CLONING AND CHARACTERIZATION OF SEIZURE-RELATED GENE, SEZ-6

Keiko Shimizu-Nishikawa¹, Kagemasa Kajiwara^{1,2} and Eiichi Sugaya^{1,2}*

¹ Department of Physiology, Kanagawa Dental College, Kanagawa 238, Japan

² Division of Neurobiology, Tsumura Institute for Oriental Medicine, School of Medicine, Keio University, Tokyo 160, Japan

Received September 27, 1995

SUMMARY SEZ-6 is a brain-specific cDNA. Its expression is increased by convulsant drug and it encodes membrane protein with five copies of short consensus repeat (SCR; complement C3b/C4b binding site) and two repeated sequences which are partially similar to CUB domain (complement C1r/s-like repeat). In this study, we analyzed the gene product of SEZ-6 by antibody raised against SEZ-6 peptide. *In vitro* translation and immunoblot revealed that SEZ-6 protein is modified post-translationally and expressed in the cerebrum and the cerebellum. To analyze the SEZ-6 expression in detail, we searched for other types of SEZ-6 cDNA by the polymerase chain reaction. Two additional types of cDNAs were isolated and sequence analysis showed that one cDNA encodes membrane protein with a different C-terminal region and the other encodes secreted protein with two SCRs and one CUB-like domain.

SEZ-6 is a brain-specific cDNA which was isolated from the cerebrum cortex cell cDNA library treated with PTZ, one of the convulsant drugs, by a differential hybridization technique (1). During screening, many clones were found to hybridized differently between the PTZ-treated and control library (2). Among these clones, SEZ-6 was characterized by increased expression with PTZ. Sequence analysis revealed that SEZ-6 encodes transmemebrane protein with multiple motifs, five SCR (reviewed in 3, 4) and two CUB-like domains (5). These motifs are often seen in the proteins of the immune system and were suggested to have roles for protein-protein interaction. Although recent studies show that not only the proteins in the immune system but two proteins in the nervous system have SCR (6) or CUB domain (7), SEZ-6 is the only cDNA in the nervous system that encodes both SCR and CUB domains. In the present study, to clarify the profile of SEZ-6 expression and gene product, we analyzed SEZ-6 protein with a specific antibody and cloned other types of SEZ-6 cDNA by RT-PCR.

<u>Abbreviations</u>: CBB, coomassie brilliant blue; ORF, open reading frame; PCR, polymerase chain reaction; PI, phosphatidylinositol; PTZ, pentylentetrazole; RT-PCR, reverse transcription-polymerase chain reaction; SCR, short consensus repeat; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

^{*} To whom correspondence should be addressed. Fax: 422-42-0532.

MATERIALS AND METHODS

Antibody production. Polyclonal antibody against to SEZ-6 protein was produced by immunizing rabbits using the synthetic peptide CTYETGETREYEVSI corresponding to the C-terminus of SEZ-6 type 1 protein. The synthetic peptide was conjugated to keyhole limpet hemocyanin and used for immunization. The antiserum was purified by an affinity column of FMP-activated cellulofine (Seikagaku Co. Ltd.) which was coupled with synthetic peptide according to the manufacture's instruction.

Immunoblot. Proteins were extracted with sample buffer from various organs or cultured cerebral cortex cells (8) of C57BL/6J mice and separated on SDS-PAGE (9). The proteins were transferred to PVDF membrane (Bio-Rad) according to Towbin et al. (10) The membranes were incubated with affinity purified antiserum (primary antibody) and peroxidase-labeled anti-rabbit antibody (secondary antibody). Immunoreaction was detected with 3, 3'-diaminobenzidine as a substrate.

In vitro transcription, translation and microsomal processing. Three types of SEZ-6 cDNA were subcloned into Bluescript SK(Stratagene), and mCAP mRNA capping kit (Stratagene) was used to synthesize SEZ-6 transcripts. Synthesized SEZ-6 RNAs were translated in the presence of [35S] methionine in a rabbit reticulocyte lysate (Amersham) with or without canine pancreatic microsomal membranes (Promega) according to the manufacture's instructions. The translated proteins were analyzed by SDS-PAGE and fluorography with salicylic acid.

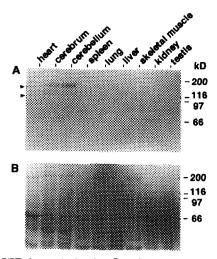
Immunoprecipitation. The proteins which were translated in the reticulocyte lysate system were diluted in NP-40 lysis buffer (10 mM Tris-HCl pH 7.4, 150 mM NaCl, 5 mM EDTA, 1% NP-40, 1% Trasylol (aprotinin), 1 mM PMSF), incubated with antiserum for 1 hr at 4 °C and then mixed with protein A sepharose (Pharmacia) for 1 hr at 4 °C. Beads were washed four times with NP-40 lysis buffer. The proteins bound to beads were eluted by boiling with sample buffer (9), subjected to SDS-PAGE and detected by fluorography.

RT-PCR. Total RNA was isolated from the cerebrum of C57BL/6J mice by the acid guanigium thiocyanate-phenol-chloroform method (11). Poly (A)+ RNA was purified from the total RNA by Dynabeads Oligo (dT)25 (Dynal A.S.). cDNA was prepared from the poly (A)+RNA using Superscript II (GIBCO BRL) and antisense primer (5'-GATGCAGGACTGCGTGTGCA-3'). cDNA was amplified by PCR using sense (5'-CCAAGCTTGGCGCGCCGCAAGCAGCAC-3') and nested antisense primer (5'-CACCTAGATGGAAACTTC-3'). The PCR conditions were 35 cycles of denaturation at 94 °C for 45 sec, annealing at 60 °C for 45 sec and extention at 72 °C for 3 min with a final extension step at 72 °C for 7 min. The PCR products were separated by electrophoresis, purified from the gel (major band at 3 kb) and cloned into Bluescript SK.

Nucleotide sequence. The nucleotide sequence was determined by Applied Biosystems 373 DNA sequencer with oligonucleotide primers and a Taq DyeDeoxy terminator cycle sequencing kit (Applied Biosystems).

RESULTS AND DISCUSSION

To analyze SEZ-6 gene product, polyclonal antibody was raised against the C-terminus of the protein which probably corresponds to the cytoplasmic domain. As shown in Fig. 1, the affinity purified antibody recognized 150 kD protein and faintly 110 kD protein in the cerebrum and the cerebellum with almost the same intensity. Previously, we demonstrated that SEZ-6 mRNA is expressed only in the brain (1). Thus, localization of SEZ-6 protein is in agreement on tissue specificity of SEZ-6 mRNA expression. We next examined localization of SEZ-6 protein in more detail using a cell culture. Cerebrum cortex cells on the 3rd day of culture which mostly consisted of neural cells (Fig. 2C) produced SEZ-6 protein (Fig. 2A, lane 2) as in the cerebrum (lane 1). However, the antibody recognized protein did not exist (Fig. 2A, lane 3) in the cells (13th day of culture) which were mainly glial cells (Fig. 2D). These findings suggested that SEZ-6 protein is produced in neural cells dominantly. In the previous study, we cloned SEZ-6 cDNA (3971 bp) with a single long ORF. The ORF predicted a protein of 977 amino acids with a



<u>Figure 1</u>. Expression of SEZ-6 protein *in vivo*. Proteins were extracted from various organs as indicated in the figure, separated on SDS-PAGE and detected by immunoblotting with affinity purified anti SEZ-6 peptide antibody (A) or stained with CBB (B). Arrowheads indicate SEZ-6 protein.

molecular weight of approximately 106 kD. There is a discrepancy between the molecular weight of the protein which recognized by the antibody and the protein calculated by the amino acid sequence. To clarify this, *in vitro* translation was carried out. The translation product from SEZ-6 cDNA showed a molecular weight of ca. 110 kD (Fig. 3A, lane 1). This translation product is certainly recognized by anti SEZ-6 peptide antibody (Fig. 3B, lane 1), although a translation product of another type of SEZ-6 cDNA (see below) which has a different C-terminus was not detected by immunoprecipetation (Fig. 3B, lane 2). These results imply the existence of post-translational modification of SEZ-6 gene product. Actually, the cDNA sequence of SEZ-6 has ten potential N-linked glycosylation sites (1). Thus, we next examined the *in vitro* processing of the translation product. Fig. 3C clearly shows a shift of molecular weight of the translation product with microsomes (compare lane 1 and lane 2), suggesting post-translational modification of SEZ-6 gene product. In the immunoblotting pattern of some samples from the cerebrum, 150 kD band was seen as a broad band around 130-180 kD, suggesting that SEZ-6 protein is processed with

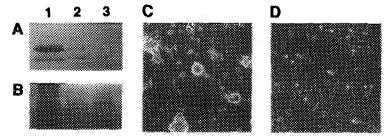


Figure 2. Expression of SEZ-6 protein in the culture. Proteins were extracted from the cerebrum (lane 1) and cultured cerebrum cortex cells for 3 days (lane 2) and 13 days (lane 3). SDS-PAGE was performed and proteins were transferred to the membrane and detected with immunoreaction (A) or stained with CBB (B). C and D show the photographs of cultured cerebrum cortex cells maintained for 3 days (C) or 13 days (D).

Figure 3. In vitro translation and post-translational modification of SEZ-6 protein (arrowheads). A. Transcripts which where synthesized from 3 types of SEZ-6 cDNA (lane 1, type 1; lane 2, type 2; lane 3, type 3) were translated in the reticulocyte lysate system. Lane 4 contains a sample translated in the absence of RNA.

- B. In vitro translation products of SEZ-6 cDNA (lane 1, type 1; lane 2, type 2) were immunoprecipitated with anti SEZ-6 peptide antibody. Same samples in Fig. 3A were used for this experiment.
- C. Type 1 SEZ-6 cDNA was transcribed in vitro and synthesized RNA was translated with (lane 2) or without (lane 1) microsomes. Lane 3 contains a sample translated with microsomes in the absence of RNA.

complicated modification in vivo (data not shown). Recently, the importance of posttranslational modification of the neural cell adhesion molecule has been indicated (reviewed in 12). The processing of SEZ-6 protein perhaps has an important role in its function.

