMA GAG GAA ATG AMT GCC AMA AMA ATC ACT TTT GAA GAG TTC ATC CTG GGC AMA CCC AGC AMA GAG GGA ATG AMT GCC AMA GMA ATC ACT TTT GAA GAG TTC TTC CTG GGC AMA CCA GCA GCT GCT AAC AAC AMA GGT CAG GGT ACC TTT GAA GAC TTT GTT CTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AMA GGAT CAG GGT ACC TTT GAA GAC TTT GTT CTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AMA GGC ACA GGC ACG GT CAC TTT GAA GAC TTT GTT CTG CCC ATG CTG CAT GCA GCA GCT GCT AAC AAC AMA GGA CAC GGC ACA GTC ATG GG GCT GAG CTC CCC GAA GGT CTG CGT GTT TTC GAC AMA GAA GGA AGC GGC ACA GTC ATG GGG GCT GAG CTC CCC GAA GGT CTG CGT GTT TTC GAC AMA GAA GGA AGC GGC ACA GTC ATG GGG GCT GAG CTC CCC GAA GGT CTG CGT GTT TTC GAC AMA GAA GAA GAA GAA GAA GTA GAA GAA GAA CTT ATG TT GTT CTG GCT GCA CAC GGG GGA GAA GTG ACA GAA GAA GTA GAA GAA CTA GTA GTA GTA GTA CTG GCT ACA CTG GGC GAG AMA GTA ACA CAA GAA GGA ATA GAA CTA GAA GAA CTT GT CTG CGT GTT TTC GAC GGC GCA GTA GAA GAA GTA CTA GTA GAA GAA CTC ATC ATC GTT TTA GAA GAC GTC CCC AAG ATA ATT CTT GCT CTG GCA GAA GTA GAA CTA ATC ATA ATA GTA ABA GAA GTC CCC AGA ATA ATA CCA ATG ATA GAA GAA GTA GTA GAA GAA CTA CTA TGT TT GTT GTG CTG GGT AC	р <b>SM</b> / 50	2-1												c	CTC	TCA	GCT 10	AAT 0	ccc	TCC
Lys Glu Pro Ala Ile Asp Leu Lys Ser Ile Lys Ile Glu Phe Ser Lys Glu Gln Gln Asp GGC CCG TCC GCT CTA CTT TTC CAA CTC TCA ATC ATG CTC TC CA CCT GAC GAA ATC ANT Ser Phe Ser Pro Asp Glu Ile Asn Ser Phe Ser Pro Asp Glu Ile Asn GAC TTC AMG GMG GCC TTC CTC CTC TTT GAC MGG ACT GGT GAT GCC AMG ATT ACC CTG AGC Asp Phe Lys Glu Ala Phe Leu Leu Phe Asp Arg Thr Gly Asp Ala Lys Ile Thr Leu Ser CAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC GAC ACA AAC GCT GAG ATC AAC AAG GCn GT GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AAG GIn Val Gly Asp Ile Val Arg Ala Leu Gly Gln Asn Pro Thr Asn Ala Glu Ile Asn Lys ATC CTG GGC AAC CCC AGC AAG GAG GAA ATG AAT GCC AAG AMA GAC ATC TTT GAA GAG TTC Ile Leu Gly Asn Pro Ser Lys Glu Glu Met Asn Ala Lys Lys Ile Thr Phe Glu Glu Phe 300 CTG CCC ATG CTG GA GCA GCT GCT AAC AAC AAG GAT CAG GAT ACC TTT GAA GAG TTC TTT Leu Pro Met Leu Gln Ala Ala Ala Asn Asn Lys Asp Gln Gly Thr Phe Glu Asp Phe Val 400 GAA GGT CTG CGC GAC GCT GCT AAC AAC AAG GAA GGA CAA GTC ATG TG GAG GCT GAG CTC CCC GlU Gly Leu Arg Val Phe Asp Lys Glu Gly Asn GCC AAC GAC AGT CAT GG GGG CT GAG CTC CCC GGT GCT GCT GTT TTC GAC AAG GAA GGC AAC GGC AAC GAC ATC ATG GGG GCT GAG CTC CCC GIU Gly Leu Arg Val Phe Asp Lys Glu Gly Asn Gly Thr Val Met Gly Ala Glu Leu Arg His Val Leu Ala Thr Leu Gly Glu Lys Met Thr Glu Glu Val Glu Glu Leu Met Lys GGT CAG GAA GCTC CAC AAC GAC TCC ATC AAC GAA GAG GAA GAT GAA GAA GAA GAA CAA GAA GAA GAA GAA GAA	AAA	GAA	CCA	GCC	ATC	GAC	CTC	AAG	AGC	ATC	AAG	ATC	GAG	TTC	TTC	AAG	GAG	CAG	CAG	GAT <sup>*</sup>
GGC CCG TCC GCT CTA CTT TTC CAA CTC TCA ATC ATG TCC TTC TCA CCT GAC GAA ATC AAT Ser Phe Ser Pro Asp Glu Ile Asn GAC TTC AMG GMG GCC TTC CTC CTC TTT GAC AGG ACT GGT GAT GCC AMG ATT ACC CTG AGC Asp Phe Lys Glu Ala Phe Leu Leu Phe Asp Arg Thr Gly Asp Ala Lys Ile Thr Leu Ser CAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AMG GGT GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AMG CGG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AMG CAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AMG CGT GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AMG CGT GGC AAC CCC AGC AAA GAG GAA ATG AAT GCC AAG AMG ATC ACT TTT GAA GAG TTC ATC CTG GGC AAC CCC AGC AAA GAG GAA ATG AAT GCC AAG AMA ATC ACT TTT GAA GAG TTC ATC CTG GGC AAC CCC AGC GAC GCT GCT AAC AAC AAG GAT CAG GGT ACC TTT GAA GAG TTC TT CG CC ATG CTG CAA GCA GCT GCT AAC AAC AAG GAT CAG GGT ACC TTT GAA GAC TTT GTT Leu Gly Asn Pro Ser Lys Glu Glu Met Asn Ala Lys Lys Ile Thr Phe Glu Glu Phe 150 GTG CCC ATG CTG GTT TTC GAC ANG GAA GGC AAC AGC ACA GTC ATG GGG GCT GAG CTC CCC GU GTY Leu AIT Ala Ala Ala Asn Asn Lys Asp Gln Gly Thr Phe Glu Asp Phe Val 350 CAT GTA CTG GCT GTT TTC GAC AMG GAA GGA AGC ACA GGC ACA GTC ATG GGG GCT GAG CTC CCC GIU GIY Leu Arg Val Phe Asp Lys Glu Gly Asn Gly Thr Val Met Gly Ala Glu Leu Arg CAT GTA CTG GCT ACA CTG GCC GAG AMA GTA ACA AAC GAG GAA GTA GAA GAA CTG ATC ATA His Val Leu Ala Thr Leu Gly Glu Lys Met Thr Glu Glu Glu Val Glu Glu Val Glu Leu Met Lys 500 GTC CAA GAA GAC CTC AAC GGC GCC ACT ACA CAC CAC ATC ATG GTA CTA AGA GAA CTG ATC ATG GTA CTG GCT AGA GAA GTC CAA AGA GGA AGA GTA CAA GAA GTA GAA GAA CTG ATC ATA His Val Leu Ala Thr Leu Gly Glu Lys Met Thr Glu Glu Glu Val Glu Glu Leu Met Lys 500 GTC CAA GAA GAC CTC AAC GGC TGC ATC AAC TAAC CAA CAA GTA GAA GTA GAA GTA GAA GTA GTA ACC TTA AGA GAA CTC CTA CC AAC CTC CCA ACC ATC ATT GTT TTA AC CG TGC ATA CAA CAA TC ATG CTT GGA CAT TGG CTC AAG GAT GAA ATA TGC CTC	Lys	Glu	Pro	Ala	Ile	Asp	Leu	Lys	Ser	Ile	Lys	Ile	Glu	Phe	Ser	Lys	Glu	Gln	Gln	Asp
Ser Phe Ser Pro Asp Glu Ile Asn 150 150 150 150 150 150 150 150	GGC	CCG	TCC	GCT	CTA	CTT	TTC	CAA	CTC	TCA	ATC	ATG	TCC	TTC	TCA	CCT	GAC	GAA	ATC	AAT
GAC TTC AAG GAG GCC TTC CTC CTC TTT GAC AGG ACT GGT GAT GCC AAG ATT ACC CTG AGC GAC TTC AAG GAG GCC TTC CTC CTC TTT GAC AGG ACT GGT GAT GCC AAG ATT ACC CTG AGC Asp Phe Lys Glu Ala Phe Leu Leu Phe Asp Arg Thr Gly Asp Ala Lys Ile Thr Leu Ser 200 CAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AAG GAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AAG GAN TAl Gly Asp Ile Val Arg Ala Leu Gly Gln Asn Pro Thr Asn Ala Glu Ile Asn Lys ATC CTG GGC AAC CCC AGC AAG GAG GAA ATG AAT GCC AAG AAG ATC ACT TTT GAA GAG TTC ATC CTG GGC AAC CCC AGC AAA GAG GAA ATG AAT GCC AAG AAG ATC ACT TTT GAA GAG TTC ATC CTG GGC AAC CCC AGC AAG AAG GAA GAA GAA GA													Ser	Phe	Ser	Pro	Asp	Glu	Ile	Asn
GAC TTC AAG GAG GCC TTC CTC CTC TTT GAC AGG ACT GGT GAT GCC AAG ATT ACC CTG AGC GAC TTC AAG GAG GCC TTC CTC CTC TTT GAC AGG ACT GGT GAT GCC AAG ATT ACC CTG AGC Asp Phe Lys Glu Ala Phe Leu Leu Phe Asp Arg Thr Gly Asp Ala Lys Ile Thr Leu Ser 200 CAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AAG GIN Val Gly Asp Ile Val Arg Ala Leu Gly GIn Asn Pro Thr Asn Ala GLU Ile Asn Lys ATC CTG GGC AAC CCC ACC AAA GAG GAA ATG AAT GCC AAG AAG ATC ACT TTT GAA GAG TTC Ile Leu Gly Asn Pro Ser Lys Glu GLu Met Asn Ala Lys Lys Ile Thr Phe Glu Glu Phe 300 CTG CCC ATG CTG CAA GCG GCT GCT AAC AAC AAG GAT CAG GAT ACC TTT GAA GAG TTC Ile Leu Gly Asn Pro Ser Lys Glu GLU Met Asn Ala Lys Lys Ile Thr Phe Glu Glu Phe 300 CTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AAG GAT CAG GAT ACC TTT GAA GAC TTT GT Leu Pro Met Leu Gln Ala Ala Ala Asn Asn Lys Asp Gln Gly Thr Phe Glu Asp Phe Val 400 GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GAA GGT CTG CGT GTT TTC GAC AAG GAA GAA GAA GAA GAA GAA GAA GA																				
GAC TTC AMG GAG GCC TTC CTC CTC TTT GAC AGG ACT GGT GAT GCC AMG ATT ACC CTG AGC Asp Phe Lys Glu Ala Phe Leu Leu Phe Asp Arg Thr Gly Asp Ala Lys Ile Thr Leu Ser 200 CAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AMC GCT GAG ATC AAC AAG GAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AMC GCT GAG ATC AAC AAG GAN GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AMC GCT GAG ATC AAC AAG GIn Val Gly Asp Ile Val Arg Ala Leu Gly GIn Asn Pro Thr Asn Ala Glu Ile Asn Lys 250 ATC CTG GGC AAC CCC ACC AMA GAG GAA ATG AAT GCC AAG AAG ATC ACT TTT GAA GAG TTC ATC CTG GGC AAC CCC ACC AMA GAG GAA ATG AAT GCC AAG AAG ATC ACT TTT GAA GAG TTC Ile Leu Gly Asn Pro Ser Lys Glu Glu Met Asn Ala Lys Lys Ile Thr Phe Glu Glu Phe 300 CTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AAG GAT CAG GGT ACC TTT GAA GAC TTT GTT Leu Pro Met Leu Gln Ala Ala Ala Ala Asn Asn Lys Asp Gln GLy Thr Phe Glu Asp Phe Val 350 GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC AAC GCC ATG CTG TG GG GCT GAG CTC CCC GIU Gly Leu Arg Val Phe Asp Lys Glu Gly Asn Gly Thr Val Met Gly Ala Glu Leu Arg 450 CAT GTA CTG GCT ACA CTG GGC GAG AAG AGG AAG GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CCC GIU Gly Leu Arg Val Phe Asp Lys Glu Gly Asn Gly Thr Val Met Gly Ala Glu Leu Arg CAT GTA CTG GCT ACA CTG GGC GAG AAG AGA GAA GAA GAA GAA GAA GA														150						
