# Characterization of a novel structural member, LukE-LukD, of the bi-component staphylococcal leucotoxins family

A. Gravet, D.A. Colin, D. Keller, R. Giradot, H. Monteil, G. Prévost\*

UPRES EA-1318, Institut de Bactériologie de la Faculté de Médecine, Université Louis Pasteur-Hôpitaux Universitaires de Strasbourg, 3 rue Koeberlé, F-67000 Strasbourg, France

Received 19 August 1998

Abstract A new member of the staphylococcal bi-component leucotoxins family, LukE (32 kDa) and LukD (34.3 kDa) has been characterized from Staphylococcus aureus strain Newman. LukE was 58-68% identical with the class S proteins, whereas LukD was 71-77% identical with the class F proteins of the family. A partial immunoreactivity with the various affinitypurified antibodies specific for the other proteins was observed. Immunoprecipitation assay and gene probing confirmed a 30% frequency among human clinical isolates, differing from the distribution of the other known leucotoxins (P < 0.005). LukE+LukD was as effective as the Panton-Valentine leucocidin for inducing dermonecrosis when injected in the rabbit skin, but not hemolytic and poorly leucotoxic compared to other leucotoxins expressed by Staphylococcus aureus.

© 1998 Federation of European Biochemical Societies.

Key words: Characterization; Bi-component leucotoxin; LukE+LukD; Ca<sup>2+</sup>/ethidium penetration; Dermonecrosis; Staphylococcus aureus

## 1. Introduction

Staphylococcal bi-component leucotoxins are composed of two distinct proteins: S-related (Slow eluted: 31-32 kDa) or F-related proteins (Fast eluted: 36-38 kDa) that are secreted separately, but act synergistically upon human polymorphonuclear cells or erythrocytes (PMNs). They constitute a family of pore-forming toxins comprising the anciently described Panton-Valentine leucocidin (PVL = LukS-PV+LukF-PV) [29], and  $\gamma$ -hemolysin, self-comprising two leucotoxic entities HlgA+HlgB and HlgC+HlgB where HlgA and HlgC are Srelated components and HlgB is an F-related component [6,22]. The previous leucocidin R and that encoded by a methicillin-resistant S. aureus are bi-component leucotoxins comparable to the latter [13,22,28]. Recently, another S-related protein, LukM, which is also coupled to a class F component, LukF-PV like (LukF'-PV in this text), was characterized [14]. Another leucotoxin, LukS-I+LukF-I was reported to be produced by Staphylococcus intermedius strains [23]. These components were shown to be predominantly constituted of  $\beta$ sheet structures [20]. Genes encoding PVL are distributed in 2% of clinical strains, strongly associated with furuncles [7,21]. y-Hemolysin encoding genes are encountered in 99% of clinical isolates of S. aureus [21]. Any combination of a given class S protein with a class F protein may be toxic [22] and generates an intense inflammatory response when injected in the rabbit skin [8] or in the pars plana of the rabbit eye [25]. Sequential binding of S components to cell membranes is a

E-mail: gprevibs@tpgnet.net

prerequisite for that of F components [5]. This provokes a  $Ca^{2+}$  influx by opening of  $Ca^{2+}$  channels [27], and the activation of target cells [15,16]. Then, the leucotoxin forms pores through the plasma membrane which are sensitive to both Na<sup>+</sup>, K<sup>+</sup>, or ethidium [10,19,27].

The aim of this work was to characterize a cross-immunoreactive product with PVL components, and to demonstrate its effective membership in this leucotoxin family.

# 2. Materials and methods

#### 2.1. Bacteria and animals

Escherichia coli XL1 Blue (Stratagene) was used as a recipient strain of the reduced genomic library cloned into the shuttle plasmid (E. coli/ S. aureus) pCU1 [2], kindly given by F. Götz (Tübingen, Germany). S. aureus Newman (NTCC 8178) and P83 strains were used for the purification of LukE/LukD and LukM/LukF'-PV [14], respectively. For the toxin-production study, 127 consecutive clinical strains from 127 individuals were collected from the routine laboratory of the Strasbourg University Hospital.

#### 2.2. Protein purification

One colony of Newman strain was grown for 8 h in 5 ml of YCP (3% (w/v) yeast extract (Oxoid), 2% (w/v) bacto-casaminoacids (Difco), 2% (w/v) sodium pyruvate (Merck), 0.25% (w/v) Na<sub>2</sub>HPO<sub>4</sub>, 0.042% (w/v) KH<sub>2</sub>PO<sub>4</sub>, pH 7.0) at 37°C. Then, the purification procedure was as described earlier [9,27], involving MonoS cation-exchange FPLC (Fast Performance Liquid Chromatography, Pharmacia, Uppsala, Sweden) and AlkylSuperose hydrophobic FPLC, respectively. Purified compounds were controlled by SDS-PAGE electrophoresis on PHAST System (Pharmacia) and then stored at -80°C at OD<sub>280nm</sub> = 1.0 in 50 mM Na-phosphate buffer, 150 mM NaCl, pH 7.0.

#### 2.3. Nucleic acids methods

DNA modifying and restriction enzymes were used as recommended by the manufacturers (New England Biolabs-Gibco-BRL). synthetic 5' <sup>32</sup>P-labelled oligonucleotide 5'-GAAAATATA/ A TGGA/TGATGGA/TGCA/TGAAGTTAT-3' whose sequence was deduced from the N-terminal sequence of LukE: NTNIENIGDGAE-VIKRTE was used as a probe for cloning lukE and  $\overline{lukD}$ . A total DNA [28] library of Newman strain reduced to HindIII DNA fragments ranging from 2.0 to 3.0 kb inserted into pCU1 was screened as previously described [23]. Universal and reverse primers and the cloning oligonucleotide were initially used for nucleotide sequence determination of double stranded DNA with T7 DNA polymerase [24]. Other primers were deduced from the previously read sequences. Total RNA from 1 ml of 10-, 13-, and 16-h cultures of Newman strain was prepared by the Fast Prep (Bio101) method [4]. Southern and Northern blots [24] were performed by using QIABRANE membrane (Qiagen, Paris). For the determination of the transcription start of lukE-lukD, total RNA of Newman strain was submitted to extension with the 5' <sup>32</sup>P-labelled oligonucleotide 5'-552-TCCTACTGCAAAC-TTGCAGCTAACA-3'-527. The reaction was carried out with 5 µg of RNA with a 1-h hybridization step at 65°C and a 45-min primer extension with an avian myeloblastosis virus reverse transcriptase (Appligene) as previously described [18]. DNA and peptide analyses were performed with DNAStar software (DNAStar, London, UK). The nucleotide sequence of lukEllukD is recorded at the EMBL/Gen-Bank library with the accession number Y13225.

<sup>\*</sup>Corresponding author. Fax: (33) (3) 88 25 11 13.

#### 2.4. Anti-LukE, -LukD, and -LukM antibodies and radial gel immunoprecipitations

Anti-LukS-PV and anti-LukF-PV affinity-purified rabbit polyclonal antibodies [9] were used for immunoblotting [24] during the purification of LukE and LukD. The rabbit polyclonal antibodies anti-LukE, anti-LukD and anti-LukM were affinity-purified [9]. All available antibodies were used in immunoprecipitation assays [9] for the detection of the production of PVL, LukE+LukD, and LukM+LukF'-PV by clinical isolates. A fresh colony of each strain was incubated in a 24well cell culture plate filled with 550 µl of YCP medium, for 16 h at 37°C under a 10% CO2 humid atmosphere with 180 rpm rotary shaking. To enhance the detection, immunoprecipitation plates were capillary dried. Proteins were precipitated for 5 min in 5% (w/v) trichloroacetic acid, and stained with 0.1% (w/v) Coomassie blue for 15 min, then washed five times with a 10% (v/v) acetic acid/30% (v/v) methanol solution, and fixed with a 10% (v/v) acetic acid/5% (v/v) glycerol solution. The  $\chi^2$  test was used to compare the different distributions of the leucotoxin-producing strains.

#### 2.5. Preparation of human polymorphonuclear cells (PMNs)

Twelve ml of J-Prep solution (TechGen) were added to 30 ml of buffy coats from healthy donors and human PMNs were prepared as previously described [27]. Human PMNs were suspended and adjusted to  $2 \times 10^6$  cells/ml with 140 mM NaCl, 5 mM KCl, 10 mM glucose, 0.1 mM ethylene glycol-bis-( $\beta$ -aminoethyl ether) N,N,N',N'-tetraacetic acid, 10 mM HEPES, 3 mM Tris-base (pH 7.3). The latter buffer was used to wash, after loading the cells with the Ca<sup>2+</sup>-specific fluorescent probe Fluo-3 (Molecular Probes, Eugene, OR, USA).

#### 2.6. Flow cytometry measurements

Assays were made with  $2 \times 10^6$  cells/ml loaded with 5  $\mu$ M Fluo-3 during 1 h at 37°C, then washed and resuspended in the absence or presence of 1 mM free Ca<sub>2+</sub>, and 25 nM ethidium bromide. Proteins were added at 2 nM. Flow cytometry measurements were performed by using a FacSort (Becton-Dickinson, Le Pont de Claix, France) equipped with a 15-mW argon laser tuned at 488 nm. Forward (FSC) and Side (SSC) light scatter dot plots acquired from 5000 leucocytes ( $2 \times 10^6$  cells/ml) were classically gated [19]. Fluo-3 fluorescence due to the calcium penetration through opening of Ca<sup>2+</sup> channels was recorded from the fluorescence light 1 ( $\lambda_{\rm Em} = 530$  nm). The fluorescence light 3 ( $\lambda_{\rm Em} = 650$  nm) was used to record the fluorescence of ethidium combined with nucleic acids, after penetrating through the leucotoxin pores. Data were calculated by Lysis 2TM software (Becton-Dickinson, France) and results expressed as mean ± S.D. of data obtained from PMNs of four different donors, at least.

#### 2.7. Activity of LukE, LukD in a rabbit skin model

Shaved New Zealand rabbits (≥2.5) were injected intradermally with 50 µl of sterile 9.55‰ (w/v) PBS Dulbecco solution with equal amounts (100, 300, 1000, 3000, 10000 ng) of class S- and class Frelated components [8,22]. Clinical observations were recorded 12, 24, and 48 h after injection of toxins. Five reaction levels reflecting the intensity of the lesions were defined according to the evolution of the inflammatory response: level 0 was the normal skin, level 1 showed a red inflammatory reaction ( $\emptyset = 10-15$  mm) without abscess formation, level 2 showed a hot, red papule with relief ( $\emptyset = 15-20$  mm, 5 mm high), level 3 showed a more intense, blue and oedematous abscess ( $\emptyset = 20-30$  mm, 10 mm high), level 4 still progressed  $(\emptyset = 30 \text{ mm}, 10-15 \text{ mm high})$  with pus collection surrounded by blue oedema, level 5 showed a lesion with a central depression and pus sweating. Data obtained from three independent series of injections were analyzed with the non-parametric test of Wilcoxon-Mann-Whitney.

# 3. Results

# 3.1. New leucotoxin characterization

Characterization of a new leucotoxin from *S. aureus* strain Newman originated by the observation in culture supernates (Fig. 1A) of materials immunologically related to HlgC, LukS-PV, LukF-PV, but not with HlgA, HlgB and LukM, when the strain was grown in YCP medium. This observation evidenced that the supernatants contained proteins related to class S and class F components, which differed from PVL, γ-hemolysin or LukM/LukF'-PV. However, the Newman strain grown in HI medium then produced detectable amounts of HlgA, HlgB, and HlgC (Fig. 1A). By using the radial gel immunoprecipitation test, the class S protein arbitrarily called LukE and the class F protein called LukD were purified by using previously described chromatographies. The class S-related LukE protein was eluted on MonoS at a NaCl concentration of 280-290 mM similar to that of LukM. However, LukE was eluted in 1 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> by hydrophobic chromatography whereas LukM was not retained. LukD and LukF-PV were eluted at 110 mM NaCl onto MonoS and at 780 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> onto alkyl-Superose. LukD and LukF-PV were produced by different strains. SDS-PAGE analysis of purified LukE and LukD confirmed apparent molecular masses similar to the other class S proteins (32 kDa) and to the other class F proteins (36 kDa), respectively (data not shown).

Affinity-purified antibodies were specific enough to distinguish amongst the related antigens (Fig. 1B), by generating continuous immunoprecipitation lines only with corresponding proteins and crossed immunoprecipitation lines between heterologous proteins. Few proteins were not precipitated by antibodies like LukE, LukS-PV and LukM by anti-HlgA antibodies, HlgC by anti-LukE antibodies, HlgA and LukM by anti-HlgC antibodies, or HlgA and LukS-PV by anti-LukM antibodies (Fig. 1B). Antigenic similarity seemed closer with LukD, HlgB, LukF-PV, LukF'-PV; LukF'-PV was not precipitated by anti-HlgB antibodies.

## 3.2. Cloning and sequencing lukE and lukD

Hybridization of the total DNA from Newman strain with the new leucotoxin-directed oligonucleotide probe (data not shown) revealed *DraI* (2.1 kb), *Eco*RI (9 kb), *Eco*RV (>10 kb), *HincII* (1.9 kb), *HindIII* (2.8 kb), *PstI* (>10 kb), *ScaI* (2.1 kb) DNA fragments differing in length from the *Eco*RI (>8.3 kb), *Eco*RV (4.2 kb), *ScaI* (3.8 kb) DNA fragments specific of the *hlg* locus, which seemed to be highly conserved in a majority of *S. aureus* isolates [21,22].

A clone containing a 2.8-kb HindIII insert was obtained and immunoblotting of the crude extract of the recombinant E. coli strain indicated the production of both LukE and LukD with apparent molecular masses comparable to those of the purified components (data not shown). Sequence analysis of a 2571-bp long nucleotide sequence overlapped two consecutive ORFs (Fig. 2) which spanned positions 511-1455 and 1457-2440, respectively. No open reading frame was found in the 500-bp preceding lukE differing from the γ-hemolysin locus. The two ORFs are separated by one thymidine as in other related genes (Fig. 2), and cotranscribed in a single mRNA expressed at the late exponential growth (Fig. 3A). Predictable -35 and -10 transcription signals and a putative ribosome binding site are reported in Fig. 2. The determination of the transcription start by the primer extension method (Fig. 3B) indicated the nucleotide 477, located 10 bases downstream the -35 box. *lukE* encodes a 314-amino acid protein with a calculated molecular mass of 35 222 Da and a calculated pI of 9.87. The sequence encoding a peptide NTNIEN..., located 28 amino acids dowstream the ATG initiation codon, corresponds exactly to the N-terminal sequence of LukE. This 28-residues long peptide bearing positivelycharged residues followed by hydrophobic amino acids corresponds to the secretion signal peptide. The secreted 286-amino acid LukE has a calculated molecular mass of  $32\,239$  Da and a calculated p*I* of 9.88. Amino acid content of the secreted LukE revealed twice more basic than acidic residues (42:23). LukE had from 58.7% to 68% identity with the other class S components [6,14,22], with no conserved peptide sequence longer than 7 residues.

The second ORF, lukD, encodes a 327-amino acid-long protein with a calculated molecular mass of 36974 Da and a calculated pI of 9.09. This nucleotide sequence is preceded by an rbs-like sequence. An inverted repeat probably involved in the termination of the transcription is located 60 bp downstream the stop codon of lukD. The 18 first sequenced N-terminal residues of the secreted LukD (AQNIT...), corre-

spond to those deduced from the nucleotide sequence (position 1535). Therefore, the signal peptide of LukD, bears similar characteristics as above, contains 26 residues, and the secreted LukD is 301 amino acids long for a calculated molecular mass of 34 302 Da with a calculated pI of 9.02. The peptide sequence identity of LukD with class F components was comprised between 71 and 77% according to the proteins. Four domains were conserved between LukD and other class F components (from residues 34 to 52, 94 to 103, 137 to 150, 166 to 183). Despite a lack of sequence identity with  $\alpha$ -toxin [12], homology was observed with both class S and class F proteins, confirming the observations described recently [11].