While cloning full length of SEZ-6 cDNA in the pervious study, we found several cDNA fragments which have different sequences from originally cloned SEZ-6 cDNA. To examine the SEZ-6 expression in more detail, we attempted to clone the other types of SEZ-6 cDNA with full length using RT-PCR. Since these cDNA fragments had a common 3' sequence, we chose two antisense primers for reverse transcription and PCR. Several sense primers were chosen for PCR from the 5' region of originally cloned cDNA sequence but only one primer located at the most 3' region could amplify cDNA fragments. The major PCR products were observed at approximately 3 kb. This band was purified, subcloned and sequenced. Two types of sequences were found other than the originally cloned cDNA. We designated original cDNA as type 1 and two newly cloned cDNAs as type 2 and type 3. Type 2 cDNA was 3013 bp and type 3 cDNA was 3226 bp in length (Fig. 5). These clones were not amplified from genomic DNA because our preliminary data indicated that SEZ-6 cDNA was spread over more than 50 kb of the genomic region. Both type 2 and type 3 cDNAs have single long ORFs and the ORFs predict proteins consisting of 992 and 605 amino acids with molecular weights of 108 kD and 65 kD, respectively. Results of in vitro translation with type 2 and type 3 cDNA are in agreement with the predicted molecular weight (Fig. 3A, lanes 2 and 3; note that the translation product of type 2 cDNA is slightly larger than that of type 1). Compared with the nucleotide sequence of type 1 cDNA, type 2 cDNA has two insertional sequences at the 3' region (Fig. 4A, underlined). This leads to a longer and different amino acid sequence in the C-terminus of type 2 protein. Type 3 cDNA has the same 3' sequence as type 2 but one additional insertion (Fig. 4B, underlined). This insertional sequence contains inframe stop codon, so that the predicted type 3 protein is much shorter than type 1 and type 2. Analysis of the amino acid sequence revealed that both proteins have several N-linked glycosylation sites (Fig. 4, boxed; ten for type 2 and nine for type 3) and a stretch of hydrophobic