GAC TTC AMG GAG GCC TTC CTC CTC TTT GAC AGG ACT GGT GAT GCC AMG ATT ACC CTG AGC APP Phe Lys Glu Ala Phe Leu Leu Phe Asp Arg Thr Gly Asp Ala Lys ILe Thr Leu Ser 200 CAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AAG Gln Val Gly Asp Ile Val Arg Ala Leu Gly Gln Asn Pro Thr Asn Ala Glu Ile Asn Lys 250 ATC CTG GGC AAC CCC AGC AAA GAG GAA ATG AAT GCC AAG AAG ATC ACT TTT GAA GAG TTC TT CTG GGC AAC CCC AGC AAA GAG GAA ATG AAT GCC AAG AAG ATC ACT TTT GAA GAG TTC TLE Leu Gly Asn Pro Ser Lys Glu Glu Met Asn Ala Lys Lys Ile Thr Phe Glu Glu Phe 300 CTG CCC ATG CTG CAA GCA GCT GCT TAAC AAC AAG GAT CAG GT ACC TTT GAA GAC TTT GT Leu Pro Met Leu Gln Ala Ala Ala Asn Asn Lys Asp Gln Gly Thr Phe Glu Asp Phe Val 400 GAA GGT CTG CGT GTT TTC GAC AAG GAA GAA GAA GAA GAA GAA GAA GA	GAC	TTC	AAG	GAG	GCC	TTC	CTC	CTC	TTT	GAC	AGG	ACT	GGT	GAT	GCC	AAG	ATT	ACC	CTG	AGC
Asp Phe Lys Glu Ala Phe Leu Leu Phe Asp Arg Thr Gly Asp Ala Lys IIe Thr Leu Ser200CAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AAGGAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AAGGIn Val Gly Asp IIe Val Arg Ala Leu Gly Gln Asn Pro Thr Asn Ala Glu IIe Asn LysATC CTG GGC AAC CCC AGC AAA GAG GAA ATG AAT GCC AAG AAG ATC ACT TTT GAA GAG TTCATC CTG GGC AAC CCC AGC AAA GAG GAA ATG AAT GCC AAG AAG ATC ACT TTT GAA GAG TTCATC CTG GGC AAC CCC AGC CAA GAG GAC ACG ACA AAG GAT CAG GGT ACC ATT TT GAA GAG TTCATC GTG CAT GTG CAA GCA GCT GCT AAC AAC AAG GAT CAG GGT ACC TTT GAA GAC TTT GTTLeu Gly Asn Pro Ser Lys Glu Glu Met Asn Ala Lys Lys IIe Thr Phe Glu Glu Glu PheCTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AAC GAG GAT CAG GGT ACC TTT GAA GAC TTT GTTLeu Pro Met Leu Gln Ala Ala Ala Asn Asn Lys Asp Gln Gly Thr Phe Glu Asp Phe Val350GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGCGAA GGT CTG CGT GTT TTC GAC AAG GAA GAC GAC GGC ACA GTC ATG GGG GCT GAG CTC CGCGAA GGT CTG CGT GTT TC GAC AAG GAA GAC GAA CGA CAG ACA GAT GAA GAA CTG ATG AAAHis Val Leu Arg Val Phe Asp Lys Glu Gly Asn Gly Thr Val Met Gly Ala Clu Leu ArgGGT CAG GAA GAC TCC AAC GGC GGC GGC AAC AAG ACA GAA CGA ACT AGA CTA AAAHis Val Leu Ala Thr Leu Gly Glu Lys Met Thr Glu Glu Glu Glu Glu Glu Ceu Met LysGGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAA CAA CAA CAA GAG GCA TTT GTA AAG CAA ATC ATG TCTGGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAA CAA CAA CAA CAA CAA CAA CA	GAC	TTC	AAG	GAG	GCC	TTC	CTC	CTC	TTT	GAC	AGG	ACT	GGT	GAT	GCC	AAG	ATT	ACC	CTG	AGC
CAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AAG GAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AAG GIn Val Gly Asp Ile Val Arg Ala Leu Gly GIn Asn Pro Thr Asn Ala Glu Ile Asn Lys ATC CTG GGC AAC CCC AGC AAA GAG GAA ATG AAT GCC AAG AAG ATC ACT TTT GAA GAG TTC ATC CTG GGC AAC CCC AGC AAA GAG GAA ATG AAT GCC AAG AAG ATC ACT TTT GAA GAG TTC Ile Leu Gly Asn Pro Ser Lys Glu Glu Met Asn Ala Lys Lys Ile Thr Phe Glu Glu Phe 300 CTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AAG GAT CAG GGT ACC TTT GAA GAG TTT GTT CTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AAG GAT CAG GGT ACC TTT GAA GAC TTT GTT Leu Pro Met Leu Gln Ala Ala Ala Asn Asn Lys Asp Gln Gly Thr Phe Glu Asp Phe Val 350 GAA GGT CTG CGT GTT TTC GAC AAG GAA GGA AGG AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GAA GGT CTG CGT GTT TTC GAC AAG GAA AGG AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GAA GGT CTG CGT GTT TTC GAC AAG GAA GGA AGG AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GIU Gly Leu Arg Val Phe Asp Lys Glu Gly Asn Gly Thr Val Met Gly Ala Glu Leu Arg CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAG GAA GTA GAA GAA CTG ATG AAA His Val Leu Ala Thr Leu Gly Glu Lys Met Thr Glu Glu Glu Val Glu Glu Leu Met Lys 500 GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT Gly Gln Glu Asp Ser Asn Gly Cys Ile Asn Thr Glu Ala Phe Val Lys His Ile Met Ser 550 GTC TAA GAG GAC CTC CAC ACG GC TC AAC TAC CAG GCA GTT TGTA AAG CAC ATC ATG TCT GGT GAG GAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAC CAA TGA CCA ATG GAT GGA GAC CTC CCC AGA TAA ATT CTT GCT CTG GCA GAT CAT CAT CAC CAT GAT GAC AAC TTA GAG GAC GTC CCC AGA TAA ATT CTT GCT CTG GCA GAT CAT CAT CAC CAA CGA GAT GAT GGT CTA GAG GAC GTC CCC AGA TAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GGG ATA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGT GAT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT CTT TTT TAA CCG GGC ATC CAT GGT GAG GAC	Asp	Pne	Lys	GIu	Ala	Phe	Leu	Leu	Phe	Asp	Arg	Thr	GIY	Asp	Ala	Lys	Ile	Thr	Leu	Ser
CAG GTC GGT GAC ATC GTT CGG GAC CTG GGA CAG GAA CCC CACA AAC GCT GAG ATC AAC AAG GIN Val Gly Asp lie Val Arg Ala Leu Gly GIN Asn Pro Thr Asn Ala Glu lie Asn Lys 250 ATC CTG GGC AAC CCC AGC AAA GAG GAA ATG GAT GCC AAG AAC CCC ATT TT GAA GAG TTC ATC CTG GGC AAC CCC AGC AAA GAG GAA ATG AAT GCC AAG AAG ATC ACT TTT GAA GAG TTC Ile Leu Gly Asn Pro Ser Lys Glu Glu Met Asn Ala Lys Lys lie Thr Phe Glu Glu Phe CTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AAG GAT CAG GGT ACC TTT GAA GAC TTT GTT CTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AAG GAT CAG GGT ACC TTT GAA GAC TTT GTT CTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AAG GAT CAG GGT ACC TTT GAA GAC TTT GTT CTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AAG GAC CAG GT CATG GGG GCT GAG CTC CGC GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GIU Gly Leu Arg Val Phe Asp Lys Glu Gly Asn Gly Thr Val Met Gly Ala Glu Leu Arg 400 CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAG GAA GTA GAA GAA CTG ATG AAA His Val Leu Ala Thr Leu Gly Glu Lys Met Thr Glu Glu Glu Val Glu Glu Leu Met Lys 500 GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC CAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC CAC GAG TAG AAA ATT CTT GCT CTG GCA GAT TTG TA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC CAC GAG ATA AATA CAA ATG AAC AAG TTG GAT AAG CAC ATC ATG TCT GGT CAA GAG GAC GTC CCC AGA TAA AATA CTA AAT AAC AAG TTG GAT CAT CAT CAA CCA GTA GGC GTC TAA GAG GAC TCC CAC AGA TA ACC CAC ACT CTT TTT TTT TAA CCG TGG AAA ACC TAA GAT CTG CCC GAG ATA AATA CTC CAA CCA CTA CTT TCT TTT TAA CCG TGG AAA CTG ACC GGA CAT TGC CTT GG GTC ACT TCC AAA ACT GAA CTT CTT CTT TTT TTA ACCG TGG AAA CTG ACC GGA CAT TGC CTC TGG G		CAG GTC GGT GAC ATC GTT CGG GCA CTG GGA CAG AAC CCC ACA AAC GCT GAG ATC AAC AAC																		
Cho Gir Gar Arc Gir Cor Cor Cor Gir Cor Cir Gor Cho Arc Cir Cir Act Cir Gir Arc Cir Gir Cho Arc Arc Arc Gir Gir Asp file Val Arg Ala Leu Gir Gir Asp Fro Thr Asp Ala Giu Ile Asp Lys 250 ATC CTG GGC AAC CCC AGC AAA GAG GAA ATG AAT GCC AAG AAG ATC ACT TTT GAA GAG TTC Ile Leu Gir Asp Pro Ser Lys Glu Glu Met Asp Ala Lys Lys Ile Thr Phe Glu Glu Phe 300 CTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AAG GAT CAG GGT ACC TTT GAA GAC TTT GTT Leu Pro Met Leu Gin Ala Ala Ala Asp Asp Lys Gln Gly Thr Phe Glu Asp Phe Val 400 GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GAA GGT CTG CGT GTT TTC GAC AAG GAA GAC AGA GAA GAA GGA AGG CAA GTA ATG GGG GCT GAG CTC CGC GAA GGT CTG CGT GTT TTC GAC AAG GAA GAA GAA AAC GAA GAA GAA GTA GAA GAA CTG ATG AAA CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAA GGA GTA GAA GAA CTG ATC AAA His Val Leu Ala Thr Leu Gly Glu Lys Met Thr Glu Glu Glu Val Glu Leu Met Lys 500 GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG ATC GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT Gly Gln Glu Asp Ser Asn Gly Cys Ile Asn Thr Glu Ala Phe Val Lys His Ile Met Ser 550 GTC TAA GAG GAC GTC CCC AGA TAA ATG AAC AAG TAG AAT TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG ATT CGG GAA AGA CCA GTA GAC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG ATT CGG GAA AGA CCA GAT GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG ATT TCG GAA AGA CCA GAT GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC TAT CTT GCT TTG GAA AGA CCA GAT GGC GTC TAA GAG GAC GTC CCC AGA ATA TGC CTC CAA ACC AC TCT CTT GCT GGG GGT TG TGG AAA CCC ATG GGA CAT TGC CTG GGC ACA TTA TGC CTC CAA ACC ACT TCT GGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA	CAG	GTC	GGT	GAC	ATC	GTT	CGG	GCA	CTG	GGA	CAG	AAC	000	ACA	AAC	GCT	GAG	ATC	AAC	AAG