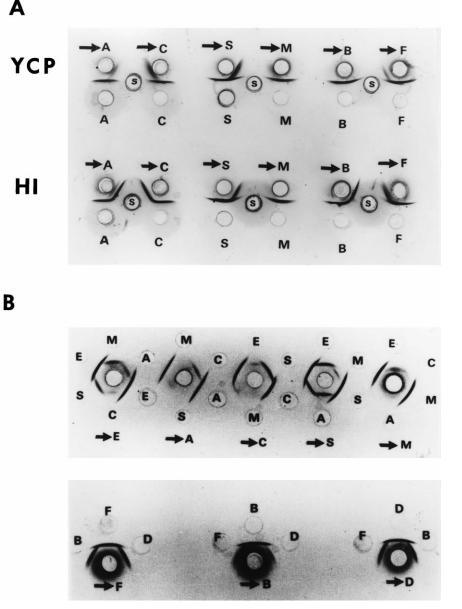


Fig. 1. Characterization of the new staphylococcal leucotoxin LukE/LukD. Panel A: Identification of materials immunologically related to bicomponent leucotoxins secreted by *S. aureus* Newman strain. Immunoprecipitation tests of culture supernatants (s) in YCP and HI media against 100 µg of affinity-purified rabbit antibodies anti-HlgA ( $\rightarrow$ A), anti-HlgB ( $\rightarrow$ B), anti-HlgC ( $\rightarrow$ C), anti-LukF-PV ( $\rightarrow$ F), anti-LukM ( $\rightarrow$ M), anti-LukS-PV ( $\rightarrow$ S) affinity purified rabbit polyclonal antibodies and 1 µg of HlgA (A), HlgB (B), HlgC (C), LukF-PV (F), LukM (M), LukS-PV (S). Panel B: Analysis by an immunoprecipitation assay stained with Coomassie blue of the antigenic relations between 1 µg of purified proteins constituting leucotoxins (HlgA: A; HlgB: B; HlgC: C; LukD: D; LukE: E; LukF-PV: F; LukM: M; LukS-PV: S), and 100 µg of the corresponding affinity-purified rabbit antibodies ( $\rightarrow$ A,  $\rightarrow$ B,...).