\sim																																					
1																																GG	CGCC	3GCC	GCAA	GCAGO	CACC
22	AT	GCG	CCA	GCA	GCCC	TGC	TGC	rcc:	TGC	CTO	CGCT	GCT	GGC	зстс	c TGG	CTC	CGG	ACTO	TCCT	AGA	GCT	CGA	TCAC	GGG	GGAA	GGTC	ATG	CAC	GGG	CAT	CAG	GGA	GACG	3GAT	GGGG	AGCTO	ACC
																																				L	
•	*****				•••••		···· ·	 -	····					~ ~	errii	-					`` \																
							• ama						. mm.			a			ames:				1001		L COUR	-				mc s 1			~c.m.				
																																				AGG AJ	
44	A	A	P	T	P E	Q	s	D	R	G	v	н	F	v	тт		P	т	L K	L	L	• н	н	₽	L	LE	E	F	L	Q	E	G	L	E	R E	E	A
280	CC	GCA	жст	GCA	CTGC	CCT	TCC	AGC	CGG	ACTO	CACC	TAC	ACAC	CTTT	ACTO	CAAG	CCCC	CTC	CCTC	CCT	CACC	MCC	AGG?	CAA	CCGC	CCCG	TC T	TAC	CAG	TCC	GAC'	TCC	AGCC	GTG	GCTG	CAGC	ACCC
87	P	Q	P	Α :	L P	•	Q	P	D	s	P	T	Н	r	T P	s	P	L	P R	L	T 1	e Q	D	ĸ	R	P V	r	T	s	P	T	P	λ	v	A A	A	P
409	ACI	CCAC	ccc	CAC	TCCA	GGG	AGAZ	ACC	стто	GAA	CCT	AGA	ATC	:AAA	cccc	CTGA	GC T	TTCT	ATCAG	ATC	STCC	TTC	стес	AGG	3CCG	AGTA	TGG	GGT	GCC	CAC	ACT	GCT	ccca	AGAG	GACA	GACCO	CAGT
																																				P	
130		~	•	п .		•			"		-	-	3			£	2	5				•	٠	J	•	5		٠	•	•	-	-	•	£		•	
																																				CTCC	
173	Ŧ	Ŧ	P	P :	s Q	λ	W	T	P	T	Q	E	G	P	G D	М	D	R	P W	V	PI	; v	М	s	K	T T	G	L	G	v	E	G	T	I	A T	s	T
667	GG'	TTC	LGGG	GAT	JACG.	AAG	AGAC	CAC	CTAC	CAC	CAT	CATI	TACC	ACT	ACTG	TCAC	CAC	AGTT	CAGCO	ACC	AGGC	CCT	GTAG	CTG	GAAT	TTCT	CAGO	CCC.	AGAC	3GGC	TC	гсто	GA1	rtcc	CCCA	GGCC	ccc
																																				λ	
	-	-	-	-	-	_	-		_	-	-	-							-																		
706				.											- m m				20 10				maa.	a			mma i					~ n m/				CCTC	
259	S	s	P	s	v	G	L	D	С	F	Y	Y	1	S	V Y	P	G	Y	3 V	E	1)	, v	E	N	1	Sj L	Q	E	G	E	T	1	т	v	E G	L	G
																																				TTTC	
302	G	P	D	P 1	L P	L	٨	N	Q	s	F	L	L	R (₫ Q	v	I	R	S P	T	Нζ	Α (Α	L	R	F Q	s	Ļ	₽	L	P	A	G	P	3 T	F	Н
1054	TT	ccac	TAC	CAA	SCC T.	ATC	TCC1	GAC	GC TG	CCA	CTT	TCCC	CCGA	CGT	CCAG	CGTA	TGG	GAT	STGAC	TGT	ACC	GTC	TCCA	ccci	AGGA	GGCA	GCGC	CCA	CTT	CA1	rrgi	rgco	ACT	GGC	TACC	AGC TO	AAG
																																				L	
	-			•		-	_	-		••		-			•		-	-													-		•	-		_	•
							a ma =								mm								a m a a														
																																				TGGC	
388	G	A	R		. т	С	L	N			Q	P	F	W	D S	Q	E	P	, c	I	,	С	G	G	V	I R	N	Α.	т	Т	G	R	I	ν :	S P	G	F
1312	CCC	GGGG	AAC	TAC	AGC A	ACA	ACC 1	CAC	CCTG	CCA	CTG	GTTC	SCTA	GAG	CCTC	CAGA	GAGC	CAG	CGGC	GCA	CTG	ACT	TTGA	AAAG	GTC	TCCC	TGGC	AGA	AGAC	GAC	GAC	AGG	CTC	ATC.	ATCC	CAAT	'GGA
431	P	G.	M	Y S	N	N	L	T	C	н	W	L	L	E I	P P	E	s	Q i	R L	Н	L	F	E	ĸ	ν	s L	A	E	D	D	D	R	L	1	I R	N	G
1441	447	TAAC	GTG	SAGO	ccc	rgo	CGGT	GT.	CGA	стс	CTA	TGAG	GTG	GAA'	TACC	TGCC	CATT	GAG	GCCT	GCTC	AGCI	YC TG	GCAG	ACAC	TTC	TTCG	TGGA	GTT	CAGI	CACT	GAC	AGC	AGT	vaca.	CAG	TGCA	ccc
																																				A	
			•			P	٠	1	J	5	•	-	٠			P			, ,	L	5 2		•	п	•		L	•	5	T	U	5	5		٠ ^	^	G
1570																																				CCCT	
1570																																				CCCT	
1570																																					
1570 517	М	λ	L	R Y	E	y	¥	ō	Q	R	н	С	Y	E	P F	٧	ĸ	Y 1	3 18	F	<u>s</u> s	s	A	P	s ·	Y P	٧	G	Ŧ	T	V	E	F	s (D :	Þ	G
1570 517 1699	M TAC	A CACC	L :	R Y	E E	A GC To	F	Q CAT	Q CAT	R CGA	H ATG	с	Y EGAC	E I	P F	V ACCC	K CCAG	Y :	ATGA	F GAC	ളി ഭ ശൂദേ	S CAG	A CCTG	P CCGA	s ·	y P GTGT	V GCAG	g KCGG(T GGAG	T SATO	V :ACA	E LGAC	P TCT	S (D GCG	P GGTG	G CTC
1570 517 1699	M TAC	A CACC	L :	R Y	E E	A GC To	F	Q CAT	Q CAT	R CGA	H ATG	c ccrc	Y EGAC	E I	P F	V ACCC	K CCAG	Y :	ATGA	F GAC	ളി ഭ ഭ ര ദ	S CAG	A CCTG	P CCGA	s ·	y P GTGT	V GCAG	g KCGG(T GGAG	T SATO	V :ACA	E LGAC	P TCT	S (D GCG	Þ	G CTC
1570 517 1699 560	M TAC Y	A CACC T	L :	R) Gagge	E E	A GCT S	F CCAT I	Q CAT	Ç TACN 1	R CGA E	H ATG	c cgrc v	Y EGAC D	E I	PF CACG	V ACCC P	K CCAG	Y I	ATGA	F GACA	S S	CAG	A CCTG C	P CCGA R	S AGCC	Y P GTGT V C	V GCAG S	G G G	T SGAG E	T SATC I	V CACA T	E AGAC D	r TCT s	S (GCA(A (GCG1	P PGTG	G CTC L
1570 517 1699 560	M TAC Y TCT	A T TCCA	L CTG	R) Gagg E (E ECGG	A GCT S AGC	F CCAT I	Q CAT I	Ç I I DOOG	R CGA E BAGG	H C C	C CGTC V GGAC	Y CGAC D CTGC	E I	P F CACG H D TGGG	V ACCC P GTGT	K CCAG Q GCAT	Y TGG. W D	ATGA E	GAC	S S	CAGO	A C C	P CCGA R GCTG	S A GCC A	Y P GTGT V C	V GCAG S SAGT	G G G	T GGAG E	T SATC I SATA	V T IGGC	E AGAC D	F TCT S	S (GCA)	GGCG1 GCG1 GV GTAC1	P CGTG V	G CTC L TTC
1570 517 1699 560	M TAC Y TCT	A T TCCA	L CTG	R) Gagg E (E ECGG	A GCT S AGC	F CCAT I	Q CAT I	Ç I I DOOG	R CGA E BAGG	H C C	C CGTC V GGAC	Y CGAC D CTGC	E I	P F CACG H D TGGG	V ACCC P GTGT	K CCAG Q GCAT	Y TGG. W D	ATGA E	GAC	S S	CAGO	A C C	P CCGA R GCTG	S A GCC A	Y P GTGT V C	V GCAG S SAGT	G G G	T GGAG E	T SATC I SATA	V T IGGC	E AGAC D	F TCT S	S (GCA)	GGCG1 GCG1 GV GTAC1	P PGTG	G CTC L TTC
1570 517 1699 560 1828 603	M TAC Y TCT S	A T TCCA	L :	GAGG	EAGGGCCGG	A S AGC	F CTTA	CAT I TGG G	CAT I SCCG R	R CGA E AGG G	H C GGCAG	C CGTC V GGAC D	Y EGAC D TGC C	E (CTCC)	P F CACG. H D TGGGG	ACCC P STGT V	K CCAG Q GCAT H	Y TOG	AATGA LE BAGGA	GACA GGACA D	S S GAGG	CAG A GCA	A C C TCAT	P CCGA R GCTG L	S A GAC D	Y P GTGT(V C ATCC(I R	GCAG S SAGT V	G G G GCTG L	T EGAG E ECGC R	T SATC I SATA I	V T GGC	E AGAC D TCT	r TCT S GGG	S (GCA(A (GAT(GGCGT	P CGGTG V CGACC T	G CTC L TTC F
1570 517 1699 560 1828 603	M TAC Y TCT S	A CACC T FCCA P	L CTG	GAGG	F ECAGGG	A S S AGC P	ECAT I ETTA Y CAGC	CAT I TGG	Q ICAT I ECCG R	R CGA E AGG G	H C GGGGG	C CGTC V GGAC D	Y CGAC D CTGC C	E CTCCLL CATCLL IN TORS	P F CACG. H D TGGG W G	ACCC P STGT V	K CCAG Q GCAT H	Y TOG. W [AATGA AATGA I E SAGGA E E	GACA GGAC	S S GAGG E F AAGG K R	CAGG	CCAT	P CCGA R GCTG L	S A GAC D	Y P GTGT(V C ATCC) I R GTCAG	S S SAGT V	G G G G CCAC	T SGAG E SCGG R	T I ATA I	T GGC G	E GAC	TCT S GGG G	S (GCA A (GAT D (GGGA	GGCGC GGCGC GGCGC GGCGCC GGCGCCCCCCCCCC	P CGGTG V CGACC T	G CTC L TTC F
1570 517 1699 560 1828 603	M TAC Y TCT S	A CACC T FCCA P	L CTG	GAGG	F ECAGGG	A S S AGC P	ECAT I ETTA Y CAGC	CAT I TGG	Q ICAT I ECCG R	R CGA E AGG G	H C GGGGG	C CGTC V GGAC D	Y CGAC D CTGC C	E CTCCLL CATCLL IN TORS	P F CACG. H D TGGG W G	ACCC P STGT V	K CCAG Q GCAT H	Y TOG. W [AATGA AATGA I E SAGGA E E	GACA GGAC	S S GAGG E F AAGG K R	CAGG	CCAT	P CCGA R GCTG L	S A GAC D	Y P GTGT(V C ATCC) I R GTCAG	S S SAGT V	G G G G CCAC	T SGAG E SCGG R	T I ATA I	T GGC G	E GAC	TCT S GGG G	S (GCA A (GAT D (GGGA	GGCGC GGCGC GGCGC GGCGCC GGCGCCCCCCCCCC	P CGGTG V CGACC T	G CTC L TTC F
1570 517 1699 560 1828 603	M TAC Y TCT S	A CACC T FCCA P	L CTG	GAGG	F ECAGGG	A S S AGC P	ECAT I ETTA Y CAGC	CAT I TGG	Q ICAT I ECCG R	R CGA E AGG G	H C GGGGG	C CGTC V GGAC D	Y CGAC D CTGC C	E CTCCLL CATCLL IN TORS	P F CACG. H D TGGG W G	ACCC P STGT V	K CCAG Q GCAT H	Y TOG. W [AATGA AATGA I E SAGGA E E	GACA GGAC	S S GAGG E F AAGG K R	CAGG	CCAT	P CCGA R GCTG L	S A GAC D	Y P GTGT(V C ATCC) I R GTCAG	S S SAGT V	G G G G CCAC	T SGAG E SCGG R	T I ATA I	T GGC G	E GAC	TCT S GGG G	S (GCA) GATO D (GGGA)	GGCGC GGCGC GGCGC GGCGCC GGCGCCCCCCCCCC	P CGGTG V CGACC T	G CTC L TTC F
1570 517 1699 560 1828 603 1957 646	TAC Y TCT S TAC Y	A T FCCA P CGAT	L : CTG L : AAC B : GGGG	GAGG	CAGGG CCGG CCGG CCGG CCGG CCGG CCGG CCG	A GCTV S A GCC P T CAA	F CATTA Y CAGC	Q CAT I S G CCCG	Ç I I SCCG R SGGT V	R E AGG G	H C GGGAG Q GGGGG	C V GGAC D CCAA	Y EGAC D ETGC C C ATAC Y	E I	P F CACG H D TGGGG H G GGGCC G P	ACCC P STGT V CCCG R	K CCAG Q GCAT H TGGC G	Y TOGEN W CACT	AATGA KAGGA E E TCAA	GACA GGACA D GCTC L	S S GAGG E F AAGG K F TTTA	CAGO	A C C TCAT M CCAT	P CCGA R GCTG L GGCC	S A GAC D GAT	Y P GTGT(V C ATCCO I R GTCAU	V GCAG S GAGT V CCAT	G G G CCAC	T GGAG E GCGC R GTTC	T I ATA I CAG	V T GGGC G TCA	E AGAC D TCT S AGAC D	F TCT S GGG G	S (GCAC A (GATC D (GGGA GGGA G 1	GGCGTACT	P CGGTG V CGACC T	G CTC L TTC F CTG L
1570 517 1699 560 1828 603 1957 646	M TAC Y TCT S TAC Y	A CACCA T FCCA P CGAT D	L CTG	R) GGAGG	CAGGG CCGG CCG CCGGG CCGGG CCGGG CCGGG CCGGG CCGGG CCGGG CCGGG CCGGG CCGGG CCG	A GCT S A GCC P T C A G T T T T T T T T T T T T T T T T T T	F CCAT I CTTA Y CAGC A	Q CAT I I ITGG G CCCG	Q TCAT I GGCGG R V ACTT	R CCGA E EAGG G CCCT L	H C GGCAG G GGGGG	C CGTC V GGAC D CCAA Q	Y EGAC D ETGC C ATAC Y	E I	P F CACG. H D TGGGG V G GGGCC G P	ACCC P STGT V CCCG R	K CCAG Q GCAT H TGGC G	Y 1 STGG: W STGG V 1 SCACT H 1	AATGA AAGGA E E TCAA K	GACA GGACA D GCTC L	S S GAGG	GCAS GCAS I CCTX S	A CCAA	P CCGA R GCTG L GGCC A	S GAC D TGG	Y P GTGT(V C ATCCC I R GTCA(V T	S SAGT V CCAT	G G G G C C C Q	T SGAG E SCGC R STTC	T I I CAG	V T GGGG G TCA	E D D CTCT S GAC D	F S G G CCT P (S GCAG	GGCAC	P CGACC T GGCG A	G CTC L TTC F CTG L GTC
