

US007070959B1

(12) United States Patent

Papadopoulos et al.

(54) MODIFIED CHIMERIC POLYPEPTIDES WITH IMPROVED PHARMACOKINETIC PROPERTIES

- (75) Inventors: Nicholas J. Papadopoulos, Lagrangeville, NY (US); Samuel Davis, New York, NY (US); George D. Yancopoulos, Yorktown Heights, NY (US)
- (73) Assignee: Regeneron Pharmaceuticals, Inc., Tarrytown, NY (US)
- (*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

This patent is subject to a terminal disclaimer.

- (21) Appl. No.: 10/009,852
- (22) PCT Filed: May 23, 2000
- (86) PCT No.: PCT/US00/14142

§ 371 (c)(1), (2), (4) Date: Dec. 6, 2001

(87) PCT Pub. No.: WO00/75319

PCT Pub. Date: Dec. 14, 2000

Related U.S. Application Data

- (60) Provisional application No. 60/138,133, filed on Jun. 8, 1999.
- (51) Int. Cl.

(2006.01)
(2006.01)
(2006.01)
(2006.01)
(2006.01)

- (52) U.S. Cl. 435/69.7; 435/70.1; 435/71.1; 435/320.1; 435/325; 435/252.3; 530/387.3; 530/350; 536/23.4

(56) References Cited

U.S. PATENT DOCUMENTS

5,712,380	А		1/1998	Kendall et al.	
6,011,003	А		1/2000	Charnock-Jones et al.	
6,100,071	\mathbf{A}	琢	8/2000	Davis-Smyth et al	435/69.7

(10) Patent No.: US 7,070,959 B1 (45) Date of Patent: *Jul. 4, 2006

45) Date of Fatent. "Jul. 4, 2000

FOREIGN P.	ATENT I	DOCUMENTS
------------	---------	-----------

WO	WO97/44453	11/1997
WO	WO98/13071	4/1998
WO	WO99/03996	1/1999

OTHER PUBLICATIONS

Herley et al. (1999), Biochem. Biophys. Res. Comm. 262: 731-738.*

Terman, B. I., et al, "Identification of a new endothelial cell growth factor receptor tyrosine kinase", Oncogene (1991) 6:1677-1683.

Terman, B.I., et al, "Identification of the KDR tyrosine kinase as a receptor for vascular endothelial cell growth factor", Biochem Biophys Res Comm (1992) 187(3):1579-1586.

Tsutsumi, Y., et al, "PEGylation of interleukin-6 effectively increases its thrombopoietic potency", Thrombosis and Haemostasis (1997) 77(1):168-173.

Dunca, R. and Spreafico, F., "Polymer Conjugates", Drug Delivery Systems (1994) 27(4):290-306.

Hileman, R.E., et al., "Glycosaminoglycan-protein interactions: definitions of consensus sites in glycosaminoglycan binding proteins", BioEssays (1998) 20:156-167.

deVries, Carlie, et al., "The *fms*-like tyrosine kinase, a receptor for vascular endothelial growth factor", Science (1992) 225:989-991.

Sharifi, J., et al., "Improving monoclonal antibody pharmacokinetics via chemical modification", Quart J Nucl Med (1998) 42:242-249.

Jensen-Pippo, K.E., et al., "Enteral bioavailability of human granulocyte colony stimulating factor conjugated with poly-(ethylene glycol)", (1996) Pharm Res 13(1):102-107.

Tanaka, K., et al., "Characterization of the extracellular domain in vascular endothelial growth factor receptor-1 (Flt-1 Tyrosine kinase)", (1997) Jpn J Cancer Res 88:767-876.

Yang, J.C., et al., "The use of polyethylene glycol-modified interleukin-2 (PEG-IL-2) in the treatment of patients with metastatic renal cell carinoma and melanoma", (1995) Cancer 76(4): 687-694.

(Continued)

Primary Examiner—Lorraine Spector Assistant Examiner—Jon M Lockard

(74) Attorney, Agent, or Firm—Valeta Gregg, Esq.

(57) ABSTRACT

Modified chimeric polypeptides with improved pharmacokinetics are disclosed. Specifically, modified chimeric Flt1 receptor polypeptides that have been modified in such a way as to improve their pharmacokinetic profile are disclosed. Also disclosed are methods of making and using the modified polypeptides including but not limited to using the modified polypeptides to decrease or inhibit plasma leakage and/or vascular permeability in a mammal.

15 Claims, 55 Drawing Sheets

OTHER PUBLICATIONS

Davis-Smyth, T., et al., 1996, "The second immunoglobulinlike domain of the VEGF tyrosine kinase receptor Flt-1 determines ligand binding and may initiate a signal transduction-cascade", The EMBO Journal 15(18):4919-4927.

Holash, J., et al. (2002) PNAS, vol. 99, No. 17, pp. 11393-11398.

Heidaran, M.A., (1990) The Journal of Biological Chemistry, vol. 265, No. 31, Issue of Nov. 5, pp. 18741-18744.

Cunningham, S.A. et al., (1997) Biochemical and Biophysical Research Communications, vol. 231, pp. 596-599.

Fuh, G. et al. (1998) The Journal of Biological Chemistry,

Vol. 273, No. 18, Issue of May 1, pp. 11197-11204.
Wiesmann, C. et al. (1997) Cell, vol. 91, pp. 695-704.
Barleon, B. et al. (1997) The Journal of Biological Chemistry, vol. 272, No. 16, pp. 10382-10388.

Davis-Smyth, T. et al. (1998) The Journal of Biological Chemistry, vol. 273, No. 6, pp. 3216-3222.

* cited by examiner









U.S. Patent

Sheet 3 of 55







Fig.6B.



Fig.7.







Fig.10A.

		1	0			20			30			4	10			50			60
	*		*	*		*		*	*		*		*	*		*		*	*
ATG	GTC	AGC	TAC	TGG	GAC	ACC	GGG	GIC	CTG	CTG	TGC	GCG	CTG	CTC	AGC	TGT	CIG	CTT	CTC
TAC	CAG	TCG	ATG	ACC	CIG	TGG	CCC	CAG	GAC	GAC	ACG	CGC	GAC	GAG	TCG	ACA	GAC	GAA	GAG
Met	Val	Ser	Tyr	Trp	Asp	Thr	Gly	Val	Leu	Leu	Cys	Ala	Leu	Leu	Ser	Cys	Leu	Leu	Leu>
			_		_														
		7	0			80			90			10	00		1	10			120
	*		*	*		*		*	*		*		*	*		*		*	*
ACA	GGA	TCT	AGT	TCA	GGT	TCA	ААА	TTA	AAA	GAT	CCT	gaa	CTG	AGT	TTA	aaa	GGC	ACC	CAG
TGT	CCT	AGA	TCA	AGT	CCA	AGT	TTT	TAA	TTT	CTA	GGA	CTT	GAC	TCA	AAT	TTT	CCG	TGG	GTC
Thr	Gly	Ser	Ser	Ser	Gly	Ser	Lys	Leu	Lys	Asp	Pro	Glu	Leu	Ser	Leu	Lys	Gly	Thr	Gln>
		13	0		1	.40			150			10	60			.70			180
	*		*	*		*		*	*		*		*	*				*	
CAC	ATC	ATG	CAA	GCA	GGC	CAG	ACA	CIG	CAT	CIC	CAA	TGC	AGG	GGG	GAA	GCA	GCC	CAT	AAA
GIG	TAG	TAC	GTT	CGT	CCG	GIC	TGT	GAC	GTA	GAG	GTT	ACG	TCC	ccc	CTT	CGT	CGG	GIA	TTT
His	Ile	Met	Gln	Ala	Gly	Gln	Thr	Leu	His	Leu	Gln	Cys	Arg	Gly	Glu	Ala	Ala	HIS	råz>
					_							~	~~						240
		19	¥0			200			210		-	2	20			230		•	240
	*		*	*		*		*					-		3/73	200		-	000
TGG	TCT	TIG	CCT	GAA	ATG	GIG	AGT	AAG	GAA	AGC	GAA	AGG	CIG	AGC	MIA	ACT	AAA mmm	ACA	000
ACC	AGA	AAC	GGA	CTT	TAC	CAC	ICA	TIC	CIT	106	CIT	100	GAC	100	TIO	TGA	Tim	Cor	21a2
зтр	ser	Leu	Pro	GIU	Met	vai	ser	ьys	Giu	Ser	GIU	ALG	Deu	Ser	116	****	DYD	Jer	nia-
		25	50			260			270			2	80			290			300
	*	23	*	*		*		*	*		*	~	*	*		*		*	*
ጥርጥ	CCA	2022		000	***	CIA	ጥጥን	TCC.	AGT	ACT	тта	ACC	TTG	AAC	ACA	GCT	CAA	GCA	AAC
JCP	CCT	TOT I	7772	000	1000	CTT	ANG	ACG	TCA	TIGA	AAT	TGG	AAC	TIG	TGT	CGA	GTT	CGT	TTG
<u>6</u>	Glv	2101	len	Cly	TAVE	Gin	Phe	Cvs	Ser	Thr	Leu	Thr	Leu	Asn	Thr	Ala	Gln	Ala	Asn>
-1-	423			973	2,2	•=		415											
		3	10			320			330			3	40			350			360
	*		*	*		*		*	*		*		*	*		*		*	*
CAC	ACT	GGC	TTC	TAC	AGC	TGC	AAA	TAT	CTA	GCT	GTA	CCI	ACT	TCA	AAG	AAC	AAG	GAA	ACA
GTG	TGA	CCG	AAG	ATG	TCG	ACG	TTT	ATA	GAT	CGA	CAT	GGZ	I TGA	AGT	TIC	TTC	TTC	CTT	TGT
His	Thr	Gly	Phe	Tyr	Ser	Cys	Lys	Тут	: Leu	Ala	Val	Pro) Thr	Ser	Lys	Lys	Lys	Glu	Thr>
		3	70			380			390)		4	100			410			420
	*		*	*		*		*	1	r	*		*	*		*		*	*
GAA	TCT	GCA	ATC	TAT	ATA	TTI	ATI	AG	GA	ACA	GG1	AG2	A CC3	TIC	GT	GAC	ATC	TAC	AGT
CTT	AGA	CGI	TAG	ATA	TAT	AAA	TAA	TC2	A CTI	I TG3	r ccz	A TC:	r GGZ	AAG	CAT	CT	: TAC	ATC	5 TCA
Glu	Ser	Ala	Ile	Туг	lle	Phe	: Ile	e Sei	r Ası	> Thi	G13	/ Arg	g Pro	o Phe	va:	L GI	1 Met	- TYI	ser>
																470			490
		4	30			440			45)			460		k.	4/0		* .	400
~ ~ ~																	- 1972	- 00	2 (2007)
GAA	ATC		GAA	ATT	A12		ATA :	AC	r GAL	n ~~~	H AG(3 GAU 7 000			2 003		2 204		, CP7
CPI	TAG	. 666	CT	144	A TA:	r GIQ	; TAC	- 16			r 100			່ນໃດ	т. 1 т.	n. 000 n. Dm	- 00	5 GC.	v Val>
GTG	1 116	PIC	GIU	116	3 176	s 111	Met	. 110		n ori	A WT	9 91	T DG	u ve.		C FL	0 CY	- nu	
			190			500			51	0			520			530			540
	*		*		k	*		*		*	*		*		*	*		*	*
ACC			• AAO	TA S		r gr	r ac	r TT	A AA	A A A	G TT	т сс	а ст	T GA	CAC	T TT	GAT	c cc	T GAT
TG	AG	GGI	TT	TA	JTG	A CA	ATG	A A A	T TT	T TT	CAA	A GG	T GA	A CT	G TG	a aa	с та	G GG	a cta
Th	Sei	Pro) Asi	1 110	e Th	r Va	1 Th	r Le	u Ly	s Ly	s Ph	e Pr	o Le	u Asj	p Th	r Le	ù 11	e Pro	o Asp>

Fig.10B.

GGA AAA CGC ATA ATC TGG GAC AGT AGA AAG GGC TTC ATC ATA TCA AAT GCA ACG TAC AAA CCT TTT GCG TAT TAG ACC CTG TCA TCT TTC CCG AAG TAG TAT AGT TTA CGT TGC ATG TTT Gly Lys Arg Ile Ile Trp Asp Ser Arg Lys Gly Phe Ile Ile Ser Asn Ala Thr Tyr Lys> * * GAA ATA GGG CTT CTG ACC TGT GAA GCA ACA GTC AAT GGG CAT TTG TAT AAG ACA AAC TAT CTT TAT CCC GAA GAC TGG ACA CTT CGT TGT CAG TTA CCC GTA AAC ATA TTC TGT TTG ATA Glu Ile Gly Leu Leu Thr Cys Glu Ala Thr Val Asn Gly His Leu Tyr Lys Thr Asn Tyr> CTC ACA CAT OGA CAA ACC AAT ACA ATC ATA GAT GTC CAA ATA AGC ACA CCA CGC CCA GTC GAG TGT GTA GCT GTT TGG TTA TGT TAG TAT CTA CAG GTT TAT TCG TGT GGT GCG GGT CAG Leu Thr His Arg Gln Thr Asn Thr Ile Ile Asp Val Gln Ile Ser Thr Pro Arg Pro Val> AAA TTA CTT AGA GGC CAT ACT CTT GTC CTC AAT TGT ACT GCT ACC ACT CCC TTG AAC ACG TTT AAT GAA TCT CCG GTA TGA GAA CAG GAG TTA ACA TGA CGA TGG TGA GGG AAC TTG TGC Lys Leu Leu Arg Gly His Thr Leu Val Leu Asn Cys Thr Ala Thr Thr Pro Leu Asn Thr> AGA GTT CAA ATG ACC TGG AGT TAC CCT GAT GAA AAA AAT AAG AGA GCT TOC GTA AGG CGA TCT CAA GIT TAC TGG ACC TCA ATG GGA CTA CTT TTT TTA TTC TCT CGA AGG CAT TCC GCT Arg Val Gln Met Thr Trp Ser Tyr Pro Asp Glu Lys Asn Lys Arg Ala Ser Val Arg Arg> CGA ATT GAC CAA AGC AAT TCC CAT GCC AAC ATA TTC TAC AGT GTT CTT ACT ATT GAC AAA GCT TAA CTG GTT TCG TTA AGG GTA OGG TTG TAT AAG ATG TCA CAA GAA TGA TAA CTG TTT Arg Ile Asp Gln Ser Asn Ser His Ala Asn Ile Phe Tyr Ser Val Leu Thr Ile Asp Lys> ATG CAG AAC AMA GAC AMA GGA CTT TAT ACT TGT CGT GTA AGG AGT GGA CCA TCA TTC AMA TAC GTC TTG TTT CTG TTT CCT GAA ATA TGA ACA GCA CAT TCC TCA CCT GGT AGT AAG TTT Met Gln Asn Lys Asp Lys Gly Leu Tyr Thr Cys Arg Val Arg Ser Gly Pro Ser Phe Lys>

TCT GTT AAC ACC TCA GTG CAT ATA TAT GAT AAA GCA GGC CCG GGC GAG CCC AAA TCT TGT AGA CAA TTG TGG AGT CAC GTA TAT ATA CTA TTT CGT CCG GGC CCG CTC GGG TTT AGA ACA Ser Val Asn Thr Ser Val His Ile Tyr Asp Lys Ala Gly Pro Gly Glu Pro Lys Ser Cys>

Fig. 10C.

 1090
 1100
 1110
 1120
 1130
 1140

 *
 *
 *
 *
 *
 *
 *
 *

 TTC CTC TTC CCC CCA AAA CCC AAG GAC ACC CTC ATG ATC TCC CGG ACC CCT GAG GTC ACA

 AAG GAG AAG GGG GGT TTT GGG TTC CTG TGG GAG TAC TAG AGG GCC TGG GGA CTC CAG TGT

 Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr>

121012201230124012501260*********GGC GTG GAG GTG CAT AAT GCC AAG ACA AAG CCG CGG GAG GAG CAG TAC AAC AGC ACG TACCCG CAC CTC CAC GTA TAA CGG TTC TGT TTC GGC GCC CTC CTC GTC ATG TTG TCG TGC ATGGly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Glu Gln Tyr Asn Ser Thr Tyr>

133013401350136013701380*********TGC AAG GTC TCC AAC AAA GCC CTC CCA GCC CCC ATC GAG AAA ACC ATC TCC AAA GCC AAAACG TTC CAG AGG TTG TTT CGG GAG GGT CGG GGG TAG CTC TTT TGG TAG AGG TTT CGG TTTCys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys>

145014601470148014901500********AAC CAG GTC AGC CTG ACC TGC CTG GTC AAA GGC TTC TAT CCC AGC GAC ATC GCC GTG GAGTTG GTC CAG TCG GAC TGG ACG GAC CAG TTT CCG AAG ATA GGG TCG CTG TAG CGG CAC CTCAsn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu>

 1510
 1520
 1530
 1540
 1550
 1560

 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *

Fig.10D.

163016401650166016701680*********AAC GTC TTC TCA TGC TCC GTG ATG CAT GAG GCT CTG CAC AAC CAC TAC ACG CAG AAG AGCTTG CAG AAG AGT ACG AGG CAC TAC GTA CTC CGA GAC GTG TTG GTG ATG TGC GTC TTC TCGAsn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser>

1690 1700 * * * * CTC TCC CTG TCT CCG GGT AAA TGA GAG AGG GAC AGA GGC CCA TTT ACT Leu Ser Leu Ser Pro Gly Lys ***>





U.S. Patent Jul. 4, 2006

Fig.13A.

		:	10			20			30			4	0			50			60
	*		*	*		*		*	*		*		*	*		*		*	*
ATG	GTC	AGC	TAC	TGG	GAC	ACC	GGG	GTC	CTG	CTG	TGC	GCG	CIG	CTC	AGC	TGT	CTG	CTT	CTC
TAC	CAG	TCG	ATG	ACC	CTG	TGG	CCC	CAG	GAC	GAC	ACG	CGC	GAC	GAG	TCG	ACA	GAC	GAA	GAG
Met	Val	Ser	Tyr	Trp	Asp	Thr	Gly	Val	Leu	Leu	Cys	Ala	Leu	Leu	Ser	Cys	Leu	Leu	Leu>
						~~			~~			10				10			120
			70	-		80			90		*	Τ¢	*			*		*	*
101	~~~	-	*		000			- 		~~~~	~~~~	~~~~	~	- 			000	200	CNG
ACA	GGA	TCT	AGT	ACA	COL	ACT	mmm	11A	mm	CON	CON	CUM	CIG	1001	220	(VIII)	000	700	CTC
TGT	CCT 01	AGA	Com	Com	CLA	Cor	111	Lou	Lim	Ann	Dro	CIN	Lou	Ser	Ten	TATE	Glv	Thr	Gln>
1111	GTÀ	Der	Ser	Der.	GIY	261	nys	Deu	цуз	лар		920	200		200	- 1-	,		
		13	30		1	40			150			16	50		;	70			180
	٠		*	*	-	*		*	*		*		*	*		*		*	*
CAC	ATC	ATG	CAA	GCA	GGC	CAG	ACA	CIG	CAT	CTC	CAA	TGC	AGG	GGG	GAA	GCA	GCC	CAT	ааа
GTG	TAG	TAC	GTT	CGT	CCG	GTC	TGT	GAC	GTA	GAG	GTT	ACG	TCC	ccc	CTT	CGT	CGG	GTA	TTT
His	Ile	Met	Gln	Ala	Gly	Gln	Thr	Leu	His	Leu	Gln	Cys	Arg	Gly	Glu	Ala	Ala	His	Lys>
		19	90		2	200			210			22	20		:	230			240
	*		*	*		*		*	*		*		*	*		*		*	*
TGG	TCT	TTG	CCT	GAA	ATG	GTG	AGT	AAG	GAA	AGC	GAA	AGG	CTG	AGC	ATA	ACT	AAA	TCT	GCC
ACC	AGA	AAC	GGA	CIT	TAC	CAC	TCA	TTC	CTT	ICG	CTT	TCC	GAC	TCG	TAT	TGA	TTT	AGA	CGG
Trp	Ser	Leu	Pro	Glu	Met	Val	Ser	Lys	Glu	Ser	Glu	Arg	Leu	Ser	He	Inr	LYS	ser	AI9>
		2	60			260			270			25	20			290			300
	*	2	*	*		*		*	2/0		*	20	*	*		*		*	*
TCT	GGA	AGA	b b tr	GGC	282	CAA	- Triner	na.	AGT	ACT	тта	ACC	TTG	AAC	ACA	GCT	CAA	GCA	AAC
ACA	CCT	TOT	TTA	CCG	TTT	GTT	AAG	ACG	TCA	TGA	AAT	TGG	AAC	TIG	TGT	CGA	GTT	CGT	TTG
Cvs	Glv	Arg	Asn	Glv	Lvs	Gln	Phe	Cvs	Ser	Thr	Leu	Thr	Leu	Asn	Thr	Ala	Gln	Ala	Asn>
								-											
		33	10		3	320			330			34	40		:	350			360
	*		*	*		*		*	*		*		*	*		*		*	*
CAC	ACT	GGC	TTC	TAC	AGC	TGC	aaa	TAT	CTA	GCT	GTA	CCT	ACT	TCA	AAG	AAG	AAG	GAA	ACA
GTG	TGA	CCG	AAG	ATG	TCG	ACG	$\mathbf{T}\mathbf{T}\mathbf{T}$	АТА	GAT	CGA	CAT	GGA	TGA	AGT	TIC	TTC	TTC	CTT	TGT
His	Thr	Gly	Phe	Туг	Ser	Cys	Lys	Tyr	Leu	Ala	Val	Pro	Thr	Ser	Lys	Lys	Lys	GIU	Thr>
		-				200			200				00			410			420
	*	2	*	*		*			390		*		*	*		*		*	*
GAP	TYCTP	GCA	200	ጥልጥ	ልጥል	TTT	አጥጥ	AGT	GAT	ACA	GGT	AGA	CCT	TIC	GTA	GAG	ATG	TAC	AGT
CTT	AGA	CGT	TAG	ATA	TAT	AAA	TAA	TCA	CTA	TGT	CCA	TCT	GGA	AAG	CAT	CTC	TAC	ATG	TCA
Glu	Ser	Ala	Ile	Tyr	Ile	Phe	Ile	Ser	Asp	Thr	Gly	Arg	Pro	Phe	Val	Glu	Met	Tyr	ser>
				•					•		-	-							
		4	30			440			450			4	60			470			480
	*		*	*		*		*	*		*		*	*		*		*	*
GAA	100	CCC	GAA	ATT	ATA	CAC	ATG	ACT	GAA	GGA	AGG	GAG	CIC	GTC	ATT	CCC	TGC	CGG	GTT
CTT	AIC									000	-	$-\pi$	~~~	~~~					- 17 B
	TAG	GGG	CTT	TAA	TAT	GTG	TAC	TGA	CTT	CCT	1CC	CIC	GAG	LAG	TAA	GGG	ACG	GCC	UNA UNA
Glu	TAG Ile	GGG Pro	CTT Glu	TAA Ile	TAT Ile	GTG His	TAC Met	TGA Thr	CTT Glu	Gly	Arg	Glu	Leu	Val	Ile	Pro	ACG Cys	GCC Arg	Val>
Glu	TAG Ile	GGG Pro	Glu	TAA Ile	TAT Ile	GTG His	TAC Met	TGA Thr	Glu	Gly	Arg	Glu	Leu 20	Val	Ile	Pro	ACG Cys	GCC Arg	Val>
Glu	ATC TAG Ile	GGG Pro	Glu 90	TAA Ile	TAT Ile	GTG His 500	TAC Met	TGA Thr	Glu 510	Gly	Arg	Glu 5	Leu 20	Val	Ile	530	ACG Cys	GCC Arg	Val>

TGC AGT GGA TTG TAG TGA CAA TGA AAT TTT TTC AAA GGT GAA CTG TGA AAC TAG GGA CTA Thr Ser Pro Asn Ile Thr Val Thr Leu Lys Lys Phe Pro Leu Asp Thr Leu Ile Pro Asp>

Fig.13B.

550 560 570 580 590 600 GGA AAA CSC ATA ATC TSG GAC AGT AGA AAG CSC TTC ATC ATA TCA AAT GCA ACG TAC AAA CCT TTT GOG TAT TAG ACC CTG TCA TCT TTC CCG AAG TAG TAT AGT TTA CGT TGC ATG TTT Gly Lys Arg Ile Ile Trp Asp Ser Arg Lys Gly Phe Ile Ile Ser Asn Ala Thr Tyr Lys> 640 650 610 620 630 660 * * * * * GAA ATA GOG CTT CTG ACC TGT GAA GCA ACA GTC AAT GGG CAT TTG TAT AAG ACA AAC TAT CTT TAT CCC GAA GAC TGG ACA CTT CGT TGT CAG TTA CCC GTA AAC ATA TTC TGT TTG ATA Glu Ile Gly Leu Leu Thr Cys Glu Ala Thr Val Asn Gly His Leu Tyr Lys Thr Asn Tyr> 670 680 690 700 710 720 ٠ * * * * * * CTC ACA CAT OGA CAA ACC AAT ACA ATC ATA GAT GTC CAA ATA AGC ACA CCA CGC CCA GTC GAG TGT GTA GCT GTT TGG TTA TGT TAG TAT CTA CAG GTT TAT TCG TGT GGT GCG GGT CAG Leu Thr His Arg Gln Thr Asn Thr Ile Ile Asp Val Gln Ile Ser Thr Pro Arg Pro Val> 760 770 780 750 730 740 * * * * * * AAA TTA CTT AGA GGC CAT ACT CTT GTC CTC AAT TGT ACT GCT ACC ACT CCC TTG AAC ACG TTT AAT GAA TCT COG GTA TGA GAA CAG GAG TTA ACA TGA CGA TGG TGA GGG AAC TTG TGC Lys Leu Leu Arg Gly His Thr Leu Val Leu Asn Cys Thr Ala Thr Thr Pro Leu Asn Thr> 790 800 810 820 830 840 * * * * * AGA GTT CAA ATG ACC TGG AGT TAC CCT GAT GAA ATT GAC CAA AGC AAT TCC CAT GCC AAC TCT CAA GTT TAC TOG ACC TCA ATG OGA CTA CTT TAA CTG GTT TOG TTA AGG GTA CGG TTG Arg Val Gln Met Thr Trp Ser Tyr Pro Asp Glu Ile Asp Gln Ser Asn Ser His Ala Asn> 890 900 850 870 880 860 * * ٠ * * * * * ATA TTC TAC AGT GTT CTT ACT ATT GAC AAA ATG CAG AAC AAA GAC AAA GGA CTT TAT ACT . TAT AAG ATG TCA CAA GAA TGA TAA CTG TTT TAC GTC TTG TTT CTG TTT CCT GAA ATA TGA Ile Phe Tyr Ser Val Leu Thr Ile Asp Lys Met Gln Asn Lys Asp Lys Gly Leu Tyr Thr> 940 950 960 920 910 930 * * * * * * * * TGT CGT GTA AGG AGT GGA CCA TCA TTC AAA TCT GTT AAC ACC TCA GTG CAT ATA TAT GAT ACA GCA CAT TCC TCA CCT GGT AGT AAG TTT AGA CAA TTG TGG AGT CAC GTA TAT ATA CTA Cys Arg Val Arg Ser Gly Pro Ser Phe Lys Ser Val Asn Thr Ser Val His Ile Tyr Asp> 970 990 1000 1010 1020 980 * * * * * AAA GCA GGC CCG GGC GAG CCC AAA TCT TGT GAC AAA ACT CAC ACA TGC CCA COG TGC CCA TTT CGT CCG GGC CCG CTC GGG TTT AGA ACA CTG TTT TGA GTG TGT ACG GGT GGC ACG GGT Lys Ala Gly Pro Gly Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro> 1030 1040 1050 1060 1070 1080 * GCA CCT GAA CTC CTG GGG GGA CCG TCA GTC TTC CTC TTC CCC CCA AAA CCC AAG GAC ACC CGT GGA CTT GAG GAC CCC CCT GGC AGT CAG AAG GAG AAG GGG GGT TTT GGG TTC CTG TGG

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Arp Thr>

Fig.13C.

115011601170118011901200*********CCT GAG GTC AAG TTC AAC TGG TAC GTG GAG GGC GTG GAG GTG CAT AAT GCC AAG ACA AAGGGA CTC CAG TTC AAG TTG ACC ATG CAC CTG CCG CAC CTC CAC GTA TTA CGG TTC TGT TTCPro Glu Val LysPhe Asn Trp Tyr Val Asp Gly Val Glu Val HisAsn Ala LysThr Lys>

121012201230124012501260********CCG CGG GAG GAG GAG CAG TAC AAC AGC ACG TAC CGT GTG GTC AGC GTC CTC ACC GTC CTG CACGGC GCC CTC CTC GTC ATG TTG TCG TGC ATG GCA CAC CAG TCG CAG GAG TGG CAG GAC GTGPro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His>

 1510
 1520
 1530
 1540
 1550
 1560

 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *
 *

Fig.13D.

		163	30		16	540		1	1650			166	50		16	570	
	*		*	*		*		*	*		*		*	*		*	
GCT	CTG	CAC	AAC	CAC	TAC	ACG	CAG	AAG	AGC	CTC	TCC	CTG	TCT	CCG	GGT	AAA	TGA
CGA	GAC	GTG	TTG	GTG	ATG	TGC	GTC	TTC	TCG	GAG	AGG	GAC	AGA	GGC	CCA	$\mathbf{T}\mathbf{T}\mathbf{T}$	ACT
Ala	Leu	His	Asn	His	Tyr	Thr	Gln	Lys	Ser	Leu	Ser	Leu	Ser	\mathbf{Pro}	Gly	Lys	***>

Fig.14A.

40 50 60 30 ATG GTC AGC TAC TGG GAC ACC GGG GTC CTG CTG TGC GCG CTG CTC AGC TGT CTG CTT CTC TAC CAG TCG ATG ACC CTG TGG CCC CAG GAC GAC ACG CGC GAC GAG TCG ACA GAC GAA GAG Met Val Ser Tyr Trp Asp Thr Gly Val Leu Leu Cys Ala Leu Leu Ser Cys Leu Leu> 120 110 70 80 90 100 ACA GGA TCT AGT TCC GGA GGT AGA CCT TTC GTA GAG ATG TAC AGT GAA ATC CCC GAA ATT TGT CCT AGA TCA AGG CCT CCA TCT GGA AAG CAT CTC TAC ATG TCA CTT TAG GGG CTT TAA Thr Gly Ser Ser Ser Gly Gly Arg Pro Phe Val Glu Met Tyr Ser Glu Ile Pro Glu Ile> 170 180 160 130 150 140 ATA CAC ATG ACT GAA GGA AGG GAG CTC GTC ATT CCC TGC CGG GTT ACG TCA CCT AAC ATC TAT GTG TAC TGA CTT CCT TCC CTC GAG CAG TAA GGG ACG GCC CAA TGC AGT GGA TTG TAG Ile His Met Thr Glu Gly Arg Glu Leu Val Ile Pro Cys Arg Val Thr Ser Pro Asn Ile> 240 220 230 190 200 210 ACT GTT ACT TTA AAA AAG TTT CCA CTT GAC ACT TTG ATC CCT GAT GGA AAA CGC ATA ATC TGA CAA TGA AAT TTT TTC AAA GGT GAA CTG TGA AAC TAG GGA CTA CCT TTT GCG TAT TAG Thr Val Thr Leu Lys Lys Phe Pro Leu Asp Thr Leu Ile Pro Asp Gly Lys Arg Ile Ile> 300 280 290 250 260 270 * TGG GAC AGT AGA AAG GGC TTC ATC ATA TCA AAT GCA ACG TAC AAA GAA ATA GGG CTT CTG ACC CTG TCA TCT TTC CCG AAG TAG TAT AGT TTA CGT TGC ATG TTT CTT TAT CCC GAA GAC Trp Asp Ser Arg Lys Gly Phe Ile Ile Ser Asn Ala Thr Tyr Lys Glu Ile Gly Leu Leu> 360 340 350 310 320 330 ACC TGT GAA GCA ACA GTC AAT GGG CAT TTG TAT AAG ACA AAC TAT CTC ACA CAT CGA CAA TOG ACA CTT CGT TGT CAG TTA CCC GTA AAC ATA TTC TGT TTG ATA GAG TGT GTA GCT GTT Thr Cys Glu Ala Thr Val Asn Gly His Leu Tyr Lys Thr Asn Tyr Leu Thr His Arg Gln> 420 410 400 390 370 380 * ACC AAT ACA ATC ATA GAT GTC CAA ATA AGC ACA CGA CGC CCA GTC AAA TTA CTT AGA GGC TGG TTA TGT TAG TAT CTA CAG GTT TAT TCG TGT GGT GCG GGT CAG TTT AAT GAA TCT CCG Thr Asn Thr Ile Ile Asp Val Gln Ile Ser Thr Pro Arg Pro Val Lys Leu Leu Arg Gly> 480 470 450 460 430 440 * ٠ * CAT ACT CTT GTC CTC AAT TGT ACT GCT ACC ACT CCC TTG AAC ACG AGA GTT CAA ATG ACC GTA TGA GAA CAG GAG TTA ACA TGA CGA TGG TGA GGG AAC TTG TGC TCT CAA GTT TAC TGG His Thr Leu Val Leu Asn Cys Thr Ala Thr Thr Pro Leu Asn Thr Arg Val Gln Met Thr> 530 540 520 490 500 510 *

TGG AGT TAC CCT GAT GAA ATT GAC CAA AGC AAT TCC CAT GCC AAC ATA TTC TAC AGT GTT ACC TCA ATG GGA CTA CTT TAA CTG GTT TCG TTA AGG GTA CGG TTG TAT AAG ATG TCA CAA Trp Ser Tyr Pro Asp Glu Ile Asp Gln Ser Asn Ser His Ala Asn Ile Phe Tyr Ser Val>

Fig.14B.

CTT ACT ATT GAC AAA ATG CAG AAC AAA GAC AAA GGA CTT TAT ACT TGT CGT GTA AGG AGT GAA TGA TAA CTG TTT TAC GTC TTG TTT CTG TTT CCT GAA ATA TGA ACA GCA CAT TCC TCA Leu Thr Ile Asp Lys Met Gln Asn Lys Asp Lys Gly Leu Tyr Thr Cys Arg Val Arg Ser> GGA CCA TCA TTC ANA TCT GTT AAC ACC TCA GTG CAT ATA TAT GAT ANA GCA GGC CCG GGC CCT GGT AGT AAG TTT AGA CAA TTG TGG AGT CAC GTA TAT ATA CTA TTT CGT COG GGC CCG Gly Pro Ser Phe Lys Ser Val Asn Thr Ser Val His Ile Tyr Asp Lys Ala Gly Pro Gly> GAG CCC AAA TCT TGT GAC AAA ACT CAC ACA TGC CCA CCG TGC CCA GCA CCT GAA CTC CTG CTC GGG TTT AGA ACA CTG TTT TGA GTG TGT ACG GGT GGC ACG GGT CGT GGA CTT GAG GAC Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu> GGG GGA CCG TCA GTC TTC CTC TTC CCC CCA AAA CCC AAG GAC ACC CTC ATG ATC TCC CGG CCC CCT GGC AGT CAG AAG GAG AAG GGG GGT TTT GGG TTC CTG TGG GAG TAC TAG AGG GCC Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg> ACC CCT GAG GTC ACA TGC GTG GTG GTG GAC GTG AGC CAC GAA GAC CCT GAG GTC AAG TTC TGG GGA CTC CAG TGT ACG CAC CAC CAC CTG CAC TCG GTG CTT CTG GGA CTC CAG TTC AAG Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe> AAC TGG TAC GTG GAC GGC GTG GAG GTG CAT AAT GCC AAG ACA AAG CCG CGG GAG GAG CAG TTG ACC ATG CAC CTG CCG CAC CTC CAC GTA TTA CGG TTC TGT TTC GGC GCC CTC CTC GTC Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln> TAC AAC AGC ACG TAC CGT GTG GTC AGC GTC CTC ACC GTC CTG CAC CAG GAC TGG CTG AAT ATG TTG TCG TGC ATG GCA CAC CAG TCG CAG GAG TGG CAG GAC GTG GTC CTG ACC GAC TTA Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn> GGC AAG GAG TAC AAG TGC AAG GTC TCC AAC AAA GCC CTC CCA GCC CCC ATC GAG AAA ACC CCG TTC CTC ATG TTC ACG TTC CAG AGG TTG TTT CGG GAG GGT CGG GGG TAG CTC TTT TGG Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr> ATC TCC AAA GCC AAA GGG CAG CCC CGA GAA CCA CAG GTG TAC ACC CTG CCC CCA TCC CGG TAG AGG TTT CGG TTT CCC GTC GGG GCT CTT GGT GTC CAC ATG TGG GAC GGG GGT AGG GCC

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg>

Fig.14C.

1140 1130 1090 1120 1100 1110 GAT GAG CTG ACC AAG AAC CAG GTC AGC CTG ACC TGC CTG GTC AAA GGC TTC TAT CCC AGC CTA CTC GAC TGG TTC TTG GTC CAG TCG GAC TGG ACG GAC CAG TTT CCG AAG ATA GGG TCG Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser> 1190 1200 1170 1180 1150 1160 . * ٠ × GAC ATC GCC GTG GAG TGG GAG AGC AAT GGG CAG CCG GAG AAC AAC TAC AAG ACC ACG CCT CTG TAG CGG CAC CTC ACC CTC TCG TTA CCC GTC GGC CTC TTG TTG ATG TTC TGG TGC GGA Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro> 1250 1260 1240 1210 1220 1230 * * ٠ * ٠ * CCC GTG CTG GAC TCC GAC GGC TCC TTC TTC CTC TAC AGC AAG CTC ACC GTG GAC AAG AGC GGG CAC GAC CTG AGG CTG CCG AGG AAG AAG GAG ATG TCG TTC GAG TGG CAC CTG TTC TCG Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser> 1300 1310 1320 1290 1270 1280 * ٠ *

AGG TGG CAG CAG GGG AAC GTC TTC TCA TGC TCC GTG ATG CAT GAG GCT CTG CAC AAC CAC TCC ACC GTC GTC CCC TTG CAG AAG AGT ACG AGG CAC TAC GTA CTC CGA GAC GTG TTG GTG Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His>

1350 1330 1340 * TAC ACG CAG AAG AGC CTC TCC CTG TCT CCG GGT AAA TGA ATG TGC GTC TTC TCG GAG AGG GAC AGA GGC CCA TTT ACT Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys ***>

*

*

Fig.15A.

ATG GTC AGC TAC TGG GAC ACC GGG GTC CTG CTG TGC GOG CTG CTC AGC TGT CTG CTT CTC TAC CAG TOG ATG ACC CTG TGG CCC CAG GAC GAC ACG CGC GAC GAG TOG ACA GAC GAA GAG Met Val Ser Tyr Trp Asp Thr Gly Val Leu Leu Cys Ala Leu Leu Ser Cys Leu Leu Leu> ACA GGA TCT AGT TCC GGA GGT AGA CCT TTC GTA GAG ATG TAC AGT GAA ATC CCC GAA ATT TGT CCT AGA TCA AGG CCT CCA TCT GGA AAG CAT CTC TAC ATG TCA CTT TAG GOG CTT TAA Thr Gly Ser Ser Ser Gly Gly Arg Pro Phe Val Glu Met Tyr Ser Glu Ile Pro Glu Ile> ATA CAC ATG ACT GAA GGA AGG GAG CTC GTC ATT CCC TGC CGG GTT ACG TCA CCT AAC ATC TAT GTG TAC TGA CTT CCT TCC CTC GAG CAG TAA GGG ACG GCC CAA TGC AGT GGA TTG TAG Ile His Met Thr Glu Gly Arg Glu Leu Val Ile Pro Cys Arg Val Thr Ser Pro Asn Ile> * ACT GTT ACT TTA AAA AAG TTT CCA CTT GAC ACT TTG ATC CCT GAT GGA AAA CGC ATA ATC TGA CAA TGA AAT TTT TTC AAA GGT GAA CTG TGA AAC TAG GGA CTA CCT TTT GCG TAT TAG Thr Val Thr Leu Lys Lys Phe Pro Leu Asp Thr Leu Ile Pro Asp Gly Lys Arg Ile Ile> TGG GAC AGT AGA AAG GGC TTC ATC ATA TCA AAT GCA ACG TAC AAA GAA ATA GGG CTT CTG ACC CTG TCA TCT TTC CCG AAG TAG TAT AGT TTA CGT TGC ATG TTT CTT TAT CCC GAA GAC Trp Asp Ser Arg Lys Gly Phe Ile Ile Ser Asn Ala Thr Tyr Lys Glu Ile Gly Leu Leu> ACC TGT GAA GCA ACA GTC AAT GOG CAT TTG TAT AAG ACA AAC TAT CTC ACA CAT CGA CAA TGG ACA CTT CGT TGT CAG TTA CCC GTA AAC ATA TTC TGT TTG ATA GAG TGT GTA GCT GTT Thr Cys Glu Ala Thr Val Asn Gly His Leu Tyr Lys Thr Asn Tyr Leu Thr His Arg Gln> ACC AAT ACA ATC ATA GAT GTC CAA ATA AGC ACA CCA CGC CCA GTC AAA TTA CTT AGA GGC TGG TTA TGT TAG TAT CTA CAG GTT TAT TOG TGT GGT GCG GGT CAG TTT AAT GAA TCT COG Thr Asn Thr Ile Ile Asp Val Gln Ile Ser Thr Pro Arg Pro Val Lys Leu Leu Arg Gly> * * CAT ACT CTT GTC CTC AAT TGT ACT GCT ACC ACT CCC TTG AAC ACG AGA GTT CAA ATG ACC GTA TGA GAA CAG GAG TTA ACA TGA CGA TGG TGA GGG AAC TTG TGC TCT CAA GTT TAC TGG His Thr Leu Val Leu Asn Cys Thr Ala Thr Thr Pro Leu Asn Thr Arg Val Gln Met Thr> * TGG AGT TAC CCT GAT GAA AAA AAT AAG AGA GCT TCC GTA AGG CGA CGA ATT GAC CAA AGC ACC TCA ATG GGA CTA CTT TTT TTA TTC TCT CGA AGG CAT TCC GCT GCT TAA CTG GTT TCG Trp Ser Tyr Pro Asp Glu Lys Asn Lys Arg Ala Ser Val Arg Arg Arg Ile Asp Gln Ser>

Fig.15B.