1	САСТТТАТСА	AGACTTCGAA	AATCATGGTT	ͲͲͲͲϪͲͲͲͲͲ	AAAGTGAAAT	TTTAGTTCAT	астааатата	70
						GTAAGCGATA		140
						TCGTTTAACG		210
						CATTATTTTT		280
						TTAAAGGTAA		350
351	GCCAAAATGA	CCGGTTTTCA	GCACGATAAA	TGAATAAATT		TAATATTTTA	TAGCATTTTA	420
					-35	⊢ St	art	
421	TTTTTTAATG	AACAGTTAAT	TTTCATAAAT	AAGTGAAATA	ATCTAGAATA	AATTTGTATT	AAAAATTTAT	490
	-10 rbs							
491	ATAGAAAGAA	AGTGAAACTT	ATGTTTAAGA	AAAAAATGTT	AGCTGCAAGT	TTGTCAGTAG	GACTGATTGC	560
1)1	11111011110111	1101011		KKML		L S V	GLIA	17
	ACCUUTACCA	TOTOCONTRO				ATATTGGTGA		630
			O E S R					40
18	P L A	<u>S P I</u>			NIE	N I G D	G A E	
						AAATGTCCAA		700
41	VIK	RTED	VSS	K K W	GVTQ	NVQ	FDF	63
701		AAAATATAAC	AAAGACGCTT	TAATTGTTAA	AATGCAAGGT	TTTATTAATT	CCAGAACTTC	770
64	V K D K	ΚΥΝ	KDA	LIVK	MQG	FIN	SRTS	87
771	ATTTTCGGAT	GTGAAGGGTA	GAGGATATGA	ATTAACTAAA	CGTCTGATTT	GGCCATTCCA	ATATAATATA	840
88	FSD	VKG	RGYE	LTK	RLI	WPFQ	Y N I	110
841	GGACTGACGA	CTAAAGATCC	AAATGTTAGC	TTAATCAATT	CAATTACCTT	ACCTAAAACC	AAAATAGAAA	910
111	GLT	ткрі	PNVS	LIN	SITI	. РКТ	KIE	133
						TCAGCACCAT		980
134		GQT	LGY	N I G G	NFQ	SAP	SIGG	157
							AGTAGACAAG	
158	N G S	FNY	SKTI	S Y T	QKS	Y V S E	V D K	180
							AAAAAATTTG	
181		K S V K	WGV	KAN	KFVT	PDG	KKF	203
1121	CGCATGATAG	ATATTTATTC	GTACAAAGTC	CAAATGGTCC	AACAGGTTCA	GCAAGAGAAT	ATTTTGCTCC	1190
204	AADR	YLF	VQS	P N G P	TGS	ARE	YFAP	227
1191	TGATAATCAA	TTGCCACCTT	TAGTTCAAAG	TGGCTTTAAT	CCATCGTTTA	TCACTACACT	ATCACATGAA	1260
228	DNQ	LPP	LVQS	GFN	PSF	ITTL	SHE	250
1261	AAAGGTTCAA	AGTTGATACG	AGTGAATTTG	AAATTTTCAT	ATGGTAGAAA	CTTAGATATT	ACATATGCGA	1330
251	KGS	KLIR		KFS	YGRN	LDI	ТҮА	273
1331	CTTTATTCCC	TAGAACTGGT	ATTTACGCAG	AAAGAAAGCA	TAATGCATTT	GTAAATAGAA	ACTTTGTAGT	1400
							ACTTTGTAGT	
274	TLFP	R T G	IYA	E R K H	N A F	VNR	N F V V	297
274 1401	T L F P TAGATATAAA	R T G GTTAATTGGA	I Y A AAACACACGA	E R K H AATTAAAGT <u>G</u>	N A F <u>AAAGGA</u> CATA	V N R ATTAATATGA	n f v v Aaattgaaaa	297 1470
274 1401 298/3	T L F P TAGATATAAA L R Y K	R T G GTTAATTGGA V N W	I Y A AAACACACGA K T H E	E R K H AATTAAAGT <u>G</u> I K V	N A F <u>AAAGGA</u> CATA K G H	V N R ATTAATATGA N stop <u>M</u>	N F V V AAATTGAAAA K I E K	297 1470 314/5
274 1401 298/3 1471	T L F P TAGATATAAA l R Y K ATTAGGCAAA	R T G GTTAATTGGA V N W TCATCAGTTG	I Y A AAACACACGA K T H E CTTCATCAAT	E R K H AATTAAAGT <u>G</u> I K V TGCACTGCTT	N A F <u>AAAGGA</u> CATA K G H TTGCTATCGA	V N R ATTAATATGA N stop <u>M</u> ATACAGTTGA	N F V V AAATTGAAAA <u>K I E K</u> TGCAGCTCAA	297 1470 314/5 1540
274 1401 298/3 1471 6	T L F P TAGATATAAA L R Y K ATTAGGCAAA <u>L G K</u>	R T G GTTAATTGGA V N W TCATCAGTTG S S V	I Y A AAACACACGA K T H E CTTCATCAAT A S S I	E R K H AATTAAAGT <u>G</u> I K V TGCACTGCTT A L L	N A F <u>AAAGGA</u> CATA K G H TTGCTATCGA L L S	V N R ATTAATATGA N stop <u>M</u> ATACAGTTGA N T V D	N F V V AAATTGAAAA <u>K I E K</u> TGCAGCTCAA <u>A</u> A Q	297 1470 314/5 1540 28
274 1401 298/3 1471 6 1541	T L F P TAGATATAAA L R Y K ATTAGGCAAA <u>L G K</u> AATATCACAC	R T G GTTAATTGGA V N W TCATCAGTTG S S V CTAAAAGAGA	I Y A AAACACACGA K T H E CTTCATCAAT <u>A S S I</u> GAAAAAAGTA	E R K H AATTAAAGT <u>G</u> I K V TGCACTGCTT <u>A L L</u> GATGACAAAA	N A F <u>AAAGGA</u> CATA K G H TTGCTATCGA L L S TCACTTTATA	V N R ATTAATATGA N stop <u>M</u> ATACAGTTGA <u>N T V D</u> CAAAACAACA	N F V V AAATTGAAAA K I E K TGCAGCTCAA A Q GCAACATCTG	297 1470 314/5 1540 28 1610
274 1401 298/3 1471 6 1541 29	T L F P TAGATATAAA L R Y K ATTAGGCAAA <u>L G K</u> AATATCACAC N I T	R T G GTTAATTGGA V N W TCATCAGTTG S S V CTAAAAGAGA P K R E	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V	E R K H AATTAAAGT <u>G</u> I K V TGCACTGCTT <u>A L L</u> GATGACAAAA D D K	N A F AAAGGACATA K G H TTGCTATCGA L L S TCACTTTATA I T L Y	V N R ATTAATATGA N stop <u>M</u> ATACAGTTGA <u>N T V D</u> CAAAACAACA K T T	N F V V AAATTGAAAA K I E K TGCAGCTCAA A A Q GCAACATCTG A T S	297 1470 314/5 1540 28 1610 51
274 1401 298/3 1471 6 1541 29	T L F P TAGATATAAA L R Y K ATTAGGCAAA <u>L G K</u> AATATCACAC N I T	R T G GTTAATTGGA V N W TCATCAGTTG S S V CTAAAAGAGA P K R E	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V	E R K H AATTAAAGT <u>G</u> I K V TGCACTGCTT <u>A L L</u> GATGACAAAA D D K	N A F AAAGGACATA K G H TTGCTATCGA L L S TCACTTTATA I T L Y	V N R ATTAATATGA N stop <u>M</u> ATACAGTTGA <u>N T V D</u> CAAAACAACA K T T	N F V V AAATTGAAAA K I E K TGCAGCTCAA A Q GCAACATCTG	297 1470 314/5 1540 28 1610 51
274 1401 298/3 1471 6 1541 29	T L F P TAGATATAAA I R Y K ATTAGGCAAA <u>L G K</u> AATATCACAC N I T ATAATGATAA	R T G GTTAATTGGA V N W TCATCAGTTG S S V CTAAAAGAGA P K R E	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V	E R K H AATTAAAGT <u>G</u> I K V TGCACTGCTT <u>A L L</u> GATGACAAAA D D K	N A F AAAGGACATA K G H TTGCTATCGA L L S TCACTTTATA I T L Y TTTCATTAAG	V N R ATTAATATGA N stop <u>M</u> ATACAGTTGA <u>N T V D</u> CAAAACAACA K T T	N F V V AAATTGAAAA K I E K TGCAGCTCAA A A Q GCAACATCTG A T S	297 1470 314/5 1540 28 1610 51
274 1401 298/3 1471 6 1541 29 1611 52	T L F P TAGATATAAA L R Y K ATTAGGCAAA <u>L G K</u> AATATCACAC N I T ATAATGATAA D N D K	R T G GTTAATTGGA V N W TCATCAGTTG S S V CTAAAAGAGA P K R E ATTGAATATT L N I	$\begin{matrix} I & Y & A \\ AAACACACAGA \\ K & T & H & E \\ CTICATCAAT \\ A & S & S & I \\ GAAAAAAGTA \\ K & K & V \\ TTTCAAATTT \\ F & Q & I \end{matrix}$	E R K H AATTAAAGTG I K V TGCACTGCTT A L L GATGACAAAA D D K TAACGTTTAA L T F N	N A F AAAGGACATA K G H TTGCTATCGA L L S TCACTTTATA I T L Y TTTCATTAAG F I K	V N R ATTAATATGA N stop <u>M</u> ATACAGTTGA N T V D CAAAACAACA K T T GATAAAAGTT D K S	N F V V AAATTGAAAA K I E K TGCAGCTCAA A A Q GCAACATCTG A T S ATGACAAAGA	297 1470 314/5 1540 28 1610 51 1680 75
274 1401 298/3 1471 6 1541 29 1611 52	T L F P TAGATATAAA L R Y K ATTAGGCAAA <u>L G K</u> AATATCACAC N I T ATAATGATAA D N D K	R T G GTTAATTGGA V N W TCATCAGTTG S S V CTAAAAGAGA P K R E ATTGAATATT L N I	$\begin{matrix} I & Y & A \\ AAACACACAGA \\ K & T & H & E \\ CTICATCAAT \\ A & S & S & I \\ GAAAAAAGTA \\ K & K & V \\ TTTCAAATTT \\ F & Q & I \end{matrix}$	E R K H AATTAAAGTG I K V TGCACTGCTT A L L GATGACAAAA D D K TAACGTTTAA L T F N	N A F AAAGGACATA K G H TTGCTATCGA L L S TCACTTTATA I T L Y TTTCATTAAG F I K	V N R ATTAATATGA N stop <u>M</u> ATACAGTTGA N T V D CAAAACAACA K T T GATAAAAGTT D K S	N F V V AAATTGAAAA K I E K TGCAGCTCAA GCAACATCTG A T S ATGACAAAGA Y D K D	297 1470 314/5 1540 28 1610 51 1680 75
274 1401 298/: 1471 6 1541 29 1611 52 1681 76	T L F P TAGATATAAA R Y K ATTAGGCAAA <u>L G K</u> AATATCACAC N I T ATAATGATAA D N D K TACGTTAGTA T L V	R T G GTTAATTGGA V N W TCATCAGTG S V CTAAAAGAGA P K R E ATTGAATATT L N I CTTAAGGCAG L K A	$\begin{bmatrix} I & Y & A \\ AAACACACGA \\ K & T & H & E \\ CTTCATCAAT \\ A & S & S & I \\ GAAAAAAGTA \\ K & K & V \\ TTTCAAATTT \\ F & Q & I \\ CCGGAAACAT \\ A & G & N & I \\ \end{bmatrix}$	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	N A F AAAGGACATA K G H TTGCTATCGA L L S TCACTTTATA I T L Y TTTCATTAAG F I K TATAAAAATT Y K N	V N R ATTAATATGA N stop M ATACAGTTGA N T V D CAAAACAACA K T T GATAAAAGTT D K S CTAATCCAAA F N P K	$\begin{array}{c cccc} N & F & V & V \\ AAATTGAAAA \\ K & I & E & K \\ TGCAGCTCAA \\ A & A & Q \\ GCAACATCTG \\ A & T & S \\ ATGACAAAGA \\ Y & D & K & D \\ AGATTACAAT \\ D & Y & N \end{array}$	297 1470 314/5 1540 28 1610 51 1680 75 1750 98
274 1401 298/: 1471 6 1541 29 1611 52 1681 76	T L F P TAGATATAAA R Y K ATTAGGCAAA <u>L G K</u> AATATCACAC N I T ATAATGATAA D N D K TACGTTAGTA T L V	R T G GTTAATTGGA V N W TCATCAGTG S V CTAAAAGAGA P K R E ATTGAATATT L N I CTTAAGGCAG L K A	$\begin{bmatrix} I & Y & A \\ AAACACACGA \\ K & T & H & E \\ CTTCATCAAT \\ A & S & S & I \\ GAAAAAAGTA \\ K & K & V \\ TTTCAAATTT \\ F & Q & I \\ CCGGAAACAT \\ A & G & N & I \\ \end{bmatrix}$	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	N A F AAAGGACATA K G H TTGCTATCGA L L S TCACTTTATA I T L Y TTTCATTAAG F I K TATAAAAATT Y K N	V N R ATTAATATGA N stop M ATACAGTTGA N T V D CAAAACAACA K T T GATAAAAGTT D K S CTAATCCAAA F N P K	N F V V AAATTGAAAA <u>K I E K</u> TGCAGCTCAA <u>Q</u> GCAACATCTG A T S ATGACAAAGA Y D K D AGATTACAAT	297 1470 314/5 1540 28 1610 51 1680 75 1750 98
274 1401 298/3 1471 6 1541 52 1611 52 1681 76 1751 99	$ \begin{array}{ccccc} T & L & F & P \\ TAGATATAAA \\ R & Y & K \\ ATTAGGCAAA \\ \underline{L} & \underline{G} & K \\ AATATCACAC \\ N & I & T \\ ATAATGATAA \\ D & N & D & K \\ TACGTTAGTA \\ T & L & V \\ TACTCACAGT \\ Y & S & Q \\ \end{array} $	R T G GTTAATTGGA V N W TCATCAGTTG S V CTAAAAGAGA P K R E ATTGAATATT L N I CTTAAGGCAG L K A TTTAATGGGG F Y W G	$\begin{matrix} I & Y & A \\ AAACACACGA \\ K & T & H & E \\ CTTCATCAAT \\ A & S & S & I \\ GAAAAAAGTA \\ K & K & V \\ TTCAAATTT \\ F & Q & I \\ CCGGAAACAT \\ A & G & N & I \\ CGGTAAGTAT \\ G & K & Y \end{matrix}$	$\begin{array}{cccc} & R & K & H \\ \text{AATTAAAGTG} \\ & I & K & V \\ \text{TGCACTGCTT} \\ & L & L \\ \text{GATGACAAAA} \\ & D & D & K \\ \text{D} & D & K \\ \text{TAACGTTTAA} \\ L & T & F & N \\ \text{TAACGTTTAA} \\ L & T & F & N \\ \text{TAACTCAGGT} \\ & N & S & G \\ \text{AATGTTTCGG} \\ & N & V & S \end{array}$	N A F AAAGGACATA K G H TTGCTATCGA L L S TCACTTTATA I T L Y TTTCATTAAG F I K TATAAAAATT Y K N TTAGTTCAGA V S S E	V N R ATTAATATGA N StOP M ATACAGTGA N T V D CAAAACAACA K T T GATAAAAGTT D K S CTAATCCAAA F N P K ATCAAATGAT S N D	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121
274 1401 298/3 1471 6 1541 29 1611 52 1681 76 1751 99 1821	$ \begin{array}{cccc} T & L & F & P \\ TAGATATAAA \\ R & Y & K \\ ATTAGGCAAA \\ L & G & K \\ AATATCACACA \\ N & I & T \\ ATAATGATAA \\ D & N & D & K \\ TACGTTAGTA \\ T & L & V \\ TACTCACAGT \\ Y & S & Q \\ TTGTTGACTA \\ \end{array} $	R T G GTTAATTGGA V N W TCATCAGTGG S S V CTAAAAGAGA P K R E ATTGAATATT L N I CTTAAGGCAG L K A TTTAATGGGG F Y W G TGCACCTAAA	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V TTTCAAATTT F Q I CCGGAAACAT A G N I CGGTAAGTAT G K Y AATCAAAATG	$\begin{array}{cccc} & R & K & H \\ \text{AATTAAAGTG} \\ & I & K & V \\ \text{TGCACTGCTT} \\ & A & L & L \\ \text{GATGACAAAA} \\ & D & D & K \\ \text{TAACGTTTAA} \\ & L & T & F & N \\ \text{TAACTTCAGGT} \\ & N & S & G \\ \text{AATGTTCGG} \\ & N & V & S \\ \text{AAGAATTCCA} \end{array}$	N A F <u>AAAGGA</u> CATA K G H TTGCTATCGA L L S TCACTTTATA I T L Y TTTCATTAAG F I K TATAAAAATT Y K N TTAGTTCAGA V S S E AGTTCAACAA	V N R ATTAATATGA N Stop <u>M</u> ATACAGTTGA N <u>T</u> V <u>D</u> CAAAACAACA K T <u>T</u> GATAAAAGTT D K S CTAATCCAAA F N P K ATCAAATGAT S N D ACATTAGGTT	N F V V AAATTGAAAA K I E K TGCAGCTCAA A Q GCAACATCTG A T S ATGACAAAGA Y D K D AGATTACAAT D Y N GCTGTAAATG A V N ATTCTTATGG	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890
274 1401 298/3 1471 6 1541 52 1681 76 1751 99 1821 122	$\begin{array}{cccccc} T & L & F & P \\ TAGATATAAA \\ L & R & Y & K \\ ATTAGGCAAA \\ L & G & K \\ AATATCACAC \\ N & I & T \\ ATAATGATAA \\ D & N & D & K \\ TACGTTAGTA \\ T & L & V \\ TACGTTAGTA \\ T & L & V \\ TACTCACAGT \\ Y & S & Q \\ TTGTTGACTA \\ V & V & D & Y \end{array}$	R T G GTTAATTGGA V N W TCATCAGTTG S S V CTAAAAGAGA P K R E ATTGAATATT L N I CTTAAGGCAG L K A TTTATTGGG F Y W G TGCACCTAAA A P K	$\begin{bmatrix} I & Y & A \\ AAACACACGA \\ K & T & H & E \\ CTTCATCAAT \\ A & S & S & I \\ GAAAAAAGTA \\ K & K & V \\ TTTCAAATT \\ F & Q & I \\ CCGGAAACAT \\ A & G & N & I \\ CGGTAAGTAT \\ G & K & Y \\ AATCAAAATG \\ N & H & N \\ \end{bmatrix}$	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	N A F AAAGGACATA K G H TTGCTATCGA L L S TCACTTTATA I T L Y TTTCATTAAG F I K TATAAAAATT Y K N TTAGTTCAGA V S S E AGTTCAACAA V Q Q	V N R ATTAATATGA N Stop M ATACAGTTGA N T V D CAAAACAACA K T T GATAAAAGTT D K S CTAATCCAAA F N P K ATCAAATGAT S N D ACATTAGGTT T L G	N F V V AAATTGAAAA K I E K TGCAGCTCAA A A Q GCAACATCTG A T S ATGACAAAGA Y D K D AGATTACAAT D Y N GCTGTAAATG A V N ATTCTTATGG Y S Y G	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890 145
274 1401 298/3 1471 6 1541 52 1681 76 1751 99 1821 122 1891	$\begin{array}{ccccc} T & L & F & P \\ TAGATATAAA \\ L & R & Y & K \\ ATTAGGCAAA \\ L & G & K \\ AATATCACAC \\ N & I & T \\ ATAATGATAA \\ D & N & D & K \\ TACGTTAGTA \\ T & L & V \\ TACGTAGTA \\ T & L & V \\ TACTCACAGT \\ Y & S & Q \\ TTGTTGACTA \\ V & V & D & Y \\ CGGAGATATT \end{array}$	R T G GTTAATTGGA V N W TCATCAGTTG S V CTAAAAGAGA P K R E ATTGATATT L N I CTTAAGGCAG L K A TTTATTGGGG F Y W G TGCACCTAAA A P K AATATATCTA	$\begin{bmatrix} Y & A \\ AAACACACGA \\ K & T & H & E \\ CTTCATCAAT \\ A & S & S & I \\ GAAAAAAGTA \\ K & K & V \\ TTTCAAATT \\ F & Q & I \\ CCGGAAACAT \\ A & G & N & I \\ CGGTAAGTAT \\ G & K & Y \\ AATCAAAATG \\ N & H & N \\ ATGGCTTATC \\ \end{bmatrix}$	$\begin{array}{ccccc} E & R & K & H \\ \text{AATTAAAGTG} \\ I & K & V \\ \text{TGCACTGCTT} \\ \hline A & L & L \\ \text{GATGACAAAA} \\ D & D & K \\ \text{TAACGTTTAA} \\ L & T & F & N \\ \text{TAATCAGGT} \\ N & S & G \\ \text{AAGGTTTCGG} \\ N & V & S \\ \text{AAGAATTCCA} \\ \text{CGGTGGATTA} \end{array}$	N A F AAAGGACATA K G H TTGCTATCGA L L S TCACTTTATA I T L Y TTTCATTAGA F I K TATAAAAATT Y K N TTAGTTCAGA V S S E AGTTCAACAA V Q Q AATGGATCAA	V N R ATTAATATGA N stop M ATACAGTTGA N T V D CAAAACAACA K T T GATAAAAGTT D K S CTAATCCAAA F N P K ATCAAATGAT S N D ACATTAGGTT T L G AAATCATTTC	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890 145 1960
274 1401 298/3 1471 6 1541 52 1681 76 1751 99 1821 122 1891 146	$ \begin{array}{ccccc} T & L & F & P \\ TAGATATAAA \\ R & Y & K \\ ATTAGGCAAA \\ L & G & K \\ AATATCACAC \\ N & I & T \\ ATAATGATAA \\ D & N & D & K \\ TACGTTAGTA \\ T & L & V \\ TACTCACAGT \\ Y & S & Q \\ TTGTTGACTA \\ V & D & Y \\ CGGAGATATT \\ G & D & I \\ \end{array} $	$\begin{array}{cccc} R & T & G \\ GTTAATTGGA \\ V & N & W \\ TCATCAGTTG \\ S & V \\ CTAAAAGAGA \\ P & K & R & E \\ ATTGAATATT \\ L & N & I \\ CTTAAGGCAG \\ L & K & A \\ TTTATGGGG \\ F & Y & W & G \\ TGCACCTAAA \\ A & P & K \\ AATATATCTA \\ N & I & S \end{array}$	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V TTTCAAATTT F Q I CCGGAAACAT A G N I CGGTAAGTAT G K Y AATCAAAATG N H N ATGGCTTATC N G L S	$\begin{array}{ccccc} E & R & K & H \\ AATTAAAGTG \\ I & K & V \\ TGCACTGCTT \\ A & L & L \\ GATGACAAAA \\ D & D & K \\ TAACGTTAA \\ L & T & F & N \\ TAATTCAGGT \\ N & S & G \\ AATGTTCCGG \\ N & V & S \\ AAGAATTCCA \\ E & E & F & Q \\ AGGTGGATTA \\ G & G & L \\ \end{array}$	N A F AAAGGACATA K G H TTGCTATCGA L L S TCACTTTATA I T L Y TTTCATTAAG F I K TATAAAAATT Y K N TTAGTTCAGA V S S E AGTTCAACAA V Q Q AATGGATCAA N G S	$ \begin{array}{c} V & N & R \\ \text{ATTAATATGA} \\ \text{N stop} \\ \text{ATACAGTGA} \\ \text{N } T & V \\ \text{D} \\ \text{CAAAACAACA} \\ \text{K} & T \\ \text{GATAAAGTT} \\ \text{D} & \text{K} \\ \text{S} \\ \text{CTAATCCAAA} \\ \text{F} & N \\ \text{P} \\ \text{K} \\ \text{ATCAATGAT} \\ \text{S} \\ \text{N} \\ \text{D} \\ \text{ACATTAGGTT} \\ \text{T} \\ \text{L} \\ \text{G} \\ \text{AATCATTTCC} \\ \text{K} \\ \text{S} \\ \text{F} \\ \text{S} $	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890 145 1960 168