Gin var Gry Asp Tie var Arg Ara Led Gry Gin Ash Pro The Yan Ara Giu Tie Ash Lys      250      ATC CTG GGC AAC CCC AGC AAA GAG GAA ATG AAT GCC AAG AAG AAG ATC ACT TTT GAA GAG TTC      Ile Leu Gly Asn Pro Ser Lys Glu Glu Met Asn Ala Lys Lys Ile Thr Phe Glu Glu Phe      CTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AAG GAT CAG GGT ACC TTT GAA GAC TTT GTT      CTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AAG GAT CAG GGT ACC TTT GAA GAC TTT GTT      CTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AAG GAT CAG GGT ACC TTT GAA GAC TTT GTT      Leu Pro Met Leu Gln Ala Ala Ala Asn Asn Lys Asp Gln Gly Thr Phe Glu Asp Phe Val      350      GAA GGT CTG CGT GTT TTC GAC AAG GAA GAC AC GCC ACA GTC ATG GGG GCT GAG CTC CGC      GU Gly Leu Arg Val Phe Asp Lys Glu Gly Asn Gly Thr Val Met Gly Ala Glu Leu Arg      ACT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAA GAA GTA GAA GTA GAA GAA CTG ATG AAA      CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAA GAA GTA GAA GAA CTG ATG AAA      CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAA GAG GAA GTA GAA GAA CTG ATG AAA      CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA CAA GAA GAA GTA GAA GAA CTG ATC AAA      His val Leu Ala Thr Leu Gly Glu Lys Met Thr Glu Glu Lu U Wal Glu Lu Met Lys      GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT      GGT CAG GAA GAC TCC AAC GGC TGC ATA AAA TAG AAA GAG TTG AAT TCG GAA AGA CCA GTA GGC      GT CAG GAA GAC TCC AAC GGC TCC AAC AAC ATC ATG TTT GTA AAG CAC ATC ATG TCT	CAG	U-1	GGT	GAC	ATC	GTT	CGG	GLA	CIG	GGA	CAG	AAC	Dec	ACA	AAC	GCT	GAG	ATC	AAC	AAG
ATC CTG GGC AAC CCC AGC AAA GAG GAA ATG AAT GAT GAC AAG AAG ATC ACT TTT GAA GAG TTC ATC CTG GGC AAC CCC AGC AAA GAG GAA ATG AAT GAC AAG AAG ATC ACT TTT GAA GAG TTC IIe Leu Gly Asn Pro Ser Lys Glu Glu Met Asn Ala Lys Lys IIe Thr Phe Glu Glu Phe 300 CTG CCC ATG CTG CAA GCA GCT GCT AAC AAA GAG GAT CAG GGT ACC TTT GAA GAG TTT GTT Leu Pro Met Leu Gln Ala Ala Ala Asn Asn Lys Asp Gln Gly Thr Phe Glu Asp Phe Val 400 GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GAA GGT CTG CGT GTT TTC GAC AAG GAA GAC AGA AGA GAG GAA GTA AGG CAA GTC ATG GGG GAA GGT CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAG GAA GTA GAA GAA CTG ATC AAA His Val Leu Ala Thr Leu Gly Glu Lys Met Thr Glu Glu Val Glu Glu Leu Arg 500 GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GIY Gln Glu Asp Ser Asn Gly Cys IIe Asn Thr Glu Ala Phe Val Lys His IIe Met Ser 500 ACC TTA GAG GAC GTC CCC AGA TAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GAT GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GAT GAC GTC GCT GAC TGG ACA ATA TGC CTC CAA ACC AC CAT TTT TTT TA ACCG TGC ATC GAT GGA CAT TGC CTT GGG GTC ACT TCC AAA ACT ACC ATC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTT GGG GTC ACT TCC AAA ACT GAA CTT TCT TTT TTA AC CGT GGA ATA TAG CAA TTA GAA ATT CTT GCT TTG GCA GGA TTG AAG ACCA TGG AAA ACT TGC CTT GG GCT GAC ATA TGC CTC CAA ACC ACT TCT GGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTT GG GTC ACT TCC AAA ACT GAA CTT TCT TTT TTT TTT TTA A	GIU	Val	GIŶ	Asp	ITE	var	Arg	ALA	Leu	GTÄ	GIN	ASN	PIO	Thr	ASI	ALA	GIU	IIe	Asn	rAa
ATC CTG GGC AAC CCC AGC AAA GAG GAA ATG AAT GCC AAG AAG ATC ACT TTT GAA GAG TTC 11e Leu Gly Asn Pro Ser Lys Glu Glu Met Asn Ala Lys Lys Ile Thr Phe Glu Glu Phe 300 CTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AAG GAT CAG GGT ACC TTT GAA GAC TTT GTT CTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AAG GAT CAG GGT ACC TTT GAA GAC TTT GTT Leu Pro Met Leu Gln Ala Ala Ala Asn Asn Lys Asp Gln Gly Thr Phe Glu Asp Phe Val 400 GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GIU Gly Leu Arg Val Phe Asp Lys Glu Gly Asn Gly Thr Val Met Gly Ala Glu Leu Arg 400 CTT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAG GAA GTA GAA GAC CTG ATG ATG AAG CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAG GAA GTA GAA GAA CTG ATG AAA His Val Leu Ala Thr Leu Gly Glu Lys Met Thr Glu Glu Glu Val Glu Glu Leu Met Lys 500 GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC CAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC CAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG GCT CT TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC Val cc 600 ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA 650 CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GCT ACT ACT TCC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GCT ACT TCC CAA AAC CTC TCT GAG ATT GAC CTC CAT GGA CAT TGC CTC TGG GCT ACT TCC CAA AAC CTC TCT GAG ATT GAC CTC CGA AAA CTG ACC GGA CAT TGC CTC TGG GCA CAC CCT CCT GTG TTC CTG GAG TTG CCC CCC G	ATC CTG GGC AAC CCC AGC AAA GAG GAA ATG AAT GCC AAG AAG ATC ACT TTT GAA GAG TTC															mme				
The Set Get Are Croc Are	ATC	CTG		AAC	CCC	AGC	222	CAG	CNA	ATC	AAT	GCC	AAG	AAG	ATC	ACT	111	GAA	GAG	TTC mmc
The bet of the bet by one of the the has here has here by by the time the of the the set of the	TIO	Len	600	Anc	Pro	Ser	T.ve	Clu	Clu	Mot	A an	Ala	Lue	Lug	TIO	The	Dho	GIN	Glu	Dhe
CTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AAC AAG GAT CAG GGT ACC TTT GAA GAC TTT GTT CTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AAG GAT CAG GGT ACC TTT GAA GAC TTT GTT Leu Pro Met Leu Gln Ala Ala Ala Asn Asn Lys Asp Gln Gly Thr Phe Glu Asp Phe Val 400 GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GAA GGT CTG CGT GTT TTC GAC AAG GAA AGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC G1U Gly Leu Arg Val Phe Asp Lys Glu Gly Asn Gly Thr Val Met Gly Ala Glu Leu Arg CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAA GAA GTA GAA GAA CTG ATG ATA GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAG GAA GTA GAA GAA CTG ATG ATA His Val Leu Ala Thr Leu Gly Glu Lys Met Thr Glu Glu Glu Val Glu Glu Leu Met Lys 500 GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GTG CAG GAA GAC TCC AAC GGC GGC TAC ATC ACT CAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GTG CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GTG CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GAA TTT GTA AAG CAC ATC ATG TCT GTG CAG GAA GAC TCC CAAC GGC TGC ATC AAC TAC AGA GAA GTG AAT TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GAT GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG GAT TAT GAA AGA CCA GAT GGC Val cc 600 ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GG GAC ATA TAG CAA ACC ACC ACC ACA TAT TTT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT CAT CAG CGG GGT TG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC CAAA ACC TCT TGA TTC TTT TCT		Deu	913	200	110	Der	519	Gru	Gru	nec	лап	AT0	цуа	719	116	1111	r ne	Gru	Gru	r ne
CTG CCC ATG CTG CAA GCA GCT GCT AAC AAC AAC AAC GAT CAG GGT ACC TTT GAA GAC TTT GTT Leu Pro Met Leu Gln Ala Ala Ala Ala Asn Asn Lys Asp Gln Gly Thr Phe Glu Asp Phe Val 400 GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GIU Gly Leu Arg Val Phe Asp Lys Glu Gly Asn Gly Thr Val Met Gly Ala Glu Leu Arg CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAA GGA GTA GAA GAA CTG ATG AAA His Val Leu Ala Thr Leu Gly Glu Lys Met Thr Glu Glu Glu Val Glu Glu Leu Met Lys GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC CAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG GTC 1/y Gln Glu Asp Ser Asn Gly Cys II & Asn Thr Glu Ala Phe Val Lys His II & Met Ser 550 GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC Val cc 600 ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA 650 CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG CAC ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCC ACC ACC CCC CAA ACC CTC GGA TTC CTG GAG TTG ACC CAC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT CTG GGG GGT TTG GG AAA CTG ACC 80	CTG	ccc	ATG	CTG	CAA	GCA	GCT	GCT	AAC	AAC	AAG	GAT	CAG	GGT	ACC	ጥጥጥ	GAA	GAC	ттт	ርጥጥ
Leu Pro Met Leu Gln Ala Ala Ala Ash Ash Lys Asp Gln Gly Thr Phe Glu Asp Phe Val 400 GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GIU Gly Leu Arg Val Phe Asp Lys Glu Gly Ash Gly Thr Val Met Gly Ala Glu Leu Arg CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAA GAA GTA GAA GAA CTG ATG AAA CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAG GAA GTA GAA GAA CTG ATG AAA His Val Leu Ala Thr Leu Gly Glu Lys Met Thr Glu Glu Glu Val Glu Glu Leu Met Lys 500 GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAC TTA GTA AAG CAC ATC ATG TCT 510 GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GAT GGC Val cc 600 ACC TAA GAG GAC GTC CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA 650 CTG GGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TTT TAA CCG TGC AAA 650 CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TTT TAA CCG TGC AAA 650 CTG GAC AGT CTG GCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA 650 ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA 650 ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA 650 CTG GGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT CTT TTT TAA CCG TGC AAC CAT 600 ACC GCA CCA CCA CCA CCA ACC TTC CAA ACC ACT TCT GGA TTG ACC TCC GAA ACT ACC 800 ACT GCC GCA CCA CCA CCC ACC ACC TCT GAG ATT CAT CAT GCC GGA GAA CTG ACC 800 ACT GCC GCA CCA CCA CCC AAC ACC TTC GAA ACT GAA CTT CAC GGG GGG TTG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC CCA ACC CTC GAG TTC CTG	CTG	CCC	ATG	CTG	CAA	GCA	GCT	GCT	AAC	AAC	AAG	GAT	CAG	GGT	ACC	TTT	GAA	GAC	TTT	GTT