		55	0		5	60			570			58	0		5	90			600
	*		*	*		*		*	*		*		*	*		*		*	*
AAT	TCC	CAT	GCC	AAC	ATA	TIC	TAC	AGT	GTT	CTT	ACT	ATT	GAC	ааа	ATG	CAG	AAC	ааа	GAC
TTA	AGG	GTA	CGG	TIG	TAT	AAG	ATG	TCA	CAA	GAA	TGA	TAA	CIG	TTT	TAC	GTC	TTG	TTT	CIG
Asn	Ser	His	Ala	Asn	Ile	Phe	Tyr	Ser	Val	Leu	Thr	Ile	Asp	Lys	Met	GIN	Asn	Lys	Asp>
		61	.0		6	20			630			64	0		e	50			660
	*		*	*		*		*	*		*		*	*		*		*	*
AAA	GGA	CTT	TAT	ACT	TGT	CGT	GTA	AGG	AGT	GGA	CCA	TCA	TTC	AAA	TCT	GTT	AAC	ACC	TCA
TTT	CCT	GAA	ATA	TGA	ACA	GCA	CAT	TCC	TCA	CCT	GGT	AGT	AAG	TTT	AGA	CAA	TIG	TGG	AGT
Lys	GTÀ	Leu	Tyr	Thr	Cys	Arg	Val	Arg	Ser	GIY	Pro	Ser	Phe	Lys	ser	Val	Asn	Thr	ser>
		67	70		e	80			690			70	0		7	10			720
	*		*	*		٠		*	*		*		*	*		*		*	*
GTG	CAT	ATA	TAT	GAT	AAA	GCA	GGC	CCG	GGC	GAG	ccc	aaa	TCT	TGT	GAC	AAA	ACT	CAC	ACA
CAC	GTA	TAT	ATA	CTA	TTT	CGT	CCG	GGC	CCG	CIC	GGG	TTT	AGA	ACA	CIG	TTT	TGA	GIG	TGT
Val	His	Ile	Tyr	Asp	Lys	Ala	Gly	Pro	Gly	Glu	Pro	Lys	Ser	Cys	Asp	Lys	Thr	His	Thr>
		73	30		7	40			750			76	50		•	770			780
	*		*	*		*		*	*		*		*	*		*		*	*
TGC	CCA	CCG	TGC	CCA	GCA	CCT	GAA	CIC	CTG	GGG	GGA	CCG	TCA	GTC	TTC	CIC	TTC	ccc	CCA
ACG	GGT	GGC	ACG	GGT	CGT	GGA	CTT	GAG	GAC	ccc	CCT	GGC	AGT	CAG	AAG	GAG	AAG	GGG	GGT
Cys	Pro	Pro	Cys	Pro	Ala	Pro	Glu	Leu	Leu	Gly	Gly	Pro	Ser	Val	Phe	Leu	Phe	Pro	Pro>
		79	90		8	300			810			82	20		1	330			840
	*		*	*		*		*	*		*		*	*		*		*	*
aaa	CCC	AAG	GAC	ACC	CTC	ATG	ATC	TCC	CGG	ACC	CCT	GAG	GIC	ACA	TGC	GIG	GTG	GTG	GAC
TTT	GGG	TTC	CTG	TGG	GAG	TAC	TAG	AGG	GCC	TGG	GGA	CIC	CAG	TGT	ACG	CAC	CAC	CAC	CTG
Lys	Pro	Lys	Asp	Thr	Leu	Met	Ile	Ser	Arg	Thr	Pro	Glu	Val	Thr	Cys	Val	VAL	Val	Asp>
		8	50		1	860			870			8	80			890			900
	*		*	*		*		*	*		*		*	*		*		*	*
GIG	AGC	CAC	GAA	GAC	CCT	GAG	GTC	AAG	TIC	AAC	TGG	TAC	GTG	GAC	GGC	GTG	GAG	GTG	CAT
CAC	TCG	GTG	CTT	CTG	GGA	CTC	CAG	TIC	aag	TIG	ACC	ATG	CAC	CTG	CCG	CAC	CIC	CAC	GTA
Val	Ser	His	Glu	Asp	Pro	Glu	Val	Lys	Phe	Asn	Trp	Tyr	Val	Asp	Gly	Val	Glu	Val	His>
		9	10			920			930			9	40			950			960
	*		*	*		*		*	*		*		*	*		*		*	*
AAT	GCC	AAG	ACA	AAG	CCG	CGG	GAG	GAG	CAG	TAC	AAC	AGC	ACG	TAC	CGT	GIG	GIC	AGC	GIC
TTA	CGG	TTC	TGT	TTC	GGC	GCC	CIC	CIC	GIC	ATG	TTG	TCG	TGC	ATG	GCA	CAC	CAG	TCG	CAG
Asn	Ala	Lys	Thr	Lys	Pro	Arg	Glu	Glu	Gln	Tyr	Asn	Ser	Thr	Tyr	Arg	Val	Val	Ser	Val>
		9	70			980			990			10	00		1	010			10 20
	*		*	*		*		*	*		*		*	*		*		*	*
CTC	ACC	GTC	CTG	CAC	CAG	GAC	TGG	CTG	AAT	GGC	AAG	GAG	TAC	AAG	TGC	AAG	GTC	TCC	AAC
GAG	TGG	CAG	GAC	GTG	GTC	CTG	ACC	GAC	TTA	CCG	TTC	CIC	ATG	TTC	ACG	TIC	CAG	AGG	TIG
Leu	Thr	Val	Leu	His	Gln	Asp	Trp	Leu	Asn	Gly	Lys	Glu	Tyr	Lys	Cys	Lys	Val	Ser	Asn>
		10	30		1	040			1050			10	60		1	070			1080
	*		*	*	-	*		*	*		*		*	*		*		*	*
AAA	GCC	CTC	CCA	GCC	ccc	ATC	GAG	AAA	ACC	ATC	TCC	Ала	GCC	ААА	GGC	CAG	ccc	CGA	GAA
TTT	CGG	GAG	GGT	CGG	GGG	TAG	CTC	TTI	TGG	TAG	AGG	TTT	CGG	TTT	· ccc	GTC	GGG	GCT	CTT

Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu>

Fig.15C.

109011001110112011301140*********CCA CAG GTG TAC ACC CTG CCC CCA TCC CGG GAT GAG CTG ACC AAG AAC CAG GTC AGC CTGGGT GTC CAC ATG TGG GAC GGG GGT AGG GCC CTA CTC GAC TGG TTC TTG GTC CAG TCG GACFrom the second second

1160 1170 1200 1150 1180 1190 ٠ * * * * * * * * ACC TGC CTG GTC AAA GGC TTC TAT CCC AGC GAC ATC GCC GTG GAG TGG GAG AGC AAT GGG TGG ACG GAC CAG TTT CCG AAG ATA GGG TCG CTG TAG CGG CAC CTC ACC CTC TCG TTA CCC Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly>

1250 1260 1210 1220 1230 1240 ٠ ٠ * * -* * * * CAG CCG GAG AAC AAC TAC AAG ACC ACG CCT CCC GTG CTG GAC TCC GAC GGC TCC TTC TTC GTC GGC CTC TTG TTG ATG TTC TGG TGC GGA GGG CAC GAC CTG AGG CTG CCG AGG AAG AAG Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe>

1300 1320 1270 1280 1290 1310 * * * * * * * * CTC TAC AGC AAG CTC ACC GTG GAC AAG AGC AGG TGG CAG CAG GGG AAC GTC TTC TCA TGC GAG ATG TOG TTC GAG TGG CAC CTG TTC TOG TCC ACC GTC CCC TTG CAG AAG AGT AOG Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys>

1330 1340 1350 1360 1370 1380 * * * * * * * * * TCC GTG ATG CAT GAG GCT CTG CAC AAC CAC TAC ACG CAG AAG AGC CTC TCC CTG TCT CCG AGG CAC TAC GTA CTC CGA GAC GTG TTG GTG ATG TGC GTC TTC TOG GAG AGG GAC AGA GGC Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro>

GGT AAA TGA CCA TIT ACT Gly Lys ***>

*

Fig.16A.

ATG GTC AGC TAC TOG GAC ACC GGG GTC CTG CTG TGC GCG CTG CTC AGC TGT CTG CTT CTC TAC CAG TCG ATG ACC CTG TGG CCC CAG GAC GAC ACG CGC GAC GAG TCG ACA GAC GAA GAG Met Val Ser Tyr Trp Asp Thr Gly Val Leu Leu Cys Ala Leu Leu Ser Cys Leu Leu Leu> ACA GGA TCT AGT TCA GGT TCA AAA TTA AAA GAT CCT GAA CTG AGT TTA AAA GGC ACC CAG TGT CCT AGA TCA AGT CCA AGT TTT AAT TTT CTA GGA CTT GAC TCA AAT TTT COG TGG GTC Thr Gly Ser Ser Ser Gly Ser Lys Leu Lys Asp Pro Glu Leu Ser Leu Lys Gly Thr Gln> CAC ATC ATG CAA GCA GGC CAG ACA CTG CAT CTC CAA TGC AGG GGG GAA GCA GCC CAT AAA CTG TAG TAC GTT CGT CCG GTC TGT GAC GTA GAG GTT ACG TCC CCC CTT CGT CGG GTA TTT His Ile Met Gln Ala Gly Gln Thr Leu His Leu Gln Cys Arg Gly Glu Ala Ala His Lys> TOG TOT TTG COT GAA ATG GTG AGT AAG GAA AGC GAA AGG CTG AGC ATA ACT AAA TOT GOO ACC AGA AAC GGA CTT TAC CAC TCA TTC CTT TCG CTT TCC GAC TCG TAT TGA TTT AGA CGG Trp Ser Leu Pro Glu Met Val Ser Lys Glu Ser Glu Arg Leu Ser Ile Thr Lys Ser Ala> TGT GGA AGA AAT GGC AAA CAA TTC TGC AGT ACT TTA ACC TTG AAC ACA GCT CAA GCA AAC ACA CCT TCT TTA CCG TTT GTT AAG ACG TCA TGA AAT TGG AAC TTG TGT CGA GTT CGT TTG Cys Gly Arg Asn Gly Lys Gln Phe Cys Ser Thr Leu Thr Leu Asn Thr Ala Gln Ala Asn> CAC ACT GGC TTC TAC AGC TGC AAA TAT CTA GCT GTA CCT ACT TCA AAG AAG AAG GAA ACA GTG TGA CCG AAG ATG TCG ACG TTT ATA GAT CGA CAT GGA TGA AGT TTC TTC CTT TGT His Thr Gly Phe Tyr Ser Cys Lys Tyr Leu Ala Val Pro Thr Ser Lys Lys Glu Thr> GAA TCT GCA ATC TAT ATA TTT ATT AGT GAT ACA GGT AGA CCT TTC GTA GAG ATG TAC AGT CTT AGA CGT TAG ATA TAT AAA TAA TCA CTA TGT CCA TCT GGA AAG CAT CTC TAC ATG TCA Glu Ser Ala Ile Tyr Ile Phe Ile Ser Asp Thr Gly Arg Pro Phe Val Glu Met Tyr Ser>

Fig.16B.

Jul. 4, 2006

550 560 580 590 600 570 GGA AAA CGC ATA ATC TGG GAC AGT AGA AAG GGC TTC ATC ATA TCA AAT GCA ACG TAC AAA CCT TTT GOG TAT TAG ACC CTG TCA TCT TTC CCG AAG TAG TAT AGT TTA CGT TGC ATG TTT Gly Lys Arg Ile Ile Trp Asp Ser Arg Lys Gly Phe Ile Ile Ser Asn Ala Thr Tyr Lys> 610 620 630 640 650 660 GAA ATA GGG CTT CTG ACC TGT GAA GCA ACA GTC AAT GGG CAT TTG TAT AAG ACA AAC TAT CTT TAT CCC GAA GAC TGG ACA CTT CGT TGT CAG TTA CCC GTA AAC ATA TTC TGT TTG ATA Glu Ile Gly Leu Leu Thr Cys Glu Ala Thr Val Asn Gly His Leu Tyr Lys Thr Asn Tyr> 670 700 710 720 680 690 CTC ACA CAT CGA CAA ACC AAT ACA ATC ATA GAT GTC CAA ATA AGC ACA CGA CGC CCA GTC GAG TGT GTA GCT GTT TGG TTA TGT TAG TAT CTA CAG GTT TAT TCG TGT GGT GCG GGT CAG Leu Thr His Arg Gln Thr Asn Thr Ile Ile Asp Val Gln Ile Ser Thr Pro Arg Pro Val> 770 780 730 740 750 760 * * AAA TTA CTT AGA GGC CAT ACT CTT GTC CTC AAT TGT ACT GCT ACC ACT CCC TTG AAC AOG TTT AAT GAA TCT COG GTA TGA GAA CAG GAG TTA ACA TGA CGA TGG TGA GGG AAC TTG TGC Lys Leu Leu Arg Gly His Thr Leu Val Leu Asn Cys Thr Ala Thr Thr Pro Leu Asn Thr> 790 830 840 820 800 810 AGA GTT CAA ATG ACC TOG AGT TAC CCT GAT GAA AAA AAT AAG AAC GCT TCC GTA AGG CGA TCT CAA GTT TAC TOG ACC TCA ATG GGA CTA CTT TTT TTA TTC TTG CGA AGG CAT TCC GCT Arg Val Gln Met Thr Trp Ser Tyr Pro Asp Glu Lys Asn Lys Asn Ala Ser Val Arg Arg> 850 860 870 880 890 900 OGA ATT GAC CAA AGC AAT TCC CAT GCC AAC ATA TTC TAC AGT GTT CTT ACT ATT GAC AAA GCT TAA CTG GTT TCG TTA AGG GTA CGG TTG TAT AAG ATG TCA CAA GAA TGA TAA CTG TTT Arg Ile Asp Gln Ser Asn Ser His Ala Asn Ile Phe Tyr Ser Val Leu Thr Ile Asp Lys> 950 960 910 920 930 940 * * * ATG CAG AAC AAA GAC AAA GGA CTT TAT ACT TGT CGT GTA AGG AGT GGA CCA TCA TTC AAA TAC GTC TTG TTT CTG TTT CCT GAA ATA TGA ACA GCA CAT TCC TCA CCT GGT AGT AAG TTT Met Gln Asn Lys Asp Lys Gly Leu Tyr Thr Cys Arg Val Arg Ser Gly Pro Ser Phe Lys> 1020 970 980 990 1000 1010 * * * * TCT GTT AAC ACC TCA GTG CAT ATA TAT GAT ANA GCA GGC CCG GGC GAG CCC AAA TCT TGT AGA CAA TTG TGG AGT CAC GTA TAT ATA CTA TTT CGT CCG GGC CCG CTC GGG TTT AGA ACA Ser Val Asn Thr Ser Val His Ile Tyr Asp Lys Ala Gly Pro Gly Glu Pro Lys Ser Cys> 1030 1040 1050 1060 1070 1080 *

GAC AAA ACT CAC ACA TGC CCA CCG TGC CCA GCA CCT GAA CTC CTG GGG GGA CCG TCA GTC CTG TTT TGA GTG TGT ACG GGT GGC ACG GGT CGT GGA CTT GAG GAC CCC CCT GGC AGT CAG Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val>

Fig.16C.