274 1401 298/3 1471 6 1541 52 1681 76 1751 99 1821 122 1891 146 1961	$\begin{array}{ccccc} T & L & F & P \\ TAGATATAAA \\ R & Y & K \\ ATTAGGCAAA \\ L & G & K \\ AATATCACAC \\ N & I & T \\ ATAATGATAA \\ D & N & D & K \\ TACGTTAGTA \\ T & L & V \\ TACTCACAGT \\ Y & S & Q \\ TTGTTGACTA \\ V & V & D & Y \\ CGGAGATATT \\ G & D & I \\ AATTATAAAC \\ \end{array}$	R T G GTTAATTGGA V N W TCATCAGTGG S S V CTAAAAGAGA P K R E ATTGAATATT L N I CTTAAGGCAG L K A TTTATTGGGG F Y W G TGCACCTAAA A P K AATATATCTA N I S AAGAAAGTTA	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V TTTCAAATTT F Q I CCGGAAACAT A G N I CGGTAAGTAT G K Y AATCAAAATG N H N ATGGCTTATC N G L S CAGAACTACG	$\begin{array}{cccc} & R & K & H \\ \text{AATTAAAGTG} \\ & I & K & V \\ & TGCACTGCTT \\ & A & L & L \\ & GATGACAAAA \\ & D & D & K \\ & TAACGTTTAA \\ & L & T & F & N \\ & TAACTCAGGT \\ & N & S & G \\ & AATGTTCCAG \\ & N & V & S \\ & AAGAATTCCA \\ & E & F & Q \\ & AGGTGGATTA \\ & G & G \\ & ATTGATAGAA \\ \end{array}$	$\begin{array}{c} N  A  F \\ \underline{AAAGGA}CATA \\ K  G  H \\ TTGCTATCGA \\ \underline{L}  \underline{L}  S \\ TCACTTTATA \\ I  T  L  Y \\ TTTCATTAAG \\ F  I  K \\ TATAAAAATT \\ Y  K  N \\ TTAGTTCAGA \\ V  S  S  E \\ AGTTCAACAA \\ V  Q  Q \\ AATGGATCAA \\ N  G  S \\ AAACAAATCA \\ \end{array}$	$\begin{array}{c} V & N & R \\ ATTAATATGA \\ N Stop M \\ ATACAGTTGA \\ N & T & V \\ D \\ CAAAACAACA \\ K & T & T \\ GATAAAAGTT \\ D & K \\ S \\ CTAATCCAAA \\ F & N & P \\ K \\ ATCAAATGAT \\ S & N \\ D \\ ACATTAGGTT \\ T & L \\ G \\ AATCATTTCC \\ K & S \\ F & S \\ TAAATCAATT \\ \end{array}$	$\begin{array}{c cccc} N & F & V & V \\ \text{AAATTGAAAA} \\ K & I & E & K \\ TGCAGCTCAA \\ \hline \\ A & T & S \\ \text{ATGACAAAGA} \\ Y & D & K & D \\ \text{AGATTACAAT} \\ D & Y & N \\ \text{GCTGTAAATG} \\ A & V & N \\ \text{ATTCTTATGG} \\ Y & S & Y & G \\ \text{AGAAACGATA} \\ E & T & I \\ \text{GGCTGGGGTG} \end{array}$	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890 145 1960 168 2030
274 1401 298/3 1471 6 1541 52 1681 76 1751 99 1821 122 1891 146 1961 169	$\begin{array}{ccccc} T & L & F & P \\ TAGATATAAA \\ L & R & Y & K \\ ATTAGGCAAA \\ L & G & K \\ AATATGACACA \\ N & I & T \\ ATAATGATAA \\ D & N & D & K \\ TACGTTAGTA \\ T & L & V \\ TACGTTAGTA \\ T & L & V \\ TACGTTAGACTA \\ V & V & D & Y \\ CGAGATATT \\ G & D & I \\ AATTATAAAC \\ N & Y & K \end{array}$	R T G GTTAATTGGA V N W TCATCAGTG S S V CTAAAAGAGA P K R E ATTGAATATT L N I CTTAAGGCAG L K A TTTATTGGG F Y W G TGCACCTAAA A P K AATATATCTA N I S AAGAAAGTTA Q E S Y	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V TTTCAAATT F Q I CCGGAAACAT A G N I CGGTAAGTAT G K Y AATCAAAATG N H N ATGGCTTATC N G L S CAGAACTACG R T T	$\begin{array}{ccccc} E & R & K & H \\ \text{AATTAAAGTG} \\ I & K & V \\ \text{TGCACTGCTT} \\ \hline A & L & L \\ \text{GATGACAAAA} \\ D & D & K \\ \text{TAACGTTTAA} \\ L & T & F & N \\ \text{TAACTCAGGT} \\ N & S & G \\ \text{AATGATTCAGGT} \\ N & S & G \\ \text{AATGATTCCA} \\ E & E & F & Q \\ \text{AGGTGGATTA} \\ G & G & L \\ \text{ATTGATAGAA} \\ I & D & R \end{array}$	$\begin{array}{cccc} N & A & F \\ \underline{AAAGGACATA} \\ K & G & H \\ TTGCTATCGA \\ \underline{L} & \underline{L} & \underline{S} \\ TCACTTTATA \\ I & T & L & Y \\ TTTCATTAAG \\ F & I & K \\ TATAAAAATT \\ Y & K & N \\ TTAGTTCAACA \\ V & S & E \\ AGTTCAACAA \\ V & Q & Q \\ AATGGATCAA \\ N & G & S \\ AAACAAATCA \\ K & T & N & H \end{array}$	$ \begin{array}{c} V & N & R \\ \text{ATTAATATGA} \\ \text{ATTAATATGA} \\ \text{N Stop M} \\ \text{ATACAGTGA} \\ \text{N T V D} \\ \text{CAAAACAACA} \\ K & T T \\ \text{GATAAAAGT } \\ \text{CTAATCAAATGAT } \\ \text{S N D} \\ \text{ACATTAGGTT } \\ \text{T L G} \\ \text{AATCATTTAC } \\ \text{K S F S} \\ \text{TAAATCAATT } \\ \text{K S I } \\ \end{array} $	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890 145 1960 168 2030 191
274 1401 298/3 1471 6 1541 52 1681 76 1751 99 1821 122 1891 146 1961 169 2031	T L F P TAGATATAAA R Y K ATTAGGCAAA L G K AATATCACAC N I T ATAATGATAA D N D K TACGTTAGTA T L V TACTCACAGT Y S Q TTGTTGACTA V V D Y CGGAGATATT G D I AATTATAAAC N Y K TTGAGGCCCA	$\begin{array}{c c} R & T & G \\ \text{GTTAATTGGA} \\ V & N & W \\ \text{TCATCAGTTG} \\ S & S & V \\ \text{CTAAAAGAGA} \\ P & K & R & E \\ \text{ATTGAATATT} \\ L & N & I \\ \text{CTTAAGGCAG} \\ L & K & A \\ \text{TTTATTGGG} \\ F & Y & W & G \\ \text{TGCACCTAAA} \\ A & P & K \\ \text{AATATATCTA} \\ \text{N } I & S \\ \text{AAGAAAGTTA} \\ Q & E & S & Y \\ \text{CAAAATTATG} \end{array}$	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V TTTCAAATT F Q I CCGGAAACAT A G N I CCGGAAACAT A G N I CGGTAAGTAT G K Y AATCAAAATG N H N ATGGCTTATC N G L S CAGAACTACG R T T AATAATGGTT	E R K H AATTAAAGT <u>G</u> I K V TGCACTGCTT <u>A L L</u> GATGACAAAA D D K TAACGTTTAA L T F N TAATTCAGGT N S G AATGTTTCAG N V S AAGAATTCCA E E F Q AGGTGGATTA G G L ATGATAGAA I D R GGGGACCATA	$\begin{array}{c} \mathrm{N}  \mathrm{A}  \mathrm{F} \\ \underline{\mathrm{AAAGGA}} \mathrm{CATA} \\ \mathrm{K}  \mathrm{G}  \mathrm{H} \\ \mathrm{TTGCTATCGA} \\ \mathrm{L}  \mathrm{L}  \mathrm{S} \\ \mathrm{TCACTTTATA} \\ \mathrm{I}  \mathrm{T}  \mathrm{L}  \mathrm{Y} \\ \mathrm{TTCATTAAG} \\ \mathrm{F}  \mathrm{I}  \mathrm{K} \\ \mathrm{TATAAAAATT} \\ \mathrm{Y}  \mathrm{K}  \mathrm{N} \\ \mathrm{TTAGTTCAGA} \\ \mathrm{V}  \mathrm{S}  \mathrm{S}  \mathrm{E} \\ \mathrm{AGTTCAACAA} \\ \mathrm{V}  \mathrm{Q}  \mathrm{Q} \\ \mathrm{AATGGATCAA} \\ \mathrm{N}  \mathrm{G}  \mathrm{S} \\ \mathrm{AAACAAATCA} \\ \mathrm{K}  \mathrm{T}  \mathrm{N}  \mathrm{H} \\ \mathrm{TGGTAGAGAT} \end{array}$	$\begin{array}{c c} V & N & R \\ ATTAATATGA \\ N Stop M \\ ATACAGTTGA \\ N & T & V & D \\ CAAAACAACA \\ K & T & T \\ GATAAAAGTT \\ D & K & S \\ CTAATCCAAA \\ F & N & P & K \\ ATCAAATGAT \\ S & N & D \\ ACATTAGGTT \\ T & L & G \\ AATCATTTC \\ K & S & F & S \\ TAAATCAATT \\ K & S & I \\ AGTTATGACC \\ \end{array}$	$\begin{array}{c cccc} N & F & V & V \\ \text{AAATTGAAAA} \\ K & I & E & K \\ \text{TGCAGCTCAA} \\ \hline \\ A & T & S \\ \text{ATGCAACATCTG} \\ A & T & S \\ \text{ATGCAAAAGA} \\ Y & D & K & D \\ \text{AGATTACAAT} \\ D & Y & N \\ \text{GCTGTAAATG} \\ A & V & N \\ \text{ATTCTTATGG} \\ Y & S & Y & G \\ \text{AGAAACGATA} \\ E & T & I \\ \text{GGCTGGGGGATA} \\ G & W & G \\ \text{CAACATATGG} \end{array}$	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890 145 1960 168 2030 191
274 1401 298/: 1471 6 1541 29 1611 52 1681 76 1751 99 1821 122 1891 1466 1961 1699 2031 192	$\begin{array}{cccccc} T & L & F & P \\ TAGATATAAA \\ R & Y & K \\ ATTAGGCAAA \\ L & G & K \\ AATATCACAC \\ N & I & T \\ ATAATGATAA \\ D & N & D & K \\ TACGTTAGTA \\ T & L & V \\ TACTCACAGT \\ Y & S & Q \\ TTGTTGACTA \\ V & D & Y \\ CGGAGATATT \\ G & D & I \\ AATTATAAAC \\ N & Y & K \\ TGGAGGCCCA \\ V & E & A \\ H \end{array}$	$\begin{array}{cccc} R & T & G \\ GTTAATTGGA \\ V & N & W \\ TCATCAGTTG \\ S & V \\ CTAAAAGAGA \\ P & K & R & E \\ ATTGAATATT \\ L & N & I \\ CTTAAGGCAG \\ L & K & A \\ TTTATTGGGG \\ F & Y & W & G \\ TGCACCTAAA \\ A & P & K \\ AATATATCTA \\ N & I & S \\ AAGAAAGTTA \\ Q & E & S & Y \\ CAAAATTATG \\ K & I & M \end{array}$	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V TTTCAAATTT F Q I CCGGAAACAT A G N I CGGTAAGATAT G K Y AATCAAAATG N H N ATGGCTTATC N G L S CAGAACTACG R T T AATAATGGTT N N G	$\begin{array}{ccccc} & R & K & H \\ \text{AATTAAAGTG} \\ & I & K & V \\ & TGCACTGCTT \\ & L & L \\ & GATGACAAAA \\ & D & D & K \\ & D & D & K \\ & TAACGTTTAA \\ & L & T & F & N \\ & TAACGTTTAA \\ & L & T & F & N \\ & TAACGTTTCAGT \\ & N & S & G \\ & AATGTTCAGT \\ & N & V & S \\ & AAGAATTCCA \\ & E & F & Q \\ & AGGTGGATTA \\ & G & G & L \\ & ATTGATAGAA \\ & I & D & R \\ & GGGGACCATA \\ & W & G & P & Y \end{array}$	$\begin{array}{c} N  A  F \\ \underline{AAAGGA}CATA \\ K  G  H \\ TTGCTATCGA \\ L  L  S \\ TCACTTTATA \\ I  T  L  Y \\ TTCATTAAG \\ F  I  K \\ TATAAAAATT \\ Y  K  N \\ TAGTTCAGA \\ V  S  S  E \\ AGTTCAACAA \\ V  Q  Q \\ AATGGATCAA \\ N  G  S \\ AAACAAATCA \\ K  T  N  H \\ TGGTAGAGAT \\ G  R  D \end{array}$	$\begin{array}{c} V & N & R \\ ATTAATATGA \\ N Stop M \\ ATACAGTTGA \\ N & T & V \\ D \\ CAAAACAACA \\ K & T & T \\ GATAAAGATT \\ D & K & S \\ CTAATCCAAA \\ F & N & P & K \\ ATCAAATGAT \\ S & N & D \\ ACATTAGGAT \\ T & L & G \\ AATCAATTTC \\ K & S & F & S \\ TAAATCAATT \\ K & S & F & S \\ TAAATCAATT \\ K & S & T \\ AGTTATGACC \\ S & Y & D \end{array}$	$\begin{array}{c cccc} N & F & V & V \\ \text{AAATTGAAAA} \\ K & I & E & K \\ TGCAGCTCAA \\ A & Q \\ GCAACATCTG \\ A & T & S \\ ATGCACAAGA \\ Y & D & K & D \\ AGATTACAAT \\ D & Y & N \\ GCTGTAAATG \\ A & V & N \\ ATTCTTATGG \\ Y & S & Y & G \\ AGAACAGATA \\ E & T & I \\ GGCTGGGGTG \\ G & W & G \\ CAACATATGG \\ P & T & Y & G \end{array}$	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890 145 1960 168 2030 191 2100 215
274 1401 298/: 1471 6 1541 29 1611 52 1681 76 1751 99 1821 122 1891 1466 1961 1699 2031 192	$ \begin{array}{ccccc} T & L & F & P \\ TAGATATAAA \\ R & Y & K \\ ATTAGGCAAA \\ L & G & K \\ AATATCACAC \\ N & I & T \\ ATAATGATAA \\ D & N & D & K \\ TACGTTAGTA \\ T & L & V \\ TACTCACAGT \\ Y & S & Q \\ TTGTTGACTA \\ V & D & Y \\ CGGAGATATT \\ G & D & I \\ AATTATAAAC \\ N & Y & K \\ TGGAGGCCCA \\ V & E & A \\ H \end{array} $	$\begin{array}{cccc} R & T & G \\ GTTAATTGGA \\ V & N & W \\ TCATCAGTTG \\ S & V \\ CTAAAAGAGA \\ P & K & R & E \\ ATTGAATATT \\ L & N & I \\ CTTAAGGCAG \\ L & K & A \\ TTTATTGGGG \\ F & Y & W & G \\ TGCACCTAAA \\ A & P & K \\ AATATATCTA \\ N & I & S \\ AAGAAAGTTA \\ Q & E & S & Y \\ CAAAATTATG \\ K & I & M \end{array}$	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V TTTCAAATTT F Q I CCGGAAACAT A G N I CGGTAAGATAT G K Y AATCAAAATG N H N ATGGCTTATC N G L S CAGAACTACG R T T AATAATGGTT N N G	$\begin{array}{ccccc} & R & K & H \\ \text{AATTAAAGTG} \\ & I & K & V \\ & TGCACTGCTT \\ & L & L \\ & GATGACAAAA \\ & D & D & K \\ & D & D & K \\ & TAACGTTTAA \\ & L & T & F & N \\ & TAACGTTTAA \\ & L & T & F & N \\ & TAACGTTTCAGT \\ & N & S & G \\ & AATGTTCAGT \\ & N & V & S \\ & AAGAATTCCA \\ & E & F & Q \\ & AGGTGGATTA \\ & G & G & L \\ & ATTGATAGAA \\ & I & D & R \\ & GGGGACCATA \\ & W & G & P & Y \end{array}$	$\begin{array}{c} N  A  F \\ \underline{AAAGGA}CATA \\ K  G  H \\ TTGCTATCGA \\ L  L  S \\ TCACTTTATA \\ I  T  L  Y \\ TTCATTAAG \\ F  I  K \\ TATAAAAATT \\ Y  K  N \\ TAGTTCAGA \\ V  S  S  E \\ AGTTCAACAA \\ V  Q  Q \\ AATGGATCAA \\ N  G  S \\ AAACAAATCA \\ K  T  N  H \\ TGGTAGAGAT \\ G  R  D \end{array}$	$\begin{array}{c} V & N & R \\ ATTAATATGA \\ N Stop M \\ ATACAGTTGA \\ N & T & V \\ D \\ CAAAACAACA \\ K & T & T \\ GATAAAGATT \\ D & K \\ S \\ CTAATCCAAA \\ F & N & P \\ K \\ ATCAAATGAT \\ S & N \\ D \\ ACATTAGGAT \\ T & L \\ G \\ ATCAATTATC \\ K \\ S & F \\ S \\ TAAATCAATT \\ K \\ S & F \\ S \\ TAAATCAATT \\ K \\ S & T \\ AGTTATGACC \\ S & Y \\ D \end{array}$	$\begin{array}{c cccc} N & F & V & V \\ \text{AAATTGAAAA} \\ K & I & E & K \\ \text{TGCAGCTCAA} \\ \hline \\ A & T & S \\ \text{ATGCAACATCTG} \\ A & T & S \\ \text{ATGCAAAAGA} \\ Y & D & K & D \\ \text{AGATTACAAT} \\ D & Y & N \\ \text{GCTGTAAATG} \\ A & V & N \\ \text{ATTCTTATGG} \\ Y & S & Y & G \\ \text{AGAAACGATA} \\ E & T & I \\ \text{GGCTGGGGTGG} \\ G & W & G \\ \text{CAACATATGG} \end{array}$	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890 145 1960 168 2030 191 2100 215
274 1401 298/3 1471 6 1541 52 1681 76 1751 99 1821 122 1891 1466 1961 2031 192 2101 216	$\begin{array}{ccccc} T & L & F & P \\ TAGATATAAA \\ L & R & Y & K \\ ATTAGGCAAA \\ L & G & K \\ AATATCACAC \\ N & I & T \\ ATAATGATAA \\ D & N & D & K \\ TACGTTAGTA \\ T & L & V \\ TACGTTAGTA \\ T & L & V \\ TACGTTAGTA \\ Y & S & Q \\ TTGTTGACTA \\ V & V & D & Y \\ CGGAGATATT \\ G & D & I \\ AATTATAAAC \\ N & Y & K \\ TTGAGGCCCA \\ V & E & A \\ TAATGAACTG \\ N & E & L \\ \end{array}$	R T G GTTAATTGGA V N W TCATCAGTG S S V CTAAAAGAGA P K R E ATTGAATATT L N I CTTAAGGCAG L K A TTTATTGGG F Y W G TGCACCTAAA A P K AATATATCTA A A P K AATATATCTAA N I S AAGAAAGTTA Q E S Y CAAAATTATG K I M TTTTTAGGCG F L G	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V TTTCAAATT F Q I CCGGAAACAT A G N I CGGTAAGTAT G K Y AATCAAAATG N H N ATGGCTTATC N G L S CAGAACTACG R T T AATAATGGTT N N G GTGACAAAAG G D K S	$\begin{array}{ccccc} E & R & K & H \\ \text{AATTAAAGTG} \\ I & K & V \\ \text{TGCACTGCTT} \\ \hline A & L & L \\ \text{GATGACAAAA} \\ D & D & K \\ \text{TAACGTTAA} \\ L & T & F & N \\ \text{TAACTCAGGT} \\ N & S & G \\ \text{AATGATTCAGGT} \\ N & S & G \\ \text{AATGATTCCA} \\ E & E & F & Q \\ \text{AGGGTGGATTA} \\ G & G & L \\ \text{ATTGATAGAA} \\ I & D & R \\ \text{GGGGACCATA} \\ W & G & P & Y \\ \text{TAGTTCAAGAT} \\ S & N \end{array}$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$ \begin{array}{c} V & N & R \\ \text{ATTAATATGA} \\ \text{ATTAATATGA} \\ \text{N Stop M} \\ \text{ATACAGTTGA} \\ \text{N T V D} \\ \text{CAAAACAACA} \\ K & T & T \\ \text{GATAAAAGTT} \\ D & K & S \\ \text{CTAATCCAAA} \\ F & N & P & K \\ \text{ATCAAATGAT} \\ S & N & D \\ \text{ACATTAGGTT} \\ T & L & G \\ \text{AATCATTTACK } S & F & S \\ \text{TAAATCAATT} \\ K & S & I \\ \text{AGTTATGACC} \\ S & Y & D \\ \text{ATTTCTTGCC} \\ \text{N F L P } \end{array} $	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890 145 1960 168 2030 191 2100 215 2170 238