<sup>350</sup> <sup>400</sup> <sup>40</sup>	Leu	Pro	Met	Leu	Gln	Ala	Ala	Ala	Asn	Asn	Lvs	Asp	Gln	Glv	Thr	Phe	Glu	Asp	Phe	Val
GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC G1U G1y Leu Arg Val Phe Asp Lys G1u G1y Asn G1y Thr Val Met G1y Ala G1u Leu Arg CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAA GAG GAA GTA GAA GAA CTG ATG AAA CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAA GAG GAA GTA GAA GAA CTG ATG AAA CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAA GAG GAA GTA GAA GAA CTA ATC AAA His Val Leu Ala Thr Leu G1y G1u Lys Met Thr G1u G1u G1u G1u G1u G1u G1u G1u Leu Met Lys 500 GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GAC TTT GTA AAG CAC ATC ATG TCT G1y G1n G1u Asp Ser Asn G1y Cys I1e Asn Thr G1u Ala Phe Val Lys His I1e Met Ser 550 GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC Val cc 600 ACC TAA GAT GTC CTG GCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT AGC CTC CAA ACC CA CTA CTT TTT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCA CCC ACC CCC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAA AAA ACC CTC TG GAA TTC ACC TCC GAA AAT AAA GCA 848	350										-1-			1			40	0.05		
GAA GGT CTG CGT GTT TTC GAC AAG GAA GGC AAC GGC ACA GTC ATG GGG GCT GAG CTC CGC GIU GIY Leu Arg Val Phe Asp Lys Glu GIY Asn GIY Thr Val Met GIY Ala Glu Leu Arg CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAA GAA GTA GAA GAA CTG ATG AAA CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAA GAA GTA GAA GAA CTG ATC AAA His Val Leu Ala Thr Leu GIY Glu Lys Met Thr Glu Glu Glu Val Glu Glu Leu Met Lys 500 GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GIY Gln Glu Asp Ser Asn GIY Cys Ile Asn Thr Glu Ala Phe Val Lys His Ile Met Ser 550 GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG GAT TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG GAT TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG GAT TCG GAA AGA CCA GTA GGC Val cc 600 ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA 650 CTG GTG GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CAT GC TGG CAT GG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CAT GC TGG CAT GG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CAT GC TGG CTC TGG GTC ACT TCC AAA ACT GAA CTT CAT CGG GGG TTG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCC ACC TCT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA AAA	GAA	GGT	CTG	CGT	GTT	TTC	GAC	AAG	GAA	GGC	AAC	GGC	ACA	GTC	ATG	GGG	GCT	GAG	CTC	CGC
GIU GIY Leu Arg Val Phe Asp Lys Glu Gly Asn Gly Thr Val Met Gly Ala Glu Leu Arg 450 CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAA GAG GAA GTA GAA GAA CTG ATG AAA CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAA GAG GAA GTA GAA GAA CTG ATC AAA His Val Leu Ala Thr Leu Gly Glu Lys Met Thr Glu Glu Glu Val Glu Glu Leu Met Lys GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GTC TAA GAG GAC GTC CCA AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GIY Gln Glu Asp Ser Asn Gly Cys 1le Asn Thr Glu Ala Phe Val Lys His Ile Met Ser 550 GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GAT GGC Val cc 600 ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA 650 CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT CAT CGG GGG TTG TGG AAA CTG ACC 600 ACC TAA GAT CTG CCT TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC 600 ACC TAA GAT CTG CTC TGG GTC ACT TCC AAA ACT GAA CTT TTT TAT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC TAC TAT TTT TCT TTT TAA CCG TGC ATC CAT 600 ACT GCC GCA CCA CCA CCA CCC TCT CTG AGA CTT CAT GAA CTT GCG GGG TTG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCA ACC TCT CTG AGA TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA AAA	GAA	GGT	CTG	CGT	GTT	TTC	GAC	AAG	GAA	GGC	AAC	GGC	ACA	GTC	ATG	GGG	GCT	GAG	CTC	CGC
CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAA GAG GAA GTA GAA GAA CTG ATG AAA CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAA GAG GAA GTA GAA GAA CTG ATG AAA His Val Leu Ala Thr Leu Gly Glu Lys Met Thr Glu Glu Glu Val Glu Glu Leu Met Lys 500 GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GJY Gln Glu Asp Ser Asn Gly Cys Ile Asn Thr Glu Ala Phe Val Lys His Ile Met Ser 550 GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG GAA TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC Val cc 600 ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT TT GCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACC GAA CTT CAA CCA TGG AAA ACC TAA GAT CTG CCT CGG GTC ACT TCC AAA ACC GAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATA TAG CTC CAT CAA CCA TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT CAA CGG GGG TTG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCA CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCA CCC CAC CAC TCT GG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA AAA AA	GlÜ	Gly	Leu	Arg	Val	Phe	Asp	Lys	Glu	Gly	Asn	Gly	Thr	Val	Met	Gly	Ala	Glu	Leu	Arg
CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAA GAG GAA GTA GAA GAA CTG ATG AAA CAT GTA CTG GCT ACA CTG GGC GAG AAG ATG ACA GAA GAA GAG GAA GTA GAA GAA CTG ATC AAA His Val Leu Ala Thr Leu Gly Glu Lys Met Thr Glu Glu Glu Glu Glu Leu Met Lys 500 GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GJy Gln Glu Asp Ser Asn Gly Cys Ile Asn Thr Glu Ala Phe Val Lys His Ile Met Ser 550 GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC Val cc 600 ACC TAA GAG GAC GTC CCC GAG ATA AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATA CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CAT GGC ACA TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT CAC GGG GGG TTG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCC CAC CCC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC CAC CCC TCT GG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAAA AAA AAA AAA AAA AAA AAA AAA				-			-	-		-		-		450		-				•
CAT GTA CTG GCT ACA CTG GGC GAG AAG AAG ATG ACA GAA GAG GAA GTA GAA GAA CTG ATC AAA His Val Leu Ala Thr Leu Gly Glu Lys Met Thr Glu Glu Glu Val Glu Glu Leu Met Lys 500 GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT Gly Gln Glu Asp Ser Asn Gly Cys Ile Asn Thr Glu Ala Phe Val Lys His Ile Met Ser 550 GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GAT GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GAT GGC Val cc 600 ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA 650 CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT CAC CGG GGG TTG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCC ACC TCT CTT GAA TTG CAG GGG TTG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCC CAC CAC TCT TTT CCT TT TCA TGT TGT GG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA AAA	CAT	GTA	CTG	GCT	ACA	CTG	GGC	GAG	AAG	ATG	ACA	GAA	GAG	GAA	GTA	GAA	GAA	CTG	ATG	AAA
His Val Leu Ala Thr Leu Gly Glu Lys Met Thr Glu Glu Glu Val Glu Glu Leu Met Lys GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GJy Gln Glu Asp Ser Asn Gly Cys Ile Asn Thr Glu Ala Phe Val Lys His Ile Met Ser 550 GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GAT GGC Val cc 600 ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA 650 CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT CAA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCA CCC CAC CTCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC CAAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA AAA	CAT	GTA	CTG	GCT	ACA	CTG	GGC	GAG	AAG	ATG	ACA	GAA	GAG	GAA	GTA	GAA	GAA	CTG	ATC	AAA
GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT Gly Gln Glu Asp Ser Asn Gly Cys Ile Asn Thr Glu Ala Phe Val Lys His Ile Met Ser 550 GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC Val cc 600 ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT CAA CGG GGG TTG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCA CCC CAC ACC TCT GTG TTC CTG GGG GGG TTG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCC CAAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC CAAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAAA AAA AAA AAA AAA AAA AAA AAA AAA	His	Val	Leu	Ala	Thr	Leu	Gly	Glu	Lys	Met	Thr	Glu	Glu	Glu	Val	Glu	Glu	Leu	Met	Lys
GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT GGT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT Gly Gln Glu Asp Ser Asn Gly Cys Ile Asn Thr Glu Ala Phe Val Lys His Ile Met Ser 550 GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC Val cc 600 ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA CCT GAG GAC TGG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA CCT GAG GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GTT TCA CGG GGG TTG TGG AAA CTG ACC ACT GCC GCA CCA CCA CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA AAA											500									
GT CAG GAA GAC TCC AAC GGC TGC ATC AAC TAC GAG GCA TTT GTA AAG CAC ATC ATG TCT Gly Gln Glu Asp Ser Asn Gly Cys Ile Asn Thr Glu Ala Phe Val Lys His Ile Met Ser 550 GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GAT GGC 400 ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA 650 CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCC CAC CAC TCT TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG ATC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG ATC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG ATC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCA CCC AAC ACC TCT GAG ATC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCA CCA ACC TCT GAG ATC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCA CCA AAC ACC TCT GAG ATC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCA CCA ACC TCT GAG ATC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCA CCC AAC ACC TCT GAG ATC CTG AGA TTG ACC TCC GAA AAT AAA GCA	GGT	CAG	GAA	GAC	TCC	AAC	GGC	TGC	ATC	AAC	TAC	GAG	GCA	TTT	GTA	AAG	CAC	ATC	ATG	TCT
GIY GIN GIU ASP SER ASH GIY CYS IIE ASH THY GIU AIA PHE VAI LYS HIS IIE MET SET 550 GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GAT GGC Val cc 600 ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA 650 CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA AAA	GGT	CAG	GAA	GAC	TCC	AAC	GGC	TGC	ATC	AAC	TAC	GAG	GCA	TTT	GTA	AAG	CAC	ATC	ATG	TCT
550 GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GAT GGC Val cc 600 ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCC CAC CCC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCC CAC CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCC CAAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA AAA	GIY	GIN	GIU	Asp	Ser	Asn	GIY	Cys	Ile	Asn	Thr	Glu	Ala	Phe	Val	Lys	His	Ile	Met	Ser
GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GTA GGC GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GAT GGC Val cc 600 ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA CCT GAG TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC ACT GCC GCA CCA CCA CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA							550	)												
GTC TAA GAG GAC GTC CCC AGA TAA CAA ATG AAC AAG TTG AAT TCG GAA AGA CCA GAT GGC Val cc 600 ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TAC TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA	GTC	TAA	GAG	GAC	GTC	CCC	AGA	TAA	CAA	ATG	AAC	AAG	TTG	AAT	TCG	GAA	AGA	CCA	GTA	GGC
Val cc 600 ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA 650 CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TAC TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC B00 ACT GCC GCA CCA CCA CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCC CAC CAC CCC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA B48 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA AAA AAA	GTC	TAA	GAG	GAC	GTC	CCC	AGA	таа	CAA	ATG	AAC	AAG	TTG	AAT	TCG	GAA	AGA	CCA	GAT	GGC
ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATA ATC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACC GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC ACT GCC GCA CCA CCA CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC CAAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCC CCA ACA ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA B48 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA	Val	œ																		
ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA ACC TAA GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAA CCA TGG AAA 650 CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACC GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC ACT GCC GCA CCA CCA CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA				600																
ACC TAR GAT CTG CCT GAG ATG AAA ATT CTT GCT CTG GCA GAT CAT CAT CAT CAA CCA TGG AAA 650 CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT ACA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA	ACC	TAA	GAT	CTG	CCT	GAG	ATG	AAA	ATT	CTT	GCT	CTG	GCA	GAT	CAT	CAT	CAA	CCA	TGG	AAA
650    CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT      CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT      GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT ACA CGG GGG TTG TGG AAA CTG ACC      GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC      ACT GCC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC      ACT GCC GCA CCA CCA CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA      ACT GCC GCA CCA CCC ACC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA      ACT GCC GCA CCA CCC CAAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA      ACT GCC GCA CCA CCC CAAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA      ACT GCC GCA CCA CCC AAC AAC ACC TCT GAG ATC CTG AGA TTG ACC TCC GAA AAT AAA GCA      ACT GCC GCA CCA CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA      TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA	ACC	TAA	GAT	CTG	CCT	GAG	ATG	AAA	ATT	CTT	GCT	CTG	GCA	GAT	CAT	CAT	CAA	CCA	TGG	AAA
CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT 750 GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA	650																_70	0		
CTG CGT GAC TGG ACA ATA TGC CTC CAA ACC CAC TAC TTT TCT TTT TAA CCG TGC ATC CAT GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC ACT GCC GCA CCA CCA CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA	CTG	CGT	GAC	TGG	ACA	ATA	TGC	CTC	CAA	ACC	CAC	TAC	TTT	TCT	TTT	TAA	CCG	TGC	ATC	CAT
GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC GGA CAT TGC CTC TGG GTC ACT TCC AAA ACT GAA CTT TCA CGG GGG TTG TGG AAA CTG ACC 800 ACT GCC GCA CCA CCA CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCC CAC ACC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA	CTG	CGT	GAC	TGG	ACA	ATA	TGC	CTC	CAA	ACC	CAC	TAC	TTT	TCT	TTT	TAA	CCG	TGC	ATC	CAT
GA CAT THE CTC THE GTC ACT THE AAA ACT GAA CTT THA CHE GHE GHE THE ACE GAA CAT THE CTC THE GTC ACT THE AAA ACT GAA CTT THA CHE GHE GHE AAA CTG ACC 800 ACT GCC GCA CCA CCA CCC AAC ACC THE GTG THE CTG AGA TTG ACC THE GHE AAA CTG ACC ACT GCC GCA CCA CCA CCC AAC ACC THE GTG THE CTG AGA TTG ACC THE GHE AAA CTG ACC 848 TCA GTA AAT CCC THE THE CAA AAA AAA AAA AAA AAA AAA AAA TCA GTA AAT CCC THE THE CAA AAA AAA AAA AAA AAA AAA AAA AAA AAA	000		-	-	-	~~~	100	<b>m</b> 00		100	~	000	m03	750	~~~		mcc		CIEC	200
ACT GCC GCA CCA CCA CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCA CCC AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA	GGA	CAT	TGC	CTC	TGG	GTC	ACT	TCC	AAA	ACT	GAA	CTT	TCA	CGG	GGG	TTG	TGG	AAA	CTG	ACC
ACT GCC GCA CCA CCA CCC AAC ACC TCT GTG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA ACT GCC GCA CCA CCC CCA AAC ACC TCT GAG TTC CTG AGA TTG ACC TCC GAA AAT AAA GCA $848$ TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA	GGA	CAT	TGC	CTC	166	GTC	ACT	TUU	AAA	AUT	200	CIT	TCA	600	999	1.1.0	1.99	~~~	C10	ALL
ACT GCC GCA CCA CCC ACC ACC TCT GAG TIC CIG AGA TIG ACC TCC GAA AAT AAA GCA 848 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA	200		<b>CC</b> 2	~~~	~~~	~~~~	220	200	mom	CTTC	,00 mmc	CTTC	101	ጥጥሮ	300	TOO	GAA	220	***	603
848 TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA AAA TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA	ACT	300	GCA	CCA	CCA	ccc	AAC	ACC	TCT	GAG	TTC	CTG	AGA	TTG	ACC	TCC	GAA	AAT	AAA	GCA
тса бта лат ССС тбт тбт сал ала ала ала ала ала ала тса бта алт ССС тбт тбт сал ала ала ала ала ала ала		300	JUN	un		000	348		101	0.00		0.0								
TCA GTA AAT CCC TGT TGT CAA AAA AAA AAA AAA AAA AAA	TCA	GTA	ААТ	CCC	TGT	TGT	CAA	АЛА	АЛА	ала	ала	ала	хаа							
	TCA	GTA	AAT	CCC	TGT	TGT	CAA	AAA	AAA	AAA	AAA	AAA	AAA							