		109	90		11	.00		- 1	110			112	20		11	.30		1	.140
	*		*	*		*		*	*		*		*	*		*		*	*
TTC:	CTTC	സ്റ	ccc	CCA	٨٨٨	000	AAG	GAC	ACC	CTC	ATG	ATC	TCC	CGG	ACC	CCT	GAG	GTC	ACA
	010	110	000	COR		000	0000	anc	7000	010	mac	m20	200	000	mmo	002	ome	010	mom
AAG	GAG	AAG	-	GGT	-	-	TIC	CIG	166	GAG	TAC	TAG	AGG	GUU	166	GGA		CAG	IGI
Phe	Leu	Phe	Pro	Pro	Lys	Pro	Lys	Asp	Thr	Leu	Met	Ile	Ser	Arg	Thr	Pro	GIU	Val	Thr>
		115	50		11	60		1	L170			118	30		11	.90		1	200
	*		*	*		*		*	*		*		*	*		*		*	*
mco	000	CIRC	000	020	omo	100	010	~~~	020	COM	GNG	CTC	244	ጥጥጥ	244	TCC	ጥልሮ	cnc.	GAC
1GC	GIG	616	GIG	GAC	GIG	AGC	CAL	GAA	GAC		GAG	910	11103	110		100	100	010	one
ACG	CAC	CAC	CAC	CIG	CAC	TCG	GIG	CTT	CIG	GGA	CIC	CAG	TIC	AAG	TIG	ACC	AIG	CAC	CIG
Cys	Val	Val	Val	Asp	Val	Ser	His	Glu	Asp	Pro	Glu	Val	Lys	Phe	Asn	Trp	Tyr	Val	Asp>
		121	LO		12	220		:	1230			124	10		12	250		1	260
	*		*	*		*		*	*		*		*	*		*		*	*
000	000	~~~	~	~~~	220	000	220	101	120	\sim	000	010	CNC	CAC	mac	220	200	200	ጥእር
GGC	GIG	GAG	GIG	CAT	MAI	GCC	AAS	ACA	MAG	000	000	GAG	040		100	nn	AGC.	700	100
CCG	CAC	CIC	CAC	GTA	TTA	CGG	TTC	TGT	TIC	GGC	GCC	CIC	CIC	GIC	AIG	TIG	TCG	IGC	AIG
Gly	Val	Glu	Val	His	Asn	Ala	Lys	Thr	Lys	Pro	Arg	Glu	Glu	Gln	Tyr	Asn	Ser	Thr	Tyr>
		127	70		12	280		:	1290			130	00		13	310		1	1320
	*		*	*		*		*	*		*		*	*		*		*	*
~~~~	~	~~~	100	000	~~~	100	~~~	~~~	~~~	~~~	020	m00	ono	<b>3</b> 3 m	000	220	CNC	ma()	MAG
0.01	GIG	GIC	AGC	GIC	CIC.	ACC	GIC	010	040	083	ORC	100	010		~~~	AA3	000	2000	0000
GCA	CAC	CAG	TCG	CAG	GAG	TGG	CAG	GAC	GIG	GIC	CIG	ACC	GAC	TTA	000	TIC	CIC:	Ale	TIC
Arg	Val	Val	Ser	Val	Leu	Thr	Val	Leu	His	Gin	Asp	Trp	Leu	Asn	GĩÃ	LYS	GIU	ıyr	rA2>
		13:	30		13	340		:	1350			130	50		1	370		:	1380
	*	133	30 *	*	13	340 *		*	1350		*	130	50 *	*	13	370 *		*	L380 *
TGC	* AAG	13: GTC	30 * TCC	* AAC	1: AAA	340 * GCC	CTC	* CCA	1350 * GCC	ccc	* ATC	13 GAG	50 * AAA	* ACC	1: ATC	370 * TCC	ааа	: *	1380 * AAA
TGC	* AAG	133 GTC CAG	30 * TCC	* AAC TTG	1: AAA TTT	340 * GCC	CTC	* CCA	1350 * GCC	CCC 666	* ATC TAG	13 GAG CTC	50 * AAA TTT	* ACC TGG	1: ATC TAG	370 * TCC AGG	ааа ттт	: * 600 000	1380 * AAA TTT
TGC ACG	* AAG TTC	133 GTC CAG	30 * TCC AGG	* AAC TTG	13 AAA TTT	340 * GCC CGG	CTC GAG	* CCA GGT	1350 * GCC CGG	CCC GGG	* ATC TAG	13 GAG CTC	50 * AAA TTT	* ACC TGG	1: ATC TAG	370 * TCC AGG	AAA TTT	* GCC CGG Ala	AAA TTT
TGC ACG Cys	* AAG TTC Lys	133 GTC CAG Val	30 * TCC AGG Ser	* AAC TTG Asn	1: AAA TTT Lys	340 * GCC CGG Ala	CTC GAG Leu	* CCA GGT Pro	1350 * GCC CGG Ala	CCC GGG Pro	* ATC TAG Ile	130 GAG CTC Glu	50 * AAA TTT Lys	* ACC TGG Thr	1: ATC TAG Ile	370 * TCC AGG Ser	aaa TTT Lys	* GCC CGG Ala	AAA TTT Lys>
TGC ACG Cys	* AAG TTC Lys	133 GTC CAG Val	30 * TCC AGG Ser	* AAC TTG Asn	13 AAA TTT Lys	340 * GCC CGG Ala	CTC GAG Leu	* CCA GGT Pro	1350 * GCC CGG Ala	CCC GGG Pŕo	* ATC TAG Ile	130 GAG CTC Glu	50 * AAA TTT Lys	* ACC TGG Thr	1: ATC TAG Ile	370 * TCC AGG Ser	AAA TTT Lys	* GCC CGG Ala	AAA TTT Lys>
TGC ACG Cys	* AAG TTC Lys	133 GTC CAG Val 139	30 * TCC AGG Ser 90	* AAC TTG Asn	1: AAA TTT Lys 1	GCC CGG Ala	CTC GAG Leu	* CCA GGT Pro	1350 * GCC CGG Ala 1410	CCC GGG Pro	* ATC TAG Ile	130 GAG CTC Glu 142	50 * AAA TTT Lys 20	* ACC TGG Thr	1: ATC TAG Ile	370 * TCC AGG Ser 430	aaa TIT Lys	* GCC CGG Ala	1380 * AAA TTT Lys>
TGC ACG Cys	* AAG TTC Lys *	133 GTC CAG Val 133	30 * TCC AGG Ser 90 *	* AAC TTG Asn	1: AAA TTT Lys 14	340 * CGG Ala 400	CTC GAG Leu	* CCA GGT Pro	1350 * GCC CGG Ala 1410 *	CCC GGG Pro	* ATC TAG Ile	130 GAG CTC Glu 143	50 * AAA TTT Lys 20 *	* ACC TGG Thr	1: ATC TAG Ile	370 * TCC AGG Ser 430 *	AAA TTT Lys	* GCC CGG Ala	1380 * AAA TTT Lys> 1440 *
TGC ACG Cys GGG	* AAG TTC Lys * CAG	13: GTC CAG Val 13: CCC	30 * TCC AGG Ser 90 * CGA	* AAC TTG Asn * GAA	1: AAA TTT Lys 1/ CCA	GCC CGG Ala 400 CAG	CTC GAG Leu GTG	* GGT Pro * TAC	1350 * GCC CGG Ala 1410 * ACC	CCC GGG Pro CTG	* ATC TAG Ile	130 GAG CTC Glu 142 CCA	50 * AAA TTT LY3 20 * TCC	* ACC TGG Thr * CGG	1: ATC TAG Ile 1 GAT	370 * TCC AGG Ser 430 * GAG	AAA TTT Lys CTG	* GCC CGG Ala *	1380 * AAA TTT Lys> 1440 * AAG
TGC ACG Cys GGG CCC	* AAG TTC Lys * CAG GTC	13: GTC CAG Val 13: CCC GGC	30 * TCC AGG Ser 90 * CGA GCT	* AAC TTG Asn * GAA	1: AAA TTT Lys 1 CCA GGT	340 * CGG Ala 400 * CAG GTC	CTC GAG Leu GTG CAC	* GGT Pro * TAC	GCC CGG Ala 1410 * ACC TGG	CCC GGG Pro CTG GAC	* TAG Ile * CCC GGG	130 GAG CTC Glu 142 CCA GGT	50 * AAA TTT Lys 20 * TCC AGG	* ACC TGG Thr * CGG GCC	1: ATC TAG Ile 1 GAT CTA	370 * TCC AGG Ser 430 * GAG CTC	AAA TTT Lys CTG GAC	* GCC CGG Ala * ACC TGG	1380 * AAA TTT Lys> 1440 * AAG TTC
TGC ACG Cys GGG CCC	* AAG TTC Lys * CAG GTC GIn	133 GTC CAG Val 133 CCC GGC Pro	30 * TCC AGG Ser 90 * CGA GCT ATG	* AAC TTG Asn * GAA CTT Glu	1: AAA TTT Lys 1 CCA GGT Pro	GCC CGG Ala 400 * CAG GTC	CTC GAG Leu GTG CAC Val	* GGT Pro * TAC ATG	1350 * GCC CGG Ala 1410 * ACC TGG	CCC GGG Pro CTG GAC Leu	* ATC TAG Ile * CCC GGG Pro	130 GAG CTC Glu 142 CCA GGT Pro	50 * AAA TTT Lys 20 * TCC AGG Ser	* ACC TGG Thr * CGG GCC Arg	1: ATC TAG Ile 1 GAT CTA ASD	370 * TCC AGG Ser 430 * GAG CTC Glu	AAA TTT Lys CTG GAC Leu	* CGG Ala * ACC TGG Thr	1380 * AAA TTT Lys> 1440 * AAG TTC Lys>
TGC ACG Cys GGC GCC Gly	* TTC Lys * CAG GTC GIn	133 GTC CAG Val 139 CCC GGG Pro	30 * TCC AGG Ser 90 * CGA GCT Arg	* AAC TTG Asn * GAA CTT Glu	1: AAA TTT Lys 1. CCA GGT Pro	GCC CGG Ala 400 * CAG GTC Gln	CTC GAG Leu GTG CAC Val	* GGT Pro * TAC ATG Tyr	1350 * CGG Ala 1410 * ACC TGG Thr	CCC GGG Pro CTG GAC Leu	* TAG Ile * CCC GGG Pro	130 GAG CTC Glu 14: CCA GGT Pro	50 * AAA TTT Lys 20 * TCC AGG Ser	* ACC TGG Thr Thr * CGG GCC Arg	1: ATC TAG Ile 1. GAT CTA Asp	370 * TCC AGG Ser 430 * GAG CTC Glu	AAA TTT Lys CTG GAC Leu	* GCC CGG Ala * ACC TGG Thr	1380 * AAA TTT Lys> 1440 * AAG TTC Lys>
TGC ACG Cys GGC CCC Gly	* TTC Lys * CAG GTC GIn	133 GTC CAG Val 133 CCC GGG Pro	30 * TCC AGG Ser 90 * CGA GCT Arg	* AAC TTG Asn * GAA CTT Glu	1: AAA TTT Lys 1 CCA GGT Pro	GCC CGG Ala 400 * CAG GTC GIn	CTC GAG Leu GTG CAC Val	* GGT Pro * TAC ATG Tyr	GCC CGG Ala 1410 * ACC TGG Thr	CCC GGG Pro CTG GAC Leu	* TAG Ile * CCC GGG Pro	130 GAG CTC Glu 142 CCA GGT Pro	50 * AAA TTT Lys 20 * TCC AGG Ser 80	* ACC TGG Thr * CGG GCC Arg	1: ATC TAG Ile 1 GAT CTA ASP	370 * TCC AGG Ser 430 * GAG CTC Glu	AAA TTT Lys CTG GAC Leu	* GCC CGG Ala * ACC TGG Thr	1380 * AAA TTT Lys> 1440 * AAG TTC Lys>
TGC ACG Cys GGG CCC Gly	* TTC Lys * CAG GTC Gln	13: GTC CAG Val 13: CCC GGG Pro 14:	30 * TCC AGG Ser 90 * CGA GCT Arg 50	* AAC TTG Asn * GAA CTT Glu	1: AAA TTT Lys 1 CCA GGT Pro 1	GCC CGG Ala 400 * CAG GTC GIn 460	CTC GAG Leu GTG CAC Val	* GGT Pro * TAC ATG Tyr	GCC CGG Ala 1410 * ACC TGG Thr 1470	CCC GGG Pro CTG GAC Leu	* ATC TAG Ile * CCC GGG Pro	130 GAG CTC Glu 142 CCA GGT Pro 14	50 * AAA TTT Lys 20 * TCC AGG Ser 80	* ACC TGG Thr * CGG GCC Arg	1: ATC TAG Ile 1 GAT CTA ASP	370 * TCC AGG Ser 430 * GAG GLu 490	AAA TTT Lys CTG GAC Leu	* GCC CGG Ala * ACC TGG Thr	1380 * AAA TTT Lys> 1440 * AAG TTC Lys> 1500
TGC ACG Cys GGG CCC Gly	* AAG TTC Lys * CAG GTC GIn *	13: GTC CAG Val 13: CCC GGG Pro 14	30 * TCC AGG Ser 90 * CGA GCT Arg 50 *	* AAC TTG Asn * GAA CTT Glu	1: AAA TTT Lys 1 CCA GGT Pro 1	GCC CGG Ala 400 CAG GTC Gln 460	CTC GAG Leu GTG CAC Val	* GGT Pro * TAC ATG Tyr	1350 * GCC CGG Ala 1410 * ACC TGG Thr 1470	CCC GGG Pro CTG GAC Leu	* ATC TAG Ile * CCC GGG Pro	130 GAG CTC Glu 142 CCA GGT Pro 14	50 * AAA TTT Lys 20 * TCC AGG Ser 80 *	* ACC TGG Thr * CGG GCC Arg	1: ATC TAG Ile 1 GAT CTA ASP 1	370 * TCC AGG Ser 430 * GAG CTC Glu 490 *	AAA TTT Lys CTG GAC Leu	* GCCC CGGG Ala * ACCC TGGG Thr *	AAA TTT Lys> 1440 * AAG TTC Lys> 1500
TGC ACG Cys GGG GCC Gly AAC	* AAG TTC Lys * CAG GTC Gln * CAG	13: GTC CAG Val 13: CCC GGG Pro 14: GTC	30 * TCC AGG Ser 90 * CGA GCT Arg 50 * AGC	* AAC TTG Asn * GAA CTT Glu * CTG	1: AAA TTT Lys 10 CCA GGT Pro 10 ACC	340 * GCC CGG Ala 400 * CAG GTC Gln 460 * TGC	CTC GAG Leu GTG CAC Val	* GGT Pro * TAC ATG Tyr *	1350 « ССС А1а 1410 * АСС ТСС ТСС ТПГ 1470	CCC GGG Pro CTG GAC Leu	* ATC TAG Ile * CCC GGG Pro * TTC	130 GAG CTC Glu 142 CCA GGT Pro 14 TAT	50 * AAA TTT Lys 20 * TCC AGG Ser 80 * CCC	* ACC TGG Thr * CGG GCC Arg * AGC	1: ATC TAG Ile 1. GAT CTA ASP 1 GAC	TCC AGG Ser 430 CTC Glu 490 *	AAA TTT Lys CTG GAC Leu GCC	* GCC CGG Ala ACC TGG Thr & GTG	1380 AAA TTT Lys> 1440 AAG TTC Lys> 1500 * GAG
TGC ACG Cys GGC GCC Gly AAC TTG	* AAG TTC Lys * CAG GTC GIn * CAG GTC	13: GTC CAG Val 13: CCC GGG Pro 14 GTC CAG	30 * TCC AGG Ser 90 * CGA GCT Arg 50 * AGC TCG	* AAC TTG Asn * GAA CTT GLu * CTG GAC	1: AAA TTT Lys 1. CCA GGT Pro 1. ACC TGG	GCC CGG Ala 400 * CAG GTC Gln 460 * TGC ACG	CTC GAG Leu GTG CAC Val CTG GAC	* CCA GGT Pro * TAC ATG Tyr * GTC CAG	1350 * GCC CGG Ala 1410 * ACC TGG Thr 1470 * AAA TTT	CCC GGG Pro CTG GAC Leu	* ATC TAG Ile * CCC GGG Pro * TTC AAG	130 GAG CTC Glu 14: CCA GGT Pro 14 TAT	50 * AAA TTT Lys 20 * TCC AGG Ser 80 * CCC GGG	* ACC TGG Thr * CGG GCC Arg * AGC TCG	11 ATC TAG Ile 1. GAT CTA Asp 1 GAC CTG	TCC AGG Ser 430 * GAG CTC Glu 490 * ATC	AAA TITT Lys CTG GAC Leu GCC	* GCC CGG Ala * ACC TGG Thr * GTG CAC	AAA TTT Lys> 1440 * AAG TTC Lys> 1500 * GAG CTC
TGC ACG Cys GGG Gly AAC TTG Asn	* AAG TTC Lys * CAG GTC GIn * CAG GTC GIn	13: GTC CAG Val 13: CCC GGG Pro 14 GTC CAG Val	30 * TCC AGG Ser 90 * CGA GCT Arg 50 * AGG CGA CGA CGA CGA CGA CGA CGA	* AAC TTG Asn * GAA CTT Glu * CTG GAC Leu	1: AAA TTTT Lys 1. CCA GGT Pro 1. ACC TGG Thr	GCC CGG Ala 400 * CAG GTC GIn 460 * TGC ACG Cys	CTC GAG Leu GTG CAC Val CTG GAC Leu	* CCA GGT Pro * TAC ATG Tyr * GTC CAG Val	1350 * GCC CGG Ala 1410 * ACC TGG Thr 1470 * AAA TTT Lys	CCC GGG Pro CTG GAC Leu GGC GGC Gly	* ATC TAG Ile CCC GGG Pro * TTC AAG	130 GAG CTC Glu 14: CCA GGT Pro 14 TAT ATA TAT	50 * AAA TTTT Lys 20 * TCC AGG Ser 80 * CCC GGG Pro	* ACC TGG Thr * CGG GCC Arg * AGC TCG Ser	1: ATC TAG Ile 1: GAT CTA Asp 1 GAC CTG Asp	TCC AGG Ser 430 * GAG GAG CTC Glu 490 * ATC TAG	AAA TTT Lys CTG GAC Leu GCC CGG Ala	* GCCC CGG Ala * ACC TGG Thr * GTG CAC Val	1380 AAA TTT Lys> 1440 * AAG TTC Lys> 1500 * GAG CTC Glu>
TGC ACG Cys GGG Gly AAC TTG Asn	* AAG TTC Lys CAG GTC GIn * CAG GTC GIn	13: GTC CAG Val 13: CCC GGG Pro 14 GTC CAG Val	30 * TCC AGG Ser 90 * CGA GCT Arg 50 * AGC TCG Ser	* AAC TTG Asn * GAA CTT Glu * CTG GAC Leu	1: AAA TTTT Lys 1. CCA GGT Pro 1 ACC TGG Thr	340 * GCC CGG Ala 400 * CAG GIN 460 * TGC ACG Cys	CTC GAG Leu GTG CAC Val CTG GAC Leu	* GGT Pro * TAC ATG Tyr \$ GTC CAG Val	1350 * GCC CGG Ala 1410 * ACC TGG Thr 1470 * 1470 * Lys	CCC GGG Pro CTG GAC Leu GGC GGC Gly	* ATC TAG Ile CCC GGG Pro * TTC AAG Phe	130 GAG CTC Glu 142 CCA GGT Pro 14 TAT ATA Tyr	50 * AAA TTT Lys 20 * TCC Ser 80 * CCC GGG GGG Pro	* ACC TGG Thr * CGG GCC Arg ACC Arg Ser	1: ATC TAG Ile 1: GAT CTA Asp 1 GAC CTG Asp	TCC AGG Ser 430 * GAG GAG CTC Glu * ATC TAG 11e	AAA TTT Lys CTG GAC Leu GCC CGG Ala	* GCC CGG Ala * ACC TGG Thr * GTG CAC Val	1380 * AAA TTT Lys> 1440 * AAG TTC Lys> 1500 * GAG CTC Glu>
TGC ACG Cys GGG Gly AAC TTG Asn	* AAG TTC Lys * CAG GTC GIn * CAG GTC GIn	13: GTC CAG Val 13: CCC GGG Pro 14 GTC CAG Val 15	30 * TCC AGG Ser 90 * CGA GCT Arg 50 * AGC TCG Ser 10	* AAC TTG Asn * GAA CTT Glu * CTG GAC Leu	1: AAAA TTTT Lys 1: CCA GGT Pro 1: ACC TGG Thr	GCC CGG Ala 400 * CAG GTC GIN 460 * TGC Cys 520	CTC GAG Leu GTG CAC Val CTG GAC Leu	* GGT Pro * TAC ATG Tyr GTC CAG Val	1350 * GCC CGG Ala 1410 * ACC TGG Thr 1470 * 1470 * Lys 1530	CCC GGG Pro CTG GAC Leu GGC GGC Gly	* ATC TAG Ile CCC GGG Pro * TTC AAG Phe	130 GAG CTC Glu 142 CCA GGT Pro 14 TAT ATA Tyr 15	50 * AAA TTT Lys 20 * TCC AGG Ser 80 * CCC GGG GGG Pro	* ACC TGG Thr * CGG GCC Arg & AGC TCG Ser	1: ATC TAG Ile 1. GAT CTA Asp 1 GAC CTG Asp 1	TCC AGG Ser 430 * GAG GAG GIU 430 * ATC TAG 11e 550	AAA TIT Lys CTG GAC Leu GCC CGG Ala	* GCC CGG Ala * ACC TGG Thr * GTG CAC Val	1380 * AAA TTT Lys> 1440 * AAG TTC Lys> 1500 * GAG CTC Glu> 1560
TGC ACG Cys CCC Gly AAC TTG ASn	* AAG TTC Lys * CAG GTC GIn * CAG GTC GIn *	13: GTC CAG Val 13: CCC GGG Pro 14 GTC CAG Val 15	30 * TCC AGG Ser 90 * CGA GCT Arg 50 * AGG Ser 10 *	* AAC TTG Asn * GAA CTT Glu * CTG GAC Leu	1: AAAA TTTT Lys 1: CCA GGT Pro 1: ACC TGG Thr 1	GCC CGG Ala 400 * CAG GTC GIn 460 * TGC Cys 520	CTC GAG Leu GTG CAC Val CTG GAC Leu	* GGT Pro * TAC ATG Tyr GTC CAG Val	1350 * GCC CGG Ala 1410 * TGG Thr 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470 * 1470	CCC GGG Pro CTG GAC Leu GGC Gly	* ATC TAG Ile CCC GGG Pro * TTC AAG Phe	130 GAG CTC Glu 142 CCA GGT Pro 14 TAT TAT TAT TAT	50 * AAA TTT Lys 20 * TCC AGG Ser 80 * CCC GGG Pro 40	* ACC TGG Thr * CGG GCC Arg * AGC TCG Ser	1: ATC TAG Ile 1. GAT CTA Asp 1 GAC CTG Asp 1	TCC AGG Ser 430 * GAG GAG GIU 490 * ATC TAG 9 Ile 550	AAA TTT Lys CTG GAC Leu GCC CGG Ala	* GCC CGG Ala * ACC TGG Thr GTG CAC Val	L380 AAA TTT Lys> 1440 AAG TTC Lys> 1500 CTC GAG CTC Glu> 1560 *
TGC ACG Cys GGG CCC Gly AAC TTG ASn	* AAG TTC Lys CAG GTC GIn * CAG GIC GIn *	13: GTC CAG Val 13: CCC GGG Pro 14 GTC CAG Val 15	30 * TCC AGG Ser 90 * CGA AGC CGA Arg 50 * AGC TCG Ser 10 *	* AAC TTG Asn * CAA CTT Glu * CTG GAC Leu	1: AAAA TTTT Lys 1. CCA GGT Pro 1. ACC TGG Thr 1	340 CGG Ala 400 CAG GTC GIN 460 * TGC Cys 520 *	CTC GAG Leu GTG CAC Val CTG GAC Leu	* CCA GGT Pro * TAC ATG Tyr * GTC CAG	1350 * GCC CGG Ala 1410 * TGG Thr 1470 * AAA TTT Lys 1530 *	CCC GGG Pro CTG GAC Leu GGC Gly	* ATC TAG Ile CCC GGG Pro * TTC AAG Phe	130 GAG CTC Glu 142 CCA GGT Pro 14 TAT TAT TAT TAT TAT	50 AAA TTT Lys 20 * TCC AGG Ser 80 * CCC GGG Pro 40 *	* ACC TGG Thr * CGG GCC Arg * AGC TCG Ser *	1: ATC TAG Ile 1. GAT CTA Asp 1 GAC CTG Asp 1	370 * AGG Ser 430 * GAG CTC Glu 490 * ATCC 550 *	AAA TIT Lys CTG GAC Leu GCC GGC Ala	* GCCC CGGG Ala * ACCC TGG Thr * GTGG CACC Val	AAA TTT Lys> 1440 * AAG TTC Lys> 1500 * GAG CTC Glu>
TGC ACG Cys Cys CCC Gly AAC TTG Asn	* AAG TTC Lys CAG GTC GIn * CAG GTC GIn *	13: GTC CAG Val 13: CCC GGG Pro 14 GTC CAG Val 15 AGC	30 * TCC AGG Ser 90 * CGA GCT Arg 50 * AGC Ser 10 * AAT	* AAC TTG Asn * GAA CTT Glu * CTG GAC Leu * GAC	1: AAAA TTTT Lys 1: CCA GGT Pro 1: ACCC TGG Thr 1 CAG	340 CGG Ala 400 CAG GTC GIN 460 * TGC Cys 520 * CGG	CTC GAG Leu GTG CAC Val CTG GAC Leu	* GGT Pro * TAC ATG Tyr * GTC CAG Val	1350 * GCC CGG Ala 1410 * ACC TGG Thr 1470 * AAA 1530 *	CCC GGG Pro CTG GAC Leu GGC GGC Gly	* ATC TAG Ile CCCC GGG Pro * TTC AAG Phe	130 GAG CTC Glu 142 CCA GGT Pro 14 TAT TAT TAT TAT TS CCA	50 * AAA TTTT Lys 20 * TCCC AGG Ser 80 * CCCC GGG GGG Pro 40 * ACG	* ACC TGG Thr * CGG GCC Arg * AGC TCG Ser *	1: ATC TAG Ile 1. GAT CTA Asp 1 GAC CTG Asp 1 CCC	TCC AGG Ser 430 * GAG Glu * ATC * ATC * TAG * STG *	AAA TIT Lys CTG GAC Leu GCC CGG Ala	* GCCC CGG Ala * ACCC TIGG GTG CACC Val	AAA TTT Lys> 1440 * AAG TTC Lys> 1500 * GAG CTC Glu> 1560 *
TGC ACG Cys GGG Gly AAC TTG ASn TGG	* AAG TTC Lys CAG GTC GIn * CAG GTC GIn * CAG	13: GTC CAG Val 13: CCC GGG Pro 14 GTC CAG Val 15 AGC	30 * TCC AGG Ser 90 * CGA GCT Arg 50 * AGC Ser 10 * AAG	* AAC TTG Asn * GAA CTT Glu * CTG GAC Leu * * CTG GAC	1: AAAA TTTT Lys 1: CCA GGT Pro 1 ACCC TGG Thr 1 CAG CAG	340 * GCC CGG Ala 400 * CAG GTC GIn * TGC ACG Cys 520 * CGG	CTC GAG Leu GTG CAC Val CTG GAC Leu GAG	* CCA GGT Pro TAC ATG Tyr CAG CAG Val * * AAC	1350 * GCC CGG Ala 1410 * ACC TGG Thr 1470 * AAA TTT Lys 1530 * AAC * *	CTC GGG Pro CTG GAC Leu GGC Gly TAC	* ATC TAG Ile CCC GGG Pro * TTC AAG Phe *	130 GAG CTC Glu 14: CCA GGT Pro 14 TAT TAT TAT TST ACC TCG	50 * AAAA TTTT Lys 20 * TCC AGG Ser 80 * CCC GGG Pro 40 *	* ACC TGG Thr * CGG GCC Arg * AGC TCG Ser * CCT GGA	1: ATC TAG Ile 1. GAT CTA Asp 1 GAC CTG Asp i CCC GO	TCC AGG Ser 430 * GAG Glu * ATC 550 * TAG 550 *	AAA TITT Lys CTG GAC Leu GCC CGG Ala	* GCCC CGG Ala ACC TGG Thr GTG CAC Val	AAA TTT Lys> 1440 * AAG TTC Lys> 1500 * GAG CTC Glu> 1560 * CTC Glu>

1600 1620 1570 1610 1580 1590 * * * * * * * * GAC GGC TCC TTC TTC CTC TAC AGC AAG CTC ACC GTG GAC AAG AGC AGG TGG CAG CAG GGG CTG CCG AGG AAG AAG GAG ATG TOG TTC GAG TGG CAC CTG TTC TOG TCC ACC GTC GTC CCC Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly>

# Fig.16D.

1680 1670 1660 1650 1630 1640 * × * * * * * * * * * AAC GTC TTC TCA TGC TCC GTG ATG CAT GAG GCT CTG CAC AAC CAC TAC ACG CAG AAG AGC TTG CAG AAG AGT ACG AGG CAC TAC GTA CTC CGA GAC GTG TTG GTG ATG TGC GTC TTC TCG Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser>

1690 1700

*

*

*

*

CTC TCC CTG TCT CCG GGT AAA TGA GAG AGG GAC AGA GGC CCA TTT ACT Leu Ser Leu Ser Pro Gly Lys ***>







F	ig.21A			>EcoR	I_site		
10	20	30	40	50	60	70	80
AAGCTTGGGCT	GCAGGTCGATCG	ACTCTAGAGG	ATCGATCCCC	GGCGAGCIC	GAATTCGCAAC	CACCATGGT	LAGCTAC
TICGAACCCG	CGTCCAGCTAGC	TGAGATCTCC	PAGCTAGGGGG	CCGCTCGAG	CTTAAGCGTTG	GIGGTACCA	FICGATG
						M V 1	SY>
							>
					>BspEI_bri	dge	
90	100	110	120	130	140	150	160
TGGGACACCGG	GGTCCTGCTGTG	CGCGCTGCTCZ	AGCTGTCTGC	TCTCACAGG	ATCTAGTTCCG	GAGGTAGACO	TTICGT
ACCCTGTGGCC	CCAGGACGACAC	GCGCGACGAGI	CGACAGACG	AGAGTGTCC	TAGATCAAGGC	CTCCATCIGO	JAAAGCA.
wbre		FLT1 SS	зсьі	י ע ד G	5 5>		
					S	G>	
						>	
						GRI	PFV>
170	180	190	200	210	220	230	240
TCTCTACATGI E M Y	CACTTTAGGGGC S E I P	TTTAATATGIC E I I H	M T E	G R E I	AGCAGTAAGGG LVIP	ACGGCCCAAT C R V	IGCAGIG T S> 57
			_HFLT1 D2_				>
250	260	270	280	290	300	310	320
CTAACATCACT	GTTACTTTAAAA	AAGTTTCCACI	TGACACTITO	ATCCCTGAT	GAAAACGCAT	AATCTGGGA	AGTAGA
GATTGTAGTGA PNIT	CAATGAAATTTT VTLK	TTCAAAGGTGA K F P I	ACTGTGAAAG D T L	TAGGGACTAC I P D	CTTTTGCGTA G K R I	TTAGACCCTO I W D	S R>
			HET.01 102				84
330	340	350	360	370	380	390	400
AAGGGCTTCAT	CATATCAAATGC	AACGTACAAAG	JAAATAGGGC7	TCTGACCIG	GAAGCAACAG	TCAATGGGC!	ATTIGIA
K G F I	ISNA	TIGCATGITIC TYK	EIGI	LTC	EAT	V N G I	ILY>
							111
			_HFLT1 D2_				>
410 TAAGACAAACT ATTCTGTTTGA K T N	420 ATCTCACACATO TAGACITGTAG Y L T H HFLT1	430 GACAAACCAAT CTGTTTGGTTA R Q T N D2	440 TACAATCATAO ATGITAGTATO TII	450 SATGTGGTTC: TACACCAAG D> >	460 NGAGTCCGTCT ACTCAGGCAGA	470 CATGGAATIC GTACCTPAAC	480 Saactat Citgata
				vv	S P S	H G I	E L>
					1101 11	51	137
						no	

100	500	510	520	530	540	550	56
490	TIGTCTTAA	ATTGTACAG	CAAGAACTGA	CTAAATGTGG	GGATIGACTI	CAACTGGG	AATACCC
JACAACCTCTTTTCG	AACAGAATT	PAACATGTO	GTTCTTGACTT	GATTTACACC	CCTAACTGAA	GTTGACCO	TTATGGG
SVGEK	LVL	м с т	ARTE	LNV	GIDE	'N W	EYF
							1
			HFLK1 D3_				
570	580	590	600	610	620	630	64
TCTTCGAAGCATCAG	CATAAGAAA	CTTGTAAAC	CGAGACCTAA	AACCCAGTCT	GGGAGTGAG	TGAAGAA	ATTTTTGA
AGAAGCTTCGTAGTC	GTATTCTT	GAACATITO	GCTCICGATT	ITTGGGTCAGA	CCCTCACTC	ACTICTI	CAAAAACT
S S K H Q	нкк	LVN	RDLI	ктұз	GSE	мкк	FL
			UPT V1 D2				
			NEUKI D3_				
650	660	670	680	690	700	710	72
CACCTTAACTATAGA	TGGTGTAAC	CCGGAGTG	ACCAAGGATTG	PACACCIGIGC	AGCATCCAG	IGGGCTGA!	GACCAA
GTGGAATTGATATCT	ACCACATTG	GCCTCACT	IGGTTCCTAAC)	ATGTGGACACG	TCGTAGGTC	ACCCGACT	ACTGGTI
ΤΙΤΙΟ	GVT	RSI	οφισ	YTCA	ASS	GLI	n T Ka 21
			HR1.81 D3				£.
			nruki 03,				
		>Srf_F	Bridge_				
		750	1 760	770	780	790	80
730 AGAACAGCACATTIG TCTTGTCGTGTAAAC K N S T F HFLK	740 TCAGGGTCC AGTCCCAGG V R V (1 D3	750 ATGAAAAG TACTTTTC H E K>	CCGGGCCGGCGA	CAAAACTCACA GTTTTGAGTGI	CATGCCCAO	CGTGCCCA GCACGGGTV	GCACCTG
730 AGAACAGCACATTTG TCTTGTCGTGTAAAC K N S T F HFLK	740 TCAGGGTCC AGTCCCAGG V R V (1 D3	750 ATGAAAAG TACTITICO H E K>	G P G>	CAAAACTCACA	ACATGCCCAO INTACGGGTG	CGTGCCCA	GCACCTG
730 AGAACAGCACATTTG TCTTGTCGTGTAAAC K N S T F HFLK	740 TCAGGGTCC AGTCCCAGG V R V (1 D3	ATGAAAAG TACTITICO H E K>	G P G>	CAAAACTCACA GTTTTGAGTGI	T C P	CGTGCCCA GCACGGGT P C P	GCACCTG CGTGGAC
730 AGAACAGCACATTTG TCTTGTOGTGTAAAC K N S T F HFLK	740 TCAGGGTCC AGTCCCAGG V R V 1 (1 D3	ATGAAAAG TACTITTCO H E K>	G P G>	CAAAACTCACA GTTTTGAGTGI K T H	T C P	CGTGCCCA GCACGGGT P C P	GCACCTG
730 AGAACAGCACATTTG TCTTGTCGTGTGAAAC K N S T F HFLK	740 TCAGGGTCC AGTCCCAGG V R V (1 D3	ATGAAAAG TACTITTCO H E K>	G P G>	CAAAACTCACA GTTTTGAGTGI K T H	T C P	CGTGCCCA GCACGGGT P C P	A P
730 AGAACAGCACATTTG TCTTGTCGTGTGTAAAC K N S T F HFLK 810	740 TCAGGGTCC AGTCCCAGG V R V (1 D3	ATGAAAAG TACTTTTCC H E K> 	G P G> D	K T H 850	T C P FCACI (A 860	CGTGCCCA GCACGGGT P C P .) 870	A P
730 AGAACAGCACATTTG TCTTGTCGTGTGTAAAC K N S T F HFLK  810 CTCCTGGGGGGACCC	740 TCAGGGTCC AGTCCCAGG V R V (1 D3 820 STCAGTCTTC	830	G P G> G P G> B B B B B CCCAAAACCCA	K T H 850 AGGACACCCTX	T C P FCAC1 (A 860 CATGATCTCC	PCP BCP BCP BCP B70 CGGACCCC GCCTGGGG	A P
730 AGAACAGCACATTTG TCTTGTCGTGTGTAAAC K N S T F HFLK 810 CTCCTGGGGGGGACCC GAGGACCCCCTGGC L L C C C P	740 TCAGGGTCC AGTCCCAGG V R V 1 D3 820 STCAGTCTTCC CAGTCAGAAG	830 CCTCTTCCCC BL F P	G P G> G P G> B40 CCCAAAACCCA GGGTTTTGGGT	K T H 850 AGGACACCTO AGGACACCTO K D T L	T C P FCAC1 (A 860 CATGATCTCC JATGATCTCC M I S	PCP BCP BCP BTO CGGACCCC GCCTGGGG RTP	A P 8 TGAGGTCAG ACTCCAG E V
730 AGAACAGCACATTTG TCTTGTCGTGTGTAAAC K N S T F HFLK 810 CTCCTGGGGGGGACCC GAGGACCCCCCTGGC L L G G P	740 TCAGGGTCC AGTCCCAGG V R V 1 D3 820 STCAGTCTTC CAGTCAGAAG S V F	830 CCTCTTCCC SGAGAAGGG L F P	G P G> G P G> B40 CCCAAAACCCA GGTTTTGGGT P K P	K T H 850 AGGACACCT TCCTGTGGGA K D T L	T C P FCAC1 (A 860 CATGATCTCC TACTAGAGG M I S	PCP ) 870 CGGACCCC CCCTGGGG R T P	A P B B B B B B B B B B B B B B B B B B B
730 AGAACAGCACATTTG TCTTGTCGTGTGTAAAC K N S T F HFLK 810 CTCCTGGGGGGGGCCCC GAGGACCCCCCTGGC L L G G P	740 TCAGGGTCC AGTCCCAGG V R V 1 D3 820 STCAGTCTTC CAGTCAGAAG S V F	830 CCTCTTTCCC SGAGAAGGG L F P	G P G> G P G> B40 CCCAAAACCCA GGGTTTTGGGT P K P FCAC1 (A)	K T H 850 AGGACACCCT TCCTGTGGGAG K D T L	T C P FCAC1 (A 860 CATGATCTCC 3TACTAGAGG M I S	P C P ) 870 CGGACCCC GCCTGGGG R T P	A P B B TGAGGTC ACTCCAG E V
730 AGAACAGCACATTTG TCTTGTCGTGTGTAAAC K N S T F HPLK 810 CTCCTGGGGGGGGCCC GAGGACCCCCCTGGG L L G G P	740 TCAGGGTCC AGTCCCAGG V R V 1 D3 820 STCAGTCTTC CAGTCAGAAG S V F	830 CCCCCTCCCC CCCCCCCCCCCCCCCCCCCCCCCCCC	G P G> G P G> B40 CCCAAAACCCA GGGTTTTGGGT P K P FCAC1 (A)	K T H 850 AGGACACCTX TCCTGTGGGAG K D T L	T C P FCAC1 (A 860 CATGATCTCCC STACTAGAGG M I S	PCP STO CGGACCCC CGGACCCC GCCTGGGG RTP 250	A P B B TGAGGTC ACTCCAG E V
730 AGAACAGCACATTTG TCTTGTCGTGTGTAAAC K N S T F HFLK 810 CTCCTGGGGGGGACCC GAGGACCCCCCTGGC L L G G P 	740 PICAGGGTCC AGTCCCAGG V R V (1 D3 820 PICAGTCTTC CAGTCAGAAG S V F 900	830 CTCTTTCCC 830 CTCTTCCCC CCACATCCCC CCACAACGGG L F P 910	G P G> G P G> B40 CCCAAAACCCA GGGTTTTGGGT P K P FCAC1 (A) 920 CTGACGTCAAC	K T H 850 AGGACACCCTX TCCTGTGGGAG K D T L 930	T C P FCAC1 (A 860 CATGATCTCC 3TACTAGAGG M I S 940 ACGTGGACGG	PCP BCP BTO BTO CGGACCCC GCCTGGGG RTP 950 CCCTGGAGG	A P 8 TGAGGTCAG ACTCCAG 9 TGCATAA
730 AGAACAGCACATTIG TCTTGTCGTGTGTAAAC K N S T F 	740 TCAGGGTCC AGTCCCAGG V R V 1 D3 820 STCAGTCTTCC CAGTCAGAAG S V F 900 ACGTGAGCCZ	830 CCTCTTCCC SGAGAAGGG L F P 910 ACGAAGACC CCTCTCGG	G P G> G P G> B40 CCCAAAACCCA GGGTTTTGGGT P K P FCAC1 (A) 920 CTGAGGTCAAG GGCTCAAGTCCA	K T H 850 AGGACACCCTV TCCTGTGGGAA K D T L 930 STTCAACTGGT CAAGTTGACCA	T C P FCAC1 (A 860 CATGATCTCC STACTAGAGG M I S 940 ACGTGGACGG	P C P ) 870 CGGACCCC GCCTGGGG R T P 950 CGTGGAGG CGCACCTCC	A P 8 TGAGGTCAG E V 9 TGCATAA ACGTATI
730 AGAACAGCACATTTG TCTTGTCGTGTGTAAAC K N S T F HFLK 810 CTCCTGGGGGGGACCC GAGGACCCCCCTGGC L L G G P 890 ATGCGTGGTGGTGGTGGTGGT TACGCACCACCACCACC	740 TCAGGGTCC AGTCCCAGG V R V 1 D3 820 STCAGTCTTC CAGTCAGAAG S V F 900 ACGTGAGCCA D V S F	ATGAAAAGX TACTTTTCC H E K> 830 CCTCTTCCCC X3AGAAGGG L F P 910 ACGAAGACC IGCTTCTGG H E D	G P G> G P G> B40 CCCAAAACCCA GGGTTTTGGGT P K P FCAC1 (A) 920 CCCAAGGTCAAG KGACTCCAGTTC P E V K	K T H 850 AGGACACCTA K D T L 930 STTCAACTGGT CAAGTTGACCA F N W	T C P FCAC1 (A 860 CATGATCTCC FTACTAGAGG M I S 940 ACGTGGACGG Y V D G	P C P B70 CGGACCCC GCCTGGGG R T P 950 CGTGGAGG CGCACCTCC V E	A P 8 TGAGGTCAG ACTCCAG E V 9 TGCATAA ACGTATI V H N
730 AGAACAGCACATTTG TCTTGTCGTGTGTAAAC K N S T F HFLK 810 CTCCTGGGGGGGGACCC GAGGACCCCCCTGGC L L G G P 890 ATGCGTGGTGGTGGTGGT TACGCACCACCACCACCACCACCACCACCACCACCACCACC	740 TCAGGGTCC AGTCCCAGG V R V (1 D3 820 STCAGTCTTC CAGTCAGAAG S V F 900 ACGTGAGCCZ IGCACTCGGT D V S F	830 CCTCTTCCC SACGAAGAG H E K> 	G P G> G P G> B40 CCCAAAACCCA GGGTTTTGGGT P K P FCAC1 (A) 920 CCTGAGGTCAAG KIACTCCAGTTC P E V K	K T H 850 AGGACACCCTX TCCTGTGGGAX K D T L 930 STTCAACTGGT CAGGTGACCA F N W	T C P FCAC1 (A 860 CATGATCTCC STACTAGAGG M I S 940 ACGTGGACGG Y V D C	P C P B C P C C P C C P C C P C C C C C C C C C C C C C C C C C C C C C C	A P B CGTGGAC A P B TGAGGTC ACTCCAG CACTCCAG CACTCCAG CACTCCAG CACTCCAG CACTCCAG CACTCCAG CACTCCAG CACTCCAG CACCTGAC CGTGGAC A P CGTGGAC A C A C A C A C A C A C A C A
730 AGAACAGCACATTIG TCTTGTCGTGTGTAAAC K N S T F HPLK 810 CTCCTGGGGGGGGCCC GAGGACCCCCCTGGC L L G G P 890 ATGCGTGGTGGTGGTGGT C V V V J 	740 TCAGGGTCC AGTCCCAGG V R V (1 D3 820 STCAGTCTTC CAGTCAGAAG S V F 900 ACGTGAGCCA TGCACTCGGT D V S F	830 CCCCTTCCC SGAGAAGAG L F P 910 ACGAAGACC RGCTTCTGG H E D	γ         γ           G         P         G           G         P         G           B40         GCCAAAACCCA           GGGTTTTGGGT         P         K           P         K         P          FCAC1 (A)         920           GGGTCCAAGGTCCAGGTCCAGTTC         P         K           P         V         K          FCAC1 (A)         920	K T H 850 AGGACACCTX TCCTGTGGGAX K D T L 930 STTCAACTGGT CAAGTTGACCA F N W	T C P FCAC1 (A 860 CATGATCTCC STACTAGAGG M I S 940 ACGTGGACGG Y V D G	P C P B C P C C P C C P C C C C C C C C C C C C C C C C C C	A P 8 TGAGGTCAG ACTCCAG E V 9 TGCATAA 2 CGTATI V H N 2
730 AGAACAGCACATTIG TCTTGTCGTGTGTAAAC K N S T F HFLK 810 CTCCTGGGGGGGACCC GAGGACCCCCCTGGC L L G G P 890 ATGCGTGGTGGTGGTGGJ TACGCACCACCACCACCACCACCACCACCACCACCACCACC	740 PICAGGGTCC AGTCCCAGG V R V (1 D3 820 PICAGTCTTC CAGTCAGAAG S V F 900 ACGTGAGCC2 PICAGTCAGACC2 D V S F 980	830 CTCTTTCCC 830 CTCTTTCCCC SCACAAGGG L F P 910 ACGAAGACCC IGCTTCTGG H E D	G P G> G P G> G P G> B40 CCCAAAACCCA GGGTTTTGGGT P K P FCAC1 (A) 920 CTGAGGTCAAG KGACTCCAGTTC P E V K FCAC1 (A) 1000	K T H 850 AGGACACCCTX TCCTGTGGGAX K D T L 930 STTCAACTGGT CAAGTTGACCAX F N W 1010	T C P FCAC1 (A 860 CATGATCTCC 3TACTAGAGG M I S 940 ACGTGGACGG TGCACCTGCC Y V D C 1020	P C P P C P B70 CGGACCCC GCCTGGGG R T P 950 CGTGGAGG CGCACCTCC V E 1030	A P 8 TGAGGTCAG ACTCCAG E V 9 TGCATAA CACGTATI V H N 2 10
730 AGAACAGCACATTIG TCTTGTCGTGTGTAAAC K N S T F HFLK 810 CTCCTGGGGGGGACCC GAGGACCCCCCTGGC L L G G P 890 ATGCGTGGTGGTGGTGG TACGCACCACCACCACCACCACCACCACCACCACCACCACC	740 PICAGGGTCC AGTCCCAGG V R V (1 D3 820 PICAGTCTTC CAGTCAGAAG S V F 900 ACGTGAGCCZ IGCACTCGGT D V S F 980 CGGGAGGAGG	830 CATGAAAGGA H E K> 830 CATCTTCCC XGAGAAGGG L F P 910 ACGAAGACC IGCTTCTGG H E D 990 CAGTACAAC	G         P         G           G         P         G           G         P         G           B40         GCCAAAACCCA           GGGTTTTGGGT         P         K           P         K         P          FCAC1 (A)         920         GCTGAGGTCAAG           GGACTCCAGTCCAGTCC         P         E         V           FCAC1 (A)         1000         CAGCACGTACCCA	K T H 850 AGGACACCCTX TCCTGTGGGAX K D T L 930 STTCAACTGGT CAAGTTGACCA F N W 1010 STGTGGTCAGC	T C P FCAC1 (A 860 CATGATCTCCC STACTAGAGG M I S 940 ACGTGGACGG TGCACCTGCC Y V D C 1020 GTCCTCACCC	P C P P C P 1) 870 CGGACCCC GCCTGGGG R T P 950 CGTGGAGG CGTGGAGG CGTCGGAGG CGCTCGGAGG CGCTCGCAC	A P 8 TGAGGTC A P 8 TGAGGTC ACTCCAG CACGTATI V H N 2 10 CCAGGAC
730 AGAACAGCACATTIG TCTTGTCGTGTGTAAAC K N S T F 	740 PICAGGGTCC CAGTCCCAGG V R V C1 D3 820 STCAGTCTTCC CAGTCAGAAG S V F 900 ACGTGAGCCA D V S F 980 CGGGAGGAGG	ATGAAAAGX TACTTTTCC H E K> 830 CCTCTTCCCC XGAGAAGAGG L F P 910 ACGAAGACC RGCTTCTGG H E D 990 CAGTACAAC STCATGTTC	G P G> G P G> B40 CCCAAAACCCA GGGTTTTGGGT P K P FCAC1 (A) 920 CTGAGGTCAAG GACTCCAGTTC P E V K FCAC1 (A) 1000 CAGCACGTACCC	K T H 850 AGGACACCTV TCCTGTGGGAA K D T L 930 STTCAACTGGT CAAGTTGACCA F N W 1010 STGTGGTCAGCC	T C P FCAC1 (A 860 CATGATCTCC FTACTAGAGG M I S 940 ACGTGGACGG TGCACCTGCC Y V D C 1020 GTCCTCACCC CAGGAGTGG	P C P P C P 1) 870 CGGACCCC GCCTGGGG R T P 950 CGTGGAGG CGTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGG CGCTGGAGGACGT CGCTGGAGGACGT CGCTGGAGGACGT CGCTGGAGGACGT CGCTGGAGGACGT CGCTGGAGGACGT CGCTGGAGGACGT CGCTGGAGGACGT CGCTGGAGGACGT CGCTGGAGGACGT CGCTGGAGGACGT CGC CGC CGC CGC CGC CGC CGC	A P 8 TGAGGTCAG A P 8 TGAGGTCAG ACTCCAG CACGTATI V H N 2 10 CCAGGACT

US 7,070,959 B1

U.S. Patent

Jul. 4, 2006

Sheet 33 of 55

# Fig.21C.

FIG	J.21C	•						
1050	1060	1070	1080	1090	1100	1110	11	20
TGAATGGCAAGGZ	GTACAAGTG	CAAGGICICC	AACAAAGCCC	ICCCAGCCCC	CATCGAGAAA	ACCATCICO	CAAAGCC	AA. TMT
ACTTACCGTTCCI	CATGTTCAC	GTTCCAGAGO	TIGHTICGGG	AGGGICGGGG . P & P	TEK	TIS	K A	 
LNGKI	SYKC	K V S	MKA					35:
			FCAC1 (A)					>
				>A:	>C_A_allot	ype		
				- C- T 3	allotime			
				×9>1_A				
1130	1140	1150	1160	1170	1180	1190	12	00
GCGCAGCCCCGA	GAACCACAGG	TGTACACCC	GCCCCCATCC	CGGGATGAGC	TGACCAAGAA	CCAGGTCA	GCCIGAC	CT
CCCGTCGGGGCT	TTGGTGTCC	ACATGTGGG	CGGGGGTAGG	GCCCTACTCG	ACTGGTTCTI	GGTCCAGT	CGGACTG	GA
GQPR	ЕРQ	νγτι	. P P S	RDE	LTKN	QV	SLT	>
							3	
			FCACI (A)					