1570 517 1699 560 1828 603 1957 646	M TAC Y TCT S TAC Y	A CACCA T FCCA P CGAT D	L CTG	R) GGAGG	CAGGG CCGG CCG CCGGG CCGGG CCGGG CCGGG CCGGG CCGGG CCGGG CCGGG CCGGG CCGGG CCG	A GCT S A GCC P T C A G T T T T T T T T T T T T T T T T T T	F CCAT I CTTA Y CAGC A	Q CAT I I ITGG G CCCG	Q TCAT I GGCGG R V ACTT	R CCGA E EAGG G CCCT L	H C GGCAG G GGGGG	C CGTC V GGAC D CCAA Q	Y EGAC D ETGC C ATAC Y	E I	P F CACG. H D TGGGG V G GGGCC G P	ACCC P STGT V CCCG R	K CCAG Q GCAT H TGGC G	Y 1 STGG: W STGG V 1 SCACT H 1	AATGA AAGGA E E TCAA K	GACA GGACA D GCTC L	S S GAGG	GCAS GCAS I CCTX S	A CCAA	P CCGA R GCTG L GGCC A	S GAC D TGG	Y P GTGT(V C ATCCC I R GTCA(V T	S SAGT V CCAT	G G G G C C C Q	T SGAG E SCGC R STTC	T I I CAG	V T GGGG G TCA	E D D CTCT S GAC D	F S G G CCT P (S GCAG	GGCAC	P CGGTG V CGACC T	G CTC L TTC F CTG L GTC
1570 517 1699 560 1828 603 1957 646	TACT S TACT Y GGT	A CACCAT TOTAL P CACAT D TACCA Y	L :	R)	EAGGGAT	S AGCO	F CCAT I CTTA Y CAGC A	Q CAT I I ATGG G CCCG R	Q CCAT I SCCG R V V ACTT	R CCGA E EAGG G CCCT L CCTT	H AATG C GGCAG Q GGGGG G TTGAG	C C GTC V GGGAC D C C AA Q Q GGGTT V	Y CGAC D CTGC C ATAC Y	E COCTOCO L COCTOCO L	P F CACG. H D TGGGG W G GGGGC P	V ACCC P STGT V CCCG R ACAC	K CCCAG Q GCAT H TGGC G ATGT	Y 1 STEER OF THE TERM OF THE	AATGA GAGGA E E TCAA K	GACA GGACA D GCTC L	S S S S S S S S S S S S S S S S S S S	S CAGG	A CCAAN	P CCCGA R CCCGCC L CCCGCCC G	S GACO	Y P GTGTC V C ATCCC ATCCC I R GTCAC V T AAGAI	V GGCAG S GAGT V CCAT I ACCC	G G G G G G G C C C C C C C C C C C C C	T GGGAG E GGGGG R F ACAG	T SATA I CAGA CAG COT P	V T GGGC G TCA S	E AGAC D STOT	F GGGG G CCTV	S (GCACCACCACCACCACCACCACCACCACCACCCACCCAC	GGCAC	P CGGTG V CGGCG A	G CTC L TTC F CTG L GTC V
1570 517 1699 560 11828 603 11957 646 689	M TACT S TACT Y GGT G ACCO	A CACCAT T TOCAM P D TOTAC Y CTAT	L :	R)	EAGGGCCGGGCCGGGATGGGATGGGGATGGGGATGGGATG	A GCCT S A GCCC P T C A C T T T T T T T T T T T T T T T T	F CCAT I CTTA Y CCAGC A TCAT I	Q CAT I TGG G CCG R	Q TCAT I GGCCG R V ACTT F	R PCGA E EAGG G CCT L PCTT F	H ATGG C C GGCAG Q TTGAG E	C CGTC V GGGAC D CCAAA Q GGGTT V ATCC	Y CGAC D CTGC C C ATAC Y CCCC P	E ATC: ATC: CGC: R E	P F CACG	ACCC P STGT V CCCG R ACAC	K CCAG Q GCAT H TGGC G ATGT C	Y TOGO	AAATGAAATGAAATGAAATGAAATGAAATGAAAGGAAGG	GACACACACACACACACACACACACACACACACACACAC	S S S S S S S S S S S S S S S S S S S	S CAGCA' A GCA' I CCCTC S TCCCC P	A CCAAN	P CCGA R R GCTG L CGGCC A CCGGC	S SAGGERAL SAGGERA SA	Y P GTGTCV C ATCCC I R FTCAC T AAGAA K N	V GCAG S GAGG V CCAT I ACCC P	G GAGA	T GGAG E GGGGG R GGTG ACAG	T SATA I CAG CCT P	V T GGGC G TCA S	E AGACO D STOCK S GACO D GCTG L	F TTCT S G G CCTN GTG	S (GCACCACCACCACCACCACCACCACCACCACCACCACCAC	D D D D D D D D D D D D D D D D D D D	P GGTG V GGGGG A GGTG V	G CTC L TTC F CTG L GTC
1570 517 1699 560 11828 603 11957 646 689	M TACT S TACT Y GGT G ACCO	A CACCAT T TOCAM P D TOTAC Y CTAT	L :	R)	EAGGGCCGGGCCGGGATGGGATGGGGATGGGGATGGGATG	A GCCT S A GCCC P T C A C T T T T T T T T T T T T T T T T	F CCAT I CTTA Y CCAGC A I I I	Q CAT I TGG G CCG R	Q TCAT I GGCCG R V ACTT F	R PCGA E FCCT L PCTT F	H ATGG C C GGCAG Q TTGAG E	C CGTC V GGGAC D CCAAA Q GGGTT V ATCC	Y CGAC D CTGC C C ATAC Y CCCC P	E ATC: ATC: CGC: R E	P F CACG	ACCC P STGT V CCCG R ACAC	K CCAG Q GCAT H TGGC G ATGT C	Y TOGO	AAATGAAATGAAATGAAATGAAATGAAATGAAAGGAAGG	GACACACACACACACACACACACACACACACACACACAC	S S S S S S S S S S S S S S S S S S S	S CAGCA' A GCA' I CCCTC S TCCCC P	A CCAAN	P CCGA R R GCTG L CGGCC A CCGGC	S SAGGERAL SAGGERA SA	Y P GTGTCV C ATCCC I R FTCAC T AAGAA K N	V GCAG S GAGG V CCAT I ACCC P	G GAGA	T GGAG E GGGGG R GGTG ACAG	T SATA I CAG CCT P	V T GGGC G TCA S	E AGACO D STOCK S GACO D GCTG L	F CGGG G CCTN GTG V	S (GCACCACCACCACCACCACCACCACCACCACCACCACCAC	D D D D D D D D D D D D D D D D D D D	P CGGTG V CGGCG A	G CTC L TTC F CTG L GTC
1570 517 1699 560 11828 603 11957 646 689	M TACT S TACT Y GGT G ACCO	A CACCAT T TOCAM P D TOTAC Y CTAT	L :	R)	EAGGGCCGGGCCGGGATGGGATGGGGATGGGGATGGGATG	A GCCT S A GCCC P T C A C T T T T T T T T T T T T T T T T	F CCAT I CTTA Y CCAGC A TCAT I	Q CAT I TGG G CCG R	Q COAT I GGGCCG R V ACTT F	R PCGA E FCCT L PCTT F	H ATGG C C GGCAG Q TTGAG E	C CGTC V GGGAC D CCAAA Q GGGTT V ATCC	Y CGAC D CTGC C C ATAC Y CCCC P	E ATC: ATC: CGC: R E	P F CACG	ACCC P STGT V CCCG R ACAC	K CCAG Q GCAT H TGGC G ATGT C	Y TOGO	AAATGAAATGAAATGAAATGAAATGAAATGAAAGGAAGG	GACACACACACACACACACACACACACACACACACACAC	S S S S S S S S S S S S S S S S S S S	S CAGCA' A GCA' I CCCTC S TCCCC P	A CCAAN	P CCGA R R GCTG L GGCC A CCGGC	S SAGGERAL SAGGERA SA	Y P GTGTCV C ATCCC I R FTCAC T AAGAA K N	V GCAG S GAGG V CCAT I ACCC P	G GAGA	T GGAG E GGGGG R GGTG ACAG	T SATA I CAG CCT P	V T GGGC G TCA S	E AGACO D STOCK S GACO D GCTG L	F CGGG G CCTN GTG V	S (GCACCACCACCACCACCACCACCACCACCACCACCACCAC	D D D D D D D D D D D D D D D D D D D	P GGTG V GGGGG A GGTG V	G CTC L TTC F CTG L GTC
1570 517 1699 560 1828 603 1957 646 689 2215 732	M TAC Y TCT S TAC Y GGT G ACC T	A CACCAT TOCAM P CGAT D TTAC Y CTAT	L : CTG	R)	F E E E E E E E E E E E E E E E E E E E	A GCT S A GCC P T CAM T T T T T T T T T T T T T T T T T T T	F CCAT I CTTA Y CAGC A ICAT I GTTA	Q CAT I I I I I I I I I I I I I I I I I I I	Q PCAT I I GGCCG R V V ACTT F	R PCGA E SAGG G CCTT F CGTT V	H AATGC C GCAG Q GGGGG G TTGAG E	C CGTC V V GGGAC D CCAA V V ATCC S	Y CGAC D CTGC C ATAC Y CCCC P	E ATC: TCAC S C CCCC R ATTC I I I I I I I I I I I I I	P F CACG. H D TGGGG G GGGGC G AACAA N N	V ACCC P STGT V CCCG R ACAC. T	K CCAG Q GCAT H TGGC G ATGT C CCAG	Y 1 TOGGE TOGG TOGG	AATGA AATGA FE FTCAA C R AGGCT C L	GGACA GGACA GGACA D GCTC L ACCCC P AAGCC S	S S S S S S S S S S S S S S S S S S S	S CAGGA I COTO S TOCO P GTG/	A CCAAM A GGAA D	P CCGA R R GCTG L GGCC A CCGGC G	S SAGCCOA A SAGC	Y P GTGTT V C ATCCC I R GTCAC V T AAGGATC K N TCATC	S SAGT V CCAT I ACCC P Q CCAA Q	G GAGA	T GGAG E E GGCGC R GGTTC F ACAG Q V	T I FATA I CAG Q CCCT P	V T GGGG G TCA S TCA S TCA S TCA S TCA S	E AGACO D TOTAL S GACO D GCTG L TTGC C C	FOR THE POST OF TH	S (GCACCACCACCACCACCACCACCACCACCACCACCACCAC	GGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC	P GGTG V GGGGG A GGTG V	G CTC L CTG L CTG V CTG V
1570 517 1699 560 1828 603 1957 646 689 732	M TAC Y TCT S TAC Y GGT G ACC T	A CACCAT THECAM P CEGAT D THAC Y CHACAC	L : CTGC CTGC CTGC CTGC CTGC CTGC CTGC CTG	R)	F ELAGGGC	A GCT S A GCC P T C A C T C T G G G T C A C T	F CCAT I CTTA Y CAGC A TCAT I GTTA Y	Q CAT I ATGG G CCG R CCA H	Q PCAT I I GGCCG R V ACTT F V AGGT V	R PCGA E SAGG G PCCT L PCTT F PGGT V	H AATG C GGCAG Q TGAG E GGGG G	C CGTC V GGGAC D CCAA Q CGGTT V ATCC S	Y COCCO	E	P F CACG. CACG. H D TGGGG H G GGGGC H G GGGGC H G H CTCAC	V ACCC P STGT V CCCGG R ACAC T	K CCAG Q GCAT H TTGGC G ATGT C CCAG	Y TEST TATE	AATGA AATGA AATGA EAGGGA C E TTCAA C L AAACCT L	GGACA GGACA D GCTC L ACCCC P AAGCC S TGAC	S S S S S S S S S S S S S S S S S S S	S CAGO A A GCA I I CCTC S TCCCC P GTG/E	A CCAAM M CCAAM D	R GCTG L GGCC A CCGGC G CCTG L	S ACGG	Y P GTGTT V C ATCCC I R GTCAO V T AAGAA K N ICATC	V GCAT I ACCC P GCCA Q GTGC	G GCGGGG G GCTC L CCCAC Q ATCJ S GAGJ	T GGAG E E GGGGG R GTTC F ACAG V NCTC	T SATO I SCAG Q SCCT P SACA T	V EACH T GGC G TCA S GAGE TCT S	E AGAC D STOTE S AGAC D TTGC C TTGC C CAT	F GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	S (GERCE CACCO	D D GGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC	P GGGTG V GGGTG V GGGATC D AGGCC	G CTC L TTC F CTG L GTC V AGT
1570 517 1699 560 1828 603 1957 646 689 732	M TAC Y TCT S TAC Y GGT G ACC T	A CACCAT THECAM P CEGAT D THAC Y CHACAC	L : CTGC CTGC CTGC CTGC CTGC CTGC CTGC CTG	R)	F ELAGGGC	A GCT S A GCC P T C A C T C T G G G T C A C T	F CCAT I CTTA Y CAGC A TCAT I GTTA Y	Q CAT I ATGG G CCG R CCA H	Q PCAT I I GGCCG R V ACTT F V AGGT V	R PCGA E SAGG G PCCT L PCTT F PGGT V	H AATG C GGCAG Q TGAG E GGGG G	C CGTC V GGGAC D CCAA Q CGGTT V ATCC S	Y COCCO	E	P F CACG. CACG. H D TGGGG H G GGGGC H G GGGGC H G H CTCAC	V ACCC P STGT V CCCGG R ACAC T	K CCAG Q GCAT H TTGGC G ATGT C CCAG	Y TATO	AATGA AATGA AATGA EAGGGA C E TTCAA C L AAACCT L	GGACA GGACA D GCTC L ACCCC P AAGCC S TGAC	S S S S S S S S S S S S S S S S S S S	S CAGO A A GCA I I CCTC S TCCCC P GTG/E	A CCAAM M CCAAM D	R GCTG L GGCC A CCGGC G CCTG L	S ACGG	Y P GTGTT V C ATCCC I R GTCAO V T AAGAA K N ICATC	V GCAT I ACCC P GCCA Q GTGC	G GCGGGG G GCTC L CCCAC Q ATCJ S GAGJ	T GGAG E E GGGGG R GTTC F ACAG V NCTC	T SATO I SCAG Q SCCT P SACA T	V EACH T GGC G TCA S GAGE TCT S	E AGAC D STOTE S AGAC D TTGC C TTGC C CAT	F GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	S (GERCE CACCOME CACCO	D D GGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC	P GGGTG V GGGTG V GGGTG V	G CTC L TTC F CTG L GTC V AGT