Fig. 2. Nucleotide sequences and predicted amino acid sequences of pSMA1-1 and pSMA2-1 cDNAs. The amino acid sequence was deduced from the sequence of cDNA using the genetic code.  $\infty$ , the ochre terminater. \*, shows the first nucleotide of the identical 741 nucleotides between  $A_1$  and  $A_2$  cDNAs.

sequencing and the restriction map of three cDNAs are shown in Fig. 1. The sequence determination was performed on both strands of the cDNA except for the 3'-terminal 219 base pairs. The inserted sequence of pSMA1-1 corresponded to amino acid residues 15 to 190 of the  $A_1$  chain determined by Matsuda et al. (22) and to the 3'-untranslated region extending to the poly(A) segment. The codon for the valine residue at the carboxy terminus of  $A_1$  chain is followed by the translational termination codon TAA. Thus the 3'-untranslated region of  $A_1$  mRNA is 315 nucleotides long (excluding the poly(A) tract), a putative poly(A) signal AATAAA is present 22 bases away from poly(A) sequence.

The insert of pSMA2-1 covered the coding sequence for amino acid residues 1 to 149 of the  $A_2$  chain (22) and the untranslated sequence for the 5'- and 3'-end regions. The translational initiator codon ATG is followed by the serine codon of the amino terminus of the  $A_2$  chain.