1210	1220	1230	1240	1250	1260	1270	12	80
	ንግግክልሞጋጥሞ ግግግ	ACCEACATE	CCGTGGAGTG	GGAGAGCAAT	GGGCAGCCGG	AGAACAAC	TACAAGA	CC
CIGGICARAGG	CINCLATECCO	TCGCTGTAG	GGCACCTCAC	CCTCTCGTTA	CCCGTCGGCC	TCTTGTTG	ATGTTCT	GG
L V K G	FYP	SDI	AVEW	ESN	GQP	ENN	Y K	T>
lvko								404
			FCAC1 (A)					>
				>T>C				
			1220	1320	1340	1350	13	360
1290	1300		1320 2720-2020	OCAACTCAC	CGTGGACAA	AGCAGGTO	GCAGCA	GG
CGCCTCCCGTGC	ACCTUCGA ACCTUCGA	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	GAAGGAGATAI	CGTTCGAGTC	GCACCTGTT	TCGTCCAC	CGTCGT	CCC
T P P V	L D S I	GSF	FLY	SKLI	VDK	SRW	V Q Q	G>
								43
			FCΔC1 (A)					`
						1420		440
1370	1380	1390	1400	1410	1420		1	240 277 a
AACGTCTTCTCA	TGCTCCGTG	ATGCATGAGG	CTCTGCACAA	CACTACACG	AGAAGAGCC	1CTCCCIGI	ACAGGCC	CAT
TIGCAGAAGAGT	ACGAGGCAC	PACGTACTCC	GAGACGIGII	JU V T		L C L	SP	G>
NVFS	C S V	мне	ALHN	n i i	Ϋ́́Ϋ́́Ϋ́́Υ			457
			FCAC1 (A	<b>`</b>				
Not site								
I								
11450								
AATGAGCGGCCGC	2							
TACTOGCOGGOG	;							
< *>								
158								
>								
<u>ا</u> ا	g.22A	•		>EcoRI	_site			
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	
10	20	30	40	50	60	70	80	
AGCTTGGGCTGCA	GGTCGATCGA	CTCTAGAGGA	TOGATCCCCC	GGCGAGCTCC	GAATTCGCAA	CACCATGGI	CAGCTAC	
TCGAACCCGACGI	CCAGCTAGCT	GAGATCTCCT	AGCTAGGGGG	CCGCTCGAG	TTAAGCGTT	GIGGTACCA	GTCGATG	
						M V 1		
						-	•	
					>BspEI_br:	idge		
90	100	110	120	130	140	150	160	
GGGACACCGGGGI	CCTGCTGTGC	GCGCTGCTCA	GCIGICIGCI	TCTCACAGG	TCTAGTICCO	GAGGTAGAC	CTTTCGT	
CCCTGTGGCCCCA	GGACGACACG	CGCGACGAGI	CGACAGACGA	AGAGTGTCC	PAGATCAAGG	CTCCATCTO	GAAAGCA	
WDTGV	LLC	ALL	SCLI	, L T G	S S>			
	FLT1 3	SIGNAL SEQ	UENCE		s	G>		
						>		
						GR	PFV>	
							31	
							>	
170	180	190	200	210	220	230	240	
1 / 1 /								
AGAGATGTACAGTO ICTCTACATGTCAO E M Y S	AAATCCCCGA TTTAGGGGCT E I P E	AATTATACAC TTAATATGTC I I H	CATGACTGAAG TACTGACTTC M T E	GAAGGGAGC CTTCCCTCG	ICGTCATTCC AGCAGTAAGG L V I P	CTGCCGGGT GACGGCCCA/ C R V	LACGTCAC ATGCAGTG T S>	
GAGATGTACAGTC CTCTACATGTCAC E M Y S	BAAATCCCCGA CTTTAGGGGCT E I P E	AATTATACAC TTAATATGTC I I H FL1	CATGACTGAAC STACTGACTTX M T E M I G DOMAJ	GAAGGGAGC CCTTCCCTCG G R E I	ICGTCATTCC AGCAGTAAGG L V I P	CTGCCGGGT GACGGCCCA/ C R V	TACGTCAC ATGCAGTG T S> 57	
AGAGATGTACAGTO ICTCTACATGTCAC E M Y S	CAAATCCCOGA TTTAGGGGCT E I P E	AATTATACAC TTAATATGTC I I H FL1	TACTGACTGAAK TACTGACTTK M T E T1 IG DOMA	GAAGGGAGC CCTTCCCTCG G R E I IN 2	ICGTCATTCC AGCAGTAAGG L V I P	CTGCCGGGT GACGGCCCA/ C R V	EACGTCAC ATGCAGTG T S> 57 320	
GAGATGTACAGTO CTCTACATGTCAC E M Y S 250	EAAATCCCCGA TTTTAGGGGCT E I P E 260	AATTATACAC TTAATATGTC I I H FL7 270	CATGACTGAAG STACTGACTIX M T E M T E M IG DOMAJ 280	GREI 290 290	CGTCATTCC AGCAGTAAGG L V I P 300	CTGCCGGGT GACGGCCCAA C R V 310 TAATCTGGG	EACGTCAC ATGCAGTG T S> 57 320 ACAGTAGA	
CTCTACAGTC CCTCTACATGTCAC E M Y S 250 CTAACATCACTGTT SATTGTACTGACAZ	EAAATCCCCGA TTTTAGGGGCT E I P E 260 FACTTTAAAAA YIGAAATTTTT	AATTATACAC TTAATATGTC I I H FL1 270 AGTTTCCAC1 TCAAAGGTG2	CATGACTGAAC STACTGACTIX M T E 11 IG DOMAJ 280 FTGACACTIX AACTGTGAAAC	GRAGGGAGC CTTCCCTCG GREI IN 2 290 SATCCCTGAT TAGGGACTAI	CGTCATTCC AGCAGTAAGG L V I P 300 3GAAAACGCA CCTTTTGCGT	CTGCCGGGT GACGGCCCA/ C R V 310 TAATCTGGG/ ATTAGACCC	EACGTCAC ATGCAGTG T S> 57 320 ACAGTAGA RGTCATCT	
CAGAGATGTACAGTO COTOTACATGTCAC E M Y S 250 CTAACATCACTGTI SATTGTAGTGACAA P N I T V	EAAATCCCCGA TTTTAGGGGCT E I P E 260 FACTTTAAAAA TGAAATTTTT T L K 1	AATTATACAC TTAATATGTC I I H FL7 270 AGTITCCACT TCAAAGGTC2 K F P I	CATGACTGAAG STACTGACTTX M T E M T E 280 PTGACACTTX AACTGTGAAAG L D T L	GRAGGGAGC CCTTCCCTCG GREI IN 2 290 SATCCCTGAT CTAGGGACTAU IPD	CCTTTTGCGT G K R	CTGCCGGGTT GACGGCCCA/ C R V 310 TAATCTGGG/ ATTAGACCCT I I W I	EACGTCAC T S> 57 320 ACAGTAGA NGTCATCT D S R> 84	
AGAGATGTACAGTO INTETACATGTCAC E M Y S 250 250 CTAACATCACTGTT SATTGTAGTGACAA P N I T V	EAAATCCCCGA TTTTAGGGGCT E I P E 260 TACTTTAAAAA ATGAAATTTTT T L K	AATTATACAC TTAATATGTC I I H FL7 270 AGTTTCCACT TCAAAGGTG2 K F P I FL7	CATGACTGAAC TACTGACTTX M T E 11 IG DOMAJ 280 FTGACACTTX AACTGTGAAAC C D T L F1 IG DOMA	GAAGGAGC CTTCOCTOG G R E I 1N 2 290 SATOCCTGAT CTAGGGACTAU I P D IN 2	AGCAGTAAGG L V I P 300 SGAAAACGCA CCTTTIGCGT G K R	CTGCCGGGTT GACGGCCCA/ C R V 310 TAATCTGGG/ ATTAGACCCI I I W I	PACGTCAC T S> 57 320 ACAGTAGA RGTCATCT D S R> 84	
CAGATGTACAGTC CTCTACATGTCAC E M Y S 250 CTAACATCACTGTT SATTGTAGTGACAT P N I T V	EAAATCCCCGA TTTTAGGGGCT E I P E 260 TACTTTAAAAA ATGAAATTTTT T L K 1	AATTATACAC TTAATATGTC I I H FL7 270 AGTITICCACT TCAAAGGTG2 K F P I FL7 350	CATGACTGAAG STACTGACTTX M T E M T E M T E 280 FTGACACTTX AACTGTGAAAG D T L F1 IG DOMA 360	GREI CTTCCCTCG GREI IN 2 290 SATCCCTGAT CTAGGGACTA IPD IN 2 370	CCTTTTCCCACCACTAAGG L V I P 300 3GAAAACGCA CCTTTTGCGT G K R 380	CTGCCGGGT GACGGCCCA/ C R V 310 TAATCTGGG/ ATTAGACCC I I W I	EACGTCAC ATGCAGTG T S> 57 320 ACAGTAGA RGTCATCT D S R> 84 > 400	
AGAGATGTACAGTO CCTCTACATGTCAC E M Y S 250 CTAACATCACTGTT SATTGTAGTGACAZ P N I T V 330 AAGGGCTTCATCAT	EAAATCCCCGA TTTTAGGGGCT E I P E 260 FACTTTAAAAA ATGAAATTTTT T L K 1 340 FATCAAATGA	AATTATACAC TTAATATGTC I I H FL7 270 AGTTTCCACT TCAAAGGTG2 K F P I FL7 350 ACGTACAAAG	CATGACTGACA TACTGACTAC M T E 11 IG DOMAJ 280 TTGACACTATA AACTGTGAAAA C D T L T1 IG DOMA 360 SAAATAGGGC	GRED CTTCCCTCG GRED 1N2 290 SATCCCTGAT TAGGGACTAN IPD IN2 370 TTCTGACCTG	CGTCATTCC AGCAGTAAGG L V I P 300 SGAAAACGCA CCTTTTGCGT G K R 380 TGAAGCAACA	CTGCCGGGT GACGGCCCA/ C R V 310 TAATCTGGG/ ATTAGACCC I I W I	EACGTCAC ATGCAGTG T S> 57 320 ACAGTAGA RGTCATCT D S R> 84 400 CATTIGTA	
AGAGATGTACAGTO CTCTACATGTCAC E M Y S 250 CTAACATCACTGTT SATTGTAGTGACAZ P N I T V 330 AAGGGCTTCATCM	EAAATCCCCGA TTTTAGGGGCT E I P E 260 TACTTTAAAAA ATGAAATTTTT T L K 1 340 TATCAAATGCA ATAGTTTACGT	AATTATACAC TTAATATGTC I I H 270 AGTTTCCACT TCAAAGGTGZ K F P I FLA 350 ACGTACAAAC TGCATGTTTX	CATGACTGAAG TACTGACTTX M T E 11 IG DOMAJ 280 TTGACACTTX AACTGTGAAAG C D T L 11 IG DOMA 360 GAAATAGGGC CTTTATCCCG	GRAGGGAGC CTTCCCTCG GREI 290 SATCCCTGAT CTAGGGACTAM IPD IN 2 370 ITCTGACCTG AAGACTGGAC	CGTCATTCC AGCAGTAAGG L V I P 300 3GAAAACGCA CCTTTTGCGT G K R 380 TGAAGCAACA ACTTCGTIGT	CTGCCGGGT GACGGCCCAA C R V 310 TAATCTGGGA ATTAGACCC I I W I	EACGTCAC ATGCAGTG T S> 57 320 ACAGTAGA RGTCATCT D S R> 84 400 CATTTGTA STAAACAT	
ITTOTACAGTO CACACATOTACAGTO E M Y S 250 TAACATCACTOTI SATTGTAGTGACAA P N I T V 330 AAGGGCTTCATCAT FTCCCGAAGTAGTA K G F I :	EAAATCCCCGA TTTTAGGGGCT E I P E 260 FACTTTAAAAA ATGAAATTTTT T L K 1 340 FATCAAATGCA ATAGTTTACGT I S N A	AATTATACAC TTAATATGTC I I H 270 AGTITCCACI TCAAAGGTGZ K F P I FIA 350 ACGTACAAAC TGCATGTTX T Y K	CATGACTGACA TACTGACTAC M T E 11 IG DOMAJ 280 PTGACACTTAC AACTGTGAAAG C D T L 11 IG DOMA 360 GAAATAGGGC CTTTATCCCG E I G D	GRED CTTCCCTCG GRED 1N2 290 SATCCCTGATA IPD IN2 370 FTCTGACCTG AGACTGGAC LLTC	CCTTTTGCGT G K R 380 CCTTTGCGT G K R 380 CCAAGCAACA ACTTCGTIGT E A T	CTGCCGGGTT GACGGCCCAA C R V 310 TAATCTGGGA ATTAGACCCT I I W I	EACGTCAC ATGCAGTG T S> 57 320 ACAGTAGA NGTCATCT D S R> 84 400 CATTIGTA STAAACAT H L Y> 111	
GAGATGTACAGTO CTCTACATGTCAC E M Y S 250 CTAACATCACTGTI SATTGTAGTGACAI S N I T V 330 AAGGGCTTCATCAI K G F I S	EAAATCCCCGA TTTTAGGGGCT E I P E 260 FACTTTAAAAA ATGAAATTTTT T L K 1 340 FATCAAATGCA ATAGTTTACGT I S N A	AATTATACAC TTAATATGTC I I H FL7 270 AGTTTCCACT TCAAAGGTG2 K F P I FL7 350 ACGTACAAAC T Y K	CATGACTGACA STACTGACTAX M T E M T E M T E M T E 280 FTGACACTAX ACTGTGACACA D T L C D C C C TTAATCCCG E I G D	SCAAGGGAGCC CCTTCOCTOG G R E I 1N 2 290 SATCCCTGAT CTAGGGACTAN I P D IN 2 370 ITCTGACCTG AAGACTOGAC L L T C	CCTTTTCCGT AGCAGTAAGG L V I P 300 3GAAAACGCA CCTTTTGCGT G K R 380 TGAAGCAACA ACTTCGTTGT E A T	CTGCCGGGT GACGGCCCAA C R V 310 TAATCTGGG2 ATTAGACCC7 I I W I	EACGTCAC ATGCAGTG T S> 57 320 ACAGTAGA NGTCATCT D S R> 84 400 CATTTGTA STAAACAT H L Y> 111 >	
AGAGATGTACAGTO CCTCTACATGTCAC E M Y S 250 CTAACATGTCACTGTI SATTGTAGTGACAA P N I T V 330 AAGGGCTTCATCAT FTCCCGAAGTAGTA K G F I S	EAAATCCCCGA TTTTAGGGGCT E I P E 260 FACTTTAAAAA ATGAAATTTTT T L K 340 FATCAAATGCA ATAGTTTACGT I S N A	AATTATACAC TTAATATGTC I I H FL7 270 AGTTTCCACT TCAAAGGTG2 K F P I FL7 350 ACGTACAAAC TGCATGTTTC T Y K FL7	CATGACTGACA TACTGACTAC M T E M T E M T E M T E M T E 280 TGACACTAC ACTGTGAAAG C D T L F1 IG DOMA SAAATAGGGC TTTATCCCG E I G D F1 IG DOMA	SGAAGGGAGC:           SCTTCOCTOGE           G         R         E           1         290           SATCCCTGATC           TAGGGACTAN           I         P           D           IN 2           370           TTCTGACCTGAC           L         L           T         C           IN 2           370           TTCTGACCTGAC           L         L           IN 2	CGTCATTCC AGCAGTAAGG L V I P 300 SGAAAACGCA CCTTTTGCGT G K R 380 TGAAGCAACA ACTTCGTIGT E A T	CTGCCGGGT GACGGCCCA/ C R V 310 TAATCTGGG/ ATTAGACCC I I W I	CACGTCAC ATGCAGTG T S> 57 320 ACAGTAGA RGTCATCT D S R> 84 400 CATTTGTA STAAACAT H L Y> 111 >	
AGAGATGTACAGTO CTCTACATGTCAC E M Y S 250 CTAACATCACTGTI SATTGTAGTGACAZ P N I T V 330 AAGGGCTTCATCMI FTCCCGAAGTAGTI K G F I 3 410	EAAATCCCCGA TTTTAGGGGCT E I P E 260 TACTTTAAAAA ATGAAATTTTT T L K 340 TATCAAATGCA ATACTTTACGT I S N A 420	AATTATACAC TTAATATGTC I I H FL7 270 AGTTTCCACT TCAAAGGTCA K F P I FL7 350 ACGTACAAAC TGCATGTTTX T Y K FL7 430	CATGACTGAAA TACTGACTTX M T E 11 IG DOMAJ 280 PTGACACTTX AACTGTGAAAA C D T L 11 IG DOMA 360 GAAATAGGGC CTTTATCCCG E I G D F1 IG DOMA 440	GRAGGGAGC         CTTCOCTOGE         GREIN         290         SATOCCTGATIC         CTAGGGACTAM         IPD         IN2         370         TTCTGACCTGAC         LLTC         IN2         370         TTCTGACCTGAC         LLTC         AGACTGGAC         IN2         450	CGTCATTCC AGCAGTAAGG L V I P 300 SGAAAACGCA CCTTTTGCGT G K R 380 TGAAGCAACA ACTTCGTIGT E A T 460	CTGCCGGGT GACGGCCCAA C R V 310 TAATCTGGGA ATTAGACCC I I W I	ACGTCAC T S> 57 320 ACAGTAGA MGTCATCT D S R> 84 400 CATTIGTA STAAACAT H L Y> 111 480	
AGAGATGTACAGTO NCTCTACATGTCAC E M Y S 250 250 250 250 250 250 250 250 250 250	EAAATCCCCGA TTTTAGGGGCT E I P E 260 TACTTTAAAAA ATGAAATTTTT T L K 1 340 TATCAAATGCA ATAGTTTACGT I S N A 420 CTCACACATCG	AATTATACAC TTAATATGTC I I H FL7 270 AGTTTCCACT TCAAAGGTG2 K F P I FL7 350 ACGTACAAAC TGCATGTTTX T Y K FL7 430 ACAAACCAA	CATGACTGACA TACTGACTAC M T E 11 IG DOMAJ 280 PTGACACTATA AACTGTGAAAG D T L 11 IG DOMA 360 GAAATAGGGC CTTTATCCCG E I G D F1 IG DOMA 440 FACAATCATAA	SGAAGGGAGC           SCTTCOCTOGE           G         R         E           IN 2           290           SATCCCTGATC           TAGGGACTAN           I         P           IN 2           370           TTCTGACCTGAC           L         L           TC           IN 2           370           TTCTGACCTGAC           L         L           T         C           IN 2           450           GATATCCAGC	CGTCATTCC AGCAGTAAGG L V I P 300 3GAAAACGCA CCTTTTGCGT G K R 380 TGAAGCAACA ACTTCGTTGT E A T 460	CTGCCGGGTT GACGGCCCAA C R V 310 TAATCTGGGA ATTAGACCCT I I W I 	CACGTCAC ATGCAGTG T S> 57 320 ACAGTAGA NGTCATCT D S R> 84 400 CATTIGTA STAAACAT H L Y> 111 480 GGAGCTGC	
GAGATGTACAGTO CTCTACATGTCAC E M Y S 250 TTAACATCACTGTT SATTGTAGTGACAA N I T V 330 AAGGGCTTCATCAM MTCCCGAAGTAGTA K G F I 3 410 TAAGAACAAACTATX	EAAATCCCCGA TTTTAGGGGCT E I P E 260 FACTTTAAAAA ATGAAATTTTT T L K 1 340 FATCAAATGCA ATAGTTTACGT I S N A 420 CTCACACATCG EAAGTGTGTAGC	AATTATACAC           TTAATATGTC           I         I          FL7           270           AGTTTCCACT           TCAAAGGTG2           K         F          FL7           350           ACGTACAAAC           T         Y           K          FL7           350           ACGTACAAAAC           T           Y           K          FL7           430           ACAAACCAAM           TGTTTGGTTT	CATGACTGACA TACTGACTAC M T E 11 IG DOMAJ 280 TTGACACTTAC ACTGTGACACTAC ACTGTGACACTAC C D T L 11 IG DOMA 360 SAAATAGGGC TTTATCCCG E I G D F1 IG DOMA 440 TACAATCATAC	SGAAGGGAGC:           SCTTCOCTOGE           G         R         E           1N         2           290         SATCCCTGAT           TAGGGACTAN         I         P           IN         2	CCGTCATTCC AGCAGTAAGG L V I P 300 3GAAAACGCA CCTTTTGCGT G K R 380 TGAAGCAACA ACTTCGTTGT E A T 460 TGTTGCCCAG ACAACGGGTC	CTGCCGGGT GACGGCCCAA C R V 310 TAATCTGGGZ ATTAGACCC7 I I W I	EACGTCAC ATGCAGTG T S> 57 320 ACAGTAGA NGTCATCT D S R> 84 400 CATTTGTA STAAACAT H L Y> 111 480 GGAGCTGC CCTCGACG	
AGAGATGTACAGTO CTCTACATGTCAC E M Y S 250 250 250 250 250 250 250 250	EAAATCCCCGA TTTTAGGGGCT E I P E 260 PACTTTAAAAA ATGAAATTTTT T L K 340 PATCAAATGCA ATAGTTTACGT I S N A 420 CTCACACATCG GAGTGTGTAGCC L T H R FUTI IG DO	AATTATACAC TTAATATGTC I I H FL7 270 AGTTTCCACT TCAAAGGTG2 K F P I FL7 350 ACGTACAAAC TGCATGTTTC T Y K FL7 430 ACAAACCAA? TGTTTGGTTT C T N MAIN 2	CATGACTGACA TACTGACTAC M T E M T E M T E M T E M T E 280 TGACACTAC ACTGTGAAAG C D T L F1 IG DOMA 360 SAAATAGGGC CTTTATCCCG E I G D F1 IG DOMA 440 FACAATCATACATACATACATACATACATACATACATACA	SCAAGGGAGCS CCTTCOCTOG G R E I 1N 2 290 SATCCCTGATC TTAGGGACTAA I P D IN 2 370 TTCTGACCTG AAGACTGGACC L L T C IN 2 450 GATATCCAGC CTATAGGTCG D>	CCTTATTCC AGCAGTAAGG L V I P 300 SGAAAACGCA CCTTTTGCGT G K R 380 TGAAGCAACA ACTTCGTTGT E A T 460 TGTTGCCCAG ACAACGGGTC	CTGCCGGGT GACGGCCCA/ C R V 310 TAATCTGGG/ ATTAGACCC I I W I	ACGTCAC T S> 57 320 ACAGTAGA RGTCATCT D S R> 84 400 CATTTIGTA STAAACAT H L Y> 111 480 GGAGCTGC CCTCGACG	
AGAGATGTACAGTO NCTCTACATGTCAC E M Y S 250 TTAACATCACTGTT SATTGTAGTGACA2 N I T V 330 AGGGCTTCATCM MTCCCGAAGTAGTA K G F I : 410 CAAGACAAACTATX NTCTGTTTGATAGACA	EAAATCCCCGA TTTTAGGGGCT E I P E 260 TACTTTAAAAA ATGAAATTTTT T L K 1 340 TATCAAATGTA TAGATTACGT I S N A 420 CTCACACATCG BAGTGTGTAGC L T H R _FLT1 IG DC	AATTATACAC TTAATATGTC I I H FL7 270 AGTTTCCACT TCAAAGGTG2 K F P I FL7 350 ACGTACAAAC TGCATGTTTC T Y K FL7 430 ACGTACAAACCAA? TGTTTGGTT Q T N MAIN 2	CATGACTGACA TACTGACTAC M T E M T E M T E M T E M T E 280 PTGACACTAC AACTGTGAAAG C D T L F1 IG DOMA 360 GAAATAGGGC CTTTATCCCG E I G D F1 IG DOMA 440 FACAATCATAA ATGTTAGTATA T I I	GAAGGGAGC CCTTCOCTOG G R E I IN 2 290 SATOCCTGAT CTAGGGACTAM I P D IN 2 370 ITCTGACCTGA AGACTOGAC L L T C IN 2 450 GATATCCAGC CTATAGGTCG D> > I Q	CGTCATTCC AGCAGTAAGG L V I P 300 SGAAAACGCA CCTTTTGCGT G K R 380 TGAAGCAACA ACTTCGTTGT E A T 460 TGTTGCCCAG ACAACGGTC L L P F	CTGCCGGGT GACGGCCCAA C R V 310 TAATCTGGGA ATTAGACCC I I W I 	EACGTCAC ATGCAGTG T S> 57 320 ACAGTAGA NGTCATCT D S R> 84 400 CATTIGTA STAAACAT H L Y> 111 480 GGAGCTGC CCTCGACG E L> 137	

Fig.22B.

	490			500			51	10		:	520			530			54	0			550	ļ		5	60
IGGTAC	GGGGA	GAAC	<b>CTC</b>	GTC	CI	CAA	CTG	CACC	GT	GIG	GGC	TGA	GTT	TAAC	TCA	GGT	GIC	ACC	TT	IGA	CTG	GGA	CTZ	ACC.	CA
CCAT	CCCCT	CTT	GAC	CAG	GAC	GTT	GACO	STGC	CA.	CAC	CCG	ACT		ATTG	AGI	ICC#		TGG	AA.	ACT	GAC		GAT	IGG ,	GT
v	GΕ	ĸ	Г	v	г	N	C	T	v	w	A	E	F	N	э	G	v	.1.	ŗ	D				5	16
							v	GFF	23	(FL	T4)	IG	DO	MAIN	3										
	570			580			59	90		(	600			610	)		62	0			630	)		б	40
GGAAC	GCAGG	CAGZ	AGC(	GGG	TA/	AGT	GGG1	rgcc	CG.	AGC	GAC	GCT	ccc	AACA	GAC	CC7	CAC	AGA	AC.	ICT	CCA	GCA	TCC	CTG	AC
CCTIC	CGTCC	GTCI	rcgo	CCC	ATT	<b>FCA</b>	ccci	ACGC	GC	TCG	CIG	CGA	GGG	TIGI	CIG	GG7	GTG	TCI	'TG/	AGA	GGI	CGI	'AGC	GAC	TG
GK	Q	AE	5 F	G	ł	K I	W V	V I	2	ΕI	R	R	S	Q Q	5 1	r 1	4 T	· F		6	S	S	T	г	1
							v	TOPI	23	(FL	T4 )	IG	DO	MAIN	13										1
											/														
	650			660			67	70			680			690	1		70	0			710	)		7	20
ATCCI	ACAAC	GTC	AGCC	AGC	ACC	GAC	CTG	GCI	FCG	TAT	GIG	TGC	AAG	GCCA	ACA	ACC	GCA	TCC	AG	CGA	TTI	CGG	GAC	GAG	CA
TAGG	IGTIG	CAG	rcga	TCG	TG	CTG	GAC	CCGI	AGC.	ATA	CAC	ACG	TTC	CGGI	TGI	TGC	CGI	'AGG	ΠC	GCI		.GCC	CTC	CTC	GT
IH	H N	v	s	Q	H	D	L	G	s	Y	v	С	к	A	N	N	G	I	Q	R	F.	R	Е	2	> 17
							17	CF	22	(181./	TA \	TG	$\mathbf{n}$	MATN	13									•	- '
							v,			(111		10													
	730			740			7	50			760			770	ł		78	80			790	)		8	00
CGAGO	ንጥ እጥ	TGT	CAI	GAA	AA'	IGG	ccc	GGGC	CGA	CAA	AAC	TCA	CAC	ATGO	CCZ	ACCC	TGC	CCA	GC	ACC	TGA	ACT	CC.	IGG	GG
									D	K	Т	н	Т	c	P	P	с	₽	A	Ē	• E	: 1	. 1	6	G> 24
										-				_FC∆0	1.	- A	ALI	OT	PE	<u> </u>					
	810			820			8	30			840			850	)		86	50			870	)		8	80
GACCO	GTCAG	TCT	ICC:	CTI	cca	ccc	CAA	AACO	CCA	AGG	ACA	ccc	TCA	TGAT	CTC	cca	GAC	200	TG	AGO	TC?	\CA7	IGC	GIG	GT
CTGG	CAGTO	AGA	AGGZ	GAA	GG	GGG	GTT	TIGO	GGT	TCC	tgt	GGG	AGI	ACTA	GAG	GGG	CTC		GAC	TCC	CAG	(GT/	1CG	CAC	CA
GΡ	s	V I	F I	F	' 1	Р	PI	K I	P	ĸ	D	T	L	M J	5	5 1	R 1	r 1	2	E	v	т	С	V	V
								-	~ 4 ~				~~	-											2
								F(	CΔC	- 1	A	ALU	013	PD_											
	890			900	)		9	10			920			930	)		94	10			950	3		9	960
GTGG	ACGTO	AGCO	CACO	GAAC	AC	CCT	GAG	GTC	AAG	TTC	AAC	TGG	TAC	GTG	AC	GGC	GIG	SAG	TG	CA	raat	IGCO	CAA	GAC	AA
CACC	TGCAC	TCG	GIG	TTC	TG	GGA	CTC	CAG	TTC	AAG	TTG	ACC	ATC	CAC	TG	CCG	CACO	TCC	CAC	GT/	ATT?	ACGO	ЗТТ	CTO	T
v	D V	s	н	Е	D	Ρ	Е	v	к	F	N	W	Y	v	D	G	v	Ε	V	н	N	А	ĸ	1	['>
																								2	29
								F	CΔC	:1 -	A	ALL	OT	PE_											
	070			001			٥	90		1	000	1		1010	n		103	20			1030	0		10	)4(
ACCC	0 / E	GCM	2020	900 7720	22	CAG	CAC	90 GTA	ccc	TGT	GGT	CAG	CG1	CCT	CAC	CGT	CCT	GCA	CCA	GGJ	ACT	- GGC'	IGA	AT	GG
CCGCC	GCCCT	CCT	CGT	CATC	TT	GTC	GTG	CAT	GGC	ACA	CCA	GTC	GCI	GGA	TG	GCA	GGA	CGT	GGT	CC'	IGA	CCG	ACT	TA	cc
( P	RE	E	Q	Y	N	S	Т	Y	R	v	v	r s	7	/ L	т	v	г	н	Q	2	D	W D	L	N	G
																									3:
								F	C∆C	-1	A	ALL	OT	PE_											

# Fig.22C.

	~ ~ ~	1050	)	~~~	10	60 0 <b>0</b> 00	<b>m</b> 00	1	070		~~~			~~~	202	200		1	200	m	מגגי		נט ממי	ACC	11 12	120 200
AG	GAG	TAC/	AG'I	GCA	AG	GTC	TCC	AAC	AAA	CCC			1600	2000	ATCC TACC	анса ~тү~	n na Trint	ACC TGG	AIC TAG	AGG	aaa TTTI	CG	.aa Stt	TCC	CGT	IGG
ĸ	E	Y	K	C	ĸ	V	S	N	ĸ	A	L	P	A	P	I	E	ĸ	Т	I	s	ĸ	A	K	G	Q	Р 3
										FCΔ	C1	~ 1	A AI	LOT	YPE											
																_	_									
														>A>	C_A	_al:	lot	ype								
												;	•G>1	<u>_</u> A_	alle	ı otyj	pe									
															1	1	-									
		113(	)		11	40		1	150			116	50		11	70		1	180			119	90	~~~	1:	200
CG CG	AGA	ACCI		GTG	TA:	CAC	CCT	GCO	ccci			GGGZ	ATGA PACT	AGCT MCGA	GAC	CAA	GAA CTT	CCA	GGI	CAG	GGZ	CT	JCT 3GA	CCC	ACC)	IC/ AGI
R	ICI E	166. P	0	V	Y Y	T	L	P	P	S	I	RI	) E	EL	T	ĸ	N	r ç	v	, s	5 1		r	C	L V	1>
		-	~	-	-	_																			:	377
			_							FC∆	C1	- 1	A AI	LOT	YPE											
		121(	)		12	20		1	230			124	10		12	50		1	260	,		12	10			8
10	GCI CGA	TCT/ AGA	ATCO	CAG	CG CC	ACA TGT	TCG AGC	CCG CGC	TGG ACC	AGT TCA	GG( CC(	GAGA CTCI	AGCA	ATG TAC	GGC) CCG' G	AGO TCG	CGG GCC P	AGA TCI E	ACA TGI N	ACT TGA	ACZ TG1 V	VAGA PTC' K	ACC TGG T	ACG TGC T	CCTO GGA0 P	JU SG P
iC	GCI CGA G	TCTI AGA F	ATCO PAGO Y I	CAG GTC SGTC	CG CC	ACA TGT D	TCG AGC I	ICCG IGGC A	TGG ACC V	AGT TCA E	GGC CCC W	GAGA CTCT E	AGCA ICGI S	ATG TAC N	GGCJ CCG' G (	AGC TCG Q	CGG GCC P	AGA TCI E	ACA TGI N	ACT TGA N	Y Y	AG PTC' K	TGG T	ACG TGC T	GGA( P	3G P 4
	GCI CGA G	TCTI AGA F	ATCO PAGO Y I	CAG SGTC ? S	CG CC	ACA TGT D	TCG AGC I	A	TGG ACC V	AGT TCA E FCA	GG( CC( W .C1	GAGA CTCT E - i	AGCA ICGT S A AI	ATG TAC N	GGCJ CCG' G (	AGO ICG Q	CGG GCC P	AGA TCI E	ACA TGT N	ACT TGA N	Y Y	VAG. PTC' K	TGG	ACG TGC T	GGA0 P	3G P 4
kG rC	GCI CGA G	TCTI AGA F	ATCO PAGO Y I	CAG GTC P S	ICG	ACA TGT D	TCG AGC I	CCG GGC A	TGG ACC V	AGT( TCA) E ΓCΔ	GG( CC( W .C1	GAG/ CTCT E - J	AGCA ICGI S A AI I>C	ATG TAC N	GGC CCG G (	AGO ICG D	CGG GCC P	AGA TCI E	ACA TGI N	ACT TGA N	Y Y	VAG. I'TC' K	TGG	ACG TGC T	CCTO GGA0 P	20 30 9 4
	GCI CGA G	TCTI AGA F	ATCO PAGO Y I	CAG GTC SGTC		ACA TGT D	TCG AGC I	CCG GGC A	TGG ACC V	AGT( TCA) E 1 FCA	GG( CC( W .C1	GAGA CTC1 E - i	AGCA ICGI S A AI I>C	AATG TTAC N	GGC		CGG GCC P	AGA TCI E	ACA TGT N	ACT TGA N	Y Y	K		ACG TGC T	GGA( P	20 30 P 4
	GCI CGA G	TCT/ AGA F	ATCO PAGO Y 1	CAG GTC SGTC S	13	ACA TGT D	TCG		1GG ACC V 3		660 000 W 01	GAG/ CTC1 E - 1 132	AGCA ICGI S A AI I>C I 20	ATG TAC N	GGC. CCG' G ( 'YPE) 13	AGO ICG 2 30	CGG GCC P		ACA TGT N	ACI TGA N	Y	13 3CA	50	ACG TGC T		36 T
	GCI CGA G	TCTI AGA F		SACO		ACA TGT D 00 TCC	TCG		TGG ACC V 310 CTC GAG	AGT TCA E 1 FCA TAT	GGK CCX W C1 AGK	GAG/ E - ; 13; CAA( GTT(	AGCA ICGT S A AI I>C I 20 3CTC 2GAG	ATG TAC N LLOT	GGC CCG G ( YPE 13 GTG CAC	AGO ICG 2 30 GAC		AGA TCI E I AGO	ACA TGI N .340 AGC	ACT TGA N STGC	ACI ATGI Y	I3 CON	50 50	ACG TGC T	I CCTV CGA4 P	36 TT 3A
	GCI CGA G CTC GAC	TCTI AGA F	ATCO PAGO Y 1 PAGO Y 1 PAGO AGGO	CAG SGTC SGTC SACC	13 60 6	ACA TGT D 00 TCC ACG S	TCG AGC I TTTC AAG	CCG GGC A 1 TTC HAAG F	ACC V 310 CTC GAG	AGT TCA E 1 FCA TAT ATA Y	GGK CCC W C1 AGK TCC	GAGA CTCT E - i 13: CAAC GTTC K	AGCA ICGI S A AI I>C I 20 3CTC CGAC	LLOI CACC T	GGC GGC GGC V PYPE 13 GTG CAC V	AGO TCG 2 30 GAC CTG D	CGG GCC P AAG TTC K		ACA TGI N 	ACT TGA N STGC ACC W	ACA Y SCA CGIV	IAG MTC' K 13 GCA CGT Q	ACC TGG T 50 GGG CCC	ACG TGC T		G P 4
	GCT CGA G CTC GAC L	TCTI AGA F 129 GAC CTG D	ATCO FAGO Y I D TCCO S	CCAG SGTC SGTC SGTC SGTC CTGC D	13 GC GC GC GC GC GC GC GC GC GC GC GC GC	ACA TGT D 00 TCC AGG S	TCG AGC I TTCC AAC F	CCCG GGCC A 1 TTTC GAAG F	TGG ACC V 310 CTC GAG L	AGT TCA E 1 FCA TAT ATA Y	GGK CCC W C1 AGK TCC S	GAGJ E - J 133 CAAC GTTC K	AGCA ICGI S A AI I>C I 20 3CTC CGAG L	LOT CACC T	GGCC CCG ^G G ( VYPE 13 GTG CAC V	AGO ICG 2 30 GAC CTG D	CGG GCC P AAC TTC K	AGA TCI E I AGC TCC S	ACA TGT N 	ACT TGA N STGC CACC W	XCI ATG7 Υ SCA4 CGT4 Q	13 GCA CGT Q	ACC TGG T 50 GGG CCC G		I GGA P 1 CGT VGCA	G P 4
	GCI CGA G CTC GAC L	TCTI AGA' F 129 GAC' CTGI D	ATCO FAGO Y 1 D TCCO AGGO S	CCAG SGTC SGTC SACC D		ACA TGT D 00 TCC AGG S	TCG AGC I TTTC AAG F	CCCG GGC A 1 TTCC AAG F	TGG ACC V 310 CTC GAG L	AGT TCA E 1 FCA TAT ATA Y FCA	GGK CCC W C1 AGK TCC S	GAGJ E - i 133 CAAG GTTC K	AGCA ICGI S A AI I>C I 20 SCTC CGAC L A AI	AATG TAC N LLOI CACC 3TGG T LLOI	GGCC CCG ^o G ( 'YPE 13 GTG CAC V V	AGC ICG 2 30 GAC CTG D	CGG GCC P AAAC TTC K	AGA TCI E I AGC TCC S	ACA TGT N 340 AGG TCC R	ACT TGA N STGC W	ACI ATG7 Y SCA4 CGT4 Q	AG MTC' K 13 GCA CGT Q	ACC TGG T 50 GGG CCC C	ACG TGC T GGAA CTTI	CCTC GGAC P 1: CCTC CCTC CCTC CCTC CCTC CCTC CCTC C	G G P 4 36 T GA
	GCI CGA G CTC GAC L	TCTJ AGA F 1299 GAC CTG D	ATCO PAGO Y 1 D D TCCO AGGO S	CCAG SGTC SGTC SGTC SGC SGC SGC D		ACA TGT D 00 TCC AGG S	TCG AGC I TTCC AAC F	CCCG GGC A 1 TTTC AAG F	TGG ACC V 310 CTC GAG L	AGT TCA E FCΔ FCΔ TAT ATA Y FCΔ		GAGJ E - i 132 CAA4 GTT( K	AGCA S A AI F>C 20 3CTC CGAC L A AI	AATG TAC N LLOI LLOI T LLOI	GGCC GGC GGC VYPE 13: GGG CAC V V	AGC ICG 2 30 GAC CTG D	CGG GCC P AAC TTC K	AGA TCI E I I I I I I I I I I I I I I I I I I	ACA TGI N 340 AGG TCC R	ACT TGA N ) GTGC CACC W	Y Y SCA GCA	AG MTC' K 13 GCA CGT Q	ACC TGG T 50 GGG CCC C	ACG TGC T GGAA CTTI ; N 	CCTC GGAC P 11 CCGTC VGCAC VGCAC	G P 4 36 TA SA
	GCI CGA G CTC GAC L	TCTJ AGA F 129 GAC CTG D	ATCC PAGK Y I D D TCCC AGGC S	CCAG SGTC SGTC SACC D		ACA TGT D 00 TCC AGG S	TCG AGC I TTCC AAC F	1 CCCG GGC A 1 TTTC SAAG F	TGG ACC V 310 CTC GAG L	AGTU TCA/ E 1 FCΔ TAT ATA Y FCΔ		GAGA E - i 132 CAAC GTTC K - i	AGCA ICGI S A AL I>C I 20 GGTC CGAC L A AJ	AATG TAC N LLOT TGG T	GGCC CCG G ( TYPE 13 GTG CAC V	AGC ICG 2 30 GAC CTG D	CGG GCC P AAC TTC K	AGA TCT E 1 AGC S	ACA TGI N 340 AGG TCC R	ACT TGA N O GTGC CACC W	Y SCA	AG MTC' K 13 GCA CGT Q		ACG TGC T GGAA CTTI N S	In the second se	G P 4 36 TT 3A S
	GCI GGA CTG GAC L	TCTI AGA F 1290 GAC CTGI D	ATCC PAGC Y 1 D D TCCC S S	CCAG SGTC P S SACC D	13 GC G G 13	ACA TGT D 00 TCC AGG S 80	TCG AGC I TTC AAC F	CCCG GGC A 1 TTTC AAG F	TGG ACC V 310 CTC GAG L 3390	AGT TCA E FCA TAT ATA Y FCA		GAGA E - i 132 CAAA GTTV K - i 144	AGCA ICGI S A AI I>C I 20 3CTC CGAC L A AI 00	AATG TTAC N LLOT CACC STGG T LLOT	GGCC GGC 13 GGTG GGTG GCAC V YYPE 14	AGC ICG 2 30 GAC CTG D	CGG GCC P AAC TTC K	AGA TCT E I AGO TCC S	ACA TGI N 340 AGG TCO R 	ACT TGA N STGC CACC W	Y Y SCA CGIV Q	AG TC' K 13 3CA CGT Q 14	ACC TGG T 50 GGG CCC GGG CCC GGG 30	ACG TGC T GGAA CTI CTI	I I I I I I I I I I I I I I I I I I I	36 27 3A 
	GCI CGA G CTC GAC L	TCTJ AGA F 1299 GAC CTG D 1377	ATCO PAGO Y 1 D D D CCC T C CCT C	CCAG SGTO SGTO SACC D D SACC	13 60 6 13 60 6 13 60 6	ACA TGT D 00 TCC AGG S 80 TGA	TCG AGC I TTCC AGC F	1 TTC A F 1 TTC TTC	TGG ACC V 310 CTC GAG L 390 GCA	AGT TCA E 1 FCA TAT ATA Y FCA		GAGA CTCT E - i 133 CAAC GTTC K - i 140 ACT	AGCA AGCA ICGI S A AI I>C I 20 3CTC CGAC L A AI 000 ACAC	NATG TAC N LLOT CACC T T LLOT	GGCC GGCG GGCG TYPE 13 GGTG CACC V TYPE 14 GAA	AGC ICG 2 30 GAC CTG D 10 GAG			ACA TGT N .340 AGC TCC R .420	ACT TGA N ) GTGC W ) ) CACC W	ACI ATGN Y BCAA CGTC Q	AG PTC' K 13 3CA CGT Q 14	ACC TGG T 50 GGG CCC 0 30 GT/	ACG TGC T KGAA CTI S N >N >N	CCTC GGAC P 1: CCTC VGCAC V V V V V V V V V V V V	36 T 3A
	GCI CGA G CTC GAC L ATC	TCTJ AGAC F GAC CTG D 1370 CTC GAG	ATCC PAGC Y 1 D D CCC S D CCC CCC CCC CCC CCC CCC CCC	CCAG SGTO SGTO SACO D SACO D SACO CTAC		ACA TGT D 00 TCC AGG S 80 TGA	TCG AGC I TTC AAG F	1 TTC A F 1 TTC AAG	TGG ACC V 310 CTC GAG L 390 GCA	AGT TCA E FCA TAT ATA Y FCA CAA		GAGJ CTCT E - J 132 CAA4 GTTV K - J 144 ACTT	AGCA ICGI S A AI I>C I 20 3CTC CGAC L A AI 00 0 ACAC	NATG TAC N LLOT CACC T LLOT	GGCC CCG ^C G ( VYPE 13 GTG CAC V V YYPE 14 GAA	AGC ICG 2 30 GAC CTG D 10 GAG CTC			ACA TGT N .340 AGG TCC R .420 XCC1 GGG2	ACT TGP N STGC CACC W		AG MTC' K 13 GCA CGT Q 14 CGG GCC P	ACC TGG T 50 GGG CCC G GGG GGG GGG GGG GGG GGG GGG	ACG TGC T GGAA CTI S N S N S N S N S N S N S N S N S N S	I: CCTV GGAA P 1: CCTV V CCAA V V IotI I I GAG CCTV *>	36 27 3A 344 36 27 3A 3A 3A 3A 3A 3A 3A 3A 3A 3A 3A 3A 3A
AG IC IC IC IC AC	GCI GAC L ATC	TCTJ AGAC F GAC CTG GAC CTG GAG CTG CGAG	ATCC FAGC Y 1 D TCCC AGGC S S O CCT SCAC	CCAG GGTO SGTO SACO D SACO D SATO CTAC M	ISG ISG ISG ISG ISG ISG ISG ISG ISG ISG	ACA TGT D 00 TCC AGG S 80 TGA XACT I E	TCG AGC I TTCC AGG F	I I I I I I I I I I I I I I I I I I I	TGG ACC V 310 CTC GAG L 390 GCA CGT CGA	AGTU TCA E 1 FCA TAT ATA Y FCA CAA GTT CAA		GAGJ E - i 132 CAA( GTT) K - i 144 ACTI TGA	AGCA ICGI S A AI I>C I 20 3CTC CGAC L A AI 00 ACAC	NATG TAC N LLOT LLOT T LLOT CGCA GCGT	GGCC CCG ^G G ( VYPE 13. GTG CAC V V YYPE 14. GAA YCTT X X X X X X X X X X X X X X X X X X	AGC ICG 2 30 GAC CTG D 10 GAG CTC S		AGA TCT E I AGO S TCO S	ACA TGT N 340 CAGO TCC R .420 CCT GG2 5 I	ACT TGA N STGC CACC W		13 GCA CGT Q 14 CGG GCC P	ACC IGG T 50 GGG CCC G G G T 30 G T 30 G T 30 G	ACG TGC T GGAA CTTI CTTI N N AAAT TTTV K 455	I I I I I I I I I I I I I I	360 360 37 3A 3 44 360 3A 360 3A 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3

CCGC GGCG



Fig.24A.

		1	LO	•		20			30			4	10			50			60
			*			*			*			~~~	*	-	200	*	0	~	*
ATG	GIC	AGC	TAC	TGG	GAC	ACC	GGG	GIC	CIG	CIG	IGC	GCG	CIG	CIC	AGC	ACA	CIG	CPT	CIC
TAC	CAG	TCG	AIG	ACC	CIG	1GG	a		GAC	GAC	ACG	N N	GAC.	GAG T.	g	<u> </u>	T.	T.	1.>
M 1	v	S	¥	w c	D	т	ি দেশ গ	יי ייו פיוי	CNIAT	. STRY	UTEN	ริ	Ъ	15	5	C	1	-	20>
_1							1.1917			1 304	201:20								
		-	70			80			90			10	00		1	10			120
			*			*			*				*			*			*
ACA	GGA	TCT	AGT	TCC	GGA	AGT	GAT	ACC	GGT	AGA	CCT	TTC	GTA	GAG	ATG	TAC	AGT	GAA	ATC
TGT	CCT	AGA	TCA	AGG	CCT	TCA	CTA	TGG	CCA	TCT	GGA	AAG	CAT	CIC	TAC	ATG	TCA	$\mathbf{CTT}$	TAG
т	G	s	s	s	G>														
21_1	FLT1	SIG	INAL	SEQ	26>														
_						S	D	т	G	R	P	F	v	Е	м	Y	S	Е	1>
						_27_			_30	_hFI	л1 1	IG D	MAT	v 2_					_40>
																			100
		13	30		1	140			150			10	60			170			180
			*			*			*	~~~			*	~~~		ŝ	0	2000	- -
ccc	GAA	ATT	ATA	CAC	ATG	ACT	GAA	GGA	AGG	GAG	CIC	GIC	ATT	m	160	~~~~	CNA	ACG ACC	ACT
GGG	CTT	TAA	TAT	GIG	TAC	TGA	CIT	CCT	TCC	CIC	GAG	CAG	TAA	666	ACG	GCC D	CAA 17	190	C>
P	Е	I	I	H	М	т	Е,	G	R	E	ىل ئىرىت	, v	Ŧ	P 55	C	ĸ	v	1	60>
41_			-	_45			I	њцг.	16	DOM				_00_					
		-	20			200			210			2	20		:	230			240
		13	*		-	200 *			*			2	*			*			*
COL	8 8 C	አጥዮ	አርጥ	CTT	аст	ጥጥል	444	AAG	ттт	CCA	CIT	GAC	ACT	TIG	ATC	CCT	GAT	GGA	AAA
CCA	7772	TAC	TCA	CAA	TGA	አልጥ	TTT	TTC	ААА	GGT	GAA	CIG	TGA	AAC	TAG	GGA	CTA	CCT	TTT
P	N	I	T	v	T	L	ĸ	ĸ	F	Р	L	D	т	L	I	Р	D	G	К>
61		-	-	65	-		h	FLT1	IG I	DOMA.	EN 2.			75					80>
		2	50		:	260			270			2	80		:	290			300
			*			*			*				*			*			*
CGC	ATA	ATC	TGG	GAC	AGT	AGA	AAG	GGC	TTC	ATC	ATA	TCA	AAT	GCA	ACG	TAC	AAA	GAA	ATA
GCG	TAT	TAG	ACC	CIG	TCA	TCT	TIC	ccc	AAG	TAG	TAT	AGT	TTA	CGT	TGC	ATG	TIT	CIT	TAT
R	I	I	W	D	S	R	ĸ	G	F	I	I	s	N	A	т	Ŷ	ĸ	Е	1005
81_				85			h	FLT1	IG	DOMA	IN 2			95					_100>
		-	10			220			220			7	40			350			360
		د.	±0			3∠U *			350			5	*			*			*
$\sim$	COM	CITC	ACC	man	CAA	CCA	202	GTC	лат	GGG	CAT	TIG	TAT	AAG	ACA	AAC	TAT	CIC	ACA
~~~~	CNA	CIG	mag	101			TOT	CAG	TTA	â	GTA	AAC	ATA	TIC	TGT	TTG	ATA	GAG	TGT
6	T.	T.	-100 T	c	E	Δ	ጥ	v	N	G	н	L	Y	к	т	N	¥	L	T>
101	5	-	*	105	-	A	ĥ	FLT1	IG	DOMA	IN 2	_	_	_115					_120>
101																			
		3	70			380			390)		4	00			410			420
			*			*			*				*			*			*
CAT	CGA	CAA	ACC	AAT	ACA	ATC	ATA	GAT	GIG	GT	CIG	AGI	, 000	TCI	CAI	GGA	ATI	GAA	CTA
GTA	GCT	GIT	' TGG	TTA	TGI	TAG	TAT	CTA	CAC	CAA	GAC	TCA	GGC	AGA	GTA	CCI	TAT	CTI	GAT
н	R	Q	т	N	т	I	I	D>											
121		hF	LT1	IG D	OMAI	N 2_		_129	_>		-		-	~		~	-		
									v	v	г	S bur	P	s r	H	G T P	T	Б	1405
									120			1E1	wr 1	a n	111 A				

