

University of Kentucky UKnowledge

Pharmaceutical Sciences Faculty Patents

Pharmaceutical Sciences

11-26-2013

High Activity Mutants of Butyrylcholinesterase for Cocaine Hydrolysis

Chang-Guo Zhan University of Kentucky, zhan@uky.edu

Fang Zheng University of Kentucky, fzhen2@uky.edu

Wenchao Yang University of Kentucky

Lin Xue University of Kentucky, lionetxue@gmail.com

Shurong Hou University of Kentucky, shuronghou@uky.edu

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Follow this and additional works at: https://uknowledge.uky.edu/ps patents



Part of the Pharmacy and Pharmaceutical Sciences Commons

Recommended Citation

Zhan, Chang-Guo; Zheng, Fang; Yang, Wenchao; Xue, Lin; and Hou, Shurong, "High Activity Mutants of Butyrylcholinesterase for Cocaine Hydrolysis" (2013). Pharmaceutical Sciences Faculty Patents. 17. https://uknowledge.uky.edu/ps_patents/17

This Patent is brought to you for free and open access by the Pharmaceutical Sciences at UKnowledge. It has been accepted for inclusion in Pharmaceutical Sciences Faculty Patents by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.



(12) United States Patent

Zhan et al.

(54) HIGH ACTIVITY MUTANTS OF **BUTYRYLCHOLINESTERASE FOR** COCAINE HYDROLYSIS

(75) Inventors: Chang-Guo Zhan, Lexington, KY (US);

Fang Zheng, Lexington, KY (US); Wenchao Yang, Wuhan (CN); Liu Xue, Lexington, KY (US); Shurong Hou,

Lexington, KY (US)

Assignee: University of Kentucky Research

Foundation, Lexington, KY (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.

(21) Appl. No.: 13/399,406

(22)Filed: Feb. 17, 2012

Related U.S. Application Data

(60) Provisional application No. 61/443,932, filed on Feb. 17, 2011.

(51)	Int. Cl.	
	C12N 9/16	(2006.01)
	C12N 9/18	(2006.01)
	A61K 38/46	(2006.01)
	C07H 21/04	(2006.01)

(52) U.S. Cl. USPC 435/196; 435/197; 424/94.6; 536/23.2

(58) Field of Classification Search

See application file for complete search history.

(56)References Cited

U.S. PATENT DOCUMENTS

2005/0136044 A1* 6/2005 Watkins et al. 424/94.6 OTHER PUBLICATIONS

Brimijoin, S. et al., Neuropsychopharmacology 2008, 33, 2715-

(10) Patent No.:

US 8,592,193 B1

(45) **Date of Patent:**

Nov. 26, 2013

Xie, W.; Altamirano, C. V.; Bartels, C. F.; Speirs, R. J.; Cashman, J. R.; Lockridge, O. Mol. Pharmacol. 1999, 55, 83.

Xue, L.; Ko, M.-C.; Tong, M.; Yang, W.; Hou, S.; Fang, L.; Liu, J.; Zheng, F.; Woods, J. H.; Tai, H.-H.; Zhan, C.-G. "Design, preparation, and characterization of high-activity mutants of human butyrylcholinesterase specific for detoxification of cocaine", Mol. Pharmacol. 2011, 79, 290-297.

Yang, W.; Pan, Y.; Fang, L.; Gao, D.; Zheng, F.; Zhan, C.-G. "Freeenergy perturbation simulation on transition states and high-activity mutants of human butyrylcholinesterase for (-)-cocaine hydrolysis", J. Phys. Chem. B 2010, 114, 10889-10896.

Yang, W.; Xue, L.; Fang, L.; Zhan, C.-G. "Characterization of a high-activity mutant of human butyrylcholinesterase against (-)cocaine", Chemico-Biological Interactions, 2010, 187, 148-152.

Zheng, F.; Yang, W.; Xue, L.; Hou, S.; Liu, J.; Zhan, C.-G. "Design of high-activity mutants of human butyrylcholinesterase against (-)cocaine: structural and energetic factors affecting the catalytic efficiency", Biochemistry 2010, 49, 9113-9119.