These sequences disagree at two positions with a previous report (22) determined by amino acid sequencing techniques. At positions 6 and 8 of the  $A_2$  chain, we identified Glu and Asn, while Matsuda et al. reported Gln and Asp, respectively. Comparisons of nucleotide sequences of pSMAl-1 and pSMA2-1

The sequences of pSMA1-1 and pSMA2-1 were aligned to examine their homology. As shown in Fig. 3, the 741 nucleotides from residues 108 to 848 of pSMA1-1 and those of 78 to 818 of pSMA2-1 agreed completely. This contains not only the sequences of the coding region corresponding to the carboxy-terminal 141 amino acid residues of the  $A_1$  and  $A_2$  chains, but also all the 3'-untranslated regions of  $A_1$  and  $A_2$ . On the other hand, the 5'terminal 108 nucleotides of pSMA1-1 and 78 nucleotides of pSMA2-1 were quite different, showing no homology between them.

### DISCUSSION

The primary structures of skeletal muscle myosin  $A_1$  and  $A_2$  chains of the rabbit (21) and those of the chicken (22) have been determined already. As shown in Fig. 3(A) the C-terminal 141 amino acid residues of chicken  $A_1$  and chicken  $A_2$  are identical. Those of rabbit  $A_1$  are identical with those of rabbit  $A_2$ . Furthermore, not only the positions of amino acid substitution but also the kinds of the substituted amino acids between chicken  $A_1$  and rabbit  $A_1$  are identical with those between chicken  $A_2$  and rabbit  $A_2$ . On the other hand, the sequences of N-terminal 50 amino acid residues are quite different between chicken  $A_1$  and rabbit  $A_1$ . The N-terminal 8 residues of chicken  $A_2$  are also different from the corresponding residues of rabbit  $A_2$ .

Fig. 2 shows the comparison of the nucleotide sequences of mRNAs of chicken skeletal muscle light chain  $A_1$  and  $A_2$ , which are



Fig. 3. Schematic diagrams of the primary structures of A1 and A2 peptides of chicken and rabbit skeletal muscle myosin (A) and nucleotide sequences of mRNAs for chicken skeletal muscle myosin A1 and A2 (B). A. the open bars indicate the common region of A1 and A2 peptides of chicken and rabbit skeletal muscle. The dashed lines in the open bars showing rabbit A1 and A2 peptides show the amino acid substitutions between chicken and rabbit skeletal muscle myosin alkali light chains. The solid and shaded bars show the A1 and A2 specific regions, respectively. The amino acid substitutions in A1 and A2 specific regions between chicken and rabbit skeletal myosin are not shown. B. The open bars and the open bars with center line indicate the identical 741 nucleotides of A1 and A2 mRNAs corresponding to common regions (with center line), respectively. The solid and shaded bars show the A1 and A2 specific nucleotide sequences of mRNAs, respectively. Dotted area indicates the A2 specific 5'-terminal amino acid coding sequence and shaded area represents the leader sequence.

schematically described in Fig. 3(B). The 3'-terminal nucleotide sequences of chicken  $A_1$  and  $A_2$  are identical, which contain not only the corresponding region to C-terminal 141 amino acid residues but also contain the following 3'-untranslated region. On the other hand, 5'-terminal nucleotide sequences of  $A_1$  mRNA corresponding to the N-terminal 36 amino acid residues of  $A_1$  light chain are different from those of the  $A_2$  mRNA corresponding to the N-terminal 8 amino acid residues of  $A_2$  peptide and the 5'-untranslated region.