Sheet 40 of 55

Fig.24B.

		43	30		4	140			450			46	50		4	70			480
TCT	GTT	GGA	GAA	AAG	СТТ	GTC	TTA	ААТ	TGT	ACA	GCA	AGA	ACT	gaa	CTA	AAT	GTG	GGG	ATT
AGA	CAA	CCT	CIT	TTC	GAA	CAG	AAT	TTA	ACA	TGT	CGT	TCT	TGA	CTT	GAT	TTA	CAC	ccc	TAA
S	v	G	Е	K	L	v	L L	N	C	T	A TNT 7	, R	Т	E 155	г	N	v	G	1>
141_				_149_					. 16	LAND	114 2	·		155				_	100-
		49	90		5	500			510			52	20		ţ	530			540
~~~			*	~	500	*	man	<b>m</b> m	*	<b>C 3</b> m	~~~	ርጉአጥ	*	አአአ	CUL	* (773)	220	<b>CC</b> A	* CAC
GAC	AAG	AAC	ACC	CTT	ATG	GGA	AGA	AGC	TTC	GTA	GIC	GIA	TTC	TTT	GAA	CAT	TTG	GCT	CIG
D	F	N	W	Е	Y	Ρ	s	s	ĸ	н	Q	Н	K	ĸ	L	v	N	R	D>
161_				_165_			_hFI	_K1 ]	GD	MAD	13_	_		_175_	··-				180>
		55	50		5	560			570			58	30		5	590			600
			*			*			*				*			*		~	*
CTA	AAA	ACC	CAG	TCT	ccc	AGT	GAG	ATG	AAG	AAA	TIT	TIG	AGC	ACC	TTA AAT	ACT	ATA	GAT	GGT CCA
L	ĸ	TGG	0	S	G	S	E	M	ĸ	ĸ	F	L	s	T	L	т	I	D	G>
181				_185_			1	FLK	IG.	DOM	NIN 3	3		_195_					_200>
		61	0			520			630			64	10			50			660
		0.	*		`	*			*				*			*			*
GTA	ACC	CGG	AGT	GAC	CAA	GGA	TTG	TAC	ACC	TGT	GCA	GCA	TCC	AGT	GGG	CTG	ATG	ACC	AAG
CAT	TGG	GCC	TCA	CIG	GTT	CCT	AAC T.	ATG	TGG T	ACA	A	CGT	AGG	1CA S	G	GAC	M	1GG T	K>
201_	T	ĸ	5	205	~	<u> </u>	h	IK1	IG I	X	DN 3_			_215_		_			220>
									~~~			-74			,	710			720
		6	/0 *		1	*			*			1	*			*			*
AAG	AAC	AGC	ACA	TIT	GTC	AGG	GIC	CAT	GAA	aag	GAC	aaa	ACT	CAC	ACA	TGC	CCA	ccc	TGC
TTC	TTG	TCG	TGT	AAA	CAG	TCC	CAG	GIA	CIT	TTC	CIG	TTT	TGA	GIG	TGT	ACG	GGT	GGC	ACG
к 221	N	S	T hFL	r n n	V 3 DOI	R	3	н	Е	231:	•								
				-						-	D	ĸ	т	н	T	С	P	Р	0
											232			_hFC	AC1 2	·			_240>
		7	30			740			750			70	50			770			780
			*			*			*				*	~~~	~~~	*			*
CCA	GCA	CCT	GAA	CIC	CIG	ŝ	GGA	COG	TCA	GIC	TIC	CIC	TIC	000	CCA	AAA	GGG	AAG	CIG
P	A	P	E	L	L	G	G	P	S	V	F	L	F	P	P	ĸ	₽	ĸ	D>
241				_245_				h	CAC:	LA_				_255_					_260>
		7	90		,	800			810			8	20			830			840
		/.	*			*			*				*			*			*
ACC	CIC	ATG	ATC	TCC	CGG	ACC	CCT	GAG	GIC	ACA	TGC	GIG	GTG	GTG	GAC	GIG	AGC	CAC	GAA
TGG	GAG	TAC	TAG	AGG	GCC	TGG	GGA	CIC	CAG	TGT	ACG	CAC	CAC	CAC	D	CAC	106	H	E>
261				_265	~	•	-	h	FCACI	LA.				_275					_280>
									070				00			000			000
		8	\$U *			400 *			8/0			8	*			*			*
GAC	CCT	GAG	GIC	AAG	TTC	AAC	TGG	TAC	GIG	GAC	GGC	GIG	GAG	GIG	CAT	AAT	GCC	AAG	ACA
CIG	GGA	CIC	CAG	TIC	AAG	TIG	ACC	ATG	CAC	CIG	œ	CAC	crc	CAC	GTA	TTA	. CGG	TIC	TGT
D 281	P	E	v	к _285	F	N	w	h	FCAC:				Е	_295		N			_300>

Fig.24C.

		91	L0		9	20			930			94	10		9	950			960
	_		*			*			*				*		_	*		~	*
AAG	CCG	CGG	GAG	GAG	CAG	TAC	AAC	AGC	ACG	TAC	CGT	GIG	GIC	AGC	GIC	CIC	ACC	GIC	CIG
TIC	ecc	GCC	CIC	CIC	GIC	AIG	TIG	100	IGC	AIG	GCA	CAC	CAG	106	CAG	GAG	m	CAG	GAC
K	Р	R	в	305	Q	Y	N	5	T.	r N	R	v	v	215	v	ъ	.1.	v	3205
301				305						. A _				212					_3202
		97	0		a	80			990			100	0		10	10		1	.020
		91	*		2	*			*			100	*			*		-	*
CAC	CAG	GAC	TGG	CTG	ሻል	GGC	AAG	GAG	TAC	AAG	TGC	AAG	GIC	TCC	AAC	ААА	GCC	CIC	CCA
GTG	GIC	CTG	ACC	GAC	TTA	COG	TIC	CIC	ATG	TTC	ACG	TIC	CAG	AGG	TIG	TTT	CGG	GAG	GGT
н	0	D	W	L	N	G	к	E	Y	ĸ	C	K	v	S	N	к	A	L	P>
321	×	-		325				1	FCAC	1 A			-	335_					_340>
		103	30		10	040		1	L0 50			106	50		10	070		:	1080
			*			*			*				×			*			*
GCC	ccc	ATC	GAG	AAA	ACC	ATC	TCC	aaa	GCC	aaa	GGG	CAG	ccc	CGA	gaa	CCA	CAG	GIG	TAC
CGG	GGG	TAG	CTC	TTT	TGG	TAG	AGG	TTT	CGG	TIT	ccc	GIC	GGG	GCT	CTT	GGT	GIC	CAC	ATG
Α	Р	I	Е	к	т	I	S	ĸ	A	к	G	Q	P	R	Е	P	Q	v	Y>
341				_345_				hI	CAC1	. A _				_355_					_360>
																			1140
		109	90		11	100		1	1110			112	20		1.	130			1140
			*		~~~~	*	~~~	-				~~~	~	200	ano	2000	m	0	cmc
ACC	CIG	ccc	CCA	TCC	CGG	GAT	GAG	CIG	ACC	AAG	AAC	CAG	GIC	AGC	CIG	ACC	ACC	CIG	CAC
TGG	GAC	GGG	GGL	AGG	GCC	CIA	cic	GAC	TGG	TIC	TIG	GIC	17	105	GAC	166	ACG C	GAC T.	UAG VA
T	L	Р	P	365	ĸ	Ъ	Е	 	-T 201/01	~	м	Ŷ	v	375	ц	1	C	5	380>
201				_305_						. ~ _				212					_0002
		111	50		13	160			1170			118	30		1	1 9 0		:	1200
		119	50 *		13	L60 *		:	1170 *			118	30 *		1	190 *		:	1200 *
ала	œc	119 TTC	50 * TAT	200	13 AGC	L60 * GAC	ATC	GCC	1170 * GIG	GAG	TGG	118 GAG	80 * AGC	ААТ	1: GGG	190 * CAG	COG	GAG	1200 * AAC
AAA TTT	222 222	119 TTC AAG	50 * TAT ATA	COC 636	13 AGC TCG	L60 * GAC CTG	ATC TAG	: ccc ccc	L170 * GIG CAC	GAG CTC	TGG ACC	118 GAG CTC	30 * AGC TCG	AAT TTA	1: GGG CCC	190 * CAG GTC	COG GGC	GAG CTC	1200 * AAC TIG
AAA TTT K	930 930 930 9	119 TTC AAG F	50 * TAT ATA Y	COC GOG P	13 AGC TCG S	L60 * GAC CTG D	ATC TAG I	CCC CCC CCC A	GIG CAC V	GAG CTC E	TGG ACC W	118 GAG CTC E	30 * AGC TCG S	AAT TTA N	1: GGG CCC G	190 * CAG GTC Q	CCG GGC P	GAG CTC E	1200 * AAC TIG N>
AAA TTT K 381_	220 220 9 9	119 TTC AAG F	50 * TAT ATA Y	COC GGG P _385_	13 AGC TCG S	GAC CTG D	ATC TAG I	ссс ссс А ы	GIG CAC V TCAC	GAG CTC E A	TGG ACC W	118 GAG CTC E	30 * AGC TCG S	AAT TTA N _395_	1 GGG CCC G	190 * CAG GTC Q	CCG GGC P	GAG CTC E	1200 * AAC TTG N> _400>
AAA TTT K 381_	339 200 9 9	119 TTC AAG F	50 * TAT ATA Y	CCC GGG P _385_	11 AGC TCG S	LEO * GAC CTG D	ATC TAG I	GCC CGG A h1	GIG CAC V FCAC	GAG CTC E A	TGG ACC W	118 GAG CTC E	AGC TCG S	AAT TTA N _395_	1: GGG CCC G	190 * CAG GTC Q	CCG GGC P	GAG CTC E	1200 * AAC TIG N> _400>
AAA TTT K 381_	329 320 9	115 TTC AAG F 12:	50 * TAT ATA Y	COC GGG P _385_	1: AGC TCG S	L60 * CTG D	ATC TAG I	GCC CGG A hi	GIG CAC V FCAC1	GAG CIC E A	TGG ACC W	118 GAG CTC E 124	AGC TCG S	аат тта N _395_	1: GGG CCC G	190 * GTC Q 250	COG GGC P	GAG CIC E	1200 * AAC TTG N> _400>
AAA TTT K 381_	2220 2000 9	115 TTC AAG F 12:	50 * TAT ATA Y	CCC GGG P _385_	1: AGC TCG S 1:	LEO CTG D 220	ATC TAG I	GCC CGG A hi	GIG CAC V FCAC1	GAG CIC E A	TGG ACC W	118 GAG CTC E 124	AGC TCG S	AAT TTA N _395_		190 * GTC Q 250	CCG GGC P	GAG CTC E	1200 * AAC TTG N> _400> 1260 *
AAA TTT K 381_	GCG G TAC	119 TTC AAG F 123 AAG	50 * TAT ATA Y	COC GGG P _385_ ACG	1: AGC TCG S 1: CCT	CTG D 220	ATC TAG I	CTG	GIG CAC V TCACI	GAG CTC E A TCC	TGG ACC W	118 GAG CTC E 124 GGC	AGC TCG S 10 TCC	AAT TTA N _395_ TTC	1: GGG CCC G 1: TTC	190 * GTC Q 250 * CTC	COG GGC P TAC	GAG CIC E AGC	1200 * AAC TTG N> _400> 1260 *
AAA TTT K 381 AAC TTG	GGC G TAC ATG	119 TTC AAG F 12: AAG TTC	50 * TAT ATA Y 10 * ACC TGG	COC GGG P _385_ ACG TGC	12 AGC TCG S 12 CCT GGA	CTG D 220	ATC TAG I GTG CAC	GCC CCG A hi CTG GAC	GIG CAC V FCACI 1230 CIG CIG	GAG CIC E A TCC AGG	TGG ACC W GAC CTG	118 GAG CTC E 124 GGC CCG	AGC TCG S 10 * TCC AGG	AAT TTA N 395_ TTC AAG	1: GGG G I: TTC AAG	190 * CAG GTC Q 250 * CTC GAG	COG GGC P TAC ATG	GAG CTC E AGC TCG	1200 * AAC TTG N> _400> 1260 * AAG TTC K>
AAA TTT K 381 AAC TTG N	GGC CCG G TAC ATG Y	119 TTC AAG F 122 AAG TTC K	50 * TAT ATA Y LO * ACC TGG T	CCC GGG P _385_ ACG TGC T	1 AGC TCG S 1 CCT GGA P	LEO CTG D 220 CCC GOG P	ATC TAG I GTG CAC V	GCC CCGG A hi CTCG GAC L	GIG CAC V FCACI 1230 GAC CIG D	GAG CIC E A TCC AGG S	TGG ACC W GAC CTG D	118 GAG CTC E 124 GGC G G	AGC TCG S 10 * TCC AGG S	AAT TTA N _395_ TTC AAG F 415	1: GGG CCC G 1: TTC AAG F	190 CAG GTC Q 250 * CTC GAG L	CCG GGC P TAC ATG Y	GAG CTC E AGC TCG S	1200 * AAC TTG N> _400> 1260 * AAG TTC K> 420>
AAA TTT K 381 AAC TIG N 401	GGC CCG G TAC ATG Y	119 TTC AAG F 12: AAG TTC K	50 * TAT ATA Y LO * ACC TGG T	COC GGG P _385_ ACG TGC T _405_	1 AGC TCG S 1 CCT GGA P	LEO * CTG D 220 * CCC GGG P	ATC TAG I GTG CAC V	GCC CGG A hi CTG GAC L hi	GIG CAC V FCACI 1230 CAC CIG D FCACI	GAG CIC E A TCC AGG S L A	TGG ACC W GAC CTG D	118 GAG CTC E 124 GGC G	AGC TCG S 10 * TCC AGG S	AAT TTA N 395_ TTC AAG F 415_	1: GGG CCC G 1: TTC AAG F	190 * GTC Q 250 * CTC GAG L	COG GGC P TAC ATG Y	GAG CIC E AGC TCG S	1200 * AAC TTG N> _400> 1260 * AAG TTC K> _420>
AAA TTT K 381 AAC TTG N 401	GGC COG G TAC ATG Y	115 TTC AAG F 12: AAG TTC K	50 * TAT ATA Y LO * ACC TGG T	COC GGG P _385_ ACG TGC T GC T _405_	11 AGC TCG S 11 CCT GGA P	L60 * CTG D 220 * CCC GGG P 280	ATC TAG I GIG CAC V	GCC CGG A hi CTG GAC L h	1170 * CAC V FCAC1 1230 * GAC CIG D FCAC1	GAG CTC E A TCC AGG S L A	TGG ACC W GAC CTG D	118 GAG CTC E 124 GGC G G 130	AGC TCG S 10 tCC AGG S	AAT TTA N 395_ TTC AAG F _415_	1 GGG CCC G 1 TTC AAG F	190 * GTC Q 250 * CTC GAG L 310	COG GGC P TAC ATG Y	GAG CTC E AGC TCG S	1200 * AAC TTG N> _400> 1260 * AAG TTC K> _420> 1320
AAA TTT K 381_ AAC TTG N 401_	GGC CCG G TAC ATG Y	115 TTC AAG F 12 AAG TTC K	50 * TAT ATA Y 10 * ACC TGG T 70 *	COC GGG P _385_ ACG TGC T _405_	1: AGC TCG S 1: CCT GGA P 1:	L60 * GAC CTG D 220 * COCC GOG P 280 *	ATC TAG I GTG CAC V	GCC CGG A hi CTG GAC L hi	1170 * GIG CAC V FCACI 1230 * GAC CIG D FCACI 1290 *	GAG CTC E A TCC AGG S LA	TGG ACC W GAC CTG D	118 GAG CTC E 124 GGC G G 130	AGC TCG S 40 * TCC AGG S 00 *	AAT TTA 395_ TTC AAG F _415_	1 GGG G 1 TTC AAG F	190 * CAG GTC Q 250 * CTC GAG L 310 *	CCG GGC P TAC ATG Y	GAG CTC E AGC TCG S	1200 * AAC TTG N> _400> 1260 * AAG TTC K> _420> 1320 *
AAA TTTT K 381_ AAC TTG N 401_ CTC	GGC CCG G TAC ATG Y	111 TTC AAG F 12: AAG TTC K 12: GTG	50 * TAT Y 10 * ACC TGG T 70 * GAC	CCC GGG P _385_ ACG TGC T _405_ AAG	1: AGC TOG S 1: CCT GGA P 1: AGC	L60 * GAC CTG D 220 * CCC GGG P 280 * AGG	ATC TAG I GTG CAC V	CTG CGG A hi CTG GAC L hi CAG	IITO * GIG CAC V FCACI IIIIO * GAC CIG D FCACI IIIIO * CAG	GAG CTC E A TCC S AGG S L A GGG	TGG ACC W GAC CTG D	118 GAG CTC E 124 GGC G G G G G G CCG G	AGC TCG S 40 * TCC AGG S 00 * TTC	AAT TTA 395_ TTC AAG F _415_ TCA	1 GGG G 1 TTC AAG F 1 TTC	190 * CAG GTC Q 250 * CTC GAG L 310 *	CCG GGC P TAC ATG Y	GAG CTC E AGC TCG S	1200 * AAC TTG N> _400> 1260 * AAG TTC K> _420> 1320 * ; CAT
AAA TTTT K 381_ AAC TTG N 401_ CTC GAG	GGC G TAC ATG Y ACC	111 TTC AAG F 12: AAG TTC K 12' GTG CAC	50 * TAT ATA Y 10 * ACC TGG T 70 * GAC CTG	CCC GGG P _385_ ACG TGC T _405_ AAG TTC	1: AGC TOG S 1: CCT GGA P 1: AGC TOG	CTG D 220 * CCC GGG P 280 * AGG TCC	ATC TAG I GTG CAC V TGG ACC	GCC CGG A bi CTG GAC L bi CAG GTC	IITO * GIG CAC V FCACI 1230 * GAC CIG D FCACI 1290 * CAG GIC	GAG CTC E A TCC AGG S L A GGG COC	TGG ACC W GAC CTG D AAC TTG	111 GAG CTC E 124 GGC CCG G G CCG G G CCG G CAG	AGC TCG S 40 * TCC AGG S 00 * TTCC AAG	AAT TTA N 395_ TTC AAG F _415_ TCA	1: GGG CCC G TTC AAG F 1 TTC AAG	190 * CAG GTC Q 250 * CTC GAG L 310 * TCC AGG	COG GGC P TAC ATG Y GTG CAC	GAG CTC E AGC TCG S ATC TAC	1200 * AAC TTG N> 400> 1260 * AAG TTC K> 420> 1320 * CAT : GTA
AAA TTTT K 381 AAC TTG N 401 CTC GAG L	GGC G TAC ATG Y ACC TGG T	111 TTC AAG F 12: AAG TTC K 12: GTG CAC V	50 * TAT ATA Y 10 * ACC TGG T GGC CTG GAC CTG D	CCCC GGG P_385 TGC T 405 AAG TTCC K	11 AGC TCG S 12 CCT GGA P 11 AGC TCG S	L60 * GAC CTG D 220 * CCC GGG P 280 * AGG TCC R	ATC TAG I GTG CAC V TGG ACC W	GCC CCG A hu CTG GAC L hu CAG GTC Q	IITO GIG CAC V TCACI I230 * GAC CIG D FCACI I290 * CAG GIC Q	GAG CTC E A GGG GGG GGG G	TGG ACC W GAC CTG D AAC TTG N	111 GAG CTC E 124 GGC G G GTC CAG V	AGC TCG S 40 * TCC AGG S 00 * TTCC AAG F	AAT TTA N 395_ TTC AAG F 415_ TCA AGT S	1: GGG CCC G TTCC AAG F 1 TGC ACG C	190 CAG GTC Q 250 * CTC GAG L 310 * TCC AGG S	CCG GGC P TAC ATG Y GTG CAC V	GAG CTC E AGC TCG S ATC TAC M	1200 * AAC TTG N> 400> 1260 * AAG TTC K> 420> 1320 * CAT CAT H>
AAA TTTT K 381 AAC TTG N 401 CTC GAG L 421	GGC CCG G TAC ATG Y ACC TGG T	111 TTC AAG F 12: AAG TTC K 12: CAC V	50 * TAT ATA Y LO * ACC TGG T GGC CTG D	CCC GGG P _385_ TGC T _405_ AAG TTC K _425_	11 AGC TCG S 11 CCT GGA P 11 AGC TCG S	L60 * GAC CTG D 220 * CCC GOG P 280 * AGG TCC R	ATC TAG I GTG CAC V TGG ACC W	GCC CGG A hi CTG GAC L hi CAG GTC Q hi	1170 * GIG CAC V FCACI 1230 * GAC CIG D FCACI 1290 * CAG GIC Q FCACI	GAG CTC E A A GGG G COC G L A	TGG ACC W GAC CTG D AAC TTG N	1118 GAG CTC E 124 GGC CCG G G T130 GTC CAG V	AGC TOG S 40 * TOC AGG S TTCC AAGG F	AAT TTA 395_ TTC AAG F 415_ TCA AGT S _435_	1: GGG G I: TTCC AAG C	190 CAG GTC Q 250 * CTC GAG L 310 * TCC AGG S	CCG GGC P TAC ATG Y GTG CAC V	GAG CTC E AGC TCG S ATC TAC M	1200 * AAC TTG N> 400> 1260 * AAG TTC K> 420> 1320 * CAT CAT H> 440>
AAA TTTT K 381 AAC TTG N 401 CTC GAG L 421	GGC G TAC ATG Y ACC TGG T	111 TTC AAG F 12: AAG TTC K 12: GTG CAC V	50 * TAT ATA Y 10 * ACC TGG T 70 * GAC CIG D	CCC GGG P _385_ TGC T _405_ AAG TTC K _425_	11 AGC TCG S 11 CCT GGA P 11 AGC TCG S	LEO * GAC CTG D 220 * CCC GOG P 280 * AGG TCC R	ATC TAG I GTG CAC V TGG ACC W	CTG GAC L hi CAG GAC CAG GTC Q	1170 * GIG CAC V FCACI 1230 * GAC CIG D FCACI 1290 * CAG GIC Q FCACI	GAG CTC E AGG S AGG S A COC G L A	TGG ACC W GAC CTG D AAC TTG N	111 GAG CTC E 124 GGC CCG G T 130 GTC CAG V	AGC TCG S 40 * TCC AGG S TTCC AAGG F	AAT TTA 395_ TTC AAG F 415_ TCA AGT S _435_	1: GGG G I TTC AAG F I TGC ACG C	190 CAG GTC Q 250 * CTC GAG L 310 * TCC S	CCG GGC P TAC ATG Y GTG CAC V	GAG CTC E AGC TCG S ATC TAC M	1200 * AAC TTG N> 400> 1260 * AAG TTC K> 420> 1320 * CAT H> 440>
AAA TTTT K 381_ AAC TTG N 401_ CTC GAG L 421_	GGC G TAC ATG Y ACC TGG T	111 TTC AAG F 12: AAG TTC K 12' GTG CAC V 13:	50 * TAT ATA Y 10 * ACC TGG T GAC CTG D 30	CCC GGG P _385_ TGC T _405_ AAG TTC K _425_	11 AGC TCG S 12 CCT GGA P 11 AGC TCG S	L60 * GAC CTG D * 220 * COC GOG P 280 * AGG TCC R 340	ATC TAG I GTG CAC V TGG ACC W	CTG CAG CTG CTG CAG CAG CAG CAG CAG	1170 * GIG CAC V FCAC1 1230 * GAC CIG D FCAC1 1290 * CAG GIC Q FCAC1	GAG CTC E A A GGG G COC G L A	TGG ACC W GAC CTG D AAC TTG N	1118 GAG CTC E 124 GGC CCG G T 130 GTC CAG V 133	AGC TCG S 40 * TCC AGG S TCC AAG F 60	AAT TTA 395_ TTC AAG F 415_ TCA AGT S _435_	1: GGG G I: TTC AAG F I TGC ACG C	190 * CAG GTC Q 250 * CTC GAG L 310 * TCC S 370	CCG GGC P TAC ATG Y GTG CAC V	GAG CTC E AGC TCG S ATC TAC M	1200 * AAC TTG N> 400> 1260 * AAG TTC K> 420> 1320 * GTA H> 440>
AAA TTTT K 381_ AAC TTG AAC TTG GAG L 421_	GCC G TAC ATG Y ACC TGG T	111 TTC AAG F 12: AAG TTC K 12: GTG CAC V 13:	50 * TAT ATA Y 10 * ACC TGG T GAC CIG D 30 *	CCCC GGG P 385_ TGC T 405_ AAG TTC K 425_	1: AGC TCG S 1: CCT GGA P 1: AGC TCG S 1: CCT	L60 GAC CTG D 220 * CCC GOG P 280 * AGG TCC R 340 *	ATC TAG I GTG CAC V TGG ACC W	GCC CGG A bi CTG GAC CAG GTC Q CAG GTC Q h	1170 * GIG CAC V FCACI 1230 * GAC CIG D FCACI 1290 * CAG GIC Q FCACI 1350 *	GAG CTC E A GGG G G G G G G COC G L A	TGG ACC W GAC CTG D AAC TTG N	1118 GAG CTC E 124 GGC G G GTC CAG V 130 GTC CAG V 131	AGC S TCCG S TCCG S TCCC AAGG S TTCC AAGG F 60	AAT TTA N 3955 TTC AAG F 4155 TCA AGT S 4355	1: GGG G II TTC AAG C I I	190 * CAG GTC Q 250 * CTC GAG L 310 * TCC AGG S 370 *	CCG GGC P TAC ATG Y GTG CAC V	GAG CTC E AGC TCG S ATCG TAC M	1200 * AAC TTG N> _400> 1260 * AAG TTC K> _420> 1320 * CAT : GTA H> _440>
AAA TTT K 381 AAC TTG N 401 CTC GAG L 421 GAG	GGC G TAC TAC TGG T GCT	111 TTC AAG F 12: AAG TTC K 12: CTG CAC V 13: CTG	50 * TAT ATA Y 10 * ACC TGG T GAC CIG D 30 *	CCCC GGG P 385 TGC T 405 AAG TTC K 425 AAC	1: AGC TCG S 1: CCT GGA P 1: AGC TCG S 1: CAC	LGO GAC CTG D 220 * CCC GOG P 280 * AGG TCC R 340 *	ATC TAG I GTG CAC V TGG ACC W	CIG CAG CIG CAG CAG CAG CAG	1170 * GIG CAC V FCACI 1230 * GAC CIG D FCACI 1290 * CAG GIC Q FCACI 1350 *	GAG CTC E A TCC S A GG G G G C CC G L A A GC G G C C C C C C C C C C C C C C	TGG ACC W GAC CTG D AAC TTG N	1118 GAG CTC E 124 GCC G G CCG G G CCG G G CCG G CCG G TCC CAG V 130 CTC	AGC S TCCG S TCCG S TCCG S TCCC AAGG S TTCC AAGG F 60 * CCIG CCIG CCIG	AAT TTA N 395_ TTC AAG F 415_ TCA AGT S _435_ TCT	1: GGG G TTC AAG F 1: TTC ACG C 1 C C 1 C C C C C C C C C C C C C	190 * CAG GTC Q 250 * CTC GAG L 310 * TCC AGG S 370 *	CCG GGC P TAC ATG Y GTG CAC V	GAG CTC E AGC TCG S ATCG TAC M	1200 * AAC TTG N> _400> 1260 * AAG TTC _420> 1320 * CAT : CAT : GTA H> _440>
AAA TTT K 381 AAC TTG N 401 CTC GAG L 421 GAG CTC F	GGC G TAC ATG Y ACC TGG T GCT CGA	111 TTC AAG F 12 AAG TTC K 12 CAC V 13 CTG GAC	50 * TAT ATA Y LO * ACC TGG TGG TGG CAC CIG D 30 * CAC GIG	CCCC GGG P 385 TGC T 405 AAG TTCC K 425 AACC	1: AGC TCG S 1: CCT GGA P 1: CCT GGA TCG S 1: CAC CAC GTG H	LGO GAC CTG D 220 * CCC GOG P 280 * AGG TCC R 340 * TAC ATG *	ATC TAG I GTG CAC V TGG ACC W ACG TGC TGC	CIG CAG CAG CIG CAG CIC CAG CIC CAG CIC CAG CIC CAG CIC CAG CIC CIC	1170 GIG CAC V FCACI 1230 CAC CIG D FCACI 1290 CAG GIC Q FCACI 1350 * AAG TTC K	GAG CTC E A GGG G COC G COC G COC G COC G COC C C C	TGG ACC W GAC CTG D AAC TTG N CTC GAG	111 GAG CTC E 12/ GCC G G CCG G G TCC CAG V 13/ TCC AGG S	AGC S TCCG S TCCG S TCCC AAGG S TTCC AAGG S TTCC AAGG S CTG GAC T	AAT TTA N 395_ TTC AAG F 415_ TCA AGT S 435_ TCT AGA S	1: GGG G TTC AAG F 1 TGC ACG C 1 CCCG GGC P	190 * CAG GTC Q 250 * CTC GAG L 310 * TCC AGG S 370 * GGT GGT GAG CTC GAG CTC CAG CAG CTC CAG CAG CTC CAG CAG CTC CAG CAG CTC CAG CAG CTC CAG CAG CTC CAG CTC CAG CAG CTC CAG CTC CAG CTC CAG CTC CAG CTC CAG CTC CAG CTC CAG CTC CAG CTC CAG CTC CAG CTC CAG CTC CAG CTC CAG CTC CAG CTC CAG CTC CAG CTC CAG CTC CTC CAG CTC CTC CAG CTC CTC CAG CTC CTC CTC CAG CTC CTC CTC CTC CTC CTC CTC CT	CCCG GGC P TAC ATG Y GTCG CAC V	GAG CTC E AGC TCG S ATCG TAC M TCA M	1200 * AAC TTG N> 400> 1260 * AAG TTC 420> 1320 * GTA H> 440>

Fig.25A. DME - Challenge + Fit1D2VEGFR3D3.FcΔC1(a) + Fit1D2Fik1D3.FcΔC1(a) + R→C + NAS + ΔB2 + A40 + Fit1 (1-3) Fc VEGF 165

Fig.25B.







Jul. 4, 2006





1D2Flk1D3.FcΔC1(a) & VEGFR1R2-FcΔC1(a)	VEGF/VEGFR1R2-FcdC1(a)	0.98	0.94	0.99	0.97 ± 0.02	
toichiometry of hVEGF165 to Flt	VEGF/FI11D2FIk1D3.FcAC1(a)	0.93	0.97	Ļ	0.96 ± 0.03	
Binding St	hVEGF165 (nM)	-	10	50	Average ± StDev	

Fig.28.

Fig.29.









Fig.35.



Fig.36. 50
GRPFVEMYSEIPEIIHMTEGRELVIPCRVTSP <u>NITVTLKKFPLDTLIPDG</u>
KRIIWDSRKGFIIS <u>N</u> ATYKEIGLLTCEATVNGHLYKTNYLTHRQTNTIID
VVLSPSHGIELSVGEKLVL <u>NC</u> TARTELNVGIDFNWEYPSSKHQHKKLVNR
DLKTQSGSEMKKFLSTLTIDGVTRSDQGLYTCAASSGLMTKKNSTFVRVH
EKGPGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVD
VSHEDPEVKFNWYVDGVEVHNAKTKPREEQY <u>N</u> STYRVVSVLTVLHQDWLN
GKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSL
TCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKS
RWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

Fig.37.









MODIFIED CHIMERIC POLYPEPTIDES WITH IMPROVED PHARMACOKINETIC PROPERTIES

The application claims priority of U.S. Provisional Appli-5 cation No. 60/138,133, filed on Jun. 8, 1999. Throughout this application various publications are referenced. The disclosures of these publications in their entireties are hereby incorporated by reference into this application.

INTRODUCTION

The field of this invention is modified polypeptides with improved pharmacokinetics. Specifically, the field of this invention relates to Flt1 receptor polypeptides that have 15 been modified in such a way as to improve their pharmacokinetic profile. The field of this invention also relates to methods of making and using the modified polypeptides including but not limited to using the modified polypeptides to decrease or inhibit plasma leakage and/or vascular permeability in a mammal.

BACKGROUND

The ability of polypeptide ligands to bind to cells and 25 thereby elicit a phenotypic response such as cell growth, survival, cell product secretion, or differentiation is often mediated through transmembrane receptors on the cells. The extracellular domain of such receptors (i.e. that portion of the receptor that is displayed on the surface of the cell) is 30 generally the most distinctive portion of the molecule, as it provides the protein with its ligand binding characteristic. Binding of a ligand to the extracellular domain generally results in signal transduction which transmits a biological signal to intracellular targets. Often, this signal transduction 35 acts via a catalytic intracellular domain. The particular array of sequence motifs of this catalytic intracellular domain determines its access to potential kinase substrates (Mohammadi, et al.,1990, Mol. Cell. Biol. 11:5068-5078; Fantl, et al., 1992, Cell 69:413-413). Examples of receptors that 40 transduce signals via catalytic intracellular domains include the receptor tyrosine kinases (RTKs) such as the Trk family of receptors which are generally limited to cells of the nervous system, the cytokine family of receptors including the tripartate CNTF receptor complex (Stahl & Yancopou- 45 los, 1994, J. Neurobio. 25:1454-1466) which is also generally limited to the cells of the nervous system, G-protein coupled receptors such as the \beta_2-adrenergic receptor found on, for instance, cardiac muscle cells, and the multimeric IgE high affinity receptor FceRI which is localized, for the 50 most part, on mast cells and basophils (Sutton & Gould, 1993, Nature 366:421-428).

All receptors identified so far appear to undergo dimerization, multimerization, or some related conformational change following ligand binding (Schlessinger, J., 1988, 55 Trend Biochem. Sci. 13:443–447; Ullrich & Schlessinger, 1990, Cell 61:203–212; Schlessinger & Ullrich, 1992, Neuron 9:383–391) and molecular interactions between dimerizing intracellular domains lead to activation of catalytic function. In some instances, such as platelet-derived growth 60 factor (PDGF), the ligand is a dimer that binds two receptor molecules (Hart, et al., 1988, Science, 240:1529–1531; Heldin, 1989, J. Biol. Chem. 264:8905–8912) while, for example, in the case of epidermal growth factor (EGF), the ligand is a monomer (Weber, et al., 1984, J. Biol. Chem. 65 259:14631–14636). In the case of the FccRI receptor, the ligand, IgE, exists bound to FccRI in a monomeric fashion 2

and only becomes activated when antigen binds to the IgE/FceRI complex and cross-links adjacent IgE molecules (Sutton & Gould, 1993, Nature 366:421–428).

Often, the tissue distribution of a particular receptor within higher organisms provides insight into the biological function of the receptor. The RTKs for some growth and differentiation factors, such as fibroblast growth factor (FGF), are widely expressed and therefore appear to play 10 some general role in tissue growth and maintenance. Members of the Trk RTK family (Glass & Yancopoulos, 1993, Trends in Cell Biol. 3:262-268) of receptors are more generally limited to cells of the nervous system, and the Nerve Growth Factor family consisting of nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), neurotrophin-3 (NT-3) and neurotrophin-4/5 (NT-4/5), which bind the Trk RTK family receptors, promote the differentiation of diverse groups of neurons in the brain and periphery (Lindsay, R. M, 1993, in Neurotrophic Factors, S. E. Loughlin & J. H. Fallon, eds., pp. 257-284, San Diego, Calif., Academic Press). Fc∈RI is localized to a very limited number of types of cells such as mast cells and basophils. Mast cells derive from bone marrow pluripotent hematopoietic stem cell lineage, but complete their maturation in the tissue following migration from the blood stream (See Janeway & Travers, 1996, in Immunobiology, 2d. Edition, M. Robertson & E. Lawrence, eds., pp. 1:3-1:4, Current Biology Ltd., London, UK, Publisher) and are involved in the allergic response. Many studies have demonstrated that the extracellular domain of a receptor provides the specific ligand binding characteristic. Furthermore, the cellular environment in which a receptor is expressed may influence the biological response exhibited upon binding of a ligand to the receptor. For example, when a neuronal cell expressing a Trk receptor is exposed to a neurotrophin which binds to that receptor, neuronal survival and differentiation results. When the same receptor is expressed by a fibroblast, exposure to the neurotrophin results in proliferation of the fibroblast (Glass, et al., 1991, Cell 66:405-413).

A class of cell-derived dimeric mitogens with selectivity for vascular endothelial cells has been identified and designated vascular endothelial cell growth factor (VEGF). VEGF has been purified from conditioned growth media of rat glioma cells [Conn et al., (1990), Proc. Natl. Acad. Sci. U.S.A., 87. pp 2628-2632]; and conditioned growth media of bovine pituitary follicle stellate cells [Ferrara and Henzel, (1989), Biochem. Biophys. Res. Comm., 161, pp. 851-858; Gozpadorowicz et al., (1989), Proc. Natl. Acad. Sci. U.S.A., 86, pp. 7311-7315] and conditioned growth medium from human U937 cells [Connolly, D. T. et al. (1989), Science, 246, pp. 1309-1312]. VEGF is a dimer with an apparent molecular mass of about 46 kDa with each subunit having an apparent molecular mass of about 23 kDa. VEGF has some structural similarities to platelet derived growth factor (PDGF), which is a mitogen for connective tissue cells but not mitogenic for vascular endothelial cells from large vessels.

The membrane-bound tyrosine kinase receptor, known as Flt, was shown to be a VEGF receptor [DeVries, C. et al., (1992), Science, 255, pp. 989–991]. The Flt receptor specifically binds VEGF which induces mitogenesis. Another form of the VEGF receptor, designated KDR, is also known to bind VEGF and induce mitogenesis. The partial cDNA sequence and nearly full length protein sequence of KDR is known as well [Terman, B. I. et al., (1991) Oncogene 6, pp. 1677–1683; Terman, B. I. et al., (1992) Biochem. Biophys. Res. Comm. 187, pp. 1579–1586].

Persistent angiogenesis may cause or exacerbate certain diseases such as psoriasis, rheumatoid arthritis, hemangiomas, angiofibromas, diabetic retinopathy and neovascular glaucoma. An inhibitor of VEGF activity would be useful as a treatment for such diseases and other VEGF-induced pathological angiogenesis and vascular permeability conditions, such as tumor vascularization. The present invention relates to a VEGF inhibitor that is based on the VEGF receptor Flt1.

Plasma leakage, a key component of inflammation, occurs 10 in a distinct subset of microvessels. In particular, in most organs plasma leakage occurs specifically in the venules. Unlike arterioles and capillaries, venules become leaky in response to numerous inflammatory mediators including histamine, bradykinin, and serotonin. One characteristic of inflammation is the plasma leakage that results from intercellular gaps that form in the endothelium of venules. Most experimental models of inflammation indicate that these intercellular gaps occur between the endothelial cells of postcapillary and collecting venules (Baluk, P., et al., Am. J. 20 Pathol. 1998 152:1463-76). It has been shown that certain lectins may be used to reveal features of focal sites of plasma leakage, endothelial gaps, and finger-like processes at endothelial cell borders in inflamed venules (Thurston, G., et al., Am. J. Physiol, 1996, 271: H2547-62). In particular, plant 25 lectins have been used to visualize morphological changes at endothelial cell borders in inflamed venules of, for example, the rat trachea. Lectins, such as conconavalin A and ricin, that bind focally to inflamed venules reveal regions of the subendothelial vessel wall exposed by gaps that correspond to sites of plasma leakage (Thurston, G., et al., Am J Physiol, 1996, 271: H2547-62).

The properties of the microvessels are dynamic. Chronic inflammatory diseases, for example, are associated with microvascular remodeling, including angiogenesis and microvessel enlargement. Microvessels can also remodel by acquiring abnormal phenotypic properties. In a murine model of chronic airway inflammation, airway capillaries acquire properties of venules, including widened vessel diameter, increased immunoreactivity for von Willebrand factor, and increased immunoreactivity for P-selectin. In addition, these remodeled vessels leak in response to inflammatory mediators, whereas vessels in the same position in the airways of normal mice do not.

Certain substances have been shown to decrease or inhibit vascular permeability and/or plasma leakage. For example, mystixins are synthetic polypeptides that have been reported to inhibit plasma leakage without blocking endothelial gap formation (Baluk, P., et al., J. Pharmacol. Exp. Ther., 1998, 284: 693–9). Also, the beta 2-adrenergic receptor agonist formoterol reduces microvascular leakage by inhibiting endothelial gap formation (Baluk, P. and McDonald, D. M., Am. J. Physiol., 1994, 266:L461–8).

The angiopoietins and members of the vascular endothe-55 lial growth factor (VEGF) family are the only growth factors thought to be largely specific for vascular endothelial cells. Targeted gene inactivation studies in mice have shown that VEGF is necessary for the early stages of vascular development and that Ang-1 is required for later stages of vascular 60 remodeling.

U.S. Pat. No. 6,011,003, issued Jan. 4, 2000, in the name of Metris Therapeutics Limited, discloses an altered, soluble form of FLT polypeptide being capable of binding to VEGF and thereby exerting an inhibitory effect thereon, the 65 polypeptide comprising five or fewer complete immunoglobulin domains.

U.S. Pat. No. 5,712,380, issued Jan. 27, 1998 and assigned to Merck & Co., discloses vascular endothelial cell growth factor (VEGF) inhibitors that are naturally occurring or recombinantly engineered soluble forms with or without a C-terminal transmembrane region of the receptor for VEGF.

Also assigned to Merck & Co. is PCT Publication No. WO 98/13071, published Apr. 2, 1998, which discloses gene therapy methodology for inhibition of primary tumor growth and metastasis by gene transfer of a nucleotide sequence encoding a soluble receptor protein which binds to VEGF.

PCT Publication No. WO 97/44453, published Nov. 27, 1997, in the name of Genentech, Inc., discloses novel chimeric VEGF receptor proteins comprising amino acid sequences derived from the vascular endothelial growth factor (VEGF) receptors Flt1 and KDR, including the murine homologue to the human KDR receptor FLK1, wherein said chimeric VEGF receptor proteins bind to VEGF and antagonize the endothelial cell proliferative and angiogenic activity thereof.

PCT Publication No. WO 97/13787, published Apr. 17, 1997, in the name of To a Gosei Co., LTD., discloses a low molecular weight VEGF inhibitor usable in the treatment of diseases accompanied by neovascularization such as solid tumors. A polypeptide containing the first immunoglobulinlike domain and the second immunoglobulin-like domain in the extracellular region of a VEGF receptor FLT but not containing the sixth immunoglobulin-like domain and the seventh immunoglobulin-like domain thereof shows a VEGF inhibitory activity.

Sharifi, J. et al., 1998, *The Quarterly Jour. of Nucl. Med.* 42:242–249, disclose that because monoclonal antibodies (MAbs) are basic, positively charged proteins, and mammalian cells are negatively charged, the electrostatic interactions between the two can create higher levels of background binding resulting in low tumor to normal organ ratios. To overcome this effect, the investigators attempted to improve MAb clearance by using various methods such as secondary agents as well as chemical and charge modifications of the MAb itself.

Jensen-Pippo, et al., 1996, Pharmaceutical Research 13:102–107, disclose that pegylation of a therapeutic protein, recombinant human granulocyte colony stimulating factor (PEG-G-CSF), results in an increase in stability and in retention of in vivo bioactivity when administered by the intraduodenal route.

Tsutsumi, et al., 1997, Thromb Haemost. 77:168–73, disclose experiments wherein the in vivo thrombopoietic activity of polyethylene glycol-modified interleukin-6 (MPEG-IL-6), in which 54% of the 14 lysine amino groups of IL-6 were coupled with PEG, was compared to that of native IL-6.

Yang, et al., 1995, Cancer 76:687–94, disclose that conjugation of polyethylene glycol to recombinant human interleukin-2 (IL-2) results in a compound, polyethylene glycolmodified IL-2 (PEG-IL-2) that retains the in vitro and in vivo activity of IL-2, but exhibits a markedly prolonged circulating half-life.

R. Duncan and F. Spreafico, *Clin. Pharmacokinet.* 27: 290–306, 296 (1994) review efforts to improve the plasma half-life of asparaginase by conjugating polyethylene glycol.

PCT International Publication No. WO 99/03996 published Jan. 28, 1999 in the name of Regeneron Pharmaceuticals, Inc. and The Regents of The University of California describes modified human noggin polypeptides having deletions of regions of basic amino acids. The modified human noggin polypeptides are described as retaining biological

activity while having reduced affinity for heparin and superior pharmacokinetics in animal sera as compared to the unmodified human noggin.

SUMMARY OF THE INVENTION

The present invention is directed to VEGF antagonists with improved pharmacokinetic properties. A preferred embodiment is an isolated nucleic acid molecule encoding a fusion polypeptide capable of binding a VEGF polypeptide 10 comprising (a) a nucleotide sequence encoding a VEGF receptor component operatively linked to (b) a nucleotide sequence encoding a multimerizing component, wherein the VEGF receptor component is the only VEGF receptor component of the fusion polypeptide and wherein the nucle-15 otide sequence of (a) consists essentially of a nucleotide sequence encoding the amino acid sequence of Ig domain 2 of the extracellular domain of a first VEGF receptor and a nucleotide sequence encoding the amino acid sequence of Ig domain 3 of the extracellular domain of a second VEGF 20 receptor.

In a further embodiment, the isolated nucleic acid of the first VEGF receptor is Flt1.

In a further embodiment, the isolated nucleic acid of the second VEGF receptor is Flk1.

In yet another embodiment, the isolated nucleic acid of the second VEGF receptor is Flt4.

In another preferred embodiment, the nucleotide sequence encoding Ig domain 2 of the extracellular domain of the first VEGF receptor is upstream of the nucleotide sequence ³⁰ encoding Ig domain 3 of the extracellular domain of the second VEGF receptor.