1570 517 1699 560 1828 603 1957 646 689 2215 732 2344 775	M TACT S TACT Y GGT G ACC T GAG	A COACCO T TOTAL P COACCO Y COTAT Y COGCAC H	L : AAC GGGG G : CAG Q CAG Q AGC S 1	R)	ELAGGE CCGG E EGGAT F CGGGT E EGGCT E EGGCT	A GCT S S A GCC P T C A C T G G G T C A C I	F CCAT I CTTA Y CAGC A TCAT I STTA Y	Q CAG	Q CONT I I I I I I I I I I I I I I I I I I I	R CGA E FOCT L CGT V CAA	H ATG C GCAG Q GGGGG G TGAG G GTT: F	C CGTC V GGAC D CCAA Q CGGTT V ATCC S P	Y CGAC D CTGC C ATAC Y CAGT S CGTG	E COCTOC L ATC: I TCAC S C CCGC R E GGAC G	P F CACG. H D TGGGG W G GGGGCC N H D TCGGG W G GGGGCC N H D TCAC N H GCAAC N T	V ACCC P STGT V CCCGG R ACAC C CTGTV	K CCAG Q GCAT H TGGC G ATGT C CCCAG Q GCAA	Y 1 TATO THE PLANT OF THE PLA	AAATGAAAATGAAAATGAAAATGAAAATGAAAATGAAAATGAAAATGAAAAATGAAAAAA	GACA	S S S S S S S S S S S S S S S S S S S	SCAGGA GCA I CCTC S TCCC P GTGA E	A CCATC M CCCAAC N D CCCAAC V	P CCGA R GCTG L CGGCC A CCGGCC G L CCTG L	S CAGGO A CAGG	Y P GTGT V C ATCCC I R GTCAC V T AAAGA I CATC S C	V GCAT I ACCC P GGCAA Q ACCCAA	G G G G G G G G G G G G G G G G G G G	T GGGAG E GGGGG R ACAG V CTC L	T SATA I CAG COT P ACA T ACC T	V GGGC G TCA S GGGGE TCT S TGC	E AGAC D CTG L CTGC C CATCH	F GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	S GOGAGGACGACGACGACGACGACGACGACGACGACGACGACG	D D D D D D D D D D D D D D D D D D D	P GGTG V GGACC T GGGTG V GGATG C GGATG D AGGCC G	G CTC L TTC F CTG CTG V AGT S
1570 517 1699 560 1828 603 1957 646 689 2215 732 2344 775	TACTY TOTAL S TACCY GGT G ACCC T GAGGE	A CACCAT TOTAL P CACACAT TOTAL	L :	R)	E E E E E E E E E E E E E E E E E E E	A GCC S S A GCC P T C A C C T C C C C	F CCAT I Y CTAGC A TCAT I STTA Y	Q CAT I I I I I I I I I I I I I I I I I I I	Q CCCC P GTG	R CGA E SAGG G CCT L CTT F CGT V CAA	H ATG C GGCAG Q GGGGG G TGAG G GTT F	C CGTC V GGAC D CCAA Q C CGTC V A TCC C P GGAA	Y CCGAC CCCC Y CCCC P CATAC CCCC CCCC CCCC CCCCC CCCCC CCCCC	E	P F CACG. H D TCGGG W G GGGGCC M AACAA N H CTCA M GCAAA A T	V ACCC P STGT V CCCCG R ACAC. T CCTGTC V	K CCAG Q GCAT H TGGC G ATGT C CCCAG Q GCAA	Y 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	AAATGAAATGAAAGGAAGGAAGGAAGGAAGGAAGGAAGG	GGACA GGACA GGACA D GCTC L ACCCC P AAGCC S TGAC	S S S S S S S S S S S S S S S S S S S	S CAGGA A GCA A GC	A CCATTON M M AGGAM D V TTGTT	P CCGA R R GCTG L CCGGC G CCTG L L CGCCTG L	S SAGGERAL COCCES	Y P GTGTT V C ATCCO ATCCO T AAAGAA K N TCATC S C GGGAC G S	V GCAT I CCAT I ACCC P GCCA Q GTGC A	G GAAG	T GGAG E GGGG R GGGGG R ACAG V CTC L	T I I I I I I I I I I I I I I I I I I I	V CACA T GGGC G TCA S GAGE TCT S TGC C	E AGAC D TCT S GAC D CTG C CAT H	F TTCT S GGGG G CCT P GTG GTG CAT H I GAT GGAT GGAT GGGAT GGAT GGGAT GGG	S GOGAGGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	D D D D D D D D D D D D D D D D D D D	P P GGTG V GGACC T GGGGGG A GGATC C CATCC	G CTC L TTC F CTG CTG V GTG V AGT S CAC
1570 517 1699 560 1828 603 1957 646 689 2215 732 2344 775	TACTY TOTAL S TACCY GGT G ACCC T GAGGE	A CACCAT TOTAL P TOTAL	L :	R)	E E E E E E E E E E E E E E E E E E E	A GCC S S A GCC P T C A C C T C C C C	F CCAT I Y CTAGC A TCAT I STTA Y	Q CAT I I I I I I I I I I I I I I I I I I I	Q CCCC P GTG	R CGA E SAGG G CCT L CTT F CGT V CAA	H ATG C GGCAG Q GGGGG G TGAG G GTT F	C CGTC V GGAC D CCAA Q C CGTC V A TCC C P GGAA	Y CCGAC CCCC Y CCCC P CATAC CCCC CCCC CCCC CCCCC CCCCC CCCCC	E	P F CACG. H D TCGGG W G GGGGCC M AACAA N H CTCA M GCAAA A T	V ACCC P STGT V CCCCG R ACAC. T CCTGTC V	K CCAG Q GCAT H TGGC G ATGT C CCCAG Q GCAA	Y 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	AAATGAAATGAAAGGAAGGAAGGAAGGAAGGAAGGAAGG	GGACA GGACA GGACA D GCTC L ACCCC P AAGCC S TGAC	S S S S S S S S S S S S S S S S S S S	S CAGGA A GCA A GC	A CCATTON M M AGGAM D V TTGTT	P CCGA R R GCTG L CCGGC G CCTG L L CGCCTG L	S SAGGERAL COCCES	Y P GTGTT V C ATCCO ATCCO T AAAGAA K N TCATC S C GGGAC G S	V GCAT I CCAT I ACCC P GCCA Q GTGC A	G GAAG	T GGAG E GGGG R GGGGG R ACAG V CTC L	T I I I I I I I I I I I I I I I I I I I	V CACA T GGGC G TCA S GAGE TCT S TGC C	E AGAC D TCT S GAC D CTG C CAT H	F TTCT S GGGG G CCT P GTG GTG CAT H I GAT GGAT GGAT GGGAT GGAT GGGAT GGG	S GOGAGGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	D D D D D D D D D D D D D D D D D D D	P GGTG V GGACC T GGGTG V GGATG C GGATG D AGGCC G	G CTC L TTC F CTG CTG V GTG V AGT S CAC
1570 517 1699 560 1828 603 1957 646 689 2215 775 2473 818	M TAC Y TC1 S TAC Y GG1 G ACC T GAG E	A CACCAT PTOCAMP DTAGGAT Y CTAT Y GCAC H CAAGG K	L :	GAACCAR F	FERRAGE FRACE	A GCT S S A GCC P P T C A Y C T G G G C A A G C C A A G C C A A G C C A A G C C A A G C C A A G C C A A G C C C A A G C C C A A G C C C A A G C C C A A G C C C A A G C C C C	F CCAT I CTTA Y CTAGC A TCAT I STTA Y CTATC CCCCC P	Q CAT I I I I I I I I I I I I I I I I I I I	Q CAT I GGGGT V ACTT F V GGGGT V GGGGT C C GGGGT C	R PCGA E FAGG G CCT L PCTT F CGA K TCT L	H ATGG C GGCAG Q GGGGG G G TGAG E GGGG G G TTGAG L	C CGTC V GGAC D C CAAA Q C C CAAA Q C C C C C C C C C C	Y CGAC D CTGC C Y CCCC P CAGT S CGTG V I CGAA Q I	E ! SCOTECO E SE S	P F CACG CACG H D TGGGG G G G G G G G G G G G	V ACCC P STGT V CCCCG R ACAC. T C CCTGTG C C	K CCCAG Q GCAT H TTGGC G ATGT C CCCAG Q GCAA	Y 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	AATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT	GGACO D GGCTC L ACCCC P AAAGCC S TGAC	S S S S S S S S S S S S S S S S S S S	S CAGGA A GCA I CCCTC S TCCCC P GTGI F F AGAA H	A CCAT M M CCAAN N D CCAAN V V TGG:	P CCGA R R GCTG L GGCCG G CCTG L L CGCTG L	S S S S S S S S S S S S S S S S S S S	Y P GTGTT V C ATCCCI I R GTCATC V T AAGAA K N ICATC GGGAC G S AGCCC G P	V GCAT I ACCC P GCCA Q ACTGA E	G GAGAGA K	T GGAG E GGGGG R ACAG Q AGTG V TCTC L	T EATA I ECCT P ECTT ECTT ECTT ECTT ECTT ECTT EC	V GACA T GGGC G TCA S TGC T C C C C C C C C C C C C C C C C C	E AGAC D CTGT CTGC C CATCH CCAC	F TTCT S TGGG G TGG G TGGG G TGGG G TGGG G TGG G	S CACCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	CCACC	P P GGTG V P GGGCG A : GGGTG V P P P P P P P P P P P P P P P P P P	G CTC L TTC F CTG CTG V GTG V AGT S
1570 517 1699 560 1828 603 1957 646 689 22215 732 2344 775 818	M TAC Y TC1 S TAC Y GG1 G ACC T GAG E	A CACCAT TOTAL TARGET AND A CACACAT	L :	R)	FERRAGE FRACE	A GCT S S A GCC S S A GCC S T C A G C T C A G C T C A G C A G C A G C A G C A G C A G C A G C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C C C	F CCAT I I CTTA Y CAGC A I CAT I I CTTA Y CAGC C CCC P	Q CAG R CCA H CCA Q CAG S CAA K GCT	Q CAT I GGCCG R V ACTT F V CCCC P GGGG C C GAA	R PCGA E FAGG CCT L CTT F CGAA K TCTC L	H AATG C GGCAG Q TGAG E GGTT F CCTTG L	C CGTC V GGGAC D CCAA Q GGGTT V ATCC S FCCC P	Y CCGAC C C ATAC Y CCGC P CAGT S CGTG Q 1	E ! SECTOR SECTION SEC	P F CACGO H D TGGGG V G GGGGC V G AACA N N N N N N N N N N N N N N N N N N N	V ACCC P STGT V CCCCG R ACAC. T C CCTGTG C C	K CCAG Q GCAT H TGGC G ATGT C CCAG Q GCAA Q CCAT H	Y P P P P P P P P P P P P P P P P P P P	AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	GGACO B GGACO B GGACO B AACCC P AAAGCC S TGAC CCCCC A CCCCCC	S S S S S S S S S S S S S S S S S S S	SCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA	A CCAAN N CCAAN D V TGGG G	R R GCTG L GGCC A CCCA CCCA	S SAGGE A SAGG	Y P GTGTT V C ATCCCI R GTCAI V T AAGAA K N TCATC GGGAC G S AGCCC GGGAC G P	V GCAT I CCAT I ACCC P CTGAC	G GAAGA	T GGAG E GGGGG R ACAG V CTC L GGGGR	T SATA I SCAGE T T SACA T CTTTL	V CACA T GGGC G TCA S GGAG E TCT C C CAC H	E AGAC D STOCK S GAC D SCTGC C CAT H CCAL	F TTCT S TGGG G TCCT P GTGG GTG GTG GTG GTG GTG GTG GTG GTG	S (GACCACACACACACACACACACACACACACACACACACA	CCACC	P P GGTG V P GGGGG A : GGGGGG A GGGGGGGGGGGGGGGGGG	G CTC L TTC F CTG L GTC V GTG V AGT S CAC H