Zheng, F.; Yang, W.; Ko, M.-C.; Liu, J.; Cho, H.; Gao, D.; Tong, M.; Tai, H.-H.; Woods, J. H.; Zhan, C.-G. "Most Efficient Cocaine Hydrolase Designed by Virtual Screening of Transition States", J. Am. Chem. Soc. 2008, 130, 12148-12155.

* cited by examiner

Primary Examiner — Anand Desai Assistant Examiner — Igbal H Chowdhury (74) Attorney, Agent, or Firm — Stites & Harbison PLLC; Mandy Wilson Decker

(57)ABSTRACT

Butyrylcholinesterase (BChE) polypeptide variants of the presently-disclosed subject matter have enhanced catalytic efficiency for (-)-cocaine, as compared to wild-type BChE. Pharmaceutical compositions of the presently-disclosed subject matter include a BChE polypeptide variant having an enhanced catalytic efficiency for (-)-cocaine. A method of the presently-disclosed subject matter for treating a cocaine-induced condition includes administering to an individual an effective amount of a BChE polypeptide variant, as disclosed herein, to lower blood cocaine concentration.

5 Claims, No Drawings

30

HIGH ACTIVITY MUTANTS OF BUTYRYLCHOLINESTERASE FOR COCAINE HYDROLYSIS

RELATED APPLICATIONS

This application claims priority from U.S. Provisional Application Ser. No. 61/443,932 filed Feb. 17, 2011, the entire disclosure of which is incorporated herein by this reference.

GOVERNMENT INTEREST

Subject matter described herein was made with government support under Grant Number R01DA013930 awarded by the National Institute on Drug Abuse (NIDA) of the National Institutes of Health (NIH). The government has certain rights in the described subject matter.

TECHNICAL FIELD

The presently-disclosed subject matter relates to butyrylcholinesterase variant polypeptides, and in particular, butyrylcholinesterase mutants having amino acid substitutions.

INTRODUCTION

Cocaine abuse is a major medical and public health problem that continues to defy treatment. The disastrous medical and social consequences of cocaine addiction, such as violent crime, loss in individual productivity, illness, and death, have made the development of an effective pharmacological treatment a high priority. However, cocaine mediates its reinforcing and toxic effects by blocking neurotransmitter reuptake and the classical pharmacodynamic approach has failed to yield small-molecule receptor/transporter antagonists due to the difficulties inherent in blocking a blocker. An alternative to pharmacodynamic approach is to interfere with the delivery of cocaine to its receptors/transporters and accelerate its metabolism in the body.

The dominant pathway for cocaine metabolism in primates is butyrylcholinesterase (BChE)-catalyzed hydrolysis at the benzoyl ester group (Scheme 1).

Scheme 1. Schematic representation of BChE-catalyzed hydrolysis at the benzoyl ester group.

Only 5% of the cocaine is deactivated through oxidation by the liver microsomal cytochrome P450 system. Cocaine hydrolysis at benzoyl ester group yields ecgonine methyl ester, whereas the oxidation produces norcocaine. The metabolite ecgonine methyl ester is a biologically inactive metabolite, whereas the metabolite norcocaine is hepatotoxic and a local anesthetic. In human beings, more than 90% of cocaine is metabolized by BChE. BChE is synthesized in the liver and widely distributed in the body, including plasma, brain, and lung. Extensive experimental studies in animals and humans demonstrate that enhancement of BChE activity by administration of exogenous enzyme substantially decreases cocaine half-life.

Enhancement of cocaine metabolism by administration of 50 BChE has been recognized to be a promising pharmacokinetic approach for treatment of cocaine abuse and dependence. However, the catalytic activity of this plasma enzyme is three orders-of-magnitude lower against the naturally occurring (-)-cocaine than that against the biologically inac-55 tive (+)-cocaine enantiomer. (+)-cocaine can be cleared from plasma in seconds and prior to partitioning into the central nervous system (CNS), whereas (-)-cocaine has a plasma half-life of approximately 45-90 minutes (for a relatively low dose of cocaine), long enough for manifestation of the CNS effects which peak in minutes. Under the overdose condition, BChE is saturated with (-)-cocaine and, thus, the plasma half-life of (-)-cocaine will be longer. Hence, BChE mutants with high activity against (-)-cocaine are highly desired for use in humans. Although some BChE mutants with increased 65 catalytic activity over wild-type BChE have previously been generated, there exists a need for mutant BChE with even higher catalytic activity.