In still another preferred embodiment, the nucleotide sequence encoding Ig domain 2 of the extracellular domain of the first VEGF receptor is downstream of the nucleotide ³⁵ sequence encoding Ig domain 3 of the extracellular domain of the second VEGF receptor.

In a preferred embodiment of the invention, the multimerizing component comprises an immunoglobulin domain.

In another embodiment, the immunoglobulin domain is ⁴⁰ selected from the group consisting of the Fc domain of IgG, the heavy chain of IgG, and the light chain of IgG.

Preferred embodiments include an isolated nucleic acid molecule comprising a nucleotide sequence encoding a modified Flt1 receptor fusion polypeptide, wherein the coding region of the nucleic acid molecule consists of a nucleotide sequence selected from the group consisting of

- (a) the nucleotide sequence set forth in FIGS. 13A-13D;
- (b) the nucleotide sequence set forth in FIGS. 14A-14C;
- (c) the nucleotide sequence set forth in FIGS. 15A-15C;
- (d) the nucleotide sequence set forth in FIGS. 16A-16D;
- (e) the nucleotide sequence set forth in FIGS. 21A-21C;
- (f) the nucleotide sequence set forth in FIGS. 22A-22C;
- (g) the nucleotide sequence set forth in FIGS. 24A-24C; and 55
- (h) a nucleotide sequence which, as a result of the degeneracy of the genetic code, differs from the nucleotide sequence of (a), (b), (c), (d), (e), (f), or (g) and which encodes a fusion polypeptide molecule having the biological activity of the modified Flt1 receptor fusion 60 polypeptide.

In a further embodiment of the invention, a fusion polypeptide is encoded by the isolated nucleic acid molecules described above.

A preferred embodiment is a composition capable of 65 binding a VEGF molecule to form a nonfunctional complex comprising a multimer of the fusion polypeptide.

Also preferred is a composition wherein the multimer is a dimer.

In yet another embodiment, the composition is in a carrier.

Another embodiment is a vector which comprises the nucleic acid molecules described above, including an expression vector comprising a the nucleic acid molecules described wherein the nucleic acid molecule is operatively linked to an expression control sequence.

Other included embodiments are a host-vector system for the production of a fusion polypeptide which comprises the expression vector, in a suitable host cell; the host-vector system wherein the suitable host cell is a bacterial cell, yeast cell, insect cell, or mammalian cell; the host-vector system wherein the suitable host cell is *E. Coli*: the host-vector system wherein the suitable host cell is a COS cell; the host-vector system wherein the suitable host cell is a CHO cell.

Another embodiment of the invention is a method of producing a fusion polypeptide which comprises growing cells of the host-vector system under conditions permitting production of the fusion polypeptide and recovering the fusion polypeptide so produced.

Additional embodiments include a fusion polypeptide encoded by the nucleic acid sequence set forth in FIGS. **10**A–**10**D or FIGS. **24**A–**24**C, which has been modified by acetylation or pegylation wherein the acetylation is accomplished with at least about a 100 fold molar excess of acetylation reagent or wherein acetylation is accomplished with a molar excess of acetylation reagent ranging from at least about a 10 fold molar excess to about a 100 fold molar excess or wherein the pegylation is 10K or 20K PEG.

A preferred embodiment includes a method of decreasing or inhibiting plasma leakage in a mammal comprising administering to the mammal the fusion polypeptide described above, including embodiments wherein the mammal is a human, the fusion polypeptide is acetylated or the fusion polypeptide is pegylated.

A further embodiments is a fusion polypeptide which specifically binds the VEGF receptor ligand VEGF.

A preferred embodiment of the invention is a method of blocking blood vessel growth in a human comprising administering an effective amount of the fusion polypeptide described above.

Also preferred is a method of inhibiting VEGF receptor ligand activity in a mammal comprising administering to the mammal an effective amount of the fusion polypeptide described above.

Preferred embodiments of these methods are wherein the $_{50}$ mammal is a human.

Further embodiments of the methods of the invention include attenuation or prevention of tumor growth in a human; attenuation or prevention of edema in a human, especially wherein the edema is brain edema; attenuation or prevention of ascites formation in a human, especially wherein the ascites is ovarian cancer-associated ascites.

Preferred embodiments of the invention include a fusion polypeptide capable of binding a VEGF polypeptide comprising (a) a VEGF receptor component operatively linked to (b) a multimerizing component, wherein the VEGF receptor component is the only VEGF receptor component in the fusion polypeptide and consists essentially of the amino acid sequence of Ig domain 2 of the extracellular domain of a first VEGF receptor and the amino acid sequence of Ig domain 3 of the extracellular domain of a second VEGF receptor.

In a further embodiment of the fusion polypeptide, the first VEGF receptor is Flt1.

In yet a further embodiment of the fusion polypeptide, the second VEGF receptor is Flk1.

Still another embodiment of the fusion polypeptide is one in which the second VEGF receptor is Flt4.

Preferred embodiments include a fusion polypeptide wherein amino acid sequence of Ig domain 2 of the extracellular domain of the first VEGF receptor is upstream of the amino acid sequence of Ig domain 3 of the extracellular domain of the second VEGF receptor and a fusion polypeptide wherein the amino acid sequence of Ig domain 2 of the extracellular domain of the first VEGF receptor is downstream of the amino acid sequence of Ig domain 3 of the extracellular domain of the second VEGF receptor.

In yet another embodiment, the fusion polypeptide multimerizing component comprises an immunoglobulin domain including an embodiment wherein the immunoglobulin domain is selected from the group consisting of the Fc domain of IgG, the heavy chain of IgG, and the light chain of IgG.

Preferred embodiments include a fusion polypeptide comprising an amino acid sequence of a modified Flt1 receptor, wherein the amino acid sequence selected from the group consisting of (a) the amino acid sequence set forth in FIGS. **13A–13D**; (b) the amino acid sequence set forth in FIGS. **14A–14C**; (c) the amino acid sequence set forth in FIGS. **15A–15C**; (d) the amino acid sequence set forth in FIGS. **16A–16D**; (e) the amino acid sequence set forth in FIGS. **21A–21C**; (f) the amino acid sequence set forth in FIGS. **22A–22C**; and (g) the amino acid sequence set forth in 30 FIGS. **24A–24C**.

Another preferred embodiment is a method of decreasing or inhibiting plasma leakage in a mammal comprising administering to the mammal the fusion polypeptide described above.

An alternative preferred embodiment is a method of inhibiting VEGF receptor ligand activity in a mammal comprising administering to the mammal an effective amount of the fusion polypeptide described above.

BRIEF DESCRIPTION OF THE FIGURES

FIG. 1. IEF gel analysis of unmodified and acetylated Flt1(1-3)-Fc proteins. Unmodified Flt1(1-3)-Fc protein is unable to enter the gel due to its >9.3 pl, whereas acetylated Flt1(1-3)-Fc is able to enter the gel and equilibrate at pl 5.2.

FIG. 2. Binding of unmodified Flt1(1-3)-Fc and acetylated Flt1(1-3)-Fc proteins to MATRIGELTM (a trademark of Becton, Dickinson & Co. relating to solubilized basement ₅₀ membrane preparations) coated plates. Unmodified Flt1(1-3)-Fc proteins binds extensive to extracellular matrix components in MATRIGELTM, whereas acetylated Flt1(1-3)-Fc does not bind.

FIG. 3. Binding of unmodified Flt1(1-3)-Fc, acetylated 55 Flt1(1-3)-Fc, and pegylated Flt1(1-3)-Fc in a BIACORETM (a trademark of Biacore AB relating to systems for protein interaction analysis)-based assay. Acetylated (columns 13–16), pegylated (columns 17–20), and heparin-treated Flt1(1-3)-Fc (columns 21–24) are each able to completely 60 compete with the Biacore chip-bound Flt1(1-3)-Fc for VEGF binding as compared to control (columns 1–4) and irrelevant protein (columns 5–8). Unmodified Flt1(1-3)-Fc (columns 5–6) appears to only partially compete with Biacore chip-bound Flt1(1-3)-Fc for VEGF binding. However, 65 washing the bound samples with 0.5M NaCl (columns 7–8) results in a binding profile similar to the modified forms of

Flt1(1-3)-Fc, indicating that the unmodified protein is exhibiting non-specific binding to the chip that can be eliminated by the salt wash.

FIG. 4. Binding of unmodified Flt1(1-3)-Fc, acetylated Flt1(1-3)-Fc, and pegylated Flt1(1-3)-Fc to VEGF in an ELISA-based assay. Both pegylated and acetylated Flt1(1-3)-Fc proteins bind to VEGF with affinities approaching that of unmodified Flt1(1-3)-Fc.

FIG. 5. Pharmacokinetic profiles of unmodified Flt1 (1-3)-Fc, acetylated Flt1(1-3)-Fc, and pegylated Flt1(1-3)-Fc. Balb/c mice (23–28 g) were injected subcutaneously with 4 mg/kg of unmodified, acetylated, or pegylated Flt1 (1-3)-Fc. The mice were tail bled at 1, 2, 4, 6, 24 hours, 2 days, and 3 days after injection of protein and the sera were assayed in a standard ELISA-based assay designed to detect Flt1(1-3)-Fc protein. The T_{max} for all of the Flt1(1-3)-Fc proteins was between the 6 hour and 24 hour time points. The C_{max} for the different proteins was as follows: Unmodified: 0.06 µg/ml—0.15 µg/ml; acetylated: 1.5 µg/ml.

FIGS. 6A–6B. IEF gel analysis of unmodified and stepacetylated Flt1(1-3)-Fc proteins. Unmodified Flt1(1-3)-Fc protein is unable to enter the gel due to its >9.3 pl, whereas most of the step-acetylated Flt1 (1-3)-Fc samples (30–100 fold excess samples) were able to migrate into the gel and equilibrate at pis ranging between 4.55–8.43, depending on the degree of acetylation.

FIG. 7. Binding of unmodified Flt1(1-3)-Fc and stepacetylated Flt1(1-3)-Fc proteins to Matrigel® coated plates. As with the irrelevant control protein, rTie2-Fc, step-acetylated Flt1(1-3)-Fc (20 and 30 fold excess samples) does not exhibit any binding to the Matrigel coated plate, whereas the non-acetylated Flt1(1-3)-Fc protein exhibits significant binding. The 10 fold excess sample shows reduced binding, but the degree of acetylation is not enough to completely block binding to extracellular matrix components.

FIG. 8. Binding of unmodified Flt1(1-3)-Fc and stepacetylated Flt1(1-3)-Fc in a Biacore-based assay. At a substoichiometric ratio (0.5 μ g/ml of either unmodified Flt1(1-3) or step-acetylated Flt1(1-3)-Fc vs. 0.2 μ g/ml VEGF), there is not enough Flt1(1-3)-Fc (either unmodified or step-acetylated) in the solution to completely bind the VEGF. At 1.0 μ g/ml, which approximates a 1:1 stoichiometric ratio, the both unmodified and step-acetylated Flt1 (1-3)-Fc are better able to compete for VEGF binding, but there is still insufficient Flt1 (1-3)-Fc protein (either unmodified or step-acetylated) to completely saturate the available VEGF. However, at 5.0 μ g/ml, which is several times greater than a 1:1 stoichiometric ratio, both the Flt1(1-3)-Fc and the step-acetylated Flt1(1-3)-Fc proteins are able to saturate the VEGF, regardless of the degree of acetylation.

FIG. 9. Pharmacokinetic profiles of unmodified Flt1(1-3)-Fc and step-acetylated Flt1 (1-3)-Fc. Balb/c mice (23–28 g) were injected subcutaneously with 4 mg/kg of unmodified or 10, 20, 40, 60 and 100 fold excess samples of stepacetylated Flt1(1-3)-Fc (3 mice for unmodified, 10, 20 and 40 fold excess samples and 2 mice for 60 and 100 fold excess samples). The mice were tail bled at 1, 2, 4, 6, 24 hours, 2 days and 3 days after injection. The sera were assayed in an ELISA-based assay designed to detect Flt1 (1-3)-Fc. The T_{max} for all of the Flt1(1-3)-Fc proteins tested was at the 6 hour time point but the C_{max} was as follows: Unmodified Flt1(1-3)-Fc: 0.06 µg/ml; 10 fold excess sample:—0.7 µg/ml, 20 fold excess sample—2 µg/ml, 40 fold excess sample—4 µg/ml, 60 fold excess sample—2 µg/ml, 100 fold excess sample—1 g/ml.

R

FIGS. **10A–10D**. Nucleic acid and deduced amino acid sequence of Flt1(1-3)-Fc.

FIG. 11. Schematic diagram of the structure of Flt1.

FIGS. **12**A and **12**B. Hydrophilicity analysis of the amino acid sequences of Ig domain 2 and Ig domain 3 of Flt1.

FIGS. **13**A–**13**D. Nucleic acid and deduced amino acid sequence of Mut1: Flt1(1- $3_{\Delta B}$)-Fc.

FIGS. **14A-14** C. Nucleic acid and deduced amino acid sequence of Mut2: $Flt1(2-3_{AB})$ -Fc.

FIGS. **15**A–**15**C. Nucleic acid and deduced amino acid 10 sequence of Mut3: Flt1(2-3)-Fc.

FIGS. **16**A–**16**D. Nucleic acid and deduced amino acid sequence of Mut4: $Flt1(1-3_{R-2N})$ -Fc.

FIG. 17. Binding of unmodified Flt1(1-3)-Fc, basic region deletion mutant Flt1(1-3)-Fc, and Flt1(1-3)_{R->N} mutant pro- 15 teins in a Biacore-based assay. At the sub-stoichiometric ratio (0.25 µg/ml Flt1(1-3)-Fc of unmodified, acetylated or genetically modified samples vs. 01. µg/ml VEGF), there is insufficient Flt1(1-3)-Fc protein to block binding of VEGF to the FIt1(1-3)-Fc immobilized on the Biacore chip. At 0.5 20 µg/ml of unmodified, acetylated or genetically modified Flt1 (1-3)-Fc proteins, the stoichiometric ratio approximates 1:1 and there is an increased ability to block VEGF binding to the Biacore chip. At 1.0 µg/ml of unmodified, acetylated or genetically modified Flt1 (1-3)-Fc proteins, which is 25 approximately a 10:1 stoichiometric ratio, the Flt1(1-3)-Fc proteins are able to block binding of VEGF to the Biacore chip, but they are not equivalent. Unmodified, acetylated, and Mut1: Flt1(1-3_{ΔB})-Fc are essentially equal in their ability to block VEGF binding, whereas Mut4: $Flt1(1-3_{R-30})$ >N)-Fc is somewhat less efficient at blocking binding

FIG. 18. Binding of unmodified Flt1(1-3)-Fc, Mut1: Flt1 ($1-3_{AB}$)-Fc, Mut2: Flt1 ($2-3_{AB}$)-Fc, and Flt1(2-3) mutant proteins to Matrigel® coated plates. Unmodified Flt1(1-3)-Fc protein binds avidly to these wells, the Mut3: Flt1(2-3)- 35 Fc protein binds somewhat more weakly, the Mut1: Flt1(1- 3_{AB})-Fc protein binds more weakly still, and the Mut2: Flt1($2-3_{AB}$)-Fc protein shows the best profile, binding more weakly than any of the other mutant proteins. The Mut4: Flt1($1-3_{R-5N}$)-Fc glycosylation mutant protein shows only 40 marginal benefit on the Matrigel assay.

FIG. 19. Binding of unmodified Flt1(1-3)-Fc, Mut1: Flt1 ($1-3_{\Delta B}$)-Fc, Mut2: Flt1($2-3_{\Delta B}$)-Fc, and Flt1(2-3) mutant proteins in an ELISA-based assay. At the concentrations tested, unmodified Flt1(1-3)-Fc, Mut1: Flt1($1-3_{\Delta B}$)-Fc, Mut2: Flt1 45 ($2-3_{\Delta B}$)-Fc, and Flt1(2-3) mutant proteins bind VEGF similarly.

FIG. **20**. Pharmacokinetic profiles of unmodified Flt1(1-3)-Fc, Mut1: Flt1 $(1-3_{AB})$ -Fc, Mut2: Flt1 $(2-3_{AB})$ -Fc, and Flt1(2-3) mutant proteins. the Cmax for these reagents was 50 as follows: Unmodified Flt1(1-3)-Fc—0.15 µg/ml; 40 fold molar excess acetylated Flt1(1-3)-Fc—1.5 µg/ml; and Mut1: Fit1(1-3_{AB})-Fc—0.7 µg/ml.

FIGS. 21A-21C. Nucleotide and deduced amino acid sequence of the modified Flt1 receptor termed Flt1D2.Flk1 55 D3.Fc Δ C1(a).

FIGS. **22**A–**22**C. Nucleotide and deduced amino acid sequence of the modified Flt1 receptor termed Flt1D2.VEGFR3D3.Fc Δ C1(a).

FIG. 23. Extracellular Matrix (ECM) Assay. The results of 60 this assay demonstrate that the Flt1D2.Flk1D3.Fc Δ C1(a) and Flt1D2.VEGFR3D3.Fc Δ C1(a) proteins are considerably less sticky to the ECM as compared to the Flt1(1-3)-Fc protein.

FIGS. **24**A–**24**C. Nucleotide and deduced amino acid 65 sequence of the modified Flt1 receptor termed VEGFR1R2-Fc Δ C1(a).

FIGS. **25**A–**25**C. Phosphorylation assay. At a 1.5 molar excess of either Flt1(1-3)-Fc, Flt1(1-3)-Fc (A40) or transient Flt1D2Flk1D3.FcΔC1(a) there is complete blockage of receptor stimulation by these three modified Flt1 receptors as compared to control media challenge. In contrast, transient Flt1D2VEGFR3D3.FcΔC1(a) does not show significant blockage at this molar excess, as compared with VEGF positive control challenge. Similar results are seen in FIG. **25**B, where the modified Fit receptors are in a 3-fold molar excess to VEGF165 ligand. In FIG. **25**C, where the modified Flt1 receptors are in a 6-fold molar excess to VEGF165 ligand, transient Flt1D2VEGFR3D3.FcΔC1(a) can now be shown to be partially blocking VEGF165-induced stimulation of cell-surface receptors.

FIGS. **26**A–**26**B. Phosphorylation assay. Detection by Western blot of tyrosine phosphorylated VEGFR2(Flk1) by VEGF165 ligand stimulation shows that cell-surface receptors are not phosphorylated by challenge samples which have VEGF165 preincubated with 1 and 2 fold molar excess (FIG. **26**A) or 3 and 4 fold molar excess (FIG. **26**B) of either transient Flt1 D2Flk1 D3.Fc Δ C1 (a), stable Flt1 D2Flk1 D3.Fc Δ C1 (a), or transient VEGFR1R2-Fc Δ C1(a). At all modified Flt1 receptor concentrations tested there is complete binding of VEGF165 ligand during the preincubation, resulting in no detectable stimulation of cell-surface receptors by unbound VEGF165 as compared to control media challenge.

FIG. **27**. MG/R2 Cell proliferation assay. The following modified Flt receptors Flt1(1-3)-Fc, Flt1D2.Flk1D3.Fc Δ C1 (a) and Flt1D2.VEGFR3D3.Fc Δ C1(a), plus an irrelevant receptor termed Tie2-Fc as a negative control, were titrated from 40 nM to 20 pM and incubated on the cells for 1 hr at 37° C. Human recombinant VEGF165 in defined media was then added to all the wells at a concentration of 1.56 nM. The negative control receptor Tie2-Fc does not block VEGF165-induced cell proliferation at any concentration whereas Flt1 D2.Flk1 D3.Fc Δ C1 (a) blocks 1.56 nM VEGF165 with a half maximal dose of 0.8 nM. Flt1(1-3)-Fc and Flt1D2.VEGFR3D3.Fc Δ C1(a) are less effective in blocking VEGF165 in this assay with a half maximal dose of ~2 nM. VEGF165 alone gives a reading of 1.2 absorbance units and the background is 0.38 absorbance units.

FIG. **28**. Biacore analysis of Binding Stoichiometry. Binding stoichiometry was calculated as a molar ratio of bound VEGF165 to the immobilized Flt1 D2Flk1 D3.Fc Δ C1 (a) or VEGFR1 R2-Fc Δ C1 (a), using the conversion factor of 1000 RU equivalent to 1 ng/ml. The results indicated binding stoichiometry of one VEGF165 dimeric molecule per one Flt1 D2Flk1 D3.Fc Δ C1 (a) or VEGFR1 R2-Fc Δ C1 (a) molecule.

FIG. 29 and FIG. 30. Size Exclusion Chromatography Stoichiometry. Flt1 D2Flk1 D3.Fc∆C1 (a) or VEGFR1 R2-Fc Δ C1 (a) at a concentration of 1 nM (estimated to be 1000 times higher than the KD of the Flt1 D2Flk1 D3.Fc\DeltaC1 (a) or VEGFR1 R2-Fc\DeltaC1 (a)/VEGF165 interaction) were mixed with varied concentrations of VEGF165. incubation, concentrations of After the free Flt1D2Flk1D3.Fc\DeltaC1(a) in solution were measured. The data shows that the addition of 1 nM VEGF165 into the Flt1D2Flk1D3.FcΔC1(a) solution completely blocks Flt1D2Flk1D3.Fc∆C1(a) binding to the VEGF165 surface. This result suggested the binding stoichiometry of one VEGF165 molecule per one Flt1 D2Flk1 D3.Fc∆C1 (a) molecule.

FIG. 31. Size Exclusion Chromatography (SEC) under native conditions. Peak #1 represents the Flt1D2Flk1D3.Fc Δ C1(a)/VEGF165 complex and peak #2

represents unbound VEGF165. Fractions eluted between 1.1 and 1.2 ml were combined and guanidinium hydrochloride (GuHCl)was added to a final concentration 4.5M to dissociate the complex.

FIG. **32**. Size Exclusion Chromatography (SEC) under ⁵ dissociative conditions. To separate the components of the receptor-ligand complex and to determine their molar ratio, 50 μ l of dissociated complex was loaded onto a SUPEROSE 12 PCTM (a trademark of Amerisham Pharmacia Biotech AG relating to sensitive and high resolving gel filtration sepa-¹⁰ rations of proteins, peptides, polynucleotides and other biomolecules in the micropreparative scale) 3.2/30 equilibrated in 6M GuHCl and eluted. Peak #1 represents Flt1D2Flk1D3.Fc Δ C1(a) and peak #2 represents VEGF 165.

15 FIG. 33, FIG. 34 and FIG. 35. Size Exclusion Chromatography (SEC) with On-Line Light Scattering. Size exclusion chromatography column with a MiniDawn on-line light scattering detector (Wyatt Technology, Santa Barbara, Calif.) and refractive index (RI) detectors (Shimadzu, Kyoto, Japan) was used to determine the molecular weight (MW) of the receptor-ligand complex. As shown in FIG. 33, the elution profile shows two peaks. Peak #1 represents the receptor-ligand complex and peak #2 represents the unbound VEGF165. MW was calculated from LS and RI 25 signals. The same procedure was used to determine MW of the individual components of the receptor-ligand complex. The results of these determinations are as follows: MW of the Flt1D2Flk1D3.FcΔC1(a)NEGF165 complex at the peak position is 157 300 (FIG. 33), the MW of VEGF165 at the peak position is 44 390 (FIG. 34) and the MW of R1R2 at the peak is 113 300 (FIG. 35).

FIG. **36**. Peptide mapping and glycosylation analysis. The disulfide structures and glycosylation sites in Flt1 D2.Flk1 D3.FcΔC1 (a) were determined by a peptide mapping ³⁵ method. There are a total of ten cysteines in Flt1D2.Flk1D3.FcΔC1(a); six of them belong to the Fc region. Cys27 is disulfide bonded to Cys76. Cys121 is disulfide bonded to Cys182. The first two cysteines in the Fc region (Cys211 and Cys214) form an intermolecular ⁴⁰ disulfide bond with the same two cysteines in another Fc chain. However, it can not be determined whether disulfide bonding is occurring between same cysteines (Cys211 to Cys211, for example) or between Cys211 and Cys214. Cys216 is disulfide bonded to Cys306. Cys 352 is disulfide bonded to Cys410.

There are five possible N-linked glycosylation sites in Flt1D2.Flk1D3.Fc Δ C1(a) and are found to be glycosylated to varying degrees. Complete glycosylation is observed at Asn33, Asn193, and Asn282. Partial glycosylation is ₅₀ observed on Asn65 and Asn120. Sites of glycosylation are highlighted by underline in the Figure.

FIG. 37. Pharmacokinetics of Flt1(1-3)-Fc (A40), Flt1 D2.Flk1 D3.Fc Δ C1 (a) and VEGFR1 R2-Fc Δ C1 (a). Balb/c mice were injected subcutaneously with 4 mg/kg of Flt1(1- 55 (A40), CHO transiently 3)-Fc expressed Flt1D2.Flk1D3.Fc Δ C1(a), CHO stably expressed Flt1D2.Flk1D3.Fc\DeltaC1(a), and CHO transiently expressed VEGFR1R2-Fc Δ C1(a). The mice were tail bled at 1, 2, 4, 6, 24 hrs, 2 days, 3 days and 6 days after injection. The sera 60 were assayed in an ELISA designed to detect FIt1(1-3)-Fc (A40), Flt1D2.Flk1D3.FcΔC1(a) or VEGFR1R2-FcΔC1(a). The Tmax for Flt1(1-3)-Fc (A40) was at 6 hrs while the Tmax for the transient and stable Flt1D2.Flk1D3.FcΔC1(a) and the transient VEGFR1R2-Fc∆C1(a) was 24 hrs. The 65 Cmax for Flt1(1-3)-Fc (A40) was 8 µg/ml, For both transients (Flt1D2.Flk1D3.Fc\DeltaC1(a) and VEGFR1R2-Fc\DeltaC1

(a)) the Cmax was 18 μg/ml and the Cmax for the stable VEGFR1R2-FcΔC1(a) was 30 μg/ml.

FIG. **38**. Pharmacokinetics of Flt1(1-3)-Fc (A40), Flt1D2.Flk1 D3.Fc Δ C1 (a) and Flt1 D2.VEGFR3D3.Fc Δ A1 (a). Balb/c mice were injected subcutaneously with 4 mg/kg of Flt1 (1-3)-Fc (A40), CHO transiently expressed Flt1D2.Flk1D3.Fc Δ C1(a) and CHO transiently expressed Flt1D2.VEGFR3D3.Fc Δ C1(a). The mice were tail bled at 1, 2, 5, 6, 7, 8, 12, 15 and 20 days after injection. The sera were assayed in an ELISA designed to detect Flt1 (1-3)-Fc, Flt1 D2.Flk1 D3.Fc Δ C1 (a) and Flt1D2.VEGFR3D3.Fc Δ C1(a). Flt1(1-3)-Fc (A40) could no longer be detected in the serum after day 5 whereas Flt1D2.Flk1D3.Fc Δ C1(a) and Flt1D2.VEGFR3D3.Fc Δ C1(a) were detectable for 15 days or more.

FIG. **39**. The Ability of Flt1D2.Flk1D3.Fc Δ C1(a) to Inhibit HT-1080 Fibrosarcoma Tumor Growth In Vivo. Every other day or 2 times per week treatment of SCID mice with Flt1D2.Flk1D3.Fc Δ C1(a) at 25 mg/Kg significantly decreases the growth of subcutaneous HT-1080 fibrosarcoma tumors.

FIG. 40. The Ability of Flt1D2.Flk1D3.Fc Δ C1(a) to Inhibit C6 Glioma Tumor Growth In Vivo. Every other day or 2 times a week treatment of SCID mice with Flt1 D2.Flk1 D3.Fc Δ C1 (a) significantly decreases the growth of subcutaneous C6 glioma tumors at doses as low as 2.5 mg/Kg.

FIG. **41**. VEGF-induced Uterine Hyperpermeability. PMSG injected subcutaneously (5 IU) to induce ovulation in prepubertal female rats results in a surge of estradiol after 2 days which in turn causes an induction of VEGF in the uterus. This induction results in hyperpermeability of the uterus and an increase in uterine wet. Subcutaneous injection of Flt1 (1-3)-Fc (A40), Flt1 D2.Flk1 D3.Fc Δ C1 (a) and Flt1D2.VEGFR3D3.Fc Δ C1(a) at 25 mg/kg at 1 hr after PMSG injection results in about a 50% inhibition of the increase in uterine wet weight.

FIGS. **42**A–**42**B. Assessment of Corpus Luteum Angiogenesis Using Progesterone as a Readout. PMSG was injected subcutaneously (5 IU) to induce ovulation in prepubertal female rats, resulting in a fully functioning corpus luteum containing a dense network of blood vessels that secretes progesterone into the blood stream to prepare the uterus for implantation. The induction of angiogenesis in the corpus luteum requires VEGF. Resting levels of progesterone are about 5 ng/ml and can be induced to 25–40 ng/ml after PMSG. Subcutaneous injection of Flt1(1-3)-Fc (A40) or Flt1D2.Flk1D3.Fc Δ C1(a) at 25 mg/kg or 5 mg/kg at 1 hr. after PMSG injection resulted in a complete inhibition of the progesterone induction on day 4.

DETAILED DESCRIPTION OF THE INVENTION

It has been a long standing problem in the art to produce a receptor based VEGF antagonist that has a pharmacokinetic profile that is appropriate for consideration of the antagonist as a therapeutic candidate. Applicants describe herein, for the first time, a chimeric polypeptide molecule, capable of antagonizing VEGF activity, that exhibits improved pharmacokinetic properties as compared to other known receptor-based VEGF antagonists. The chimeric polypeptide molecules described herein thus provide for the first time appropriate molecules for use in therapies in which antagonism of VEGF is a desired result. The present invention provides for novel chimeric polypeptide molecules formed by fusing a modified extracellular ligand binding domain of the Flt1 receptor to the Fc region of IgG.

The extracellular ligand binding domain is defined as the 5 portion of a receptor that, in its native conformation in the cell membrane, is oriented extracellularly where it can contact with its cognate ligand. The extracellular ligand binding domain does not include the hydrophobic amino acids associated with the receptor's transmembrane domain 10 or any amino acids associated with the receptor's intracellular domain. Generally, the intracellular or cytoplasmic domain of a receptor is usually composed of positively charged or polar amino acids (i.e. lysine, arginine, histidine, glutamic acid, aspartic acid). The preceding 15-30, pre- 15 dominantly hydrophobic or apolar amino acids (i.e. leucine, valine, isoleucine, and phenylalanine) comprise the transmembrane domain. The extracellular domain comprises the amino acids that precede the hydrophobic transmembrane stretch of amino acids. Usually the transmembrane domain 20 is flanked by positively charged or polar amino acids such as lysine or arginine. von Heijne has published detailed rules that are commonly referred to by skilled artisans when determining which amino acids of a given receptor belong to the extracellular, transmembrane, or intracellular domains 25 (See von Heijne, 1995, BioEssays 17:25-30). Alternatively, websites on the Internet, have become available to provide protein chemists with information about making predictions about protein domains.

The present invention provides for the construction of 30 nucleic acid molecules encoding chimeric polypeptide molecules that are inserted into a vector that is able to express the chimeric polypeptide molecules when introduced into an appropriate host cell. Appropriate host cells include, but are not limited to, bacterial cells, yeast cells, insect cells, and 35 mammalian cells. Any of the methods known to one skilled in the art for the insertion of DNA fragments into a vector may be used to construct expression vectors encoding the chimeric polypeptide molecules under control of transcriptional/translational control signals. These methods may include in vitro recombinant DNA and synthetic techniques and in vivo recombinations (genetic recombination) (See Sambrook, et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory; Current Protocols in Molecular Biology, Eds. Ausubel, et al., Greene Publ. 45 Assoc., Wiley-Interscience, NY).

Expression of nucleic acid molecules encoding the chimeric polypeptide molecules may be regulated by a second nucleic acid sequence so that the chimeric polypeptide molecule is expressed in a host transformed with the recom- 50 binant DNA molecule. For example, expression of the chimeric polypeptide molecules described herein may be controlled by any promoter/enhancer element known in the art. Promoters which may be used to control expression of the chimeric polypeptide molecules include, but are not 5 limited to, the long terminal repeat as described in Squinto et al., (1991, Cell 65:1-20); the SV40 early promoter region (Bernoist and Chambon, 1981, Nature 290:304-310), the CMV promoter, the M-MuLV 5' terminal repeat the promoter contained in the 3' long terminal repeat of Rous 60 sarcoma virus (Yamamoto, et al., 1980, Cell 22:787-797), the herpes thymidine kinase promoter (Wagner et al., 1981, Proc. Natl. Acad. Sci. U.S.A. 78:144-1445), the regulatory sequences of the metallothionine gene (Brinster et al., 1982, Nature 296:39-42); prokaryotic expression vectors such as 65 the β-lactamase promoter (Villa-Kamaroff, et al., 1978, Proc. Natl. Acad. Sci. U.S.A. 75:3727-3731), or the tac

promoter (DeBoer, et al., 1983, Proc. Natl. Acad. Sci. U.S.A. 80:21-25, see also "Useful proteins from recombinant bacteria" in Scientific American, 1980, 242:74-94); promoter elements from yeast or other fungi such as the Gal 4 promoter, the ADH (alcohol dehydrogenase) promoter, PGK (phosphoglycerol kinase) promoter, alkaline phosphatase promoter, and the following animal transcriptional control regions, which exhibit tissue specificity and have been utilized in transgenic animals: elastase I gene control region which is active in pancreatic acinar cells (Swift et al., 1984, Cell 38:639-646; Ornitz et al., 1986, Cold Spring Harbor Symp. Quant. Biol. 50:399-409; MacDonald, 1987, Hepatology Z:425-515); insulin gene control region which is active in pancreatic beta cells (Hanahan, 1985, Nature 315:115-122), immunoglobulin gene control region which is active in lymphoid cells (Grosschedl et al., 1984, Cell 38:647-658; Adames et al., 1985, Nature 318:533-538; Alexander et al., 1987, Mol. Cell. Biol. 7:1436-1444), mouse mammary tumor virus control region which is active in testicular, breast, lymphoid and mast cells (Leder et al., 1986, Cell 45:485-495), albumin gene control region which is active in liver (Pinkert et al., 1987, Genes and Devel. 1:268-276), alpha-fetoprotein gene control region which is active in liver (Krumlauf et al., 1985, Mol. Cell. Biol. 5:1639-1648; Hammer et al., 1987, Science 235:53-58); alpha 1-antitrypsin gene control region which is active in the liver (Kelsey et al, 1987, Genes and Devel. 1:161-171), beta-globin gene control region which is active in myeloid cells (Mogram et al., 1985, Nature 315:338-340; Kollias et al., 1986, Cell 46:89-94); myelin basic protein gene control region which is active in oligodendrocyte cells in the brain (Readhead et al., 1987, Cell 48:703-712); myosin light chain-2 gene control region which is active in skeletal muscle (Shani, 1985, Nature 314:283-286), and gonadotropic releasing hormone gene control region which is active in the hypothalamus (Mason et al., 1986, Science 234:1372-1378).

Thus, according to the invention, expression vectors capable of being replicated in a bacterial or eukaryotic host comprising chimeric polypeptide molecule-encoding nucleic acid as described herein, are used to transfect the host and thereby direct expression of such nucleic acids to produce the chimeric polypeptide molecules, which may then be recovered in a biologically active form. As used herein, a biologically active form includes a form capable of binding to VEGF.

Expression vectors containing the chimeric nucleic acid molecules described herein can be identified by three general approaches: (a) DNA-DNA hybridization, (b) presence or absence of "marker" gene functions, and (c) expression of inserted sequences. In the first approach, the presence of a foreign gene inserted in an expression vector can be detected by DNA-DNA hybridization using probes comprising sequences that are homologous to the inserted chimeric polypeptide molecule sequences. In the second approach, the recombinant vector/host system can be identified and selected based upon the presence or absence of certain "marker" gene functions (e.g., thymidine kinase activity, resistance to antibiotics, transformation phenotype, occlusion body formation in baculovirus, etc.) caused by the insertion of foreign genes in the vector. For example, if the chimeric polypeptide molecule DNA sequence is inserted within the marker gene sequence of the vector, recombinants containing the insert can be identified by the absence of the marker gene function. In the third approach, recombinant expression vectors can be identified by assaying the foreign gene product expressed by the recombinant. Such assays can

be based, for example, on the physical or functional properties of the chimeric polypeptide molecules. Cells of the present invention may transiently or, preferably, constitutively and permanently express the chimeric polypeptide molecules.

The chimeric polypeptide molecules may be purified by any technique which allows for the subsequent formation of a stable, biologically active chimeric polypeptide molecule. For example, and not by way of limitation, the factors may be recovered from cells either as soluble proteins or as 10 inclusion bodies, from which they may be extracted quantitatively by 8M guanidinium hydrochloride and dialysis (see, for example, Builder, et al., U.S. Pat. No. 5,663,304). In order to further purify the factors, conventional ion exchange chromatography, hydrophobic interaction chroma-15 tography, reverse phase chromatography or gel filtration may be used.

In one embodiment of the invention, the nucleotide sequence encoding the first component is upstream of the nucleotide sequence encoding the second component. In 20 another embodiment of the invention, the nucleotide sequence encoding the first component is downstream of the nucleotide sequence encoding the second component. Further embodiments of the invention may be prepared in which the order of the first, second and third fusion polypeptide 25 components are rearranged. For example, if the nucleotide sequence encoding the first component is designated 1, the nucleotide sequence encoding the second component is designated 2, and the nucleotide sequence of the third component is designated 3, then the order of the components 30 in the isolated nucleic acid of the invention as read from 5' to 3' may be any of the following six combinations: 1,2,3; 1,3,2; 2,1,3; 2,3,1; 3,1,2; or 3,2,1.

The present invention also has diagnostic and therapeutic utilities. In particular embodiments of the invention, meth-35 ods of detecting aberrancies in the function or expression of the chimeric polypeptide molecules described herein may be used in the diagnosis of disorders. In other embodiments, manipulation of the chimeric polypeptide molecules or agonists or antagonists which bind the chimeric polypeptide 40 molecules may be used in the treatment of diseases. In further embodiments, the chimeric polypeptide molecule is utilized as an agent to block the binding of a binding agent to its target.

By way of example, but not limitation, the method of the 45 invention may be useful in treating clinical conditions that are characterized by vascular permeability, edema or inflammation such as brain edema associated with injury, stroke or tumor; edema associated with inflammatory disorders such as psoriasis or arthritis, including rheumatoid arthritis; 50 asthma; generalized edema associated with burns; ascites and pleural effusion associated with tumors, inflammation or trauma; chronic airway inflammation; capillary leak syndrome; sepsis; kidney disease associated with increased leakage of protein; and eye disorders such as age related 55 macular degeneration and diabetic retinopathy.

An amino acid sequence analysis of Flt1(1-3)-Fc revealed the presence of an unusually high number (46) of the basic amino acid residue lysine. An IEF analysis of Flt1(1-3)-Fc showed that this protein has pl greater than 9.3, confirming 60 the prediction that the protein is very basic. It was hypothesized that the basic nature of Flt1 (1-3)-Fc protein was causing it to bind to extracellular matrix components and that this interaction might be the cause of the extremely short detectable circulating serum half-life exhibited by Flt1 65 (1-3)-Fc when injected into mice. In order to test this hypothesis, Flt1(1-3)-Fc protein was acetylated at the lysine

residues to reduce the basic charge. Acetylated Flt1(1-3)-Fc was then tested in the assays described infra.

The following examples are offered by way of illustration and not by way of limitation.

EXAMPLES

Example 1

Expression of Flt1(1-3)-Fc Protein in CHO K1 Cells

Using standard molecular biology techniques (see e.g., Molecular Cloning, A Laboratory Manual (Sambrook, et al., Cold Spring Harbor Laboratory), Current Protocols in Molecular Biology (Eds. Ausubel, et al., Greene Publ. Assoc., Wiley-Interscience, NY), the gene encoding Flt1(1-3)-Fc was inserted into the expression vector pEE14.1 (Lonza Biologics, pic) at a multiple cloning site downstream of the CMV promoter. CHO K1 cells were transfected with the pEE14.1/Flt1(1-3)-Fc DNA construct using lipofectamine (Gaithersburg, Md.). The transfected CHO K1 cells were grown in glutamine-free DMEM (JRH, Kansas City, Mo.) containing 25 µM methionine sulfoximine (MSX) from Sigma Inc., St. Louis, Mo., and high recombinant protein expressors were obtained by screening the CHO K1 cell supernatants from over 100 hand-picked colony isolates using a standard immunoassay which captures and detects human Fc. The selected hand-picked clone was amplified in the presence of 100 µM MSX followed by a second round of screening of the amplified clones. The highest producing clone had a specific productivity of recombinant Flt1(1-3)-Fc protein of 55 pg/cell/day.

The selected clone was expanded in 225 cm² T-flasks (Corning, Acton, Mass.) and then into 8.5L roller bottles (Corning, Acton, Mass.) using the cell culture media described supra. Cells were removed from the roller bottles by standard trypsinization and put into 3.5L of suspension medium. The suspension medium is comprised of glutamine-free ISCHO medium (Irvine Scientific, Santa Ana, Calif.) containing 5% fetal bovine serum (FBS from Hyclone Labs, Logan, Utah), 10 µM MSX and GS supplement (JRH Scientific, Kansas City, Mo.) in a 5L Celligen bioreactor (New Brunswick Scientific, New Brunswick, N.J.) at a density of 0.3×106 cells/mL. After the cells reached a density of 3.6×10⁶/mL and were adapted to suspension they were transferred to a 60L bioreactor (ABEC, Allentown, Pa.) at a density of 0.5×106 cells/mL in 20L of ISCHO medium with 5% fetal bovine serum. After two days an additional 20L of ISCHO+5% fetal bovine serum was added to the bioreactor. The cells were allowed to grow for an additional two days reaching a final density of 3.1×106 cells/mL, and a final Flt1(1-3)-Fc concentration at harvest was 95 mg/L. At harvest the cells were removed by tangential flow filtration using 0.45 µm Prostak Filters (Millipore, Inc., Bedford, Mass.).