1570 517 1699 560 1828 603 1957 646 689 22215 732 2344 775 818	M TAC Y TC1 S TAC Y GG1 G ACC T GAG E	A CACCAT TOTAL TARGET AND A CACACAT TOTAL TARGET AND A CACACAT A C	L :	R)	FERRAGE FRACE	A GCT S S A GCC S S A GCC S T C A G C T C A G C T C A G C A G C A G C A G C A G C A G C A G C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C C C	F CCAT I CTTA Y CAGC A TCAT I GTTA Y FATC S CCCC P	Q CAG R CCA H CCA Q CAG S CAA K GCT	Q CAT I GGCCG R V ACTT F V CCCC P GGGG C C GAA	R PCGA E FAGG CCT L CTT F CGAA K TCTC L	H AATG C GGCAG Q TGAG E GGTT F CCTTG L	C CGTC V GGGAC D CCAA Q GGGTT V ATCC S FCCC P	Y CCGAC C C ATAC Y CCGC P CAGT S CGTG Q 1	E ! SECTOR SECTION SEC	P F CACGO H D TGGGG V G GGGGC V G AACA N N N N N N N N N N N N N N N N N N N	V ACCC P STGT V CCCCG R ACAC. T C CCTGTG C C	K CCAG Q GCAT H TGGC G ATGT C CCAG Q GCAA Q CCAT H	Y P P P P P P P P P P P P P P P P P P P	AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	GGACO B GGACO B GGACO B AACCC P AAAGCC S TGAC CCCCC A CCCCCC	S S S S S S S S S S S S S S S S S S S	SCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA	A CCAAN N CCAAN D V TGGG G	R R GCTG L GGCC A CCCA CCCA	S SAGGE A SAGG	Y P GTGTT V C ATCCCI R GTCAI V T AAGAA K N TCATC GGGAC G S AGCCC GGGAC G P	V GCAT I CCAT I ACCC P CTGAC	G GAAGA	T GGAG E GGGGG R ACAG V CTC L GGGGR	T SATA I SCAGE T T SACA T CTTTL	V CACA T GGGC G TCA S GGAG E TCT C C CAC H	E AGAC D STOCK S GAC D SCTGC C CAT H CCAL	F TTCT S TGGG G TCCT P GTGG GTG GTG GTG GTG GTG GTG GTG GTG	S (GACCACCACCACCACCACCACCACCACCACCACCACCACC	CCACC	P P GGTG V P GGGCG A : GGGTG V P P P P P P P P P P P P P P P P P P	G CTC L TTC F CTG L GTC V GTG V AGT S CAC H
1570 517 1699 560 1828 603 1957 646 689 22215 732 2344 775 818	M TAC Y TC1 S TAC Y GG1 G ACC T GAG E	A CACCAT TOTAL TARGET A CACCACAT TOTAL TARGET A CACCACACACACACACACACACACACACACACACAC	L :	R)	FERRAGE FRACE	A GCT S S A GCC S S A GCC S T C A G C T C A G C T C A G C A G C A G C A G C A G C A G C A G C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C A G C C C C	F CCAT I CTTA Y CAGC A TCAT I GTTA Y FATC S CCCC P	Q CAG R CCA H CCA Q CAG S CAA K GCT	Q CAT I GGCCG R V ACTT F V CCCC P GGGG C C GAA	R PCGA E FAGG CCT L CTT F CGAA K TCTC L	H AATG C GGCAG Q TGAG E GGTT F CCTTG L	C CGTC V GGGAC D CCAA Q GGGTT V ATCC S FCCC P	Y CCGAC C C ATAC Y CCGC P CAGT S CGTG Q 1	E ! SECTOR SECTION SEC	P F CACGO H D TGGGG V G GGGGC V G AACA N N N N N N N N N N N N N N N N N N N	V ACCC P STGT V CCCCG R ACAC. T C CCTGTG C C	K CCAG Q GCAT H TGGC G ATGT C CCAG Q GCAA Q CCAT H	Y P P P P P P P P P P P P P P P P P P P	AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	GGACO D GGTC L ACCCC P AAAGCC S TGAC	S S S S S S S S S S S S S S S S S S S	SCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA	A CCAAN N CCAAN D V TGGG G	R R GCTG L GGCC A CCCA CCCA	S SAGGE A SAGG	Y P GTGTT V C ATCCCI R GTCAI V T AAGAA K N TCATC GGGAC G S AGCCC GGGAC G P	V GCAT I CCAT I ACCC P CTGAC	G GAAGA	T GGAG E GGGGG R ACAG V CTC L GGGGR	T SATA I SCAGE T T SACA T CTTTL	V CACA T GGGC G TCA S GGAG E TCT C C CAC H	E AGAC D STOCK S GAC D SCTGC C CAT H CCAL	F TTCT S TGGG G TCCT P GTGG GTG GTG GTG GTG GTG GTG GTG GTG	S (GACCACCACCACCACCACCACCACCACCACCACCACCACC	CCACC	P P GGTG V P GGGGG A : GGGGGG A GGGGGGGGGGGGGGGGGG	G CTC L TTC F CTG L GTC V GTG V AGT S CAC H
1570 517 1699 560 1828 603 1957 646 2086 689 2215 732 2344 775 2473 818 8602 861	M TACE Y TOTE S TACE Y GGT G ACCC T GAGGE CCC F	A CACC T T CAA P CAAC Y T TACC Y T TACC Y T TACC H CAAC K CTCC S	L :	R)	E E E E E E E E E E E E E E E E E E E	A GCCT S S A GCC P T CA T T T T T T T T T T T T T T T T T	F CCAT I I CTTA Y CAGO A Y CATCAT I STTA Y CCCCC P V CCCCC P V	O CAG R CCAG R CCAG CCAG S CCAA K GCT L	Q CCAT I GGCCG R GGGT V ACTT F AGGT C C GGGGT C C GGAAA K	R POGGA E FOCTT F FOGT V CAA K TCTC G G G	H LATG C GCAG Q GGGG G TGAG E CTTG L CCAG	C CGTC V GGAA Q GGTT V ATCC S FCCC P GGAA E	Y CCGAC C ATAC Y CCCC P CGTGC V CAATAC S CGTGC V CAATAC CGTGC V CGTGC CGTG	E ! COCCE CO COCCE CO COCCE CO	P F CACG. H D TGGGG M G GGGGGCC M G GGGGCC M N N N N N N CTCAC	V ACCC P STGT V CCCGGR ACAC. T CCTGTG V	K CCAG Q GCAT H TGGC G ATGT C CCAG Q GCAA H F GCCAT H F F	Y : STORY ST	AACCC AACCC AACCC AACCC AACCC AACCC	GGACA GGACA D GGACA D GGACA ACCCC P AAAGCC S CGCCC A CTCCG S	S S S S S S S S S S S S S S S S S S S	S CCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	A CCCTG C C C C C C C C C C C C C C C C C	P CCGA R R GCTG L GCCTG L CCCA P	S SAGGER S S	Y P GTGTT V C ATCCC I R GTGTCAAC V T AAAGAAC K N TCATC GGGAC G G GGGAC F I	V GCAT I ACCC P GCCA Q GTGC A CTGA C	G GGGGG G G G G G G G G G G G G G G G	T GGGAG E GGGGG R ACAG Q AGTG V CCTC L AGGGGR R A AGGGGAR	T SATO I CAG Q CCT P AACA T ACC T CTT CTT GCCC A	V CACA T GGGC G TCA S TCT C CAC H TCT S	E AGAC D STORY OF THE COLOR OF	F TTCT S GGG G CCT P GTG CATTCH I CATTCH D GGATC CGATC	S (CACCACCACACACACACACACACACACACACACACACA	D D GGGGGG V D TACE TO GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	P P P P P P P P P P P P P P P P P P P	G CTC L TTC F CTG L GTC V GTG V AGT S CAC H
1570 517 1699 560 1828 603 1957 646 2086 689 2215 732 2344 775 4473 818 8602 861	M TAC Y TC1 S TAC Y GG1 G ACC T GAG E CCC F	A CACC T T CACA P CACA	L : CTGC CTGC CTGC	R) GAGG GAGG CAAGG CCAAGG CCAAGG	E E E E E E E E E E E E E E E E E E E	A SCT S S S S S S S S S S S S S S S S S S	F CCAT I I CTTA Y CAGGO P V AGGC.	Q CAG S CAA K GCT L	Q CAT I GGCCG R GGGT V ACTT F AGGT C C GGGGT K TGCC	R CCGA E CCTT F CCTT V CAA K TCTC G G CCCCC	H LATG C C C C C C C C C C C C C	C CGTC V GGAA Q GGTT V ATCC S FCCC P GGAA E	Y CCGACC C ATACC Y CCCC P CGGGGC C CGGGCC CGGCCC CGCCC CGGCCC CGCCC CGCC CGCCC CGCCC CGCCC CGCCC CGCCC CGCC CGCCC CGCCC CGCC CCC CGCC CGCCC CGCC	E ! COGGE	P F CACG. H D TGGGG H G GGGGCC H G GGGGCC H CC H CC CC CC CC CC	V ACCC P STGT V CCCCG R ACAC. T CCCCG C CCCCG C CCCCG V TTGCC	K CCAG Q GCAT H TGGC G ATGT C CCAG Q GCAA Q CCAT H TCAC	Y : : : : : : : : : : : : : : : : : : :	AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	GGACA GGACA GGACA GGACA GGACA GGACA ACCCC P AAAGCC S TGACC A CTCG S TGACC A CTCG S TGACC	S S S S S S S S S S S S S S S S S S S	SCAGO A GCA I CCTX S TCCX P GTG/A B GT	A CCCAPA CCCAPA M CCCAPA N AGGAN V AGGAN D TGGG G TGAC D	P CCGA R R GCTG L GCCTG L CCCA P	S SAGGE A SAGG	Y P GTGT V C ATCCC I R GTCAC I R ANGAL K N TCATC S C GGGAC S P CCCAT P I GCCAT	V GCAT I ACCC P GCCA Q GTGC A CTGA CTGA CTGA CTGA CTGA CTGA CT	G GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	T GGGAG R GGTGC R ACAG V CTC L GGGTA A GGTTA A CTGGTG A	T I I CAGG COTT I CTT I	V CACA T GGGC G TCA S TGC C CAC H TCT S	E AGAC D CTGC C CATCH CCAC	F TOTOT S S S S S S S S S S S S S S S S S	S COACOA CACCA CAC	GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	P P P P P P P P P P P P P P P P P P P	G CTC L TTC F CTG L GTC V GTG V AGT S CAC H
1570 517 1699 560 1828 603 1957 646 2086 689 2215 732 2344 775 4473 818 8602 861	M TAC Y TC1 S TAC Y GG1 G ACC T GAG E CCC F	A CACC T T CACA P C CACA C CACA C	L : CTGC CTGC CTGC	R) GAGG GAGG CAAGG CCAAGG CCAAGG	E E E E E E E E E E E E E E E E E E E	A SCT S S S S S S S S S S S S S S S S S S	F CCAT I I CTTA Y CAGGO P V AGGC.	Q CAG S CAA K GCT L	Q CAT I GGCCG R GGGT V ACTT F AGGT C C GGGGT K TGCC	R CCGA E CCTT F CCTT V CAA K TCTC G G CCCCC	H LATG C C C C C C C C C C C C C	C CGTC V GGAA Q GGTT V ATCC S FCCC P GGAA E	Y CCGACC C ATACC Y CCCC P CGGGGC C CGGGCC CGGCCC CGCCC CGGCCC CGCCC CGCC CGCCC CGCCC CGCCC CGCCC CGCCC CGCC CGCCC CGCCC CGCC CCC CGCC CGCCC CGCC	E ! COGGE	P F CACG. H D TGGGG H G GGGGCC H G GGGGCC H CC H CC CC CC CC CC	V ACCC P STGT V CCCCG R ACAC. T CCCCG C CCCCG C CCCCG V TTGCC	K CCAG Q GCAT H TGGC G ATGT C CCAG Q GCAA Q CCAT H TCAC	Y : : : : : : : : : : : : : : : : : : :	AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	GGACA GGACA GGACA GGACA GGACA GGACA ACCCC P AAAGCC S TGACC A CTCG S TGACC A CTCG S TGACC	S S S S S S S S S S S S S S S S S S S	SCAGO A GCA I CCTX S TCCX P GTG/A B GT	A CCCAPA CCCAPA M CCCAPA N AGGAN V AGGAN D TGGG G TGAC D	P CCGA R R GCTG L GCCTG L CCCA P	S SAGGE SAGG	Y P GTGT V C ATCCC I R GTCAC I R ANGAL K N TCATC S C GGGAC S P CCCAT P I GCCAT	V GCAT I ACCC P GCCA Q GTGC A CTGA CTGA CTGA CTGA CTGA CTGA CT	G GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	T GGGAG R GGTGC R ACAG V CTC L GGGTA A GGTTA A CTGGTG A	T I I CAGG COTT I CTT I	V CACA T GGGC G TCA S TGC C CAC H TCT S	E AGAC D CTGC C CATCH CCAC	F TOTOT S S S S S S S S S S S S S S S S S	S COACOA CACCA CAC	GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	P P P P P P P P P P P P P P P P P P P	G CTC L TTC F CTG L GTC V GTG V AGT S CAC H