What is the relationship between the genes for the A1 and A2 proteins? There are two possibilities. One is that the A, and A, peptides are encoded by two distinct genes. The other is that while the 5'-terminal regions of the A1 and A2 mRNAs are encoded in two separate DNA segments, their 3'-terminal 741 nucleotides are encoded in a common DNA segment on one gene. The total identity of the relatively long, 3'-terminal sequences of the mRNA, suggests the latter possibility. Two mechanisms have been reported for the phenomenon that genes for two or many kinds of proteins contain one DNA segment in common. One is the gene recombination that takes place on immunoglobulin genes (33, 34), and the other is RNA-RNA splicing that has been shown in the expression of either secreted and membrane types of immunoglobulin  $\mu$  chain and  $\delta$  chain, (35, 36) or  $\alpha$ -amylase mRNAs in two different tissues; the salivary gland and the liver (37). Experiments to distinguish between these two possibilities in the case of chicken myosin light chain  $A_1$  and  $A_2$  are now in progress.

It is known that chicken myosin  $A_2$  chain is not synthesized in the embryonic muscle. It will be interesting to know the mechanism by which the pattern of gene expression changes from only the  $A_1$  peptide to the  $A_1$  plus  $A_2$  peptides during the course of skeletal muscle differentiation.

For the case of the myosin alkali light chains, heart and gizzard muscles are composed of one alkali light chain homologous to skeletal muscle  $A_1$  and  $A_2$  light chains, respectively (23). It is also of interest as to why heart and gizzard muscles express only one type of alkali light chain, while skeletal muscles express both types.

During the preparation of this manuscript, Robert et al. (38)

reported the cloning of cDNA for an alkali light chain of mouse skeletal muscle myosin and determined the partial nucleotide sequences of its 3'-untranslated region.

REFERENCES:

- 1. Weeds, A.G. and Pope, B. (1977) J. Mol. Biol. 111, 129-157.
- 2. Coleman, J.R. and Coleman, A.W. (1968) J. Cell. Physiol. 72, 19-34.
- Reporter, M.C., Konigsberg, I.R. and Strehler, B.L. (1963) Exp. Cell Res. <u>30</u>, 410-420.
  Shainberg, A., Yagil, G. and Yaffe, D. (1971) Dev. Biol. <u>25</u>,
- 1-29.
- 5. Paterson, B.M. and Strohman, R.C. (1972) Dev. Biol. 29, 113-138.
- 6. Chi, J.C., Rubinstein, N., Strahs, K. and Holtzer, H. (1975) J. Cell Biol. 67, 523-557.
- 7. Carmon, Y., Neuman, S. and Yaffe, D. (1978) Cell, 14, 393-401.
- 8. Chi, J.C., Fellini, S.A. and Holtzer, H. (1975) Proc. Natl. Acad. Sci. USA 72, 4999-5003. 9. Masaki, T. and Yoshizaki, C. (1974) J. Biochem. 76, 123-131.
- 10. Streter, F.A., Balient, M. and Gregely, J. (1975) Dev. Biol. 46, 317-325.
- 11. Lowey, S. and Risky, D. (1971) Nature, (London) <u>234</u>, 81-85.
- 12. Sarker, S., Sreter, F.A. and Gergely, J. (1971) Proc. Natl. Acad. Sci, USA 68, 946-950.
- 13. Yaffe, D. and Dym, H. (1972) Cold Spring Harbor Symp. Quant, Biol. 37, 543-547.
- 14. Low, R.B., Vournakis, J.N. and Rich, A. (1971) Biochemistry, 10, 1813-1818.
- Whalen, R.G., Sell, S.M., Butler-Browne, G.S., Schwartz, K., Bouveret, P. and Pinset-Härsträm (1981) Nature, 292, 805-809.
- 16. Arnold, H.H., Siddiqui, M.A.Q. (1979) Biochemistry, 18, 5641-5647.
- 17. Katcoff, D., Nudel, U., Zevir-Sonkein, D., Carmon, Y., Shani, M., Lehrach, H., Frischauf, A.M. and Yaffe, D. (1980) Proc. Natl. Acad. Sci. USA 77, 960-964.
- 18. Medford, R.M., Wydro, R.M., Nguyen, H.T. and Nadal-Ginard, B. (1980) Proc. Natl. Acad. Sci. USA 77, 5749-5753.
- 19. Nedel, U., Katcoff, D., Carmon, Y., Zenin-Sonkin, D., Levi, A., Shaul, Y., Shani, M. and Yaffe, D. (1980) Nucleic Acids Res. 8, 2133-2146.
- 20. Robbens, J., Freyer, G.A., Chisholm, D. and Gilliam, T.C. (1982) J. Biol. Chem. <u>257</u>, 549-556. 21. Frand, G. and Weeds, H.G. (1974) Eur. J. Biochem. <u>44</u>, 317-334. 22. Matsuda, G., Maita, T. and Umegane, T. (1981) FEBS Lett. <u>126</u>,
- 111-113.
- Matsuda, G., Maita, T., Kato, Y., Chen, J.I. and Umegane, T. (1981) FEBS Lett. <u>135</u>, 232-236.
  Kennedy, D.S. and Heywood, S.M. (1976) FEBS Lett. <u>72</u>, 314-318.
- 25. Nabeshima, Y. and Ogata, K. (1980) Eur. J. Biochem. 107, 323-329.
- 26. Duquid, J.R., Steiner, D.F. and Chick, W.L. (1976) Proc. Natl. Acad. Sci. USA 73, 3539-3542.
- 27. Pelham, H.R.B., Jackson, J. (1976) Eur. J. Biochem. 67, 247-256.

- 28. Roychoundhury, R., Jay, E. and Wu, R. (1976) Nucleic Acids Res. 3, 101-116.
- 29. Grunstein, M. and Hogness, D.S. (1975) Proc. Natl. Acad. Sci. USA <u>72</u>, 3961-3965.
- 30. Hoeijmakers, J.H.J., Bosst, P., Van Den Burg, J., Weissmann, C. and Cross, G.A.M. (1980) Gene, 8, 391-417.
- 31. Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.
- 32. Staynov, D.Z., Pinder, J.C. and Gratzer, W.B. (1972) Nature New Biol. 235, 108-110.
- 33. Brack, C., Hirama, M., Schuller, R. and Tonegawa, S. (1978) Cell, <u>15</u>, 1-14.
- 34. Sakano, H., Huppi, D., Heinrich, G. and Tonegawa, S. (1979) Nature, <u>280</u>, 288-294.
- 35. Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R. and Hood, L. (1980) Cell, <u>20</u>, 313-319.
- 36. Maki, R., Roeder, W., Traunecker, A., Raschke, W. and Tonegawa, S. (1981) Cell, <u>25</u>, 353-365.
- 37. Young, R.A., Hagenbuchle, O. and Schibler, U. (1981) Cell, <u>23</u>, 451-458.
- Robert, B., Weydert, A., Caravatti, M., Minty, A., Cohen, A., Daubas, Ph., Gros, F. and Buckingham, M. (1982) Proc. Natl. Acad. Sci. USA <u>79</u>, 2437-2441.