Example 2

Purification of Flt1(1-3)-Fc Protein Obtained from CHO K1 Cells

Flt1 (1-3)-Fc protein was initially purified by affinity 65 chromatography. A Protein A column was used to bind, with high specificity, the Fc portion of the molecule. This affinitypurified protein was then concentrated and passed over a SEC column. The protein was then eluted into the formulation buffer. The following describes these procedures in detail.

Materials and Methods

All chemicals were obtained from J. T. Baker, Phillipsburg, N.J. with the exception of PBS, which was obtained as a 10× concentrate from Life Technologies, Gaithersburg, Md. Protein A Fast Flow and SUPERDEX200TM (a trademark of Amerisham Pharmacia Biosciences relating to a prepacked column for high performance of gel filtration of protein. DNA fragments and other biomolecules) preparation grade resins were obtained from Pharmacia, Piscataway, N.J. Equipment and membranes for protein concentration were obtained from Millipore, Bedford, Mass.

Approximately 40 L of 0.45 µm-filtered CHO conditioned media containing Flt1(1-3)-Fc protein was applied to a 290 mL Protein A Fast Flow column (10 cm diameter) that had been equilibrated with PBS. The column was washed with PBS containing 350 mM NaCl and 0.02% CHAPS and the bound protein was eluted with 20 mM Citric Acid containing 10 mM Na2HPO4. The single peak in the elution was collected and its pH was raised to neutrality with 1M NaOH. The eluate fractions was concentrated to approximately 9 mg/mL using 10K regenerated cellulose membranes by both tangential flow filtration and by stirred cell concentration. To remove aggregates and other contaminants, the concentrated protein was applied to a column packed with Superdex 200 preparation grade resin (10 cm×55 cm) and run in PBS containing 5% glycerol. The main peak fractions were 30 pooled, sterile filtered, aliquoted and stored at -80° C.

Example 3

Acetylation of Flt1(1-3)-Fc Protein

Two milligrams of Flt1(1-3)-Fc protein were acetylated as described in the instruction manual provided with the sulfo-NHS-acetate modification kit (Pierce Chemical Co., Rockford, Ill., Cat.#26777).

Example 4

Characterization of Acetylated Flt1(1-3)-Fc Protein

(a.) IEF analysis

Flt1(1-3)-Fc and acetylated Flt1(1-3)-Fc were analyzed by standard IEF analysis. As shown in FIG. 1, Flt1(1-3)-Fc protein is not able to migrate into the gel and therefore must have a pl greater than 9.3, the highest pl in the standard. However, acetylated Flt1(1-3)-Fc is able to migrate into the gel and equilibrate at a pl of approximately 5.2. This result demonstrates that acetylation reduces the net positive charge of the protein and therefore its pl considerably.

(b.) Binding to Extracellular Matrix Components

To test for binding to extracellular matrix components, Flt1 (1-3)-Fc and acetylated Flt1 (1-3)-Fc where tested in an assay designed to mimic the interaction with extracellular matrix components. In this assay, 96-well tissue culture 60 plates are coated with Matrigel (Biocoat MATRIGEL®) matrix thin layer 96 well plate, Catalog #40607, Becton Dickinson Labware, Bedford, Mass.). The plates are incubated with varying concentrations of either Flt1(1-3)-Fc, acetylated Flt1 (1-3)-Fc, or rTic2-Fc (an irrelevant control) 65 protein are added to the wells. The plates are incubated for 1-2 hours at either room temperature or 37° C. degrees and

then detection of bound proteins is accomplished by adding a secondary alkaline phosphatase-conjugated anti-human Fc antibody to the wells. Finally, alkaline phosphatase substrate is added to the wells and optical density is measured. FIG. **2** shows the results of this assay. Like the irrelevant control protein rTie2-Fc, acetylated Flt1(1-3)-Fc does not exhibit any binding to the Matrigel coated plate, whereas the non-acetylated Flt1 (1-3)-Fc protein exhibits significant binding. This result indicates that acetylation of basic amino acid residues is an effective way to interfere with the charge interactions that exist between positively charged proteins and the negatively charged extracellular matrix components they are exposed to in vivo.

Example 5

Pegylation of Flt1(1-3)-Fc Protein

Although pegylation (polyethylene glycol—PEG) of proteins has been shown to increase their in vivo potency by enhancing stability and bioavailability while minimizing immunogenicity (see references cited supra), it is counterintuitive that pegylating molecules that are too large to be filtered by the kidney glomeruli would improve their pharmacokinetic properties. Without being bound by theory, Applicants postulated that pegylation of the Flt1(1-3)-Fc molecules could improve the pharmacokinetic properties, so possibly not by altering the positive charge or by decreasing the pl of Flt1(1-3)-Fc, but rather by physically shielding the positive charges from interacting with the extracellular matrix. Applicants decided to attempt to improve the pharmacokinetic properties of Flt1(1-3)-Fc molecules by attach-35 ing strands of 20K PEGs as described infra.

Materials and Methods

Purified Flt1(1-3)-Fc derived from CHO cells (see supra) was used in the following pegylation experiments. Func-40 tionalized PEGs were obtained from Shearwater Polymers, Huntsville, Ala.; Bicine from Sigma, St Louis, Mo.; Superose 6 column from Pharmacia, Piscataway, N.J.; PBS as a 10x concentrate from Life Technologies, Gaithersburg, Md.; Glycerol from J. T. Baker, Phillipsburg, N.J.; and Bis-Tris
⁴⁵ precast gels from Novex, CA.

20K PEG strands functionalized with amine-specific terminal moieties were used in small-scale reaction studies that were set-up to evaluate different reaction conditions in which the PEG:protein stoichiometry was varied. Based on these reactions and the analyses of samples on standard SDS-PAGE, Flt1(1-3)-Fc at a concentration of 1.5 mg/mL was reacted at pH 8.1 with 20K SPA-PEG (PEG succinimidyl propionate) molecules at a PEG-to-Flt1(1-3)-Fc mono-55 mer molar ratio of 1:6. The reaction was allowed to proceed at 8° C. overnight. For initial purification, the reaction products were applied to a 10 mm×30 cm Superose 6 column equilibrated with PBS containing 5% Glycerol. The column appeared to separate pegylated Flt1(1-3)-Fc molecules based on the extent of pegylation. Fractions corresponding to what appeared to be primarily mono-pegylated and di-pegylated dimeric Flt1 (1-3)-Fc, as judged by banding patterns on reducing and non-reducing SDS-PAGE gels were pooled. The protein concentration was determined by measuring absorbance at 280 nm. The pegylated Flt1(1-3)-Fc protein was sterile filtered, aliquoted and stored at -40° C.

Example 6

Binding of Unmodified, Acetylated, and Pegylated Flt1(1-3)-Fc in a Biacore-Based Assay

Unmodified, acetylated, and pegylated Flt1(1-3)-Fc proteins were tested in a Biacore-based assay to evaluate their ability to bind to the Flt1 ligand, VEGF. In this assay, unmodified Flt1(1-3)-Fc protein was immobilized on the surface of a Biacore chip (see Biacore Instruction Manual, 10 Pharmacia, Inc., Piscataway, N.J., for standard procedures) and a sample containing 0.2 µg/ml VEGF and either unmodified Flt1(1-3)-Fc, acetylated Flt1(1-3)-Fc or pegylated Flt1 (1-3)-Fc (each at 25 µg/ml) was passed over the Ftl1(1-3)-Fc-coated chip. To minimize the effects of non- 15 specific binding, the bound samples were washed with a 0.5M NaCl wash. In one sample, unmodified Flt1(1-3)-Fc was mixed with heparin. Heparin is a negatively charged molecule and the Fit1(1-3)-Fc protein is a positively charged molecule, so when the two molecules are mixed together, 20 they should interact through their respective charges. This essentially neutralizes Flt1 (1-3)-Fc's inherent positive charge making the molecule behave as if it has been chemically or genetically modified so as to reduce its charge and its tendency to bind via charge interactions. As shown in 25 FIG. 3, acetylated (columns 13-16), pegylated (columns 17-20), and heparin-treated Flt1(1-3)-Fc (columns 21-24) are each able to completely compete with the Biacore chip-bound Flt1(1-3)-Fc for VEGF binding as compared to control (columns 1-4) and irrelevant protein (columns 5-8). 30 Unmodified Flt1(1-3)-Fc (columns 5-6) appeared to only partially compete with Biacore chip-bound Flt1(1-3)-Fc for VEGF binding. However, washing the bound samples with 0.5M NaCl (columns 7-8) resulted in a binding profile similar to the modified forms of Flt1(1-3)-Fc, indicating that 35 the unmodified protein was exhibiting non-specific binding to the chip that could be eliminated by the salt wash.

Example 7

Binding of Unmodified, Acetylated, and Pegylated Flt1(1-3)-Fc in an ELISA-Based Assay

Unmodified, acetylated, and pegylated Flt1 (1-3)-Fc proteins were tested in a standard ELISA-based assay to evalu-45 ate their ability to bind the Flt1 receptor ligand VEGF. As shown in FIG. **4**, both pegylated and acetylated Flt1(1-3)-Fc proteins are capable of binding to VEG F, demonstrating that modifying the protein either by pegylation or acetylation does not destroy its ability to bind its ligand. 50

Example 8

Pharmacokinetic Analysis of Unmodified Flt1(1-3)-Fc, Acetylated Flt1(1-3)-Fc, and Pegylated Flt1(1-3)-Fc

In vivo experiments were designed to assess the pharmacokinetic profiles of unmodified Flt1(1-3)-Fc, acetylated Flt1(1-3)-Fc, and pegylated Flt1(1-3)-Fc protein. Balb/c 60 mice (23-28 g; 3 mice/group) were injected subcutaneously with 4 mg/kg of unmodified, acetylated, or pegylated Flt1 (1-3)-Fc. The mice were tail bled at 1, 2, 4, 6, 24 hours, 2 days, and 3 days after injection of protein. The sera were assayed in a standard ELISA-based assay designed to detect 65 Flt1(1-3)-Fc protein. Briefly, the assay involves coating an ELISA plate with VEGF, binding the unmodified, acety-

lated, or pegylated Flt1(1-3)-Fc-containing sera, and reporting with an anti-Fc antibody linked to alkaline phosphatase. As shown in FIG. 5, the Tmax for all of the Flt1(1-3)-Fc proteins was between the 6 hour and 24 hour time points. The Cmax for the different proteins was as follows: Unmodified: $0.06\mu/ml-0.15 \mu g/ml$; acetylated: $1.5 \mu g/ml-4.0 \mu g/ml$; and pegylated: approximately 5 $\mu g/ml$.

Example 9

Step-Acetylation of Flt1(1-3)-Fc

To determine what minimal amount of acetylation is necessary to eliminate binding to extracellular matrix components, an experiment was designed that acetylated the Flt1(1-3)-Fc protein in a step-wise fashion by using increasing amounts of molar excess of acetylation reagent in the acetylation reaction mixture. The range of molar excess was as follows: 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100 moles of acetylation reagent per 1 mole of Flt1(1-3)-Fc monomer. The reactions were performed as detailed in the instruction manual provided with the sulfo-NHS-Acetate modification kit (Pierce Chemical Co., Rockford, Ill., Cat.# 26777).

Example 10

Characterization of Step-Acetylated Flt1(1-3)-Fc

(a.) IEF analysis

40

Unmodified Flt1(1-3)-Fc and step-acetylated Flt1(1-3)-Fc proteins were analyzed by standard IEF analysis. As shown in FIGS. **6**A–**6**B, unmodified Flt1(1-3)-Fc protein was not able to migrate into the gel due to its extremely high pl (greater than 9.3). However, most of the step-acetylated Flt1(1-3)-Fc samples (30-100 fold molar excess samples) were able to migrate into the gel and equilibrate at pis ranging between 4.55–8.43, depending on the degree of acetylation can change the positive charge of the protein in a dose-dependent manner and that reduction of the pl can be controlled by controlling the degree of acetylation.

(b.) Binding of Step-Acetylated Flt1(1-3)-Fc to Extracellular Matrix Components

To test for binding to extracellular matrix components, Flt1(1-3)-Fc and step-acetylated Flt1(1-3)-Fc where tested in the above-described assay designed to mimic the interaction with extracellular matrix components. Varying con-50 centrations of either unmodified Flt1 (1-3)-Fc. step-acetvlated Flt1(1-3)-Fc (10, 20, and 30 fold molar excess samples), or rTie2-Fc (an irrelevant control) protein were added to the wells. The plates were incubated for 1-2 hours at room temperature or 37° C. and then detection of bound 55 proteins was accomplished by adding a secondary alkaline phosphatase-conjugated anti-human Fc antibody to the wells. Alkaline phosphatase substrate was subsequently added to the wells and optical density measured. FIG. 7 shows the results of this assay. Like the irrelevant control protein rTie2-Fc, step-acetylated Flt1(1-3)-Fc (20 and 30 fold molar excess samples) did not exhibit any significant binding to the Matrigel coated plate, whereas the nonacetylated Flt1 (1-3)-Fc protein exhibited significant binding. The binding is saturable, indicating that the Flt1(1-3)-Fc protein may be binding to specific sites, rather than a more general charge-mediated interaction that might not be saturable. The 10 fold molar excess sample showed reduced

50

binding, but the degree of acetylation was not enough to completely block binding to extracellular matrix components. The 20 fold molar excess and higher samples displayed no detectable binding, despite the fact that by IEF analysis (FIGS. 6A and 6B) the lower molar excess samples 5 still had a large net positive charge. This result demonstrates that it is not necessary to completely acetylate all available basic amino acids in order to eliminate binding to extracellular matrix components.

(c.) Binding of Step-Acetylated Flt1(1-3)-Fc in a Biacore- 10 Based Assay

Unmodified and step-acetylated Flt1(1-3)-Fc proteins where tested in a Biacore-based assay to evaluate their ability to bind to the Flt1 ligand, VEGF. In this assay, unmodified Flt1(1-3)-Fc protein (0.5, 1.0, or 5.0 µg/ml) was immobilized on the surface of a Biacore chip (see Biacore Instruction Manual, Pharmacia, Inc., Piscataway, N.J., for standard procedures) and a solution containing 0.2 µg/ml VEGF and either unmodified Flt1(1-3)-Fc (at either 0.5, 1.0, or 5.0 µg/ml) or 10 different step-acetylated Flt1 (1-3)-Fc samples (at 0.5, 1.0, or 5.0 µg/ml each) were passed over the Flt1 (1-3)-Fc-coated chip. As shown in FIG. 8, at a substoichiometric ratio (0.5 µg/ml of either unmodified Flt1(1-3) or step-acetylated Flt1(1-3)-Fc vs. 0.2 µg/ml VEGF), there is not enough Flt1(1-3)-Fc (either unmodified or step-acetylated) in the solution to completely bind the VEGF. At 1.0 µg/ml, which approximates a 1:1 stoichiometric ratio, both unmodified and step-acetylated Flt1(1-3)-Fc are better able to compete for VEGF binding, but there is still insufficient Flt1 (1-3)-Fc protein (either unmodified or step-acetylated) to completely bind the available VEGF. However, at 5.0 µg/ml, which is several times greater than a 1:1 stoichiometric ratio, both the Flt1 (1-3)-Fc and the step-acetylated Flt1(1-3)-Fc proteins are able to bind the 35 VEGF, regardless of the degree of acetylation. This clearly demonstrates that acetvlation does not alter Flt1(1-3)-Fc's ability to bind VEGF.

(d.) Pharmacokinetic Analysis of Step-Acetylated Flt1(1-3)-Fc

In vivo experiments were designed to assess the pharmacokinetic profiles of unmodified Flt1(1-3)-Fc and stepacetylated Flt1(1-3)-Fc protein. Balb/c mice (23-28 g) were injected subcutaneously with 4 mg/kg of unmodified or 10, 20, 40, 60 and 100 fold molar excess samples of step-45 acetylated Flt1 (1-3)-Fc (3 mice for unmodified, 0.10, 20 and 40 fold molar excess samples and 2 mice for 60 and 100 fold molar excess samples). The mice were tail bled at 1, 2, 4, 6, 24 hours, 2 days and 3 days after injection. The sera were assayed in an ELISA-based assay designed to detect Flt1 (1-3)-Fc (described supra). FIG. 9 details the results of this study. The Tmax for all of the Flt1(1-3)-Fc proteins tested was at the 6 hour time point but the Cmax was as follows: Unmodified Flt1(1-3)-Fc: 0.06 µg/ml; 10 fold molar excess sample:-0.7 µg/ml, 20 fold molar excess sample-2 µg/ml, 40 fold molar excess sample-4 µg/ml, 60 fold molar excess sample-2 µg/ml, 100 fold molar excess sample-1 µg/ml. This results demonstrates that acetylation or pegylation of Flt1(1-3)-Fc significantly improves its pharmacokinetic profile.

Example 11

Construction of Flt1(1-3)-Fc Basic Region Deletion Mutant Designated Mut1: Flt1(1-3AB)-Fc

Based on the observation that acetylated Flt1(1-3)-Fc, which has a pl below 6, has much better pharmacokinetics 22

than the highly positive unmodified Flt1(1-3)-Fc (pl>9.3), it was asked whether the difference in pharmacokinetics could be attributed to the net charge of the protein, which made it stick to negatively charged extracellular matrix components, or whether there were perhaps specific locations on the surface of the Flt1 (1-3)-Fc protein that constituted specific binding sites for extracellular matrix components. For example, many proteins are known to have heparin binding sites, often consisting of a cluster of basic residues. Sometimes these residues are found in a cluster on the primary sequence of the protein; some of the literature has identified "consensus sequences" for such heparin binding sites (see for example Hileman, et al., 1998, Bloessays 20(2):156-67). In other cases, the known crystal structure of a protein reveals a cluster of positively charged residues on the surface of a protein, but the residues come from different regions of the primary sequence and are only brought together when the protein folds into its tertiary structure. Thus it is difficult to deduce whether an isolated amino acid residue forms part of a cluster of basic residues on the surface of the protein. However, if there is a cluster of positively charged amino acid residues in the primary sequence, it is not unreasonable to surmise that the residues are spatially close to one another and might therefore be part of an extracellular matrix component binding site. Flt1 receptor has been studied extensively and various domains have been described (see for example Tanaka et al., 1997, Jpn. J. Cancer Res 88:867-876). Referring to the nucleic acid and amino acid sequence set forth in FIGS. 10A-10D of this application, one can identify the signal sequence for secretion which is located at the beginning of the sequence and extends to the glycine coded for by nucleotides 76-78. The mature protein begins with Ser-Lys-Leu-Lys, starting at nucleotide 79 of the nucleic acid sequence. Flt1Ig domain 1 extends from nucleotide 79 to 393, ending with the amino acids Ser-Asp-Thr. Flt1 Ig domain 2 extends from nucleotide 394 to 687 (encoding Gly-Arg-Pro to Asn-Thr-lle), and Flt1 Ig domain 3 extends from nucleotides 688 to 996 (encoding lie-Asp-Val to Asp-Lys-Ala). There is a bridging amino acid sequence, Gly-Pro-Gly, encoded by nucleotides 997-1005, followed by the nucleotide sequence encoding human Fc (nucleotides 1006-1701 or amino acids Glu-Pro-Lys to Pro-

Gly-Lys-stop). A more detailed analysis of the Flt1 amino acid sequence reveals that there is a cluster, namely, amino acid residues 272-281 (KNKRASVRR) of FIGS. 10A-10D, in which 6 out of 10 amino acid residues are basic. This sequence is located in Flt1 Ig domain 3 of the receptor (see FIG. 11), which is not itself essential for binding of VEGF ligand, but which confers a higher affinity binding to ligand. An alignment of the sequence of Ig domain 3 with that of Ig domain 2 reveals that in this region, there is very poor alignment between the two Ig domains, and that there are about 10 additional amino acids in Ig domain 3. An analysis of the hydrophilicity profiles MACVECTOR™, a trademark of Accelrys relating to computer software) of these two domains clearly indicates the presence of a hydrophilic region in the protein (FIGS. 12A-12B). These observations raised the possibility that the actual three dimensional con-60 formation of Flt1 Ig domain 3 allowed for some type of protrusion that is not in Flt1 Ig domain 2. To test this hypothesis, the 10 additional amino acids were deleted and the resulting protein was tested to see whether the deletion would affect the pharmacokinetics favorably without seriously compromising the affinity of the receptor for VEGF. This DNA construct, which was constructed using standard

molecular biology techniques (see e.g., Molecular Cloning,

20

A Laboratory Manual (Sambrook, et al., Cold Spring Harbor Laboratory), Current Protocols in Molecular Biology (Eds. Ausubel, et al., Greene Publ. Assoc., Wiley-Interscience, NY) in the mammalian expression vector pMT21 (Genetics Institute, Inc., Cambridge, Mass.), is referred to as Mutl: 5 Flt1($1-3_{AB}$)-Fc. The Mutl: Flt1($1-3_{AB}$)-Fc construct was derived from Flt1(1-3)-Fc by deletion of nucleotides 814-843 (set forth in FIGS. 10A–10D), which deletes the highly basic 10-amino acid residue sequence Lys-Asn-Lys-Arg-Ala-Ser-Val-Arg-Arg-Arg from Flt1 Ig domain 3.

The final DNA construct was sequence-verified using an ABI 373A DNA sequencer and Taq Dideoxy Terminator Cycle Sequencing Kit (Applied Biosystems, Inc., Foster City, Calif.). The sequence of Mut1: $Flt1(1-3_{\Delta B})$ -Fc is set forth in FIGS. 13A–13D.

Example 12

Construction of Flt1(1-3)-Fc Basic Region Deletion Mutant Designated Mut2: Flt1(2-3_{AB})-Fc

A second deletion mutant construct, designated Mut2: Flt1(2-3_{AB})-Fc, was derived from the Mut1: Flt1(1-3_{AB})-Fc construct by deletion of Flt1 Ig domain 1 encoded by nucleotides 79-393 (see FIGS. 10A-10D); for convenience, 25 nucleotides 73-78 (TCA GGT) were changed to TCC GGA. This introduced a restriction site (BspE1) without altering the associated amino acid sequence, Ser-Glv. This DNA construct, which was constructed using standard molecular biology techniques (see e.g., Molecular Cloning, A Labora- 30 tory Manual (Sambrook, et al., Cold Spring Harbor Laboratory), Current Protocols in Molecular Biology (Eds. Ausubel, et al., Greene Publ. Assoc., Wiley-Interscience, NY) in the mammalian expression vector pMT21 (Genetics Institute, Inc., Cambridge, Mass.), was also sequence-veri- 35 fied using an ABI 373A DNA sequencer and Taq Dideoxy Terminator Cycle Sequencing Kit (Applied Biosystems, Inc., Foster City, Calif.). The sequence of Mut2: Flt1(2-3AB)-Fc is set forth in FIGS. 14A-14C.

Example 13

Construction of Flt1(1-3)-Fc Deletion Mutant Designated Mut3: Flt1(2-3)-Fc.

A third deletion mutate construct, designated Mut3: Flt1 (2-3)-Fc, was constructed the same way as the Mut2: Flt1 ($2-3_{\Delta B}$)-Fc construct, except that Flt1 Ig domain 3 was left intact (the basic region amino acids were not deleted). The construct was constructed using standard molecular biology 50 techniques and the final construct was sequence-verified as described supra. The sequence of Mut3: Flt1(2-3)-Fc is set forth in FIGS. **15**A–**15**C.

Example 14

Construction of Flt(1-3)-Fc Basic Region N-Glycosylation Mutant Designated Mut4: Flt1 $(1-3_{R-2N})$ -Fc.

A final construct was made in which a N-glycosylation 60 site was introduced into the middle of the basic region of Flt1 Ig domain 3. This construct was designated Mut4: Flt1($1-3_{R->N}$)-Fc and was made by changing nucleotides 824-825 from GA to AC, consequently changing the coded Arg residue (AGA) into an Asn residue (MC) (see FIGS. 65 **10**A–**10**D). The resulting amino acid sequence is therefore changed from Arg-Ala-Ser to Asn-Ala-Ser, which matches

the canonical signal (Asn-Xxx-Ser/Thr) for the addition of a N-glycosylation site at the Asn residue. The sequence of Mut4: $Flt1(1-3_{R-N})$ -Fc is set forth in FIGS. **16**A–**16**D.

Example 15

Characterization of Acetylated Flt1(1-3)-Fc. Mut1: Flt1(1-3_{$\Delta\beta$})-Fc, and Mut4: Flt1(1-3_{R->N})-FcMutants.</sub>

(a.) Binding to Extracellular Matrix Components

To determine whether the three modified proteins were more or less likely to have improved pharmacokinetic properties, Matrigel coated 96-well dishes (as described supra) were incubated with varying concentrations of the mutant proteins and detected with anti-human Fc/alkalinephosphatase conjugated antibodies. As shown in FIG. 18, this experiment showed that while the unmodified Flt1(1-3)-Fc protein could bind avidly to these wells, the Mut3: Flt1(2-3)-Fc protein bound somewhat more weakly, the Mut1: Flt1(1-3_{AB})-Fc protein bound more weakly still, and the Mut2: Flt1(2-3 $_{\Delta B}$)-FC protein showed the best profile, binding more weakly than any of the other mutant proteins. The Mut4: Flt1 (1-3_{R->N})-Fc glycosylation mutant protein showed only marginal benefit on the Matrigel assay. These results confirm the hypothesis that a linear sequence of positive amino acids can be deleted from the primary sequence resulting in a decrease in charge interaction with extracellular matrix components.

(b.) Binding of Mut1: Flt1 $(1-3_{\Delta B})$ -Fc and Mut4: Flt1 $(1-3_{R-N})$ -Fc in a Biacore-Based Assay

Unmodified and acetylated Flt1(1-3)-Fc and genetically modified Mut1: Flt1(1-3 $_{\Delta B}$)-Fc and Mut4: Flt1(1-3 $_{R->N}$)-Fc proteins where tested in a Biacore-based assay to evaluate their ability to bind to the Flt1 ligand, VEGF. In this assay, unmodified Flt1(1-3)-Fc protein (0.25, 0.5, or 1.0 µg/ml) was immobilized on the surface of a Biacore chip (see Biacore Instruction Manual, Pharmacia, Inc., Piscataway, 40 N.J., for standard procedures) and a solution containing 0.1 µg/ml VEGF and either purified or COS cell supernatant containing unmodified Flt1(1-3)-Fc (at approximately (0.25, 0.5, or 1.0 µg/ml), purified acetylated Flt1(1-3)-Fc (at (0.25, 0.5, or 1.0 µg/ml), COS cell supernatant containing Mut1: Flt1(1-3_{ΔB})-Fc (at approximately (0.25, 0.5, or 1.0 µg/ml), or COS cell supernatant containing Mut4: Flt1(1-3_{R->N})-Fc (at approximately (0.25, 0.5, or 1.0 µg/ml) were passed over the Flt1(1-3)-Fc-coated chip. As shown in FIG. 17, at the sub-stoichiometric ratio (0.25 µg/ml Flt1 (1-3)-Fc of unmodified, acetylated or genetically modified samples vs. 01. µg/ml VEGF), there is insufficient Flt1(1-3)-Fc protein to block binding of VEGF to the Flt1(1-3)-Fc immobilized on the Biacore chip. At 0.5 µg/ml of unmodified, acetylated or genetically modified Flt1(1-3)-Fc proteins, the stoichiomet-55 ric ratio approximates 1:1 and there is an increased ability to block VEGF binding to the Biacore chip. At 1.0 ug/ml of unmodified, acetylated or genetically modified Flt1(1-3)-Fc proteins, which is approximately a 10:1 stoichiometric ratio, the Flt1(1-3)-Fc proteins are able to block binding of VEGF to the Biacore chip, but they are not equivalent. Unmodified, acetylated, and Mut1: Flt1(1-3_{AB})-Fc are essentially equal in their ability to block VEGF binding, whereas Mut4: Flt1(1- $3_{R_{r} > N}$)-Fc is somewhat less efficient at blocking binding. These results confirm the hypothesis that it is possible to reduce the non-specific binding of a positively charged molecule by genetically removing a linear sequence of predominantly negatively charged amino acids.

(c.) Binding of Mut1: Flt1(1-3 $_{\Delta B}$)-Fc. Mut2: Flt1(2-3 $_{\Delta B}$)-Fc, Mut3: Flt1(2-3)-Fc, and in an ELISA-Based Assay

To determine whether the three mutant proteins could bind the Flt1 ligand VEGF, binding experiments were done in which 96-well plates coated with VEGF were incubated 5 with varying concentrations of the respective mutant protein, and after washing, the amount bound was detected by incubating with an alkaline phosphatase conjugated antihuman Fc antibody and quantitated calorimetrically by the addition of an appropriate alkaline phosphatase substrate. As 10 shown in FIG. 19, this experiment showed that all the mutant proteins could bind VEGF similarly, at the concentrations tested.

Example 16

Pharmacokinetic Analysis of Acetylated Flt1(1-3)-Fc, Mut1: Flt1(1- 3_{AB})-Fc, and unmodified Flt1(1-3)-Fc

In vivo experiments were designed to assess the pharmacokinetic profiles of unmodified Flt1(1-3)-Fc, Mut1: Flt1(1- $3_{\Delta B}$)-Fc, and 40 fold molar excess acetylated Flt1(1-3)-Fc protein. Balb/c mice (25-30 g) were injected subcutaneously with 4 mg/kg of unmodified Flt1(1-3)-Fc, 40 fold molar 25 excess acetylated Flt1(1-3)-Fc, and Mut1: Flt1(1-3AB)-Fc proteins (4 mice each). These mice were tail bled at 1, 2, 4, 6, 24 hours, 2 days, 3 days, and 5 days after injection. The sera were assayed in an ELISA designed to detect Flt1(1-3)-Fc protein which involves coating an ELISA plate with 30 VEGF, binding the Flt1(1-3)-Fc and reporting with an anti-Fc antibody linked to alkaline phosphatase. As shown in FIG. 20, the Cmax for these reagents was as follows: Unmodified Flt1(1-3)-Fc-0.15 µg/ml; 40 fold molar excess acetylated Flt1(1-3)-Fc-1.5 µg/ml; and Mut1: Flt1(1-3AB)- 35 Fc-0.7 µg/ml.

Example 17

Modified Fit1 Receptor Vector Construction

The rationale for constructing modified versions of the Flt1 receptor (also known as VEGFR1) was based on the observation that the protein sequence of Flt1 was highly basic, and was therefore likely to stick to extracellular 45 domain 2 fused directly to the beginning of Flk1 Ig domain matrix (ECM). The highly basic nature of Flt1 probably explains why unmodified Flt1 (1-3)-Fc (described supra) has poor pharmacokinetics that make it difficult to use as a therapeutic agent. As described supra, the chemically modified form of 40 fold molar excess acetylated Flt1 (1-3)-Fc, 50 hereinafter termed A40, exhibited a greatly improved pharmacokinetic (PK) profile over the non-acetylated Flt1(1-3)-Fc. Therefore, attempts were made to engineer DNA molecules that could be used to recombinantly express modified forms of a Flt1 receptor molecule that would possess the 55 improved PK profile exhibited by A40 and still maintain the ability to bind tightly to VEGF.

It is known in the literature that the first Ig domain of Flt1 (which has a net charge of +5 at neutral pH) is not essential for tight binding to VEGF, so this domain was deleted. The 60 third Ig domain (having a net charge of +11) is not essential for binding, but confers higher affinity for VEGF than the second Ig domain, so instead of deleting it entirely, it was replaced with the equivalent domains of the Flt1 receptor relatives Flk1 (also known as VEGFR2) and Flt4 (also 65 known as VEGFR3). These chimeric molecules (denoted R1R2 (Flt.D2.Flk1D3.Fc\DeltaC1(a) and VEGFR1 R2-Fc\DeltaC1

(a) and R1 R3 (Flt1 D2.VEGFR3D3-Fc∆C1 (a) and VEGFR1R3-Fc∆C1(a) respectively, wherein R1 and Flt1D2=Ig domain 2 of Flt1 (VEGFR1); R2 and Flk1D3=Ig domain 3 of Flk1 (VEGFR2); and R3 and VEGFR3D3=Ig domain 3 of Flt4 (VEGFR3)) were much less sticky to ECM, as judged by an in vitro ECM binding assay as described infra, had greatly improved PK as described infra. In addition, these molecules were able to bind VEGF tightly as described infra and block phosphorylation of the native Flk1 receptor expressed in endothelial cells as described infra.

(a) Construction of the Expression Plasmid pFlt1D2, Flk1D3,Fc∆C1(a)

Expression plasmids pMT21.Flt1(1-3).Fc (6519 bp) and 15 pMT21.Flk-1(1-3).Fc (5230 bp) are plasmids that encode ampicillin resistance and Fc-tagged versions of Ig domains 1-3 of human Fit1 and human Flk1, respectively. These plasmids were used to construct a DNA fragment consisting of a fusion of 1 g domain 2 of Flt1 with Ig domain 3 of Flk1, using PCR amplification of the respective Ig domains fol-20 lowed by further rounds of PCR to achieve fusion of the two domains into a single fragment. For Ig domain 2 of Flt1, the 5' and 3' amplification primers were as follows:

5': bsp/flt1D2 (5'-GACTAGCAGTCCGGAGGTAGAC-CTTTCGTAGAGATG-3')

Flt1D2-Flk1D3.as (5'-CGGACTCAGAACCACATC-31: TATGATTGTATTGGT-3')

The 5' amplification primer encodes a BspE1 restriction enzyme site upstream of Ig domain 2 of Fit1, defined by the amino acid sequence GRPFVEM (corresponding to amino acids 27-33 of FIGS. 21A-21C). The 3' primer encodes the reverse complement of the 3' end of Flt1 Ig domain 2 fused directly to the 5' beginning of Flk1 Ig domain 3, with the fusion point defined as TIID of Fit1 (corresponding to amino acids 123-126 of FIGS. 21A-21C) and continuing into VVLS (corresponding to amino acids 127-130 of FIGS. 21A-21C) of Fik1.

For Ig domain 3 of Flk1, the 5' and 3' amplification primers were as follows:

5': Flt1D2-Flk1D3.s (5'-ACMTCATAGATGTGGTTCT-GAGTCCGTCTCATG G-3')

3': Flk1D3/apa/srf.as (5'-GATMTGCCCGGGC-CCTTTTCATGGACCCTGAC AAATG-3')

The 5' amplification primer encodes the end of Flt1 Ig 3, as described above. The 3' amplification primer encodes the end of Flk1 Ig domain 3, defined by the amino acids VRVHEK (corresponding to amino acids 223-228 of FIGS. 21A-21C), followed by a bridging sequence that includes a recognition sequence for the restriction enzyme Srf1, and encodes the amino acids GPG. The bridging sequence corresponds to amino acids 229-231 of FIGS. 21A-21C.

After a round of PCR amplification to produce the individual domains, the products were combined in a tube and subjected to a further round of PCR with the primers bsp/flt1D2 and Flk1 D3/apa/srf.as (described supra) to produce the fusion product. This PCR product was subsequently digested with the restriction enzymes BspEI and SmaI and the resulting 614 bp fragment was subcloned into the BspEI to SrfI restriction sites of the vector pMT21/ΔB2.Fc, to create the plasmid pMT21/Flt1D2.Flk1D3.Fc. The nucleotide sequence of the Flt1D2-Flk1D3 gene fusion insert was verified by standard sequence analysis. This plasmid was then digested with the restriction enzymes EcoRI and SrfI and the resulting 702 bp fragment was transferred into the EcoRI to SrfI restriction sites of the plasmid pFlt1(1-3)B2- $Fc\Delta C1(a)$ to produce the plasmid pFlt1D2.Flk1D3.Fc $\Delta C1$

(a). The complete DNA and deduced amino acid sequences of the Flt1D2.Flk1D3.Fc Δ C1(a) chimeric molecule is set forth in FIGS. **21**A–**21**C.

(b) Construction of the Expression Plasmid $pFlt1D2VEGFR3D3Fc\DeltaC1(a)$

The expression plasmid pMT21.Flt1(1-3).Fc (6519 bp) encodes ampicillin resistance and an Fc-tagged version of Ig domains 1-3 of human Fit1 receptor. This plasmid was used to produce a DNA fragment containing Ig domain 2 of Flt1 by PCR. RNA from the cell line HEL921.7 was used to produce Ig domain 3 of Flk1, using standard RT-PCR methodology. A further round of PCR amplification was used to achieve fusion of the two Ig domains into a single fused fragment. For Ig domain 2 of Flt1, the 5' and 3' amplification primers were as follows: 15

5': bsp/ftt1D2 (5'-GACTAGCAGTCCGGAGGTAGAC-CTTTCGTAGAGATG-3')

3': Flt1D2.VEGFR3D3.as(TTCCTGGGCAACAGCTG-GATATCTATGATTGTATTGGT)

The 5' amplification primer encodes a BspE1 restriction site upstream of Ig domain 2 of Flt1, defined by the amino acid sequence GRPFVEM (corresponding to amino acids 27-33 of FIGS. **22**A–**22**C). The 3' amplification primer encodes the reverse complement of the end of Flt1 Ig domain 2 fused directly to the beginning of VEGFR3 Ig domain 3, with the fusion point defined as TIID of Fit1 (corresponding to amino acids 123-126 of FIGS. **22**A–**22**C) and continuing into IQLL of VEGFR3 (corresponding to amino acids 127-130 of FIGS. **22**A–**22**C). For Ig domain 3 of VEGFR3, the 5' and 3' primers used for RT-PCR were as follows:

5': R3D3.s (ATCCAGCTG1TGCCCAGGAAGTCGCTGG-AGCTGCTGGTA)

3': R3D3.as (ATTTTCATGCACAATGACCTCGGT- 35 GCTCTCCCGAAATCG)

Both the 5' and 3' amplification primers match the sequence of VEGFR3. The 296 bp amplification product of this RT-PCR reaction was isolated by standard techniques and subjected to a second round of PCR to add suitable 40 sequences to allow for fusion of the Flt1D2 with the Flk1D3 domains and fusion of the Flk1D3 and Fc domains via a GPG bridge (see below). The amplification primers were as follows:

5':Flt1 D2.VEGFR3D3.s (TCATAGATATCCAGCT- 4 GTrGCCCAGGMGTCGCTGGAG)

3': VEGFR3D3/srf.as (GATMTGCCCGGGCCATTTTCAT-GCACMTGACCTCGGT)

The 5' amplification primer encodes the 3' end of Flt1 Ig domain 2 fused directly to the beginning (5' end) of 50 VEGFR3 Ig domain 3, as described above. The 3' amplification primer encodes the 3' end of VEGFR3 Ig domain 3, defined by the amino acids VIVHEN (corresponding to amino acids 221-226 of FIGS. **22**A–**22**C), followed by a bridging sequence that includes a recognition sequence for 55 Srf1, and encodes the amino acids 227-229 of FIGS. **22**A–**22**C.

After one round (for Flt1 Ig domain 2) or two rounds (for Flt4 Ig domain 3) of PCR to produce the individual Ig 60 domains, the PCR products were combined in a tube and subjected to a further round of PCR amplification with the amplification primers bsp/flt1D2 and VEGFR3D3/srf.as described supra, to produce the fusion product. This PCR product was subsequently digested with the restriction 65 enzymes BspEI and Smal and the resulting 625 bp fragment was subcloned into the BspEI to SrfI restriction sites of the

vector pMT21/Flt1AB2.Fc (described supra), to create the plasmid pMT21/Flt1D2.VEGFR3D3.Fc. The sequence of the Flt1D2-VEGFR3D3 gene fusion insert was verified by standard sequence analysis. This plasmid was then digested with the restriction enzymes EcoRI and SrfI and the resulting 693 bp fragment was subcloned into the EcoRI to SrfI restriction sites of the plasmid pFlt1(1-3) Δ B2-FC Δ C1 (a) to produce the plasmid designated pFlt1D2.VEGFR3D3.FcΔC1(a). The complete DNA deduced acid sequence of the amino Flt1D2.VEGFR3D3.Fc\DeltaC1(a) chimeric molecule is set forth in FIGS. 22A-22C.

Example 18

Extracellular Matrix Binding (ECM) Binding Assay

ECM-coated plates (Becton Dickinson catalog # 35-4607) were rehydrated with warm DME supplemented with glutamine (2 mM), 100U penicillin, 100U streptomycin, and 10% BCS for at least 1 hr. before adding samples. The plates were then incubated for 1 hr. at room temperature with varying concentrations of Flt1D2.Flk1 D3.Fc\DeltaC1 (a) and Flt1D2.VEGFR3D3.Fc\DeltaC1(a) starting at 10 nM with subsequent 2-fold dilutions in PBS plus 10% BCS. The plates were then washed 3 times with PBS plus 0.1% Triton-X and incubated with alkaline phosphatase-conjugated anti-human Fc antibody (Promega, 1:4000 in PBS plus 10% BCS) for 1 hr. at room temperature. The plates were then washed 4 times with PBS 0.1% Triton-X and alkaline phosphatase buffer/pNPP solution (Sigma) was added for color development. Plates were read at I=405-570 nm. The results of this experiment are shown in FIG. 23 and demonstrate that the Flt1D2.Flk1 D3.FcΔC1(a) and Flt1D2.VEGFR3D3.FcΔC1 (a) proteins are considerably less sticky to the ECM as compared to the Flt1(1-3)-Fc protein.