1570 517 1699 560 1828 603 1957 646 689 2215 732 2344 775 4473 818 1602 861	TACTY GGT G GAGGE CCC F TTCC F CGT R	A CACCO TO TOO TO TOO TO TOO TO TOO TO TOO TO	L : CTGC CTGC CTGC CTGC CTGC CTGC CTGC CTG	R) GRACO GRAC	E EAGGE E E E E E E E E E E E E E E E E	A AGCOMENT TO TO THE AGGORAL A AGCOMENT TO THE AGGORAL A AGCOMENT TO THE AGGORAL A AGCOMENT A AGCOMENT TO THE AGGORAL A AG	FOR THE PROPERTY OF THE PROPER	O CAG CAA CCAA CCAA CCAA CAA CAA C	Q CCC P CCC	R CGAA CCTT CTT CAA K TCTC CGGGG	H LATGE C GCAG G GTTT E CCAG Q CTCC S	C CGTC V GGAC D CCAA Q CGGAA C CGGAA E CGGAA E CGGAA E CGGAA E CGGAA CGG	Y CGAC C C ATAC C ATAC Y CCCC P CAGAC S CAGAC S CAGC A 1	E 1 1 COCCECCION CONTROL CONTR	CACGO A	V ACCC P STGT V CCCG R ACAC T CCTGTC V CGTGC V CTGCT A	CCAG CCAG CCAG CCAG CCAT CCAG CCCAG CCAG CCCAG CCCCAG CCCCAG CCCAG CCCCAG CCCCAG CCCAG CCCCAG CCCCAG CCCCAG CCCCAG CCCCAG CC	TATOGOGO F	AAATGAAATGAAATGAAATGAAAATGAAAATGAAAAAAAA	GACAC GGACA D GGTC L ACCC P AAAGC S TGAC C C C C C C C C C C C C C C C C C C	GAGGE FAMAGE FAMAGE INTERPRETATION OF THE PROPERTY OF THE PROP	S CAGE A GAME B GGAGA S COLOR S	A CCAPT M CCAP	R GCTG L GGCC G GCTG L GCTG L GCCA P ATTG	S S S S S S S S S S S S S S S S S S S	Y P GTGTT V C ATCCC ATCC	V GCAT V CCAT I ACCC P GCCA Q CTGA CTGA CTGA CTGA CTGA CTGA CTGA V	G G G G G G G G G G G G G G G G G G G	T GGAG R GTTC R ACAG V CTC R GCT A CTG L	T SATA I CAG Q CCT P ACA T ACC T GCC X	T GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	E AGAC D TOTAL S GAC D CTG C C C AT H C C AG G C C G G C G G G C G G G G G G	FOR TOTAL SECTION OF THE POST	S CACCACCACCACCACCACCACCACCACCACCACCACCAC	E D GGCGC GGCGC STACC SGCAC GAGGCAC TCTA Y TCTA	P GGGGG V GGGCG A GGGTG V GGATC CAACC R CAACC F 5	G CTC L TTC F CTG CTG V GTG V AGT S CAC H GGCC GTG GTC GTG GTC GTG GTC GTG GTG GT
1570 517 1699 560 1828 603 1957 646 2086 689 2215 732 2344 775 4473 818 8002 861	TACTY GGT G GAGGE CCC F TTCC F CGT R	TOTAL PERSON OF THE PROPERTY P	L : CTGC CAGG CAGG CAGG CAGG CAGG CAGG CAGG	GRACCO A FACTOR OF THE PROPERTY OF THE PROPERT	E E E E E E E E E E E E E E E E E E E	A GCCT S A GCCCT P T CAN T T T T T T T T T T T T T T T T T T T	FEED TO A TOTAL TO	O CAG	Q CAN COLOR R	R CGAA CCTT F CGAA K TCTT L GGGGG	H ATGA C GGGA G GGGA E GGGGA E C C C C C C C C C C C C	C CGTC V GGAC D CCAA O C CGAA C CGAA C CAGT S C	Y CGAC C ATAC C P CGTGC C AGAA CGAA A A CGAA CG	E 1 1 COCTO	P F CACG. H D TGGGG GGGCCC H GGGGCCC H GGGGCCC H GGGGCCC H GGGGCCC H GGGAAC C C C H AAATC C C C H AAATC C C C	V ACCC P STGT V CCCG R ACAC C CTGT V V CGTGG C V CTGC A	K CCAG Q GCAT H TGGC G ATGT C CCAG Q GCAAT H TCACG H TTACGC H TTACGC H TTACGC TTACGC H TTACGC TTACCC TTACCC TTACCC TTACCC TTACC	Y TOO TOO TO TOO TO TOO TO TOO TO TOO TO	AAATGAAATGAAAATGAAAAATGAAAAAAAAAAAAAAA	GGACA GG	GAGGE FAAGGE FAA	S CAGO A A GAM	A CCATT M CCAAM N CCAAM D CCAA	R GCCGA R GCCCG G GCCCG L GCCCG L GCCCG L GCCCA P ATTG	S SAGGE A SAGG	Y P GTGTT V C ATCCC ATCC	V GCAT I ACCC P GCCA Q GTGA CTGA CTGA CTGA CC CTTA	G GCTG CCAC Q Q ATC! S GAG! R CATT! I GAAGE K TAGE L L TGAGE	T GGAGG R GGTG V CTC L GGGG R GGTT L GGGGT L GGGGT L GGGGT L GGGT L GGGGT L GGGGGT L GGGGGT L GGGGT L GGGGGT L GGGGGT L GGGGGT L GGGGT L GGGGGT L	T SATA I CAG Q CCT P ACA T ACC T GCC A GGG GGG GGG	V CACA T GGGGE TCA S GGAGE TCT S TGC C CAC G G TCT TCT TCT TCT TCT TCT TCT TCT TC	E AGAC D CTG C CAT H CCAT H CCAT G G C CTG C C C C C C C C C C C C C C	FTCTT S TCCTT	S CACCACCACCACCACCACCACCACCACCACCACCACCAC	GOOD TOTAL TOTAL Y CAGG	P P P P P P P P P P P P P P P P P P P	G CTC L TTC F CTG L GTC V GTG V AGT S CAC H
1570 517 1699 560 1828 603 1957 646 2086 689 2215 732 2344 775 4473 818 8002 861	TACTY GGT G GAGGE CCC F TTCC F CGT R	TOTAL PERSON OF THE PROPERTY P	L : CTGC CAGG CAGG CAGG CAGG CAGG CAGG CAGG	GRACCO A FACTOR OF THE PROPERTY OF THE PROPERT	E E E E E E E E E E E E E E E E E E E	A GCCT S A GCCCT P T CAN T T T T T T T T T T T T T T T T T T T	FEED TO A TOTAL TO	O CAG	Q CAN COLOR R	R CGAA CCTT F CGAA K TCTT L GGGGG	H ATGA C GGCAA C GGGCA G GTTGAA E CCTTC CCTCC CCTCC S CCCGA	C CGTC V GGAC D CCAA O C CGAA C CGAA C CAGT S C	Y CGAC C ATAC C P CGTGC C AGAA CGAA A A CGATC CGCC A CGATC CGAT	E 1 1 COCTO	P F CACG. H D TGGGG GGGCCC H GGGGCCC H GGGGCCC H GGGGCCC H GGGGCCC H GGGAAC C C C H AAATC C C C H AAATC C C C	V ACCC P STGT V CCCG R ACAC C CTGT V V CGTGG C V CTGC A	K CCAG Q GCAT H TGGC G ATGT C CCAG Q GCAAT H TCACG H TTACGC H TTACGC H TTACGC TTACGC H TTACGC TTACCC TTACCC TTACCC TTACCC TTACC	Y TOO TOO TO TOO TO TOO TO TOO TO TOO TO	AAATGAAATGAAAATGAAAAATGAAAAAAAAAAAAAAA	GGACA GG	GAGGE FAAGGE FAA	S CAGO A A GAM	A CCATT M CCAAM N CCAAM D CCAA	R GCCGA R GCCCG G GCCCG L GCCCG L GCCCG L GCCCA P ATTG	S SAGGE A SAGG	Y P GTGTT V C ATCCC ATCC	V GCAT I ACCC P GCCA Q GTGA CTGA CTGA CTGA CC CTTA	G GCTG CCAC Q Q ATC! S GAG! R CATT! I GAAGE K TAGE L L TGAGE	T GGAGG R GGTG V CTC L GGGG R GGTT L GGGG R CTG L GGGGL L GGGGGL L GGGGL L GGGGL L GGGGGL L GGGGL L GGGGGL L GGGGGL L GGGGGL L GGGGGL L GGGGGG	T SATA I CAG Q CCT P ACA T ACC T GCC A GGG GGG GGG	V CACA T GGGGE TCA S GGAGE TCT S TGC C CAC G G TCT TCT TCT TCT TCT TCT TCT TCT TC	E AGAC D CTG C CAT H CCAT H CCAT G G C CTG C C C C C C C C C C C C C C	FTCTTS CCTM GGTG CCATM GGATC GGA	S CACCACCACCACCACCACCACCACCACCACCACCACCAC	GOOD TOTAL TOTAL Y CAGG	P GGGGG V GGGCG A GGGTG V GGATC CAACC R CAACC F 5	G CTC L TTC F CTG L GTC V GTG V AGT S CAC H
1570 517 1699 560 1828 603 1957 646 2086 689 2215 732 2344 775 4473 818 8002 861	TACTY GGT G GAGGE CCC F TTCC F CGT R	TOTAL PERSON OF THE PROPERTY P	L : CTGC CAGG CAGG CAGG CAGG CAGG CAGG CAGG	GRACCO A FACTOR OF THE PROPERTY OF THE PROPERT	E E E E E E E E E E E E E E E E E E E	A GCCT S A GCCCT P T CAN T T T T T T T T T T T T T T T T T T T	FEED TO A TOTAL TO	O CAG	Q CAT F A ACT	R CGAA CCTT F CGAA K TCTT L GGGGG	H ATGA C GGGA G GGGA E GGGGA E C C C C C C C C C C C C	C CGTC V GGAC D CCAA O C CGAA C CGAA C CAGT S C	Y CGAC C ATAC C P CGTGC C AGAA CGAA A A CGATC CGCC A CGATC CGAT	E 1 1 COCTO	P F CACG. H D TGGGG GGGCCC H GGGGCCC H GGGGCCC H GGGGCCC H GGGGCCC H GGGAAC C C C H AAATC C C C H AAATC C C C	V ACCC P STGT V CCCG R ACAC C CTGT V V CGTGG C V CTGC A	K CCAG Q GCAT H TGGC G ATGT C CCAG Q GCAAT H TCACG H TTACGC H TTACGC H TTACGC TTACGC H TTACGC TTACCC TTACCC TTACCC TTACCC TTACC	Y TOO TOO TO TOO TO TOO TO TOO TO TOO TO	AAATGAAATGAAAATGAAAAATGAAAAAAAAAAAAAAA	GGACA GG	GAGGE FAAGGE FAA	S CAGO A A GAM	A CCATT M CCAAM N CCAAM D CCAA	R GCCGA R GCCCG G GCCCG L GCCCG L GCCCG L GCCCA P ATTG	S SAGGE A SAGG	Y P GTGTT V C ATCCC ATCC	V GCAT I ACCC P GCCA Q GTGA CTGA CTGA CTGA CC CTTA	G GCTG CCAC Q Q ATC! S GAG! R CATT! I GAAGE K TAGE L L TGAGE	T GGAGG R GGTG V CTC L GGGG R GGTT L GGGG R CTG L GGGGL L GGGGGL L GGGGL L GGGGL L GGGGGL L GGGGL L GGGGGL L GGGGGL L GGGGGL L GGGGGL L GGGGGG	T SATA I CAG Q CCT P ACA T ACC T GCC A GGG GGG GGG	V CACA T GGGGE TCA S GGAGE TCT S TGC C CAC G G TCT TCT TCT TCT TCT TCT TCT TCT TC	E AGAC D CTG C CAT H CCAT H CCAT G G C CTG C C C C C C C C C C C C C C	FTCTTS CCTM GGTG CCATM GGATC GGA	S CACCACCACCACCACCACCACCACCACCACCACCACCAC	GOOD TOTAL TOTAL Y CAGG	P P P P P P P P P P P P P P P P P P P	G CTC L TTC F CTG L GTC V GTG V AGT S CAC H
1570 517 1699 560 1828 603 1957 646 2086 689 2215 732 2344 775 4473 818 8002 861	TACTY GGT G ACCC T GAGGE CCCC F TTCC R AGAR	A CACCAT TOCAM P CGAT D TTACCY STAT Y SCAC B CACCAT CCAC S TAGC F	L SCAGO	GAAGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGAAGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGAAGGGAAGGAGGAAGGAGGAAGGAGGAAGGAGGAAGGAGGAAGGAGGAAGGAGGAAGGAGGAAGGAAGGAGGAAGGAGGAAGGAGA	E E E E E E E E E E E E E E E E E E E	A SCT S S S S S S S S S S S S S S S S S S	EAGC A FOR TOTAL	Q CAG CAAG CAA CC P GCA Q	Q CAT F A ACT	R CGAA CCTT F CGAA K TCTT L GGGGG	H ATGA C GGGA G GGGA E GGGGA E C C C C C C C C C C C C	C CGTC V GGAC D CCAA O C CGAA C CGAA C CAGT S C	Y CGAC C ATAC C P CGTGC C AGAA CGAA A A CGATC CGCC A CGATC CGAT	E 1 1 COTTO	P F CACG. H D TGGGG GGGCCC H GGGGCCC H GGGGCCC H GGGGCCC H GGGGCCC H GGGAAC C C C H AAATC C C C H AAATC C C C	V ACCC P STGT V CCCG R ACAC C CTGT V V CGTGG C V CTGC A	K CCAG Q GCAT H TGGC G ATGT C CCAG Q GCAAT H TCACG H TTACGC H TTACGC H TTACGC TTACGC H TTACGC TTACCC TTACCC TTACCC TTACCC TTACC	Y TOO TOO TO TOO TO TOO TO TOO TO TOO TO	AAATGAAATGAAAATGAAAAATGAAAAAAAAAAAAAAA	GGACA GG	GAGGE FAAGGE FAA	S CAGO A A GAM	A CCATT M CCAAM N CCAAM D CCAA	R GCCGA R GCCCG G GCCCG L GCCCG L GCCCG L GCCCA P ATTG	S SAGGE A SAGG	Y P GTGTT V C ATCCC ATCC	V GCAT I ACCC P GCCA Q GTGA CTGA CTGA CTGA CC CTTA	G GCTG CCAC Q Q ATC! S GAG! R CATT! I GAAGE K TAGE L L TGAGE	T GGAGG R GGTG V CTC L GGGG R GGTT L GGGG R CTG L GGGGL L GGGGGL L GGGGL L GGGGL L GGGGGL L GGGGL L GGGGGL L GGGGGL L GGGGGL L GGGGGL L GGGGGG	T SATA I CAG Q CCT P ACA T ACC T GCC A GGG GGG GGG	V CACA T GGGGE TCA S GGAGE TCT S TGC C CAC G G TCT TCT TCT TCT TCT TCT TCT TCT TC	E AGAC D CTG C CAT H CCAT H CCAT G G C CTG C C C C C C C C C C C C C C	FTCTTS CCTM GGTG CCATM GGATC GGA	S CACCACCACCACCACCACCACCACCACCACCACCACCAC	GOOD TOTAL TOTAL Y CAGG	P P P P P P P P P P P P P P P P P P P	G CTC L TTC F CTG L GTC V GTG V AGT S CAC H