274 1401 298/3 1471 6 1541 52 1681 76 1751 99 1821 122 1891 1466 1961 2031 192 2101 216	$\begin{array}{ccccc} T & L & F & P \\ TAGATATAAA \\ L & R & Y & K \\ ATTAGGCAAA \\ L & G & K \\ AATATCACAC \\ N & I & T \\ ATAATGATAA \\ D & N & D & K \\ TACGTTAGTA \\ T & L & V \\ TACGTTAGTA \\ T & L & V \\ TACGTTAGTA \\ Y & S & Q \\ TTGTTGACTA \\ V & V & D & Y \\ CGGAGATATT \\ G & D & I \\ AATTATAAAC \\ N & Y & K \\ TTGAGGCCCA \\ V & E & A \\ TAATGAACTG \\ N & E & L \\ \end{array}$	R T G GTTAATTGGA V N W TCATCAGTG S S V CTAAAAGAGA P K R E ATTGAATATT L N I CTTAAGGCAG L K A TTTATTGGG F Y W G TGCACCTAAA A P K AATATATCTA A A P K AATATATCTAA N I S AAGAAAGTTA Q E S Y CAAAATTATG K I M TTTTTAGGCG F L G	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V TTTCAAATT F Q I CCGGAAACAT A G N I CGGTAAGTAT G K Y AATCAAAATG N H N ATGGCTTATC N G L S CAGAACTACG R T T AATAATGGTT N N G GTGACAAAAG G D K S	$\begin{array}{ccccc} E & R & K & H \\ \text{AATTAAAGTG} \\ I & K & V \\ \text{TGCACTGCTT} \\ \hline A & L & L \\ \text{GATGACAAAA} \\ D & D & K \\ \text{TAACGTTAA} \\ L & T & F & N \\ \text{TAACTCAGGT} \\ N & S & G \\ \text{AATGATTCAGGT} \\ N & S & G \\ \text{AATGATTCCA} \\ E & E & F & Q \\ \text{AGGGTGGATTA} \\ G & G & L \\ \text{ATTGATAGAA} \\ I & D & R \\ \text{GGGGACCATA} \\ W & G & P & Y \\ \text{TAGTTCAAGAT} \\ S & N \end{array}$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$ \begin{array}{c} V & N & R \\ \text{ATTAATATGA} \\ \text{ATTAATATGA} \\ \text{N Stop M} \\ \text{ATACAGTTGA} \\ \text{N T V D} \\ \text{CAAAACAACA} \\ K & T & T \\ \text{GATAAAAGTT} \\ D & K & S \\ \text{CTAATCCAAA} \\ F & N & P & K \\ \text{ATCAAATGAT} \\ S & N & D \\ \text{ACATTAGGTT} \\ T & L & G \\ \text{AATCATTTACK } S & F & S \\ \text{TAAATCAATT} \\ K & S & I \\ \text{AGTTATGACC} \\ S & Y & D \\ \text{ATTTCTTGCC} \\ \text{N F L P } \end{array} $	$\begin{array}{c cccc} N & F & V & V \\ \text{AAATTGAAAA} \\ K & I & E & K \\ TGCAGCTCAA \\ A & Q \\ GCAACATCTG \\ A & T & S \\ ATGACAAAGA \\ Y & D & K & D \\ AGATTACAAT \\ D & Y & N \\ GCTGTAAATG \\ A & V & N \\ ATTCTTATGG \\ Y & S & Y & G \\ AGAAACGATA \\ E & T & I \\ GGCTGGGGTG \\ G & W & G \\ CAACATATGG \\ P & T & Y & G \\ AACACATCAA \\ \end{array}$	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890 145 1960 168 2030 191 2100 215 2170 238
$\begin{array}{c} 274\\ 1401\\ 1401/2\\ 1471\\ 6\\ 1541\\ 29\\ 1611\\ 52\\ 1681\\ 1751\\ 122\\ 1891\\ 142\\ 1891\\ 146\\ 1961\\ 169\\ 2031\\ 192\\ 2101\\ 2102\\ 2101\\ 2161\\ 2171\\ \end{array}$	T L F P TAGATATAAA L R Y K ATTAGGCAAA L G K AATATCACACA N I T ATAATGATAA D N D K TACGTTAGTA T L V TACGTTAGTA T L V TACTCACAGT Y S Q TTGTTGACTAA V V D Y CGGAGATATT G D I AATTATAAAC N Y K TTGAGGCCCA V E A H TAATGAACTG N E L ATCCCTTTAT	R T G GTTAATTGGA V N W TCATCAGTG S S V CTAAAAGAGA P K R E ATTGAATATT L N I CTTAAGGCAG L K A TTTATTGGG F Y W G TGCACCTAAA A P K AATATATCTA A A P K AATATATCTAA N I S AAGAAAGTTA Q E S Y CAAAATTATG K I M TTTTTAGGCG F L G	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V TTTCAAATT F Q I CCGGAAACAT A G N I CCGGAAACAT A G N I CCGGAAACAT G K Y AATCAAAATG N H N ATGGCTTATC N G L S CAGAACTACG R T T AATAATGGTT N N G GTGACAAAAG G D K S TAACTTTAAC	E R K H AATTAAAGT $\underline{G}$ I K V TGCACTGCTT $\underline{A}$ L L GATGACAAAA D D K TAACGTTTAA L T F N TAATTCAGGT N S G AATGTTTCAGGT N V S AAGAATTCCA E E F Q AAGGAGGACTAA G G L ATTGATAGAA I D R GGGGACCATA W G P Y TAGTTCAAAA S N CCAGAATTTA	N A F AAAGGACATA K G H TTGCTATCGA L L S TCACTTTATA I T L Y TTTCATTAAG F I K TATAAAAATT Y K N TTAGTTCAGA V S S E AGTCCAACAA V Q Q AATGGATCAA N G S AAACAAATCA K T N H TGGTAGAGAT G R D GCTGGTCAACA	$\begin{array}{c c} V & N & R \\ \text{ATTAATATGA} \\ \text{ATTAATATGA} \\ \text{N stop } \\ \text{ATACAGTTGA} \\ \text{N } T & V & D \\ \text{CAAAACAACA} \\ K & T & T \\ \text{GATAAAAGTT} \\ D & K & S \\ \text{CTAATCCAAA} \\ \text{F } N & P & K \\ \text{ATCAAATGAT} \\ \text{S } N & D \\ \text{ACATTAGGTT} \\ T & L & G \\ \text{AATCATTTTC} \\ K & S & F & S \\ \text{TAAATCAATT} \\ K & S & I \\ \text{AGTTATGACC} \\ \text{S } Y & D \\ \text{ATTCTTGGCC} \\ \text{N } F & L & P \\ \text{TTCTCATAAA} \end{array}$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890 145 1960 168 2030 191 2100 215 2170 238
$\begin{array}{c} 274\\ 1401\\ 298/3\\ 1471\\ 6\\ 1541\\ 299\\ 1611\\ 52\\ 1681\\ 766\\ 1751\\ 99\\ 1821\\ 142\\ 1961\\ 1496\\ 1961\\ 192\\ 2031\\ 192\\ 2031\\ 192\\ 2101\\ 2166\\ 2171\\ 239\\ 231\\ 192\\ 192\\ 192\\ 192\\ 192\\ 192\\ 192\\ 19$	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	$\begin{array}{c} {\rm R} {\rm T} {\rm G} {\rm GTTAATTGGA} \\ {\rm V} {\rm N} {\rm W} \\ {\rm TCATCAGTTG} \\ {\rm S} {\rm S} {\rm V} \\ {\rm CTAAAAGAGA} \\ {\rm P} {\rm K} {\rm R} {\rm E} \\ {\rm ATTGAATATT} \\ {\rm L} {\rm N} {\rm I} \\ {\rm CTTAAGGCAG} \\ {\rm L} {\rm K} {\rm A} \\ {\rm TTTATGGGG} \\ {\rm F} {\rm Y} {\rm W} {\rm G} \\ {\rm TGCACCTAAA} \\ {\rm A} {\rm P} {\rm K} \\ {\rm AATAATATCA} \\ {\rm N} {\rm I} {\rm S} \\ {\rm AAGAAAGTTA} \\ {\rm Q} {\rm E} {\rm S} {\rm Y} \\ {\rm CAAAATATAT} \\ {\rm CAAAATATATG} \\ {\rm K} {\rm I} {\rm M} \\ {\rm TTTTTAGGCG} \\ {\rm F} {\rm L} {\rm G} \\ {\rm TGGCCCTGG} \\ {\rm L} {\rm A} {\rm R} {\rm G} \end{array} \\ \end{array}$	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V TTCCAAATTT F Q I CCGGAAACAT A G N I CGGTAACATAT G K Y AATCAAAATG N H N ATGGCTTATC N G L S CAGAACTACG R T T AATAATGGTT N N G GTGACAAAAG G D K S TAACTTACAC N F N	$\begin{array}{cccc} & R & K & H \\ \text{AATTAAAGTG} \\ & I & K & V \\ & TGCACTGCTT \\ & L & L \\ & GATGACAAAA \\ & D & D & K \\ & TAACGTTAA \\ & L & T & F & N \\ & TAACGTTTAA \\ & L & T & F & N \\ & TAATTCAGGT \\ & N & S & G \\ & AATGTTCAGG \\ & N & V & S \\ & AAGAATTCCA \\ & E & F & Q \\ & AGGTGGATTA \\ & G & G & L \\ & ATTGATAGAA \\ & I & D & R \\ & GGGGACCATA \\ & W & G & P & Y \\ & TAGTTCAAAT \\ & S & S & N \\ & CCAGAATTA \\ & F & F \end{array}$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\begin{array}{c} V & N & R \\ ATTAATATGA \\ N StOP M \\ ATACAGTTGA \\ N & T & V \\ D \\ CAAAACAACA \\ K & T & T \\ GATAAAGATT \\ D & K \\ S \\ CTAATCCAAA \\ F & N & P \\ K \\ S \\ CTAATCCAAA \\ F & N \\ D \\ ACATTAGGAT \\ T & L \\ G \\ ATCAATGAT \\ T & L \\ S \\ T \\ A \\ CAATAGAT \\ T \\ K \\ S \\ F \\ S \\ TAAATCAATT \\ K \\ S \\ T \\ A \\ T \\ T \\ C \\ T \\ C \\ T \\ C \\ T \\ C \\ T \\ T$	$\begin{array}{c c} N & F & V & V \\ \text{AAATTGAAAA} \\ K & I & E & K \\ TGCAGCTCAA \\ \hline \\ A & T & S \\ \\ GCAACATCTG \\ A & T & S \\ \\ A & C & C \\ A & T & C \\ A & C & C \\ C & C & C \\ A & C & C \\ C & C & C \\ C & C & C \\ C & C &$	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890 145 1960 168 2030 191 2100 215 2170 238 2240 261
$\begin{array}{c} 274\\ 1401\\ 298/3\\ 1471\\ 6\\ 1541\\ 299\\ 1611\\ 52\\ 1681\\ 1751\\ 99\\ 1821\\ 122\\ 1891\\ 146\\ 1961\\ 169\\ 2031\\ 192\\ 2101\\ 2161\\ 2171\\ 239\\ 2241\\ \end{array}$	$\begin{array}{ccccc} T & L & F & P \\ TAGATATAAA \\ R & Y & K \\ ATTAGGCAAA \\ L & G & K \\ AATATCACACA \\ N & I & T \\ ATAATGATAA \\ D & N & D & K \\ TACGCTAGTA \\ T & L & V \\ TACTCACAGT \\ Y & S & Q \\ TTGTTGACTA \\ V & D & Y \\ CGGAGATATT \\ G & D & I \\ AATTATAAAC \\ N & Y & K \\ TTGAGGCCCA \\ V & E & A & H \\ TAATGAACTG \\ N & E & L \\ ATCCCTTTAT \\ I & P & L \\ \end{array}$	R T G GTTAATTGGA V N W TCATCAGTGG S S V CTAAAAGAGA P K R E ATTGAATATT L N I CTTAAGGCAG L K A TTTATTGGGG F Y W G TGCACCTAAA A P K AATATATCTA N I S AAGAAAGTTA Q E S Y CAAAATTATG K I M TTTTTTAGGCG F L G TGGCGCGTGG L A R G TAAAATCAAA	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V TTTCAAATTT F Q I CCGGAAACAT A G N I CGGTAAGTAT G K Y AATCAAAATG N H N ATGGCTTATC N G L S CAGAACTACG R T T AATAATGGTT N N G GTGACAAAAG G D K S TAACTTAAC N F N GTAACTTACC	E R K H AATTAAAGTG I K V TGCACTGCTT A L L GATGACAAAA D D K TAACGTTTAA L T F N TAATTCAGGT N S G AATGTTCAGGT N V S AAGAATTCCA E E F Q AGGTGGACTAA G G L ATTGATAGAA I D R GGGGACCATA W G P Y TAGTTCAAAT S S N CCAGAATTTA P E F AAAGAGAAAT	$\begin{array}{c} \mathrm{N}  \mathrm{A}  \mathrm{F} \\ \underline{\text{AAAGGACATA}} \\ \mathrm{K}  \mathrm{G}  \mathrm{H} \\ \mathrm{TTGCTATCGA} \\ \mathrm{L}  \mathrm{L}  \mathrm{S} \\ \mathrm{TCACTTTATA} \\ \mathrm{I}  \mathrm{T}  \mathrm{L}  \mathrm{Y} \\ \mathrm{TTCATTAAG} \\ \mathrm{F}  \mathrm{I}  \mathrm{K} \\ \mathrm{TATAAAAATT} \\ \mathrm{Y}  \mathrm{K}  \mathrm{N} \\ \mathrm{TTAGTTCAGA} \\ \mathrm{V}  \mathrm{S}  \mathrm{S}  \mathrm{E} \\ \mathrm{AGTTCAACAA} \\ \mathrm{V}  \mathrm{Q}  \mathrm{Q} \\ \mathrm{AATGGATCAA} \\ \mathrm{N}  \mathrm{G}  \mathrm{S} \\ \mathrm{AAACAAATCA} \\ \mathrm{K}  \mathrm{T}  \mathrm{N}  \mathrm{H} \\ \mathrm{TGGTAGAGAT} \\ \mathrm{G}  \mathrm{R}  \mathrm{D} \\ \mathrm{GCTGGTCAAA} \\ \mathrm{A}  \mathrm{G}  \mathrm{Q} \\ \mathrm{TAGCGATCA} \\ \mathrm{I}  \mathrm{S}  \mathrm{V}  \mathrm{L} \\ \mathrm{GGATAGATAT} \\ \mathrm{I}  \mathrm{S}  \mathrm{V}  \mathrm{L} \end{array}$	$\vee$ N R ATTAATATGA N Stop M ATACAGTTGA N T $\vee$ D CAAAACAACA K T T GATAAAAGTT D K S CTAATCCAAA F N P K ATCAAATGAT S N D ACATTAGGTT T L G AATCATTTCC K S F S TAAATCAATT K S I AGTTATGACC S Y D ATTTCTTGCC N F L P TTCTCATAAA S H K ACTAATCAAT	$\begin{array}{c c} N & F & V & V \\ \text{AAATTGAAAA} \\ K & I & E & K \\ TGCAGCTCAA \\ \hline \\ A & T & S \\ TGCAACATCTG \\ A & T & S \\ TGCAACAACA \\ A & V \\ ATGCAAAGA \\ Y & D & K & D \\ AGATTACAAT \\ D & Y & N \\ ATTCTTATGG \\ Y & S & Y & G \\ AGAAACGATA \\ E & T & I \\ GGCTGGGGTG \\ G & W & G \\ CAACATATGG \\ P & T & Y & G \\ AACACATATGG \\ P & T & Y & G \\ AACACATATGA \\ T & H & Q \\ CTTTTTGATA \\ L & F & D \\ GGAATCGATC \\ \end{array}$	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890 145 1960 168 2030 191 2100 215 2170 238 2240 261 2310
274 1401 298/3 1471 6 1541 299 1611 52 1681 766 1751 122 1891 1426 1991 2031 14961 1992 2031 2101 2162 2171 2399 2241 262	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	$\begin{array}{cccc} R & T & G \\ \text{GTTAATTGGA} \\ V & N & W \\ \text{TCATCAGTG} \\ \textbf{S} & \textbf{S} & V \\ \text{CTAAAAGAGA} \\ P & K & R & E \\ \text{ATTGAATATT} \\ L & N & I \\ \text{CTTAAGGCAG} \\ L & K & A \\ \text{TTTATTGGGG} \\ F & Y & W & G \\ \text{TGCACCTAAA} \\ A & P & K \\ \text{AATATATCAA} \\ \text{AAGAAAGTTA} \\ \text{Q} & \textbf{E} & \textbf{S} & Y \\ \text{CAAAATTAGGCG} \\ K & I & M \\ \text{TTTTAGGCG} \\ F & L & G \\ \text{TGGCGCGCGTGG} \\ L & A & R & G \\ \text{TAAAATCAAA} \\ \text{K} & I & K \\ \end{array}$	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V TTTCAAATTT F Q I CCGGAAACATT G K Y AATCAAAATG N H N ATGGCTTATC N G L S CAGAACTACG R T T AATAATGGTT N N G GTGACAAAAG G D K S TAACTTTACC V T Y	E R K H AATTAAAGT $\underline{I}$ K W TGCACTGCTT $\underline{A}$ L L GATGACAAAA D D K TAACGTTTAA L T F N TAATTCAGGT N S G AATGATTCAGGT N S G AATGATTCAG E E F Q AGGGGACTAA G G L ATTGATAGAA I D R GGGGACCATA W G P Y TAGTCCAATT S S N CCAGAATTCA P E F AAAGAGAAAT Q R E M	N A F <u>AAAGGACATA</u> K G H TTGCTATCGA <u>L L S</u> TCACTTTATA I T L Y TTTCATTAAG F I K TATAAAAATT Y K N TTAGTTCAGA V Q Q AATGGATCAA N G S AAACAAATCA K T N H TGGTAGAGAAT G R D GCTGGTCAAA A G Q TAAGCGTACT I S V L GGATAGATAT D R Y	$\begin{array}{c} V & N & R \\ ATTAATATGA \\ N Stop M \\ ATACAGTTGA \\ N & T & V & D \\ CAAAACAACA \\ K & T & T \\ GATAAAAGTT \\ D & K & S \\ CTAATCCAAA \\ F & N & P & K \\ ATCAAATGAT \\ S & N & D \\ ACATTAGGTT \\ T & L & G \\ AATCATTTCC \\ K & S & F & S \\ TAAATCAATT \\ K & S & I \\ AGTTATGACC \\ S & Y & D \\ ATTCTCTGCC \\ N & F & L & P \\ TTCTCATAAA \\ S & H & K \\ ACTAATCAAT \\ T & N & Q \\ \end{array}$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890 145 1960 168 2030 191 2100 215 2170 238 2240 261 2310 285
274 1401 298/3 1471 52 1541 52 1681 52 1681 751 99 91821 146 1751 199 1821 146 1961 192 2031 192 2031 2121 2241	T L F P TAGATATAAA R Y K ATTAGGCAAA L G K AATATGCACAC N I T ATAATGATAA D N D K TACTCACAGT Y S Q TTGTTGACTA V V D Y CGGAGATATT G D I AATTATAAAC N Y K TTGAGGCCCA V E A H TAGGGCCCA V E A H TAGGGCCCA V E A H TAGGGCCCA CA K K S ACACTGGGTT	$\begin{array}{c} \mathbf{R}  \mathbf{T}  \mathbf{G} \\ \mathbf{GTTAATTGA} \\ \mathbf{V}  \mathbf{N}  \mathbf{W} \\ \mathbf{TCATCAGTTG} \\ \mathbf{S}  \mathbf{S}  \mathbf{V} \\ \mathbf{CTAAAAGAGA} \\ \mathbf{P}  \mathbf{K}  \mathbf{R}  \mathbf{E} \\ \mathbf{ATTGAATATT} \\ \mathbf{L}  \mathbf{N}  \mathbf{I} \\ \mathbf{CTTAAGGCAG} \\ \mathbf{L}  \mathbf{K}  \mathbf{A} \\ \mathbf{TTTATTGGGG} \\ \mathbf{F}  \mathbf{Y}  \mathbf{W}  \mathbf{G} \\ \mathbf{TGCACTAAA} \\ \mathbf{A}  \mathbf{P}  \mathbf{K} \\ \mathbf{AATATATCTA} \\ \mathbf{N}  \mathbf{I}  \mathbf{S} \\ \mathbf{AAGAAAGTTAA} \\ \mathbf{Q}  \mathbf{E}  \mathbf{S}  \mathbf{Y} \\ \mathbf{CAAAATTATG} \\ \mathbf{K}  \mathbf{I}  \mathbf{M} \\ \mathbf{TTTTTAGGCG} \\ \mathbf{F}  \mathbf{L}  \mathbf{G} \\ \mathbf{TGGCGCGTGG} \\ \mathbf{L}  \mathbf{A}  \mathbf{R}  \mathbf{G} \\ \mathbf{TAAAAATCAAA} \\ \mathbf{K}  \mathbf{I}  \mathbf{K} \\ \mathbf{GGTAATTAACT} \\ \end{array}$	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V TTCAAATTT F Q I CCGGAAACAT A G N I CCGTAAGTAT G K Y AATCAAAATG N H N ATGGCTTATC N G L S CAGAACTACG R T T AATAATGGTT N N G GTGACAAAAG G D K S TAACTTTAC N F N GTAACTTACC V T Y ACAAAAATCA	E R K H AATTAAAGTG I K V TGCACTGCTT A L L GATGACAAAA D D K TAACGTTTAA L T F N TAACGTTTAA L T F N TAATTCAGGT N V S AAGATTCGA AAGTGGATTA G G L ATGGACGATTA G G L ATGGACATTA G G L ATGGACATTA M G P Y TAGTCAAAA I D R GGGGACCATA W G P Y TAGTCCAAT S S N CCAGAATTTA P E F AAAGAGAAAT Q R E M AAATACAGTA	N A F AAAGGACATA K G H TIGCTATCGA L L S TCACTITATA I T L Y TITCATTAAG F I K TATAAAAATT Y K N TTAGTTCAGA V S S E AGTTCAACAA N G S AAAGGATCAA K T N H TGGTAGGATA G R D GCTGGTCAAA A G Q TAAGCGTACT I S V L GGATAGATAT D R Y	$ \begin{array}{c} \mathbb{V}  \mathbb{N}  \mathbb{R} \\ \mathrm{ATTAATATGA} \\ \mathrm{ATTAATATGA} \\ \mathrm{N}  \mathrm{stop}  \mathbb{M} \\ \mathrm{ATACAGTTGA} \\ \mathrm{N}  \mathbb{T}  \mathbb{V}  \mathbb{D} \\ \mathrm{CAAAACAACA} \\ \mathrm{K}  \mathbb{T}  \mathbb{T} \\ \mathrm{GATAAAGTT} \\ \mathrm{GATAAAGAT} \\ \mathrm{F}  \mathbb{N}  \mathbb{P} \\ \mathrm{K} \\ \mathrm{ATCAATGAAT} \\ \mathrm{K}  \mathbb{S}  \mathbb{F} \\ \mathrm{ATCAATTAGGAT} \\ \mathrm{T}  \mathbb{L}  \mathbb{G} \\ \mathrm{AATCATTTTC} \\ \mathrm{K}  \mathbb{S}  \mathbb{F}  \mathbb{S} \\ \mathrm{TAAATCAATT} \\ \mathrm{K}  \mathbb{S}  \mathbb{F}  \mathbb{S} \\ \mathrm{TAAATCAATT} \\ \mathrm{K}  \mathbb{S}  \mathbb{F} \\ \mathrm{AGTTAATGACC} \\ \mathrm{S}  \mathbb{Y}  \mathbb{D} \\ \mathrm{ATTCTTGCCC} \\ \mathrm{N}  \mathbb{F}  \mathbb{L}  \mathbb{P} \\ \mathrm{TTCTCATAAAA} \\ \mathrm{S}  \mathrm{H}  \mathbb{K} \\ \mathrm{ACTAATCAAT} \\ \mathrm{T}  \mathbb{N}  \mathbb{Q} \\ \mathrm{CTACTTATGAC} \\ \end{array} $	$\begin{array}{c cccc} N & F & V & V \\ AAATTGAAAA \\ K & I & E & K \\ TGCAGCTCAA \\ A & A & Q \\ GCAACATCTG \\ A & T & S \\ ATGACAAAGA \\ Y & D & K & D \\ AGATTACAAT \\ D & Y & N \\ GCTGTAAATG \\ A & V & N \\ ATTCTTATGG \\ Y & S & Y & G \\ AGAAACGATA \\ E & T & I \\ GGCTGGGGTG \\ G & W & G \\ CAACATATGG \\ P & T & Y & G \\ AACACATCAA \\ T & H & Q \\ CTTTTTGATA \\ L & F & D \\ GGAACGATC \\ W & N & R & L \\ AGTTGACTGG \\ \end{array}$	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890 145 1960 168 2030 191 2100 215 2170 238 2240 261 2310 285 2380