Example 19

Transient Expression of pFlt1D2.Flk1D3.Fc∆C1(a) in CHO-K1 (E1A) Cells

nplification primers were as (TCATAGATATCCAGCT-GAG) ACCCGGGCCATTTTCATincodes the 3' end of FIt1 Ig the beginning (5' end) of 5¹⁰ tig domain 3, TVHEN (corresponding to **22A-22C**), followed by a a recognition sequence for 5¹⁰ A large scale (2L) culture of *E. coli* DH10B cells carrying the pFlt1D2.Flk1D3.Fc Δ C1(a) plasmid described supra in Example 17(a) was grown overnight in Terrific Broth (TB) plus 100 µg/ml ampicillin. The next day, the plasmid DNA was extracted using a QIAgen ENDOFREETM, a trademark of Quigen relating to a Megaprep kit following the manufacturer's protocol. The concentration of the purified plasmid DNA was determined by standard techniques using a UV spectrophotometer and fluorometer. The plasmid DNA was verified by standard restriction enzyme digest fragments corresponded to the predicted sizes when analyzed on a 1% agarose gel.

Forty 15 cm petri plates were seeded with CHO-K1/E1A cells at a density of 4×10^6 cells/plate. Plating media was Gibco Ham's F-12 supplemented with 10% HYCLONETM (a trademark of Hyclone Laboratories relating to chemical products for scientific research, medical research, and the production of pharmaceuticals) Fetal Bovine Serum (FBS), 100U penicillin/100U streptomycin and glutamine (2 mM). The following day each plate of cells was transfected with 6 µg of the pFlt1D2.Flk1D3.Fc Δ C1(a) plasmid DNA using Gibco Optimem and Gibco Lipofectamine in 12 ml volume, following the manufacturer's protocol. Four hours after adding the transfection mix to the cells, 12 ml/plate of

60

Optimem supplemented with 10% FBS was added. Plates were incubated at 37° C. in a 5% CO_2 incubator overnight. The following day the media was removed from each plate and 25 ml expression media (Gibco CHO-S-SFM II supplemented with glutamine (2 mM) and 1 mM sodium butyrate) 5 was added. The plates were incubated at 37° C. for 3 days. After 3 days of incubation, the media was aspirated from each plate and centrifuged at 400 rpm in a swinging bucket rotor to pellet cells. The supernatant was decanted into sterile 1 L bottles and purification of the expressed protein 10 was performed as described infra.

Example 20

Construction pVEGFR1 R2-Fc∆C1(a) Expression Vector

The pVEGFR1R2.Fc Δ C1(a) expression plasmid was constructed by insertion of DNA encoding amino acids SDT (corresponding to amino acids 27-29 of FIGS. **24A**-**24**C) ²⁰ between Flt1d2-Flk1 d3-Fc Δ C1 (a) amino acids 26 and 27 of FIGS. **21A**-**21**C (GG) and removal of DNA encoding amino acids GPG corresponding to amino acids 229-231 of Figure. The SDT amino acid sequence is native to the Flt1 receptor and was added back in to decrease the likelihood of heterogeneous N-terminal processing. The GPG (bridging sequence) was removed so that the Flt1 and Flk1 Ig domains were fused directly to one another. The complete DNA and deduced amino acid sequences of the pVEGFR1R2.Fc Δ C1 (a) chimeric molecule is set forth in FIGS. **24A**-**24**C.

Example 21

Cell Culture Process Used to Produce Modified Flt1 Receptors

(a) Cell Culture Process Used to Produce $Flt1D2.Flk1D3.Fc\Delta C1(a)$

The process for production of Flt1D2.Flk1D3.Fc Δ C1(a) protein using the expression plasmid ₄₀ pFlt1D2.Flk1D3.Fc Δ C1(a) described supra in Example 1 involves suspension culture of recombinant Chinese hamster ovary (CHO K1/E1A) cells which constitutively express the protein product. The cells are grown in bioreactors and the protein product is isolated and purified by affinity and size exclusion chromatography. The process is provided in greater detail below.

Cell Expansion

Two confluent T-225 cm² flasks containing the Flt1D2.Flk1D3.Fc Δ C1(a) expressing cell line were ⁵⁰ expanded by passaging cells into eight T-225 cm² flasks in medium (GMEM+10% serum, GIBCO) and incubated at 37° C. and 5% CO₂. When the flasks approached confluence (approximately 3 to 4 days) the cells were detached using trypsin. Fresh medium was added to protect the cells from ⁵⁵ further exposure to the trypsin. The cells were centrifuged and resuspended in fresh medium then transferred to eight 850 cm² roller bottles and incubated at 37° C. and 5% CO₂ until confluent.

Suspension Culture in Bioreactors

Cells grown in roller bottles were trypsinized to detach them from the surface and washed with suspension culture medium. The cells are aseptically transferred to a 5L bioreactor (New Brunswick Celligen Plus) where the cells are grown in 3.5L of suspension culture. The suspension culture ⁶⁵ medium was a glutamine-free low glucose modification of IS-CHO (Irvine Scientific) to which 5% fetal bovine serum

(Hyclone), GS supplement (Life Technologies) and 25 μ M methionine sulfoximine (Sigma) was added. The pH was controlled at 7.2 by addition of carbon dioxide to the inlet gas or by addition of a liquid solution of sodium carbonate to the bioreactor. Dissolved oxygen level was maintained at 30% of saturation by addition of oxygen or nitrogen to the inlet gas and temperature controlled at 37° C. When a density of 4×10⁶ cells/mL was reached the cells were transferred to a 40L bioreactor containing the same medium and setpoints for controlling the bioreactor. The temperature setpoint was reduced to 34° C. to slow cell growth and increase the relative rate of protein expression.

(b) Cell Culture Process Used to Produce $Flt1D2.VEGFR3D3.Fc\DeltaC1(a)$

The same methodologies as described supra for Flt1D2.Flk1D3. $Fc\Delta C1(a)$ were used to produce Flt1D2.VEGFR3D3. $Fc\Delta C1(a)$.

Example 22

Harvest and Purification of Modified Flt1 Receptors

(a) Harvest and Purification of Flt1D2.Flk1D3.FcΔC1(a)

The product protein was aseptically harvested from the bioreactor while retaining cells using Millipore Prostak tangential-flow filtration modules and a low-shear mechanical pump (Fristam). Fresh medium was added to the bioreactor to replace that removed during the harvest filtration. Approximately 40L of harvest filtrate was then loaded onto a 400 mL column containing Protein A SEPHAROSE™ (a trademark of Amersham Biosciences relating to signal transduction and cell trafficking) resin. After loading the resin was washed with buffer containing 10 mM sodium phosphate, 500 mM sodium chloride, pH 7.2 to remove any unbound contaminating proteins. Flt1D2.Flk1D3.Fc\DeltaC1(a) 35 protein was eluted with a pH 3.0 citrate buffer. The eluted protein was neutralized by addition of Tris base and frozen at -20° C.

Several frozen lots of Flt1D2.Flk1D3.Fc Δ C1 (a) protein from the Protein A step above were thawed, pooled and concentrated using a Millipore 30 kD nominal molecular weight cutoff (NMWCO) tangential flow filtration membrane. The protein was transferred to a stirred cell concentrator (Millipore) and further concentrated to 30 mg/mL using a 30 kD NMWCO membrane. The concentrated protein was loaded onto a size exclusion column packed with Superdex 200 resin (Amersham Pharmacia) that was equilibrated with phosphate buffered saline plus 5% glycerol. The same buffer was used to run the column. The fractions corresponding to Flt1D2.Flk1D3.Fc Δ C1(a) dimer were pooled, sterile filtered through a 0.22 micron filter, aliquoted and frozen.

(b) Harvest and Purification of Flt1D2.VEGFR3D3.FcΔC1(a)

The same methodologies as described supra for Flt1D2.Flk1D3.Fc Δ C1(a) were used to harvest and purify Flt1D2.VEGFR3D3.Fc Δ C1(a).

Example 23

Phosphorylation Assay for Transiently Expressed VEGFR2

Primary human umbilical vein endothelial cells (HU-VECs), passage 4-6, were starved for 2 hrs in serum-free DME high glucose media. Samples containing 40 ng/ml (1 nM) human VEGF165, which is a ligand for the VEGF receptors Flt1, Flk1 and Flt4(VEGFR3) were prepared and
were preincubated for 1 hr. at room temperature with varying amounts of the modified Flt1 receptors Flt1(1-3)-Fc, Flt1(1-3)-Fc (A40), Flt1 D2Flk1 D3.Fc\DeltaC1 (a) and Flt1 D2VEGFR3D3.Fc\DeltaC1 (a) in serum-free DME-high glucose media containing 0.1% BSA. Cells were challenged for 5 minutes with the samples prepared above +/-VEGF165, followed by whole cell lysis using complete lysis buffer. Cell lysates were immunoprecipitated with an antibody directed against the C-terminus of VEGFR2 receptor. The immunoprecipitated lysates were loaded onto 4-12% SDS-PAGE 10 Novex gel and then transferred to PVDF membrane using standard transfer methodologies. Detection of phosphorylated VEGFR2 was done by immunoblotting with the antiphospho Tyrosine mAb called 4G10 (UBI) and developed using ECL-reagent (Amersham). FIGS. 25A-25C and 26A-26B show the results of this experiment. FIGS. 25A-25C reveals that detection by Western blot of tyrosine phosphorylated VEGFR2(Flk1) by VEGF165 ligand stimulation shows that cell-surface receptors are phosphorylated to varying levels depending on which modified FIt1 receptor 20 is used during the preincubations with VEGF. As is seen in FIG. 25A, at a 1.5 molar excess of either Flt1(1-3)-Fc, Flt1(1-3)-Fc (A40) or transient FltlD2Flk1D3.FcΔC1(a) there is complete blockage of receptor stimulation by these three modified Flt1 receptors as compared to control media 25 challenge. In contrast, transient Filt1D2VEGFR3D3.Fc∆C1 (a) does not show significant blockage at this molar excess, as compared with VEGF positive control challenge. Similar results are seen in FIG. 25B, where the modified Flt receptors are in a 3-fold molar excess to VEGF165 ligand. In FIG. 30 25C, where the modified Flt1 receptors are in a 6-fold molar VEGF165 ligand, transient excess to Flt1D2VEGFR3D3.Fc∆C1(a) can now be shown to be partially blocking VEGF165-induced stimulation of cell-surface receptors. 35

In FIGS. 26A-26B, detection by Western blot of tyrosine phosphorylated VEGFR2(Flk1) by VEGF165 ligand stimulation shows that cell-surface receptors are not phosphorylated by challenge samples which have VEGF165 preincubated with 1 and 2 fold molar excess (FIG. 26A) or 3 and 4 40 fold molar excess (FIG. 26B) of either transient Flt1D2Flk1 D3.Fc\DeltaC1 (a), stable Flt1 D2Flk1 D3.Fc\DeltaC1 (a), or transient VEGFR1R2-Fc∆C1(a). At all modified Flt1 receptor concentrations tested there is complete binding of VEGF165 ligand during the preincubation, resulting in no detectable 45 stimulation of cell-surface receptors by unbound VEGF165 as compared to control media challenge.

Example 24

Cell Proliferation Bioassay

The test cell population is MG87 cells that have been stably transfected with a expression plasmid that contains a DNA insert encoding the VEGFR2(Flk1) extracellular 55 domain fused to the TrkB intracellular kinase domain, thus producing a chimeric molecule. The reason the TrkB intracellular kinase domain was used rather than the native VEGFR2(Flk1) intracellular kinase domain is that the intracellular kinase domain of VEGFR2(Flk1) does not cause a 60 strong proliferative response when stimulated by VEGF165 in these cells. It is known that MG87 cells containing full length TrkB receptor give a robust proliferative response when stimulated with BDNF, so the TrkB intracellular kinase domain was engineered to replace the intracellular 65 kinase domain of VEGFR2(Flk1) to take advantage of this proliferative response capability.

5×103 cells/well were plated in a 96 well plate and allowed to settle for 2 hrs at 37° C. The following modified Flt receptors Flt1(1-3)-Fc, Flt1 D2.Flk1 D3.Fc\DeltaC1 (a) and Flt1 D2.VEGFR3D3.Fc\DeltaC1 (a), plus an irrelevant receptor termed Tie2-Fc as a negative control, were titrated from 40 nM to 20 pM and incubated on the cells for 1 hr at 37° C. Human recombinant VEGF165 in defined media was then added to all the wells at a concentration of 1.56 nM. The plates were incubated for 72 hrs at 37° C. and then MTS (Owen's reagent, Promega) added and the plates were incubated for an additional for 4 hrs. Finally, the plates were read on a spectrophotometer at 450/570 nm. The results of this experiment are shown in FIG. 27. The control receptor Tie2-Fc does not block VEGF165-induced cell proliferation at any concentration whereas Flt1D2.Flk1D3.FcΔC1(a) blocks 1.56 nM VEGF165 with a half maximal dose of 0.8 nM. Flt1(1-3)-Fc and Flt1D2.VEGFR3D3.Fc\DeltaC1(a) are less effective in blocking VEGF165 in this assay with a half maximal dose of ~2 nM. VEGF165 alone gives a reading of 1.2 absorbance units and the background is 0.38 absorbance units.

Example 25

Binding Stoichiometry of Modified Flt Receptors to VEGF165

(a) BIAcore Analysis

US 7.070.959 B1

The stoichiometry of Flt1 D2Flk1D3.Fc∆C1 (a) and VEGFR1 R2-Fc∆C1 (a) interaction with human VEGF165 was determined by measuring either the level of VEGF saturation binding to the Flt1D2Flk1D3.Fc\DeltaC1(a) or VEGFR1R2-Fc∆C1(a) surfaces or measuring concentration of VEGF165 needed to completely prevent binding of Flt1D2Flk1D3.Fc Δ C1(a) or VEGFR1R2-Fc Δ C1(a) to VEGF BIAcore chip surface.

Modified Flt receptors Flt1D2Flk1D3.Fc\DeltaC1(a) and VEGFR1R2-Fc∆C1(a), were captured with an anti-Fc specific antibody that was first immobilized on a Biacore chip (BIACORE) using amine-coupling chemistry. A blank antibody surface was used as a negative control. VEGF165 was injected at a concentration of 1 nM, 10 nM, and 50 nM over the Flt1D2Flk1D3.Fc Δ C1(a) and VEGFR1R2-Fc Δ C1(a) surfaces at 10 µl/min for one hour. A real-time binding signal was recorded and saturation binding was achieved at the end of each injection. Binding stoichiometry was calculated as a molar ratio of bound VEGF165 to the immobilized Flt1D2Flk1D3.Fc Δ C1 (a) or VEGFR1R2-Fc Δ C1 (a), using 50 the conversion factor of 1000 RU equivalent to 1 ng/ml. The results indicated binding stoichiometry of one VEGF165 dimeric molecule per one Flt1D2Flk1 D3.Fc\DeltaC1 (a) or VEGFR1 R2-Fc∆C1 (a) molecule (FIG. 28).

In solution, Flt1D2Flk1D3.FcAC1(a) or VEGFR1R2-Fc∆C1(a) at a concentration of 1 nM (estimated to be 1000 times higher than the KD of the Flt1 D2Flk1 D3.Fc Δ C1 (a) or VEGFR1 R2-Fc\DeltaC1 (a)/VEGF165 interaction) were mixed with varied concentrations of VEGF165. After one hour incubation, concentrations of the free Flt1D2Flk1D3.Fc Δ C1(a) in solution were measured as a binding signal to an amine-coupled VEGF165 surface. A calibration was used curve to convert the Flt1D2Flk1D3.FcΔC1(a) BIAcore binding signal to its molar concentration. The data showed that the addition of 1 nM VEGF165 into the Flt1D2Flk1 D3.Fc\DeltaC1 (a) solution completely blocked Flt1D2Flk1D3.Fc\DeltaC1(a) binding to the VEGF165 surface. This result suggested the binding sto-

ichiometry of one VEGF165 molecule per one Flt1D2Flk1D3.Fc Δ C1(a) molecule (FIG. **29** and FIG. **30**). When the concentration of Flt1D2Flk1D3.Fc Δ C1(a) was plotted as a function of added concentration of VEGF165, the slope of the linear portion was—1.06 for 5 Flt1D2Flk1D3.Fc Δ C1 (a) and—1.07 for VEGFR1 R2-Fc Δ C1 (a). The magnitude of the slope, very close to negative one, was indicative that one molecule of VEGF165 bound to one molecule of either Flt1D2Flk1D3.Fc Δ C1 (a) or VEGFR1 R2-Fc Δ C1 (a).

(b) Size Exclusion Chromatography

Flt1D2Flk1D3.Fc∆C1 (a) was mixed with a 3-fold excess of VEGF165 and the receptor-ligand complex was purified using a Pharmacia Superose 6 size exclusion chromatography column. The receptor-ligand complex was then incubated in a buffer containing 6M guanidine hydrochloride in order to dissociate it into its component proteins. Flt1D2Flk1D3.Fc∆C1(a) was separated from VEGF165 using Superose 6 size exclusion chromatography column run in 6M guanidium chloride. In order to determine complex stoichiometry, several injections of Flt1D2Flk1D3.FcΔC1 (a) and VEGF165 were made and peak height or peak integrated intensity was plotted as a function of the concentration of injected protein. The calibration was done under 25 condition identical to one used in separating components of Flt1D2Flk1D3.Fc\DeltaC1 (a)/VEGF complex. Quantification of the Flt1D2Flk1D3.Fc∆C1(a)/VEGF complex composition was based on the calibration curves. The results of this experiment are set forth in FIG. 28, which shows the ratio of VEGF165 to Flt1D2Flk1D3.Fc∆C1(a) in a complex to be 1:1.

Example 26

Determination of the Binding Stoichiometry of Flt1D2Flk1D3.Fc∆C1(a)/VEGF165 Complex by Size Exclusion Chromatography

Flt1D2Flk1 D3.FcΔC1 (a)/VEGF165 Complex Preparation 40 VEGF165 (concentration=3.61 mg/ml) was mixed with CHO cell transiently expressed Flt1 D2.Flk1 D3.FcΔC1 (a) (concentration=0.9 mg/ml) in molar ratio of 3:1 (VEGF165: Flt1D2.Flk1D3.FcΔC1(a)) and incubated overnight at 4° C.

(a) Size Exclusion Chromatography (SEC) Under Native $^{\rm 45}$ Conditions

To separate the complex from excess of unbound VEGF165, 50 μ l of the complex was loaded on a Pharmacia Superose 12 PC 3.2/30 which was equilibrated in PBS buffer. The sample was eluted with the same buffer at flow ⁵⁰ rate 40 μ l/min. at room temperature. The results of this SEC are shown in FIG. **31**. Peak #1 represents the complex and peak #2 represents unbound VEGF165. Fractions eluted between 1.1 and 1.2 ml were combined and guanidinium hydrochloride (GuHCI)was added to a final concentration ⁵⁵ 4.5M to dissociate the complex.

(b) Size Exclusion Chromatography (SEC) Under Dissociative Conditions

To separate the components of the receptor-ligand complex and to determine their molar ratio, 50 μ l of dissociated complex as described supra was loaded onto a Superose 12 PC 3.2/30 equilibrated in 6M GuHCl and eluted with the same solution at a flow rate 40 μ l/min. at room temperature. The results of this SEC are shown in FIG. **32**. Peak #1 65 represents Flt1D2Flk1D3.Fc Δ C1(a) and peak #2 represents VEGF165.

(c) Calculation of Flt1D2Flk1D3.Fc Δ C1(a):VEGF165 Complex Stoichiometry

The stoichiometry of the receptor-ligand complex was determined from the peak area or the peak height of the Concentrations of VEGF165 components. and Flt1D2Flk1D3.Fc\DeltaC1(a) corresponding to the peak height or peak area, respectively, were obtained from the standard curves for VEGF165 and Flt1D2Flk1D3.Fc∆C1(a). To 10 obtain a standard curve, four different concentrations (0.04 mg/ml-0.3 mg/ml) of either component were injected onto a Pharmacia Superose 12 PC 3.2/30 column equilibrated in 6M guanidinium chloride and eluted with the same solution at flow rate 40 µl/min. at room temperature. The standard curve was obtained by plotting peak area or peak height vs protein concentration. The molar ratio of VEGF165: Flt1D2Flk1D3.Fc\DeltaC1(a) determined from the peak area of the components was 1.16. The molar ratio of VEGF165: Flt1D2Flk1D3.Fc\DeltaC1(a) determined from the peak height of the components was 1.10.

Example 27

Determination of the Stoichiometry of the F1t1D2F1k1D3.Fc∆C1(a)/VEGF165 Complex by Size Exclusion Chromatography with On-Line Light Scattering

Complex preparation

night at 4° C.

VEGF165 was mixed with CHO transiently expressed Flt1D2.Flk1D3.FcΔC1(a) protein in molar ratio of 3:1 35 (VEGF165:Flt1 D2Flk1 D3.FcΔC1 (a)) and incubated over-

(a) Size Exclusion Chromatography (SEC) with On-Line Light Scattering

Size exclusion chromatography column with a MiniDawn on-line light scattering detector (Wyatt Technology, Santa Barbara, Calif.) and refractive index (R1) detectors (Shimadzu, Kyoto, Japan) was used to determine the molecular weight (MW) of the receptor-ligand complex. Samples were injected onto a Superose 12 HR 10/30 column (Pharmacia) equilibrated in PBS buffer and eluted with the same buffer at flow rate 0.5 ml/min. at room temperature. As shown in FIG. 33, the elution profile shows two peaks. Peak #1 represents the receptor-ligand complex and peak #2 represents the unbound VEGF165. MW was calculated from LS and R1 signals. The same procedure was used to determine MW of the individual components of the receptor-ligand complex. The results of these determinations are as follows: MW of the Flt1D2Flk1D3.FcΔC1(a)NEGF165 complex at the peak position is 157 300 (FIG. 33), the MW of VEGF165 at the peak position is 44 390 (FIG. 34) and the MW of R1R2 at the peak is 113 300 (FIG. 35).

These data indicated that the stoichiometry of the Flt1D2Flk1D3.Fc Δ C1(a)NEGF complex is 1:1 as its corresponds to the sum of molecular weights for Flt1D2Flk1D3.Fc Δ C1(a) and VEGF165. Importantly, this method conclusively proved that the Flt1D2Flk1D3.Fc Δ C1 (a)/VEGF165 complex was indeed composed of only one molecule of VEGF165 ligand and only one molecule of the Flt1D2Flk1D3.Fc Δ C1(a).

Example 28

Peptide Mapping of Flt1D2.Flk1D3.Fc∆C1(a)

The disulfide structures and glycosylation sites in ⁵ Flt1D2.Flk1D3.Fc Δ C1(a) were determined by a peptide mapping method. In this method, the protein was first cleaved with trypsin. Tryptic fragments were analyzed and identified by HPLC coupled with mass spectrometry, in ¹⁰ addition to an N-terminal sequencing technique. Reduction of the tryptic digest was employed to help identify disulfidebond-containing fragments. Treatment of the tryptic digest with PNGase F (Glyko, Novato, Calif.) was employed to help identify fragments with N-linked glycosylation sites. ¹⁵ The results are summarized in the accompanying FIG. **36**.

of ten There are a total cysteines in Flt1D2.Flk1D3.Fc∆C1(a); six of them belong to the Fc region. Cys27 has been confirmed to be disulfide bonded to Cys76. Cys121 is confirmed to be disulfide bonded to Cys $^{-20}$ 182. The first two cysteines in the Fc region (Cys211 and Cys214) form an intermolecular disulfide bond with the same two cysteines in another Fc chain. However, because these two cysteines can not be separated enzymatically from 25 each other, it can not be determined whether disulfide bonding is occurring between same cysteines (Cys211 to Cys211, for example) or between Cys211 and Cys214. Cys216 is confirmed to be disulfide bonded to Cys306. Cys 352 is confirmed to be disulfide bonded to Cys410.

There are five possible N-linked glycosylation sites in Flt1D2.Flk1D3.Fc Δ C1(a). All five of them are found to be glycosylated to varying degrees. Complete glycosylation was observed at Asn33 (amino acid sequence NIT), Asn193 (amino acid sequence NST), and Asn282 (amino acid ³⁵ sequence NST). In addition, partial glycosylation is observed on Asn65 and Asn120. Sites of glycosylation are highlighted by underline in the FIG. **36**.

Example 29

Pharmacokinetic Analysis of Modified Flt Receptors

(a) Pharmacokinetic Analysis of Flt1(1-3)-Fc (A40), Flt1D2.Flk1D3.Fc Δ C1(a) and VEGFR1R2-Fc Δ C1(a)

Balb/c mice (25-30 g) were injected subcutaneously with 4 mg/kg of Flt1(1-3)-Fc (A40), CHO transiently expressed Flt1D2.Flk1D3.FcAC1 (a), CHO stably expressed 50 Flt1D2.Flk1D3.Fc\DeltaC1(a), and CHO transiently expressed VEGFR1R2-Fc Δ C1(a). The mice were tail bled at 1, 2, 4, 6, 24 hrs, 2 days, 3 days and 6 days after injection. The sera were assayed in an ELISA designed to detect Flt1(1-3)-Fc (A40), Flt1D2.Flk1D3.FcΔC1(a) or VEGFR1R2-FcΔC1(a). 55 The ELISA involves coating an ELISA plate with VEGF165, binding the detect Flt1(1-3)-Fc (A40), Flt1 D2.Flk1 D3.Fc\DeltaC1 (a) or VEGFR1 R2-Fc\DeltaC1 (a) and reporting with an anti-Fc antibody linked to horse radish peroxidase. The results of this experiments are shown in 60 FIG. 37. The T_{max} for Flt1(1-3)-Fc (A40) was at 6 hrs while the T_{max} for the transient and stable Flt1D2.Flk1 D3.Fc Δ C1 (a) and the transient VEGFR1 R2-FcΔC1 (a) was 24 hrs. The Cmax for Flt1(1-3)-Fc (A40) was 8 µg/ml. For both transients (Flt1D2.Flk1D3.FcΔC1(a) and VEGFR1 R2-FcΔC1 (a)) the 65 C_{max} was 18 µg/ml and the C_{max} for the stable VEGFR1 R2-Fc Δ C1 (a) was 30 µg/ml.

(b) Pharmacokinetic Analysis of Ftl1(1-3)-Fc (A40). Flt1D2.Flk1D3.FcΔC1(a) and Flt1 D2.VEGFR3D3.FcΔC1 (a)

Balbic mice (25-30 g) were injected subcutaneously with 4 mg/kg of Flt1(1-3)-Fc (A40), CHO transiently expressed Flt1 D2.Flk1 D3.Fc\DeltaC1 (a) and CHO transiently expressed Flt1D2.VEGFR3D3.FcΔC1(a). The mice were tail bled at 1, 2, 5, 6, 7, 8, 12, 15 and 20 days after injection. The sera were assayed in an ELISA designed to detect Flt1 (1-3)-Fc, Flt1D2.Flk1 D3.FcΔC1 (a) and Flt1 D2.VEGFR3D3.FcΔC1 (a). The ELISA involves coating an ELISA plate with 165, binding the Flt1(1-3)-Fc, F1t1 D2.Flk1 D3.Fc∆C1 (a) or Fit1 D2.VEGFR3D3.Fc\DeltaC1 (a) and reporting with an anti-Fc antibody linked to horse radish peroxidase. Flt1(1-3)-Fc (A40) could no longer be detected in the serum after day 5 whereas. Flt1D2.Flk1 D3.Fc∆C1 (a) and Flt1D2.VEGFR3D3.Fc Δ C1 (a) were detectable for 15 days or more. The results of this experiment are shown in FIG. 38.

Example 30

Evaluation of the Ability of Flt1D2.Flk1D3.FcΔC1(a) to Inhibit Tumor Growth In Vivo

To evaluate the ability of Flt1D2.Flk1D3.Fc∆C1(a) to inhibit tumor growth in vivo a model in which tumor cell suspensions are implanted subcutaneously on the right flank of male severe combined immunodeficiency (SCID) mice was employed. Two cell lines, the human HT-1080 fibrosarcoma cell line (ATCC accession no. CCL-121) and the rat C6 glioma cell line (ATCC accession no. CCL-107), each of which exhibit distinctly different morphologies and growth characteristics, were used in the assay. The first dose of Flt1D2.Flk1D3.FcΔC1(a) (at 25 mg/Kg or as indicated in FIGS. 39 and 40) was given on the day of tumor implantation. Animals subsequently received subcutaneous injections of Flt1(1-3)-Fc (A40), Flt1D2.Flk1D3.FcAC1(a) or vehicle either every other day (EOD) or two times per week (2x/wk) for a period of 2 weeks. After 2 weeks, animals were perfused with fixative, tumors were removed and samples were blinded. Tumor volume was determined by measuring the length and width of visible subcutaneous tumors. Both of Flt1 (1-3)-Fc (A40) and Flt1D2.Flk1D3.FcΔC1(a) significantly reduced the growth 45 of tumors formed by HT-1080 and C6 cells. The results of these experiments are shown in FIG. 39 and FIG. 40.

Example 31

The Effect of VEGF165 and Modified Fit Receptors in Female Reproductive System

The stereotypic pattern of vascular remodeling which occur in the uterus and ovary over the course of the reproductive cycle has been well characterized, making these tissues particularly well suited to the study of mechanisms which regulate angiogenesis, vascular remodeling and vascular regression. Indeed, in situ hybridization studies in the reproductive tissues provided the first clear evidence that VEGF acts as a mediator of physiological angiogenesis in mature rodents, as well as humans and non-human primates (Phillips et al, 1990; Ravindranath et al, 1992; Shweiki et al, 1993; Kamat et al, 1995). As cyclic angiogenesis and vascular remodeling are prominent features of the normal ovary and uterus, it is not surprising that abnormal blood vessel growth and/or vascular dysfunction have been found to characterize many pathological conditions which affect these organs. Furthermore, these pathogenic vascular abnormalities are thought to be caused or perpetuated by the dysregulated expression of one or more angiogenic or antiangiogenic factors, most prominently VEGF.

For example, abnormal angiogenesis is characteristic of polycystic ovary disease, endometriosis and endometrial carcinoma, and in each case VEGF is over expressed in the affected tissue (Kamat et al, 1995; Shifren et al, 1996; Guidi et al, 1996; Donnez et al, 1998). Overexpression of VEGF is also thought to play a pathogenic role in the establishment of systemic vascular hyperpermeability in ovarian hyperstimulation syndrome (McClure et al, 1994; Levin et al, 1998) and preeclampsia (Baker et al, 1995; Sharkey et al, 1996). In addition, VEGF has been implicated as the per- 15 meability factor responsible for the production of ascites associated with ovarian carcinoma and other tumors (Senger et al, 1983; Boocock et al, 1995). Agents which effectively neutralize the biological actions of VEGF can reasonably be anticipated to be of therapeutic benefit in the above and 20 related conditions.

Angiogenesis and vascular remodeling are also hallmarks of blastocyst implantation and placental development (Findlay, 1986). VEGF is strongly expressed both in the maternal decidua and in embryonic trophoblasts, where it is thought 25 to first stimulate expansion and hyperpermeability of the uterine vasculature during the peri-implantation period and subsequently mediate formation of both the maternal and embryonic components of the placental vasculature (Shweiki et al, 1993; Cullinan-Bove and Koos, 1993; Chakraborty et al, 1995; Das et at, 1997). VEGF is also required for luteal angiogenesis and associated progesterone secretion necessary to prepare the uterus for implantation (Ferrara et al, 1998). Thus, agents which inhibit the biological actions of VEGF may prove to be useful as contraceptive 35 agents (by preventing implantation), or as an abortifacients in the early stages of gestation. The latter application might find particular use as a non-surgical intervention for the termination of ectopic pregnancies.

While the expression of VEGF receptors is largely con- 40 fined to the vascular endothelium in normal reproductive tissues, Flt1 is also expressed by trophoblasts in the placenta in both humans and animals (Clark et al, 1996; He et al, 1999) where it has been proposed to play a role in trophoblast invasion. Interestingly, both Flt1 and KDR (Flk1) are 45 kD PEG and tested in balb/c mice for their pharmacokinetic expressed by choriocarcinoma cell line BeWo (Charnock-Jones et al, 1994), and VEGF has been shown to promote DNA synthesis and tyrosine phosphorylation of MAP kinase in these cells. Furthermore, primary and metastatic ovarian carcinomas not only to express high levels of VEGF, but-in 50 addition to the vascular endothelium-the tumor cells themselves express KDR and/or Flt1 (Boocock et al, 1995). These findings suggest that VEGF may not only be critically involved in the generation and maintenance of tumor vasculature, but that at least in some tumors of reproductive 55 origin VEGF may subserve an autocrine role, directly supporting the survival and proliferation of the tumor cells. Thus agents which block the actions of VEGF may have particularly beneficial applications to the treatment of tumors of reproductive origin.

Methods and Results

(a) Assessment of VEGF-Induced Uterine Hyperpermeability

Pregnant mare's serum gonadotrophin (PMSG) was 65 injected subcutaneously (5 IU) to induce ovulation in prepubertal female rats. This results in a surge of estradiol after

2 days which in turn causes an induction of VEGF in the uterus, it is reported that this induction results in hyperpermeability of the uterus and an increase in uterine wet weight 6 hrs. later and, therefore, could potentially be blocked by the modified Flt receptors Flt1 (1-3)-Fc (A40), F1t1 D2.Flk1D3.Fc Δ C1(a) and Flt1D2.VEGFR3D3. Fc Δ C1(a). In this in vivo model, the normal weight of the rat uterus is about 50 mg and this can be induced to 300-350 mg by PMSG. Desiccation of the tissue re veals that this is all water 10 weight. Subcutaneous injection of Flt1(1-3)-Fc (A40),

Flt1D2.Flk1D3.Fc Δ C1(a) and FltID2.VEGFR3D3.Fc\DeltaC1(a) at 25 mg/kg at 1 hr. after PMSG injection results in about a 50% inhibition of the increase in uterine wet weight. Increasing the dose of modified Flt receptor does not further reduce the increase in wet weight suggesting that there is a VEGF-independent component to this model. The results of this experiment are shown in FIG. 41.

(a) Assessment of Corpus Luteum Angiogenesis Using Progesterone as a Readout

Pregnant mare's serum gonadotrophin (PMSG) is injected subcutaneously (5 IU) to induce ovulation in prepubertal female rats. This results in a fully functioning corpus luteum containing a dense network of blood vessels after 4 days that allows for the secretion of progesterone into the blood stream in order to prepare the uterus for implantation. The induction of angiogenesis in the corpus luteum requires VEGF; therefore, blocking VEGF would result in a lack of new blood vessels and thus a lack of progesterone secreted into the blood stream. In this in vivo model, resting levels of progesterone are about 5 ng/ml and this can be induced to a level of 25-40 ng/ml after PMSG. Subcutaneous injection of Flt1(1-3)-Fc (A40) or Flt1D2.Flk1D3.FcaC1(a) at 25 mg/kg or 5 mg/kg at 1 hr. after PMSG injection results in a complete inhibition of the progesterone induction on day 4. The results of this experiment are shown in FIGS. 42A-42B.

Example 33

Pharmacokinetic Analysis of Flt1(1-3)-Fc (A40) and Pegylated Flt1(1-3)-Fc

Flt1(1-3)-Fc was PEGylated with either 10 kD PEG or 20 profile. Both PEGylated forms of Flt1(1-3)-Fc were found to have much better PK profiles than Flt1(1-3)-Fc (A40), with the Tmax occurring at 24 hrs. for the PEGylated molecules as opposed to 6 hrs. for Flt1(1-3)-Fc (A40).

Example 34

VEGF165 ELISA to Test Affinity of Modified Flt1 Receptor Variants

10 pM of VEGF165 was incubated overnight at room temperature with modified Flt1 receptor variants ranging from 160 pM to 0.1 pM. The modified Flt1 receptor variants used in this experiment were Flt1(1-3)-Fc, Flt1(1-3)-Fc 60 (A40), transiently expressed Flt1D2Flk1D3.Fc-ΔC1 (a), transiently expressed Flt1 D2VEFGFR3D3-Fc\DeltaC1(a), Flt1- $(1-3_{NAS})$ -Fc, Flt1 $(1-3_{R\rightarrow C})$ -Fc and Tie2-Fc. Flt1 $(1-3_{NAS})$ -Fc is a modified version of Flt1(1-3)-Fc in which the highly basic amino acid sequence KNKRASVRRR is replaced by NASVNGSR, resulting in the incorporation of two new glycosylation sites and a net reduction of five positive charges, both with the purpose of reducing the unfavorable effects of this sequence on PK. $Flt1(1-3_{R->C})$ -Fc is a modification in which a single arginine (R) residue within the same basic amino acid sequence is changed to a cysteine (C) (KNKRASVRRR_KNKCASVRRR) to allow for pegylation at that residue, which could then shield the basic region 5 from exerting its unfavorable effects on PK. After incubation the solution was transferred to a plate containing a capture antibody for VEGF165 (R&D). The amount of free VEGF165 was then determined using an antibody to report free VEGF165. This showed that the modified Flt1 receptor 10 variant with the highest affinity for VEGF165 (determined as the lowest amount of free VEGF165) was Flt1D2Flk1 D3.Fc\DeltaC1 (a), followed by Flt1 (1-3)-Fc and Flt1(1-3)-Fc (A40) and then by $Flt1(1-3_{R->C})$ -Fc, $Flt1(1-3_{NAs})$ -Fc and Flt1D2VEFGFR3D3-FcΔC1(a). Tie2Fc has no affinity for 15 VEGF165.

The invention claimed is:

1. An isolated nucleic acid molecule encoding a fusion protein capable of binding vascular endothelial growth factor (VEGF), consisting of

(a) a (VEGF) receptor component having immunoglobulin-like (Ig) domains consisting of an Ig domain 2 of a first VEGF receptor human Flt1 and Ig domain 3 of a second VEGF receptor human Flk1; and

(b) a multimerizing component.

2. The isolated nucleic acid molecule of claim **1**, wherein the nucleotide sequence encoding a first VEGF receptor component is upstream of the nucleotide sequence encoding a second VEGF receptor component.

3. The isolated nucleic acid molecule of claim **1**, wherein 30 the nucleotide sequence encoding a first VEGF receptor component is downstream of the nucleotide sequence encoding a second VEGF receptor component.

4. The isolated nucleic acid of claim **1**, wherein the multimerizing component comprises an immunoglobulin 35 domain.

5. The isolated nucleic acid of claim 4, wherein the immunoglobulin domain is selected from the group consisting of the Fc domain of IgG, and the heavy chain of IgG.

6. The isolated nucleic acid molecule of claim 1, com- 40 prising a nucleic acid sequence selected from:

- (a) SEQ ID NO: 15; and
- (b) nucleic acid sequences which, as a result of the degeneracy of the genetic code, differ from the nucleic acid sequence of SEQ ID NO: 15.

7. An expression vector comprising a nucleic acid molecule encoding a fusion protein capable of binding vascular 40

endothellal growth factor (VEGF), wherein the fusion protein consists of immunoglobulin-like (Ig) domain 2 of VEGF receptor human Flt1, Ig domain 3 of VEGF receptor human Flk1, and a multimerizing component.

8. A host-vector system for the production of a fusion polypeptide comprising an expression vector encoding a fusion protein capable of binding vascular endothelial growth factor (VEGF), wherein the fusion protein consists of immunoglobulin-like (Ig) domain 2 of VEGF receptor human Flk1, Ig domain 3 of VEGF receptor human Flk1, and a multimerizing component, In a suitable isolated host cell.

9. The host-vector system of claim 8, wherein the host cell is a bacterial cell, yeast cell, insect cell, or mammalian cell.

10. The host-vector system of claim 9, wherein the host cell is selected from the group consisting of *E. coli* and CHO.

11. A method of producing a fusion polypeptide, comprising growing cells of the host-vector system of claim $\mathbf{8}$, under conditions permitting production of the fusion polypeptide and recovering the fusion polypeptide so produced.

 An isolated nucleic acid molecule encoding a fusion protein capable of binding vascular endothelial growth factor (VEGF), consisting of

- (a) a VEGF receptor component having immunoglobulinlike (Ig) domains consisting of an Ig domain 2 of a first VEGF receptor human Flt1 and Ig domain 3 of a second VEGF receptor human Flt-4: and
- (b) a multimerizing component.

13. The isolated nucleic acid molecule of claim **12**, wherein the multimerizing component is chosen from the Fc domain of IgG and the heavy chain of IgG.

14. An isolated nucleic acid molecule consisting of a nucleotide sequence encoding immunoglobulin-like (Ig) domain 2 of a first vascular endothelial growth factor (VEGF) receptor upstream of a nucleotide sequence encoding Ig domain 3 of a second VEGF receptor and a nucleotide sequence encoding a multimerizing component, wherein the nucleic acid sequence is SEQ ID NO:15.

 The isolated nucleic acid molecule of claim 14 encoding a fusion protein comprising the amino acid sequence of 45 SEQ ID NO:16.

* * * *