386

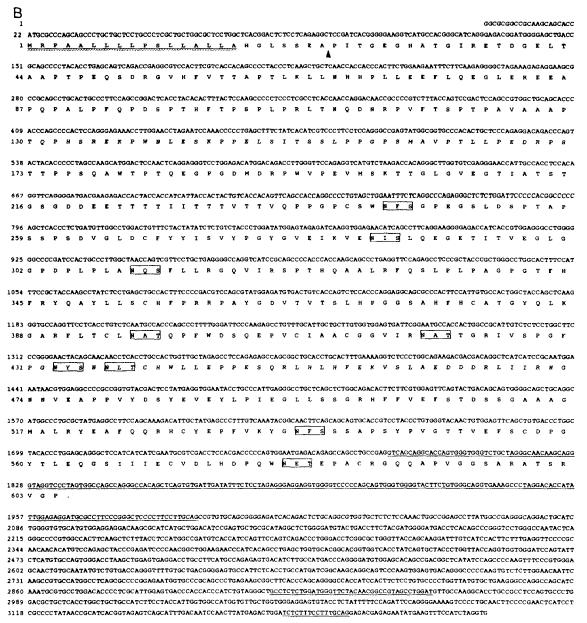


Figure 4. Nucleotide sequences of type 2 (A) and type 3 (B) SEZ-6 cDNA. Different sequences from type 1 are underlined. The deduced amino acid sequences of ORF are indicated below the nucleotide sequences by a single letter code. The putative signal peptide is underlined twice and the potential cleavage site is indicated by an arrowhead. Potential N-linked glycosylation sites are boxed. The nucleotide sequences have been deposited in DDBJ under accession number D64009 for type 2 and D64010 for type 3.

amino acids in N-terminus which probably works as a signal sequence. Although type 2 protein has a cluster of hydrophobic amino acids near the C-terminal region, type 3 protein lacked this sequence. These findings suggest that type 2 protein is a transmembrane protein like type 1 but type 3 is not. Some membrane proteins of neural cells have a transmembrane form and a PI-linked form (13, 14). The C-terminus stretches of moderately hydrophobic amino acids are

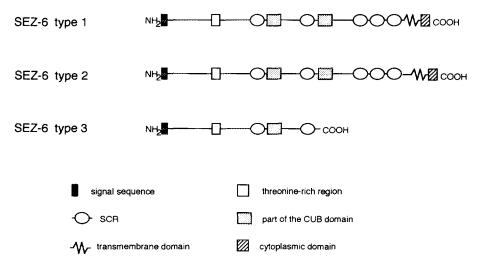


Figure 5. Schematic drawing of the domain structure of 3 types of SEZ-6 protein.

characteristic of proteins attached to the membrane by a PI anchor (15). There was no such sequence in the SEZ-6 type 3 sequence, suggesting type 3 is likely to be a secreted protein. Fig. 5 shows a summary of the domain structure of type 1, 2 and 3 SEZ-6 proteins. As shown in Fig. 5, SEZ-6 encodes two motifs, SCR and CUB-like domain (1). Type 1 and type 2 proteins have five SCRs and two CUB-like domains. On the other hand, type 3, shorter secreted protein, has two SCRs and one CUB-like domain. SCR is known as a characteristic structure of the superfamily of complement C3b/4b binding proteins. However, not only complement-related proteins but many non-complement proteins have SCRs. These proteins which have SCRs thought to be involved in protein-protein interaction (3, 4). Furthermore, CUB domain, another motif in the SEZ-6 protein, is also suggested to have a role in protein-protein interaction (5). Recent studies show that two molecules of neural cells have SCR or CUB domain. A5 antigen (neuropilin) with two CUB domains is expressed on the surface of both nerve fibers and visual centers. Drosophila secreted protein Hikaru genki with three or four SCRs is produced by subsets of the cells in the central nervous system. Both proteins are thought to be necessary for the formation of a precise neural network (6, 7, 16). SEZ-6 protein is the only molecule in the nervous system so far studied which has both SCR and CUB domain. Functions of SEZ-6 protein in vivo may be similar to such molecules. In the present study, we cloned two additional SEZ-6 cDNA which were probably produced by alternative splicing; one encodes the transmembrane protein and the other encodes the secreted type. The functional difference of these three types of SEZ-6 protein should be clarified.

ACKNOWLEDGMENTS

This work was supported in part by a research grant to K. S.-N. from Terumo Life Science Foundation. We thank Ms. Naoko Yamaguchi and Ms. Ritsuko Ishii for their assistance and Dr. Minoru Kimura for valuable advice.

REFERENCES

- 1. Shimizu-Nishikawa, K., Kajiwara, K., Kimura, M., Katsuki, M., and Sugaya, E. (1995) Mol. Brain Res. 28, 201-210.
- Kajiwara, K., Sugaya, E., Kimura, M., Katsuki, M., Nagasawa, H., Yuyama, N., Tsuda, T., Sugaya, A., Motoki, M., Ookura, T., and Shimizu-Nishikawa, K. (1995) Brain Res. 671, 170-174.
- Kristensen, T., D'eustachio, p., Ogata, R. T., Chung, L. P., Reid, K. B. M., and Tack, B. F. (1987) Fed. Proc. (1987) 46, 2463-2469.
- 4. Hourcade, D., Holers, V. M., and Atkinson, J. P. (1989) Adv. Immunol., 45, 381-416.
- 5. Bork, P. and Beckmann, G. (1993) J. Mol. Biol. 231, 539-545.
- 6. Hoshino, M., Matsuzaki, F., Nabeshima, Y., and Hama, C. (1993) Neuron 10, 395-407.
- 7. Takagi, S., Hirata, T., Agata, K., Eguchi, G., and Fujisawa, H. (1991) Neuron 7, 295-307.
- 8. Asou, H., Brunngraber, E. G., and Jeug, I. (1983) J. Histochem. Cytochem. 31, 1375-1379.
- 9. Laemmli, U. K. (1970) Nature 227, 680-685.
- 10. Towbin, H., Staehelin, T., and Goedon, J. (1979) Proc. Natl. Acad. Sci. 76, 4350-4354.
- 11. Chomczynski, P., and Sacchi, N. (1987) Anal. Biochem. 162, 156-159.
- 12. Seki, T., and Arai, Y. (1993) Neurosci. Res. 17, 265-290.
- Hemperly, J. J., Edelman, G. M., and Cunningham, B. A. (1986) Proc. Natl. Acad. Sci. USA 83, 9822-9826.
- 14. Grenningloh, G., Rehm, E. J., and Goodman, C. S. (1991) Cell 67, 45-57.
- 15. Ferguson, M. A. J. (1988) Ann. Rev. Biochem. 57, 285-320.
- 16. Takagi, S., Kasuya, Y., Shimizu, M., Matsuura, T., Tsuboi, M., Kawakami, A., and Fujisawa, H. (1995) Dev. Biol. 170, 207-222.