$\begin{array}{c} 274\\ 1401\\ 298/3\\ 1471\\ 6\\ 1541\\ 299\\ 1611\\ 52\\ 1681\\ 1751\\ 99\\ 1821\\ 142\\ 192\\ 2031\\ 146\\ 1961\\ 192\\ 2031\\ 192\\ 2031\\ 2101\\ 2161\\ 2171\\ 229\\ 2241\\ 262\\ 2311\\ 286\end{array}$	$\begin{array}{ccccc} T & L & F & P \\ TAGATATAAA \\ R & Y & K \\ ATTAGGCAAA \\ L & G & K \\ AATATCACAC \\ N & I & T \\ TATAATGATAA \\ D & N & D & K \\ TACGTTAGTA \\ T & L & V \\ TACTCACAGT \\ Y & S & Q \\ TTGTTGACTA \\ V & D & Y \\ CGGAGATATT \\ G & D & I \\ AATTATATAAC \\ N & Y & K \\ TGGAGGCCA \\ V & C & A & H \\ TAATGAACTG \\ N & E & L \\ ATCCCTTTAT \\ I & P & L \\ CAAAAAAATC \\ T & K & K \\ S & ACACTGGTT \\ H & W \\ V \end{array}$	R T G GTTAATTGGA V N W TCATCAGTTG S V CTAAAAGAGA P K R E ATTGAATATT L N I CTTAAGGCAG L K A TTTATTGGGG F Y W G TGCACCTAAA A P K AATAATATCAA N I S AAGAAAGTTA Q E S Y CAAAATTATG K I M TTTTTAGGCG F L G TGGCGCGTGG L A R G TAAAATCAAA K I K GGTAATAACT G N N	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V TTCCAAATTT F Q I CCGGAAACAT A G N I CGGTAACATAT G K Y AATCAAAATG N H N ATGGCTTATC N G L S CAGAACTACG R T T AATAATGGTT N N G GTGACAAAAG G D K S TAACTTAAC N F N GTAACTTACC V T Y ACAAAAATCA Y K N Q	E R K H AATTAAAGTG I K V TGCACTGCTT A L L GATGACAAAA D D K TAACGTTTAA L T F N TAATTCAGGT N S G AATGTTTCAGT N V S AAGAATTCCA E E F Q AGGTGGATTA G G L ATTGATAGAA I D R GGGGACCATA W G P Y TAGTTCAAAT S S N CCAGAATTTA F E F AAAGAGAAAT Q R E M AAATACAGTA N T V	N A F AAAGGACATA K G H TTGCTATCGA L L S TCACTTTATA I T L Y TTCATTATAG F I K TATAAAAATT Y K N TTAGTTCAGA V S S E AGTTCAACAA V S S E AGTTCAACAA V Q Q Q AATGGATCAACAA K T N H TGGTAGAGAT G R D GCTGGTCAAA A G Q TAAGCGTACT I S V L GGATAGATAT D R Y ACGTTACACAT T F T	$\begin{array}{c c} V & N & R \\ ATTAATATGA \\ N StOP M \\ ATACAGTTGA \\ N & T & D \\ CAAAACAACA \\ K & T & T \\ GATAAAGATT \\ D & K & S \\ CTAATCCAAA \\ F & N & P & K \\ ATCAAATGAT \\ S & N & D \\ ACATTAGGAT \\ T & L & G \\ ATCATTAGGAT \\ T & L & G \\ ATCATTATGAC \\ S & F & S \\ TAAATCAATT \\ K & S & F & S \\ TAAATCAATT \\ K & S & F & S \\ TAAATCAATT \\ K & S & T \\ AGTTATGACC \\ S & Y & D \\ ATTTCTTGCC \\ N & F & L & P \\ TTCCTCATAAA \\ S & T & N & Q \\ CTACTTATGAS \\ S & T & Y & E \\ \end{array}$	$\begin{array}{c c} N & F & V & V \\ \text{AAATTGAAAA} \\ K & I & E & K \\ TGCAGCTCAA \\ \hline \\ A & T & S \\ \\ GCAACATCTG \\ A & T & S \\ \\ A & C & C \\ A & T & S \\ \\ A & C & C \\ A & T & C \\ A & T & C \\ A & C & C \\ C & C & C \\ C & C & C \\ C & C &$	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890 145 1960 168 2030 191 2100 215 2170 238 2240 261 2310 285 2380 308
$\begin{array}{c} 274\\ 1401\\ 298/3\\ 1471\\ 6\\ 1541\\ 299\\ 1611\\ 52\\ 1681\\ 766\\ 1751\\ 99\\ 1821\\ 122\\ 1891\\ 146\\ 1961\\ 1962\\ 2031\\ 192\\ 2031\\ 192\\ 2101\\ 2166\\ 2171\\ 229\\ 2241\\ 262\\ 2311\\ 286\\ 2381\\ \end{array}$	T L F P TAGATATAAA R Y K ATTAGGCAAA L G K AATATCACACA N I T ATAATGATAA D N D K TACGTTAGTA T L V TACTCACAGT Y S Q TTGTTGACTA V V D Y CGGAGATATT G D I AATTATAAAC N Y K TTGAGGCCCA V E A H TAATGAACTG N E L ATCCCTTTAT I P L CAAAAAAATC T K K S ACACTGGGTT H W V CAAAACATAC	R T G GTTAATTGGA V N W TCATCAGTGG S S V CTAAAAGAGA P K R E ATTGAATATT L N I CTTAAGGCAG L K A TTTATTGGGG F Y W G TGCACCTAAA A P K AATATATCTA N I S AAGAAAGTTA Q E S Y CAAAATTATG K I M TTTTTTAGGCG F L G TGGCGCGTGGG L A R G TAAAATCAAA K I K GGTAATAACT	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V TTTCAAATTT F Q I CCGGAAACAT A G N I CGGTAAGTAT G K Y AATCAAAATG N H N ATGGCTTATC N G L S CAGAACTACG R T T AATAATGGTT N N G GTGACAAAAG G D K S TAACTTAAC N F N GTAACTTACC V T Y ACAAAAATCA Y K N Q AATCGGTACG	E R K H AATTAAAGTG I K V TGCACTGCTT A L L GATGACAAAA D D K TAACGTTTAA L T F N TAATTCAGGT N S G AATGTTCAGGT N V S AAGAATTCCA E E F Q AGGTGGACTAA G G L ATTGATAGAA I D R GGGGACCATA W G P Y TAGTTCAAAT S S N CCAGAATTTA P E F AAAGAGAAAT Q R E M AAATACAGTA	N A F <u>AAAGCACATA</u> K G H TIGCTATCGA L L S TCACTTTATA I T L Y TTTCATTAAG F I K TATAAAAATT Y K N TTAGTTCAGA V S S E AGTTCAACAA V Q Q AATGGATCAA N G S AAACAAATCA K T N H TGGTAGAGAT G R D GCTGGTCAAA A G Q TAAGCGTACT I S V L GGATAGATAT D R Y ACGTTTACAT AAACTAATCC	V N R ATTAATATGA N Stop M ATACAGTTGA N T V D CAAAACAACA K T T GATAAAAGTT D K S CTAATCCAAA F N P K ATCAAATGAT S N D ACATTAGGTT T L G AATCATTTCC K S F S TAAATCAATT K S I AGTTATGACC S Y D ATTTCTTGCC N F L P TTCTCATAAA S H K ACTAATCAATT T N Q CTACTTATGA S T Y E TGGAGTATAA	$\begin{array}{c c} N & F & V & V \\ \text{AAATTGAAAA} \\ K & I & E & K \\ TGCAGCTCAA \\ \hline \\ A & T & S \\ \\ \text{GCAACATCTG} \\ A & T & S \\ \\ \text{ATGACAAAGA} \\ Y & D & K & D \\ \\ \text{AGATTACAAT} \\ D & Y & N \\ \\ \text{GCTGTAAATG} \\ A & V & N \\ \\ \text{ATTCTTATGG} \\ Y & S & Y & G \\ \\ \text{AGAAACAATATGG} \\ F & T & I \\ \\ \text{GGCTGGGGTG } \\ G & W & G \\ \\ \text{CAACATATGG} \\ P & T & Y & G \\ \\ \text{AACACATATGG} \\ P & T & Y & G \\ \\ \text{AACACATATGG} \\ P & T & Y & G \\ \\ \text{AACACATATGG} \\ P & T & Y & G \\ \\ \text{AACACATATGA } \\ T & H & Q \\ \\ \text{CTTTTGATA } \\ \\ \text{GGAATCGATC } \\ \\ \text{M N R L } \\ \\ \text{AGTTGACTGG } \\ V & D & W \\ \\ \end{array}$	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890 145 1960 168 2030 191 2100 215 2170 238 2240 261 2310 285 2380 308 2450
274 1401 298/3 1471 6 1541 299 1611 52 1681 766 1751 122 1891 1426 2031 1469 2031 2162 2101 2166 2171 240 2241 262 2311 2862 2381 309	$\begin{array}{cccc} T & L & F & P \\ TAGATATAAA \\ R & Y & K \\ ATTAGGCAAA \\ L & G & K \\ AATATGACACA \\ N & I & T \\ ATAATGATAA \\ D & N & D & K \\ TACGTTAGTA \\ T & L & V \\ TACGTTAGTA \\ T & L & V \\ TACGTTAGTA \\ Y & S & Q \\ TTGTTGACTA \\ V & V & D & Y \\ CGGAGATAT \\ Y & S & Q \\ TTGTTGACTA \\ V & V & D & Y \\ CGGAGATATATAAAC \\ N & Y & K \\ TTGAGGCCCA \\ V & K \\ TGAGGGCCA \\ V & E & A \\ H & K \\ CAAAAAAACT \\ T & K & K \\ S \\ ACACTGGGT \\ H & W \\ CAAAACATAC \\ Q & N \\ I \\ \end{array}$	R T G GTTAATTGGA V N W TCATCAGTGG S S V CTAAAAGAGA P K R E ATTGAATATT L N I CTTAAGGCAG L K A TTTATTGGGG F Y W G TGCACCTAAA A P K AATATATCKA N I S AAGAAAGTTA Q E S Y CAAAATTAGGCG F L G TGCGCGCGTGG L A R G TAAAATCAAA K I K GGTAATAACT G N N TGTTAAAATT	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V TTTCAAATTT F Q I CCGGAAACAT A G N I CGGTAAGTAT G K Y AATCAAAATG N H N ATGGCTTATC N G L S CAGAACTACG G D K S TAACTATAGGT G D K S TAACTATACC V T Y ACAAAAATCA Y K N Q AATCGGTAGG I G T	E R K H AATTAAAGT $\underline{I}$ K W TGCACTGCTT $\underline{A}$ L L GATGACAAAA D D K TAACGTTTAA L T F N TAATTCAGGT N S G AATGATTCAGGT N V S AAGAATTCA E E F Q AGGTGGATTA G G L ATTGATAGAA I D R GGGGACCATA W G P Y TAGTCCAAGAAT P E F AAGGAGAAAT Q R E M AAATACAGTA N T V GATTCTAAGG D S K	N A F <u>AAAGGACATA</u> K G H TTGCTATCGA <u>L L S</u> TCACTTTATA I T L Y TTTCATTAAG F I K TATAAAAATT Y K N TTAGTTCAGA V Q Q AATGGATCAA N G S AAACAAATCA K T N H TGGTAGAGATAT G R D GCTGGTCAAA A G Q TAAGCGTACT I S V L GGATAGATAT D R Y ACGTTTACAT T F T AAACTAATCC E T N P	$ \begin{array}{c} V & N & R \\ \text{ATTAATATGA} \\ \text{ATTAATATGA} \\ \text{N stop M} \\ \text{ATACAGTGA} \\ \text{N T V D} \\ \text{CAAAACAACA} \\ K & T T \\ \text{GATAAAAGTT} \\ D K S \\ \text{CTAATCCAAA} \\ \text{F N P K } \\ \text{ATCAAATGAT} \\ S N D \\ \text{ACATTAGGTT} \\ T L G \\ \text{AATCATTAGGTT} \\ T L G \\ \text{AATCATTAGGTT} \\ K S I \\ \text{AGTAATCAATT} \\ K S I \\ \text{AGTATATGACC } \\ \text{S Y D} \\ \text{ATTCTTGCC } \\ \text{N F L P } \\ \text{TTCTCATAAA } \\ \text{S H K } \\ \text{ACTAATCAATT } \\ \text{ACTAATCAAT } \\ \text{S H K } \\ \text{ACTAATCAAT } \\ \text{ACTAATCAAT } \\ \text{S H K } \\ \text{ACTAATCAAT } \\ \text{ACTAATCAAT } \\ \text{ACTAATCAAT } \\ \text{S H K } \\ \text{ACTAATCAAT } \\ \text{ACTACTAATCAAT } \\ \text{ACTAATCAAT } \\ \text{ACTACTAATCAAT } \\ \text{ACTACTACTACAAT } \\ \text{ACTACTACTACAAT } \\ \text{ACTACTACTACAAT } \\ \text{ACTACTACTACTACAAT } \\ \text{ACTACTACTACTACAAT } \\ \text{ACTACTACTACTACAAT } \\ \text{ACTACTACTACTACAAT } \\ \text{ACTACTACTACAAT } \\ \text{ACTACTACTACTACAAT } \\ \text{ACTACTACTACTACTACAAT } \\ \text{ACTACTACTACTACAAT } \\ \text{ACTACTACTACTACTACAAT } \\ \text{ACTACTACTACTACTACTACAAT } \\ \text{ACTACTACTACTACTACAAT } \\ \text{ACTACTACTACTACTACTACAAT } \\ \text{ACTACTACTACTACTACTACTACAAT } \\ \text{ACTACTACTACTACTACTACTACAAT } \\ ACTACTACTACTACTACTACTACTAC$	N F V V AAATTGAAAA K I E K TGCAGCTCAA A A Q GCAACATCTG A T S ATGACAAAGA Y D K D AGATTACAAT D Y N ATTCTTATGG Y S Y G AGAAACGATA E T I GGCTGGGGTG G W G CAACATCAA T H Q CTTTTTGATA L F D GGAATCGATC W N R L AGTTGACTGG V D W CTGTTAGAAA	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890 145 1960 2030 191 2100 215 2170 238 2240 261 2310 285 2380 308 2450 327
274 1401 298/3 1471 6 1541 52 1681 52 1681 751 99 1821 146 1961 192 2031 192 2101 262 2171 239 2241 262 2381 262 2381 262 2381 262 2381 309 2451	T L F P TAGATATAAA R Y K ATTAGGCAAA L G K AATATGCACAC N I T ATAATGATAA D N D K TACTCACAGT Y S Q TTGTTGACTA V V D Y CGGAGATATT G D I AATTATAAAC N Y K TTGAGGCCCA V E A H TAGGGCCCA V E A H TAAGAGGCCT I P L CAAAAAAATA C K K S ACACTGGGTT H W V CAAAACATAC Q N I TAGAGGAGTT	R T G GTTAATTGGA V N W TCATCAGTTG S S V CTAAAAGAGA P K R E ATTGAATATT L N I CTTAAGGCAG F Y W G TGCACCTAAA A P K AATATATCTA N I S AAGAAAGTTA Q E S Y CAAAATTATG F L G TGGCGCGTGG L A R G TAGCACCTGA K I M TTTTTAGGCG F L G TGGCGCGTGG L A R G TAGAAATCAAC K I K GGTAATAACT G N N TGTTAAAATC	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V TTCAAATTT F Q I CCGGAAACAT A G N I CCGGAAACAT G K Y AATCAAAATG N H N ATGGCTTATC N G L S CAGAACTACG R T T AATAATGGTT N N G GTGACAAAAG G D K S TAACTTTAAC N F N GTAACTTACC V T Y ACAAAAATCA Y K N Q AATCGGTACG I G T TTTAAGAGTT	E R K H AATTAAAGTG I K V TGCACTGCTT A L L GATGACAAAA D D K TAACGTTTAA L T F N TAATTCAGGT N S G AATGTTCGG N V S AAGGTGATTTCGG A C L ATGGACATTAA G G L ATGGACATTA G G L ATGGACATTA G G L ATGGACATTAA G G L ATGGACATTAA B C P Y TAGTCAAAA N CCAGAATTTA P E F AAAGAGAAATTA Q R E M AAATACAGTA N T V GATTCTAAAG D S K AAACAAAAAAA	N A F AAAGGACATA K G H TIGCTATCGA L L S TCACTITATA I T L Y TITCATTAAG F I K TATAAAAATT Y K N TTAGTTCAGA V S S E AGTTCAACAA V Q Q AATGGATCAA N G S AAACAAATCA K T N H TGGTAGGATA G R D GCTGGTCAAA A G Q TAAGCGTACT I S V L GGATAGATAT D R Y ACGTITACAA T F T AAACTAATCC	$ \begin{array}{c} V & N & R \\ ATTAATATGA \\ N StOP M \\ ATACAGTTGA \\ N & T & V & D \\ CAAAACAACA \\ K & T & T \\ GATAAAGTT \\ D & K & S \\ CTAATCCAAA \\ F & N & P & K \\ ATCAAATGAT \\ S & N & D \\ ACATTAGGAT \\ T & L & G \\ AATCATTTCC \\ K & S & F & S \\ TAAATCAATT \\ K & S & I \\ AGTTATGACC \\ S & Y & D \\ ATTCCTTAGCC \\ N & F & L & P \\ TTCTCATAAA \\ S & H & K \\ ACTAATCAAT \\ T & N & Q \\ CTACTTATGAC \\ S & T & Y & E \\ TGGAGTATAA \\ G & V STQT \\ \end{array} $	$\begin{array}{c c} N & F & V & V \\ \text{AAATTGAAAA} \\ K & I & E & K \\ TGCAGCTCAA \\ \hline \\ A & T & S \\ \\ \text{GCAACATCTG} \\ A & T & S \\ \\ \text{ATGACAAAGA} \\ Y & D & K & D \\ \\ \text{AGATTACAAT} \\ D & Y & N \\ \\ \text{GCTGTAAATG} \\ A & V & N \\ \\ \text{ATTCTTATGG} \\ Y & S & Y & G \\ \\ \text{AGAAACAATATGG} \\ F & T & I \\ \\ \text{GGCTGGGGTG } \\ G & W & G \\ \\ \text{CAACATATGG} \\ P & T & Y & G \\ \\ \text{AACACATATGG} \\ P & T & Y & G \\ \\ \text{AACACATATGG} \\ P & T & Y & G \\ \\ \text{AACACATATGG} \\ P & T & Y & G \\ \\ \text{AACACATATGA } \\ T & H & Q \\ \\ \text{CTTTTGATA } \\ \\ \text{GGAATCGATC } \\ \\ \text{M N R L } \\ \\ \text{AGTTGACTGG } \\ V & D & W \\ \\ \end{array}$	297 1470 314/5 1540 28 1610 51 1750 98 1820 121 1890 145 1960 168 2030 191 2100 215 2170 238 2240 261 2310 285 2380 308 2450 307 2520
274 1401 298/3 1471 6 1541 52 1681 52 1681 751 99 1821 146 1961 192 2031 192 2101 262 2171 239 2241 262 2381 262 2381 262 2381 262 2381 309 2451	$\begin{array}{cccc} T & L & F & P \\ TAGATATAAA \\ R & Y & K \\ ATTAGGCAAA \\ L & G & K \\ AATATGACACA \\ N & I & T \\ ATAATGATAA \\ D & N & D & K \\ TACGTTAGTA \\ T & L & V \\ TACGTTAGTA \\ T & L & V \\ TACGTTAGTA \\ Y & S & Q \\ TTGTTGACTA \\ V & V & D & Y \\ CGGAGATAT \\ Y & S & Q \\ TTGTTGACTA \\ V & V & D & Y \\ CGGAGATATATAAAC \\ N & Y & K \\ TTGAGGCCCA \\ V & K \\ TGAGGGCCA \\ V & E & A \\ H & K \\ CAAAAAAACT \\ T & K & K \\ S \\ ACACTGGGT \\ H & W \\ CAAAACATAC \\ Q & N \\ I \\ \end{array}$	R T G GTTAATTGGA V N W TCATCAGTTG S S V CTAAAAGAGA P K R E ATTGAATATT L N I CTTAAGGCAG F Y W G TGCACCTAAA A P K AATATATCTA N I S AAGAAAGTTA Q E S Y CAAAATTATG F L G TGGCGCGTGG L A R G TAGCACCTGA K I M TTTTTAGGCG F L G TGGCGCGTGG L A R G TAGAAATCAAC K I K GGTAATAACT G N N TGTTAAAATC	I Y A AAACACACGA K T H E CTTCATCAAT A S S I GAAAAAAGTA K K V TTCAAATTT F Q I CCGGAAACAT A G N I CCGGAAACAT G K Y AATCAAAATG N H N ATGGCTTATC N G L S CAGAACTACG R T T AATAATGGTT N N G GTGACAAAAG G D K S TAACTTTAAC N F N GTAACTTACC V T Y ACAAAAATCA Y K N Q AATCGGTACG I G T TTTAAGAGTT	E R K H AATTAAAGTG I K V TGCACTGCTT A L L GATGACAAAA D D K TAACGTTTAA L T F N TAATTCAGGT N S G AATGTTCGG N V S AAGGTGATTTCGG A C L ATGGACATTAA G G L ATGGACATTA G G L ATGGACATTA G G L ATGGACATTAA G G L ATGGACATTAA B C P Y TAGTCAAAA N CCAGAATTTA P E F AAAGAGAAATTA Q R E M AAATACAGTA N T V GATTCTAAAG D S K AAACAAAAAAA	N A F AAAGGACATA K G H TIGCTATCGA L L S TCACTITATA I T L Y TITCATTAAG F I K TATAAAAATT Y K N TTAGTTCAGA V S S E AGTTCAACAA V Q Q AATGGATCAA N G S AAACAAATCA K T N H TGGTAGGATA G R D GCTGGTCAAA A G Q TAAGCGTACT I S V L GGATAGATAT D R Y ACGTITACAA T F T AAACTAATCC	$ \begin{array}{c} V & N & R \\ ATTAATATGA \\ N StOP M \\ ATACAGTTGA \\ N & T & V & D \\ CAAAACAACA \\ K & T & T \\ GATAAAGTT \\ D & K & S \\ CTAATCCAAA \\ F & N & P & K \\ ATCAAATGAT \\ S & N & D \\ ACATTAGGAT \\ T & L & G \\ AATCATTTCC \\ K & S & F & S \\ TAAATCAATT \\ K & S & I \\ AGTTATGACC \\ S & Y & D \\ ATTCCTTAGCC \\ N & F & L & P \\ TTCTCATAAA \\ S & H & K \\ ACTAATCAAT \\ T & N & Q \\ CTACTTATGAC \\ S & T & Y & E \\ TGGAGTATAA \\ G & V STQT \\ TATTTGTTT \\ \end{array} $	N F V V AAATTGAAAA K I E K TGCAGCTCAA A A Q GCAACATCTG A T S ATGACAAAGA Y D K D AGATTACAAT D Y N ATTCTTATGG Y S Y G AGAAACGATA E T I GGCTGGGGTG G W G CAACATCAA T H Q CTTTTTGATA L F D GGAATCGATC W N R L AGTTGACTGG V D W CTGTTAGAAA	297 1470 314/5 1540 28 1610 51 1680 75 1750 98 1820 121 1890 145 1960 2030 191 2100 215 2170 238 2240 261 2310 285 2380 308 2450 327

Fig. 2. Nucleotide sequence (EMBL/GenBank Y13225) of *lukE* (positions 511–1455) followed by that of *lukD* (positions 1457–2440) and the deduced peptide sequences from *S. aureus* Newman strain. The stop codon of *lukE* is located 1 base upstream the ATG codon of *lukD*. Signal peptides of the two proteins LukE and LukD are underlined. Putative consensus sequences (i.e. ribosome binding site (rbs) and promotor sequences -35 and -10) and the start point of the messenger RNA are mentioned. An inverted repeat sequence probably used as a transcription terminator (positions 2500–2521) is indicated by arrows.

# 3.3. LukE+LukD and LukM production in 127 S. aureus clinical isolates

For the study of the production of LukE+LukD among clinical isolates, attention was given to obtain significant series of isolates for each kind of sample. The immunoprecipitation test revealed no isolate producing LukM. Only one produced the PVL, which is in agreement with the previously described results [7,21]. However, 40 isolates (31.5%) produced LukE+

LukD. This was confirmed by hybridization of *Hin*dIII-restricted DNAs with the *lukE*-oligonucleotide probe used for the mRNA mapping (data not shown). The hybridized DNA fragments were 3.1 and > 8.3 kb long, according to the strains. No positive isolate was found among urine-originated samples, whereas a similar distribution concerned respiratory (6:29), skin samples (9:31), infected materials (8:19) and blood cultures (4:19). Most of the LukE+LukD-producing

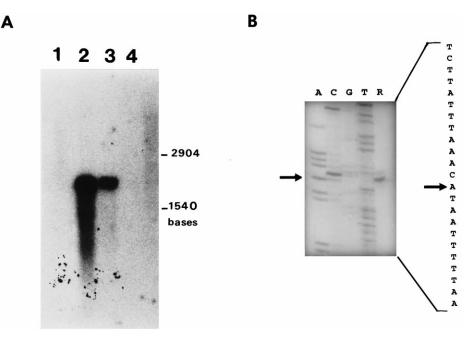


Fig. 3. Cotranscription of *lukE* and *lukD*, and determination of the corresponding mRNA start. A: Northern blot analysis of the total RNAs from *S. aureus* ATCC 49775 and Newman strains (YCP medium) with the labelled *lukE/lukD* genes. Lane 1: Total RNA from  $10^8$  CFU of ATCC 49775 after 12 h of culture. Lanes 2, 3, 4: Total RNA from  $10^8$  CFU of Newman after 12, 9, 6 h of culture, respectively. Lengths of the 16 s (1540 bases) and the 23 s (2904 bases) RNAs from *S. aureus* were obtained from formaldehyde-agarose gel stained with ethidium bromide. B: Primer extension indicated nucleotide A477 (T on the coding strand) as the mRNA start of *lukE/lukD*. Lanes A, C, G, T: Sequencing reactions; lane R, primer extension.

strains were isolated as pure cultures from diarrhea (77%) and were methicillin-resistant strains.

## 3.4. Biological activity

Such strains as Newman produce two leucotoxins among  $\gamma$ hemolysin, PVL, and LukE+LukD. Class S-class F protein combinations were tested in hemolytic assays on rabbit (RRBC) and human (HRBC) red blood cells [22], on human PMNs and in a dermonecrosis assay in rabbit skin. None of the leucotoxin proteins alone induces any biological activity which excluded that they were contaminated. Neither LukE when combined with one of the class S components (i.e. LukS-PV, HlgA, HlgC, LukM), nor LukD combined with one of the class F components (i.e. LukF-PV, HlgB, LukF'-PV) harbored any biological activity.

3.4.1. In vitro assays. On HRBC, only the specific activities of LukE+HlgB =  $7 \times 10^2 \pm 0.5 \times 10^2$ , HlgA+LukD =  $10^3 \pm 0.5 \times 10^2$  and HlgA+LukF'-PV =  $10^4 \pm 2 \times 10^3$  U/mg were quantifiable but largely smaller than that of HlgA+HlgB  $(5 \times 10^7 \pm 2 \times 10^6$  U/mg). On RRBC, the specific activities of HlgA+LukD =  $3 \times 10^5 \pm 10^4$ , and HlgA+LukF'-PV =  $3 \times 10^6$   $\pm 10^5$  were significant, whereas the other LukE-, LukD-, LukM-, LukF'-PV-derived combinations remained basal compared to other leucotoxins [22].

The influx of Ca<sup>2+</sup> in human PMNs provoked by the primary opening of  $Ca^{2+}$  channels [27] consecutive to leucotoxin activity was tested in the presence of 1 mM free Ca<sup>2+</sup> for the 20 class S/class F protein combinations. The most potent couple HlgA+HlgB was assumed at 1000 arbitrary units of Fluo3-fluorescence from loaded PMNs and was compared with fluorescence intensities obtained with other pairs (Table 1). Cells were not lysed and the remaining lymphocytes were not susceptible to them. According to the calculated S.D. values, two kinds of results were obtained: (i) those presenting an S.D. value lower than the mean of assays; and (ii) those with a higher S.D. The first kind of results corresponded to pairs which were active with all the tested donors. HlgA was active in combination with all the F components, but the HlgA+LukD pair was the only combination with LukD presenting an activity. HlgA+LukF'-PV and LukS-PV+LukF-PV were as potent as HlgA+HlgB. HlgC was as active with HlgB as with LukF'-PV but less active with LukF-PV and LukD.

Table 1

Biological activities of the 20 possible leucotoxin combinations in S. aureus on the opening of Ca<sup>2+</sup> channels of human PMNS

1	1	1 6						
HlgB	LukD	LukF-PV	LukF'-PV					
1000	349±138	$664 \pm 163$	$907 \pm 108$					
$497 \pm 288$	$130 \pm 245$	$183 \pm 177$	$493 \pm 246$					
$504 \pm 242$	$100 \pm 186$	$972 \pm 183$	$672 \pm 207$					
$167 \pm 180$	99±166	$276 \pm 197$	$295 \pm 198$					
$310 \pm 277$	$90 \pm 167$	$136 \pm 211$	$166 \pm 221$					
	$ \begin{array}{r} 1000\\ 497 \pm 288\\ 504 \pm 242\\ 167 \pm 180 \end{array} $	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	$1000$ $349 \pm 138$ $664 \pm 163$ $497 \pm 288$ $130 \pm 245$ $183 \pm 177$ $504 \pm 242$ $100 \pm 186$ $972 \pm 183$ $167 \pm 180$ $99 \pm 166$ $276 \pm 197$					

Data were obtained from the analysis of 5000 purified cells, previously charged with Fluo-3. The most lytic couple HlgA+HlgB was referred as 1000 arbitrary fluorescence units. Data are the mean  $\pm$  S.D. of 4 independent experiments.

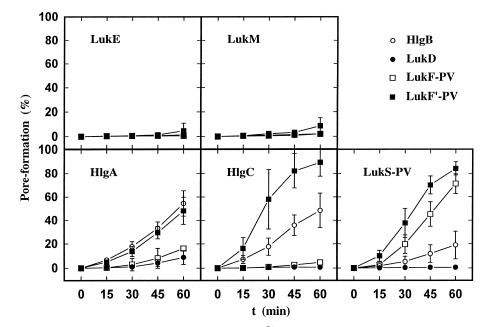


Fig. 4. The pore formation by the leucotoxins, in the absence of 1 mM  $Ca^{2+}$  out of human PMNs, was recorded by the entry of ethidium in 5000 purified cells, previously charged with Fluo-3. Data were obtained on FacSort (Becton-Dickinson), as the percentage of maximal cell lysis obtained after addition of 0.2% (w/v) of Triton X-100, and as the mean  $\pm$  S.D. of 4 independent experiments. Twenty combinations of the mentioned class S components together with one class F component ( $\bigcirc$ : HlgB,  $\bullet$ : LukD,  $\square$ : LukF-PV,  $\blacksquare$ : LukF'-PV) are presented.

LukM presented the same activity with LukF'-PV or with LukF-PV. LukE+HlgB was representative of the second kind of results where leucotoxins appeared randomly active according to the donor. For a series of four donors, only one possessed PMNs more or less sensitive to these pairs. In the absence of Ca<sup>2+</sup>, pore formation was studied by the entry of ethidium (Fig. 4). HlgA was the only S component presenting a little activity with LukD and all other class F components. LukF'-PV was the only class F component presenting activity with all class S proteins, although weak with LukE and LukM. LukF'-PV was the most potent F component with HlgC and with LukS-PV. The high S.D. value obtained for HlgC+LukF'-PV, particularly after a 30-min incubation, was due to variability of the lag-time preceding pore formation, according to the donor. In the presence of calcium, only few combinations as HlgA+HlgB or LukS-PV+LukF-PV formed pores and couples including LukE, LukM and LukD were devoid of activity (data not shown, [29]).

3.4.2. In vivo assays. Intradermal injections of leucotoxins in rabbit skin involving LukE, LukD and LukM, or LukF'-PV may induce a dermonecrotic reaction, according to the combinations. No variation was observed for rabbits receiving the same doses of leucotoxins (P < 0.05). Table 2 indicates that LukE+LukD was as potent as PVL, but the couples LukE+LukF-PV, LukM+LukD, LukE+LukF-PV, LukS-PV+LukD, HlgA+LukD, HlgC+LukD led to dermonecrotic lesions according to the doses and the delay after injections. Pairs containing HlgB, i.e. LukE+HlgB and LukM+HlgB were significantly less effective with no pus collection whatever the dose and the time following injections (P < 0.05). The

Table 2

Time course of dermonecrotic reactions after the intradermal injections of doses of functional leucotoxins including either LukE, or LukD or LukM proteins, and PVL as a reference [22]

Doses (ng)	100	100		300		1000			3000			10 000			
Time (h)	12	24	48	12	24	48	12	24	48	12	24	48	12	24	48
LukE+LukD	2	2	2	3	4	4	3	5	5	4	5	5	4	5	5
LukE+LukF-PV	1	1	2	2	3	3	3	4	5	3	5	5	4	5	5
LukE+HlgB	1	1	1	1	2	1	2	3	3	2	3	3	3	3	3
LukE+LukF'-PV	1	2	2	2	4	4	3	4	5	3	5	5	4	5	5
LukM+LukF'-PV	1	1	2	1	2	2	2	3	4	2	4	4	2	4	4
LukM+LukD	1	1	1	2	2	2	3	4	4	3	4	4	4	4	4
LukM+LukF-PV	1	1	1	1	2	3	2	3	4	2	4	4	3	4	5
LukM+HlgB	1	1	0	1	1	1	1	2	1	1	3	3	2	3	3
LukS+LukD	1	1	1	1	1	1	1	2	3	2	4	4	2	4	4
PVL	1	2	2	3	4	4	3	4	5	4	5	5	4	5	5
LukS-PV+LukF'-PV	1	2	2	2	3	3	2	4	4	3	4	5	4	4	5
HlgA+LukD	1	1	1	1	2	2	1	3	3	3	4	5	4	5	5
HlgA+LukF'-PV	1	1	1	1	1	2	2	2	3	2	3	4	2	3	4
HlgC+LukD	1	2	0	2	2	1	2	2	3	3	4	4	3	4	4
HlgC+LukF'-PV	1	1	1	1	1	2	2	2	3	2	3	3	2	3	4

Scores were attributed as described in Section 2.

reactions were always evolutive in the first 24 h, with slow healing from 5 days to 2 weeks, as previously reported [9,24].

# 4. Discussion

The new leucotoxin LukE+LukD characterized in this work remained undiscovered for two reasons. First, many of the LukE+LukD-producing S. aureus strains are able to produce simultaneously the  $\gamma$ -hemolysin whose components exhibit similar molecular masses and which cannot be distinguished by SDS-PAGE. The development of a defined medium like YCP medium was essential to determine the existence of LukE+LukD (Fig. 1A). Second, despite a marked dermonecrotic activity on rabbit skin, LukE+LukD is not strongly hemolytic or leucotoxic on human blood cells, as currently investigated for the other leucotoxins. HlgA+LukD was a Ca<sup>2+</sup> channel inducer, whereas LukE+HlgB appeared only working according to the donors (Table 1). These observations suggest either different cell receptors or different ways in the processes of oligomerization for bi-component leucotoxins. Study of their binding sites should discard one of these hypotheses.

Partial antigenic communities between class S and class F proteins (Fig. 1B) and sequence identities of LukE/LukD with the other bi-component leucotoxins indicate that it belongs to this family of toxins. The 30% distribution of clinical strains of *S. aureus* producing LukE+LukD was significantly different (P < 0.005) from those of PVL- and  $\gamma$ -hemolysin producers [8,21].

Staphylococcal α-toxin and bi-component leucotoxins were suggested to possess similar structures allowing pore formation, especially for the rim and for the transmembrane stem (residues 100-150 of LukE, LukD) domains [11,26]. However, these toxins have functional differences. Class S components bind first to membrane ligands followed by secondary binding of class F components [5,19,20]. Privileged cell targeting is observed for PVL only active on cells derived from the metamyelocyte [10], whereas HlgA+HlgB and other leucotoxins [24] are more or less active on both PMNs and erythrocytes, or may permeate synthetic membranes [20]. a-Toxin is known to be active upon erythrocytes, lymphocytes, keratinocytes and fibroblasts, but rather not on PMNs [1,3]. Thr30 of HlgA and Thr<sup>28</sup> of HlgC were important for the secondary binding of class F proteins [20]. The two cited residues and His<sup>35</sup> from  $\alpha$ -toxin also critical for its oligomerization [3,11,26], may be aligned. Therefore, oligomerization of these toxins may be mediated by differing residue interactions. Bicomponent leucotoxins induce Ca2+-channels [27], and form pores sensitive to monovalent cations. Other pore-forming toxins like  $\alpha$ -toxin or aerolysin were not reported as calcium inducers [1,3,17]. Na<sup>+</sup>, K<sup>+</sup>, and ethidium were exchanged through the pores formed by leucotoxins [27], whereas those formed by  $\alpha$ -toxin [3] and aerolysin [1] excluded ethidium and propidium.

Bi-component leucotoxins belong with  $\alpha$ -toxin, aerolysin, protective antigen of the anthrax toxin [17] to a super-family of  $\beta$ -sheet structured pore-forming toxins. Those toxins previously named synergohymenotropic toxins [28] should be recalled bi-component leucotoxins because they differ from toxins as A/B bacteriocins and diphtheria toxin that acquire their biological activity by a synergy with membranes. Acknowledgements: We thank F. Götz (Tübingen, Germany), T.J. Foster (Dublin, Ireland), Y. Kamio (Sendai, Japan), for kind gifts of the plasmid pCU1, the *S. aureus* Newman and P83 strains, respectively. Thanks are addressed to A. Staub from the IGBMC, Strasbourg, for protein sequencing. This work was supported by Grant EA1318 from the Direction de la Recherche et des Etudes Doctorales (DRED).

# References

- Abrami, L., Fivaz, M., Glauser, P.-E., Parton, R.G. and van der Goot, F.G. (1998) J. Cell Biol. 140, 525–540.
- [2] Augustin, J., Rosenstein, R., Wieland, B., Schneider, V., Schnell, N., Engelke, G., Entian, K.D. and Götz, F. (1992) Eur. J. Biochem. 204, 1149–1154.
- [3] Bhakdi, S., Bayley, H., Valeva, A., Walev, I., Walker, B., Weller, U., Kehoe, M. and Palmer, M. (1996) Arch. Microbiol. 165, 73– 79.
- [4] Cheung, A.L., Eberhardt, K.J. and Fischetti, A.A. (1994) Anal. Biochem. 222, 511–514.
- [5] Colin, D.A., Mazurier, I., Sire, S. and Finck-Barbançon, V. (1994) Infect. Immun. 62, 3184–3188.
- [6] Cooney, J., Kienle, Z., Foster, T. and O'Toole, P. (1993) Infect. Immun. 61, 768–771.
- [7] Couppié, P., Cribier, B., Prévost, G., Grosshans, E. and Piémont, Y. (1994) Arch. Dermatol. 130, 1208–1209.
- [8] Cribier, B., Prévost, G., Couppié, P., Finck, B.V., Grosshans, E. and Piémont, Y. (1992) Dermatology 185, 175–180.
- [9] Finck-Barbancon, V., Prévost, G. and Piémont, Y. (1991) Res. Microbiol. 142, 75–85.
- [10] Finck-Barbancon, V., Duportail, G., Meunier, O. and Colin, D.A. (1993) Biochim. Biophys. Acta 1182, 275–282.
- [11] Gouaux, J.E., Hobaugh, M. and Song, L. (1998) Protein Sci. 6, 2631–2635.
- [12] Gray, G.S. and Kehoe, M. (1984) Infect. Immun. 46, 615– 618.
- [13] Kamio, Y., Rahman, A., Nariya, H., Ozawa, T. and Izaki, K. (1993) FEBS Lett. 321, 15–18.
- [14] Kaneko, J., Muramoto, K. and Kamio, Y. (1997) Biosci. Biotech. Biochem. 61, 541–544.
- [15] König, B., Prévost, G., Piémont, Y. and König, W. (1995)
   J. Infect. Dis. 171, 607–613.
- [16] König, B., Prévost, G. and König, W. (1997) J. Med. Microbiol. 46, 479–485.
- [17] Lesieur, C., Vecsey-Semjen, B., Abrami, L., Fivaz, M. and van der Goot, F.G. (1997) Mol. Membr. Biol. 14, 45–64.
- [18] Mc Knight, S.L. and Kingsburry, R. (1982) Science 217, 316-324.
- [19] Meunier, O., Falkenrodt, A., Monteil, H. and Colin, D.A. (1995) Cytometry 21, 241–247.
- [20] Meunier, O. et al. (1997) Biochim. Biophys. Acta 1326, 275– 286.
- [21] Prévost, G., Couppié, P., Prévost, P., Gayet, S., Petiau, P., Cribier, B., Monteil, H. and Piémont, Y. (1995) J. Med. Microbiol. 42, 237–245.
- [22] Prévost, G., Cribier, B., Couppié, P., Petiau, P., Supersac, G., Finck-Barbançon, V., Monteil, H. and Piémont, Y. (1995) Infect. Immun. 63, 4121–4129.
- [23] Prévost, G., Bouakham, T., Piémont, Y. and Monteil, H. (1995) FEBS Lett. 376, 135–140.
- [24] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- [25] Siqueira, J.A., Speeg-Schatz, C., Freitas, F.I.S., Sahel, J., Monteil, H. and Prévost, G. (1997) J. Med. Microbiol. 46, 486– 494.
- [26] Song, L., Hobaugh, M.R., Shustak, C., Cheley, S., Bayley, H. and Gouaux, J.E. (1996) Science 274, 1859–1866.
- [27] Staali, L., Monteil, H. and Colin, D.A. (1998) J. Membr. Biol. 162, 209–216.
- [28] Supersac, G., Prévost, G. and Piémont, Y. (1993) Infect. Immun. 61, 580–587.
- [29] Woodin, A.M. (1960) Biochem. J. 75, 158-165.