SUMMARY

The presently-disclosed subject matter meets some or all of the above-identified needs, as will become evident to those of ordinary skill in the art after a study of information provided 5 in this document.

This Summary describes several embodiments of the presently-disclosed subject matter, and in many cases lists variations and permutations of these embodiments. This Summary is merely exemplary of the numerous and varied embodiments. Mention of one or more representative features of a given embodiment is likewise exemplary. Such an embodiment can typically exist with or without the feature(s) mentioned; likewise, those features can be applied to other 15 embodiments of the presently-disclosed subject matter, whether listed in this Summary or not. To avoid excessive repetition, this Summary does not list or suggest all possible combinations of such features.

The presently-disclosed subject matter includes butyryl- 20 cholinesterase (BChE) polypeptide variants. In some embodiments the amino acid sequence of the BChE polypeptide variant includes an amino acid sequence selected from the group consisting of: SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, as set forth herein. The presently-disclosed subject matter further includes nucleic acid molecules that encodes a BChE polypeptide variant, including the nucleic acid sequence selected from the group consisting of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 30 43, and 45.

The presently-disclosed subject matter further includes a pharmaceutical composition that includes a butyrylcholinesterase polypeptide variant and a suitable pharmaceutical carrier.

The presently-disclosed subject matter further includes a method of treating a cocaine-induced condition, which includes administering to an individual an effective amount of BChE polypeptide variant or functional fragment thereof, or tide variant or functional fragment thereof, as described herein, to lower blood cocaine concentration. In some embodiments, the BChE polypeptide variant exhibits a onehundred-fold or more increase in cocaine hydrolysis catalytic efficiency compared to wild-type butyrylcholinesterase.

The presently-disclosed subject matter further includes a method of treating a cocaine-induced condition, which includes administering to an individual an effective amount of a BChE nucleotide variant, i.e., a nucleotide molecule encoding a BChE polypeptide variant or functional fragment 50 thereof as disclosed herein.

BRIEF DESCRIPTION OF THE SEQUENCE LISTING

SEQ ID NO: 1 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 2;

SEQ ID NO: 2 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227S, 60 S287G, A328W, and Y332G;

SEQ ID NO: 3 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 4;

SEQ ID NO: 4 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substi- 65 tutions, as compared to wild type BChE: A199S, F227T, S287G, A328W, and Y332G;

4

SEQ ID NO: 5 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 6;

SEQ ID NO: 6 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227M, S287G, A328W, and Y332G;

SEO ID NO: 7 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 8;

SEQ ID NO: 8 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227C, S287G, A328W, and Y332G;

SEQ ID NO: 9 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO:

SEQ ID NO: 10 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227A, P285N, S287G, A328W, and Y332G;

SEQ ID NO: 11 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 12;

SEQ ID NO: 12 is an amino acid sequence encoding a 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, and 46 25 BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227P, P285A, S287G, A328W, and Y332G;

> SEQ ID NO: 13 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 14;

> SEQ ID NO: 14 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227S, P285Q, S287G, A328W, and Y332G;

> SEQ ID NO: 15 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID

SEO ID NO: 16 is an amino acid sequence encoding a a pharmaceutical composition comprising a BChE polypep- 40 BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S. F227S, P285S, S287G, A328W, and Y332G;

> SEQ ID NO: 17 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 18;

> SEQ ID NO: 18 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227S, P285G, S287G, A328W, and Y332G;

> SEQ ID NO: 19 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 20;

> SEQ ID NO: 20 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227P, P285S, L286M, S287G, A328W, and Y332G;

> SEQ ID NO: 21 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 22:

SEQ ID NO: 22 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227A, P285S, S287G, A328W, and E441D;

SEQ ID NO: 23 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 24;

SEQ ID NO: 24 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227A, P285A, S287G, A328W, and E441D;

SEQ ID NO: 25 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 26:

SEQ ID NO: 26 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227P, L286M, S287G, A328W, and Y332G;

SEQ ID NO: 27 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID $_{15}$ NO: 28;

SEQ ID NO: 28 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227G, P285A, S287G, A328W, and Y332G;

SEQ ID NO: 29 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 30:

SEQ ID NO: 30 is an amino acid sequence encoding a ²⁵ BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227G, P285G, S287G, A328W, and Y332G;

SEQ ID NO: 31 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 32;

SEQ ID NO: 32 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, ³⁵ F227G, P285Q, S287G, A328W, and Y332G;

SEQ ID NO: 33 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 34;

SEQ ID NO: 34 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227G, P285S, S287G, A328W, and Y332G;

SEQ ID NO: 35 is a nucleotide sequence encoding a 45 butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 36:

SEQ ID NO: 36 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227A, P285E, S287G, A328W, and Y332G;

SEQ ID NO: 37 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 38;

SEQ ID NO: 38 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227P, P285N, S287G, A328W, and Y332G;

SEQ ID NO: 39 is a nucleotide sequence encoding a ⁶⁰ butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 40:

SEQ ID NO: 40 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227S, P285A, S287G, A328W, and Y332G;

6

SEQ ID NO: 41 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 42:

SEQ ID NO: 42 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227S, P285N, S287G, A328W, and Y332G;

SEQ ID NO: 43 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 44:

SEQ ID NO: 44 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227S, L286M, S287G, A328W, and Y332G;

SEQ ID NO: 45 is a nucleotide sequence encoding a butyrylcholinesterase (BChE) polypeptide variant of SEQ ID NO: 46;

SEQ ID NO: 46 is an amino acid sequence encoding a BChE polypeptide variant having the following amino acid substitutions, as compared to wild type BChE: A199S, F227G, L286M, S287G, A328W, and Y332G;

DESCRIPTION OF EXEMPLARY EMBODIMENTS

The details of one or more embodiments of the presently-disclosed subject matter are set forth in this document. Modifications to embodiments described in this document, and other embodiments, will be evident to those of ordinary skill in the art after a study of the information provided in this document. The information provided in this document, and particularly the specific details of the described exemplary embodiments, is provided primarily for clearness of understanding and no unnecessary limitations are to be understood therefrom. In case of conflict, the specification of this document, including definitions, will control.

The presently-disclosed subject matter includes butyrylcholinesterase (BChE) polypeptide variants. The BChE polypeptide variants disclosed herein each have enhanced catalytic efficiency for (-)-cocaine, as compared to wild-type BChE. The presently-disclosed subject matter further includes a pharmaceutical composition including a butyrylcholinesterase polypeptide variant, as described herein, and a suitable pharmaceutical carrier. The presently-disclosed subject matter further includes a method of treating a cocaineinduced condition comprising administering to an individual an effective amount of a butyrylcholinesterase polypeptide variant of functional fragement there of, or a BChE nucleotide variant, i.e., a nucleotide molecule encoding a BChE polypeptide variant or functional fragment thereof, as disclosed herein, to lower blood cocaine concentration. A cocaine-induced condition includes conditions resulting from the administration/use of cocaine, including, for example, treatment for an addiction to cocaine. For example, a polypeptide or nucleotide, as described herein, could be administered prior to the use of cocaine as part of an addiction treatment strategy.

In some embodiments, the BChE polypeptide variant is selected from a BChE polypeptide variants set forth in Table 1. Table 1 also includes the SEQ ID NOs associated with the identified BChE polypeptide variants, as well as a summary of the approximate fold increase in catalytic efficiency against (–)-cocaine for the identified BChE polypeptide variants, as compared to wild-type BChE.

TABLE 1

	BChE Polypeptide Variants and Associated SEQ ID NOs														
Variant				Amino A	cid Subst	itution				N.A. SEQ ID	A.A. SEQ ID				
No.	199	227	285	286	287	328	332	441	C.E.*	NO:	NO:				
1	A199S	F227S	_	_	S287G	A328W	Y332G	_	2750	1	2				
2	A199S	F227T	_	_	S287G	A328W	Y332G	_	1780	3	4				
3	A199S	F227M	_	_	S287G	A328W	Y332G	_	1310	5	6				
4	A199S	F227C	_	_	S287G	A328W	Y332G	_	1290	7	8				
5	A199S	F227A	P285N	_	S287G	A328W	Y332G	_	2790	9	10				
6	A199S	F227P	P285A	_	S287G	A328W	Y332G	_	2810	11	12				
7	A199S	F227S	P285Q	_	S287G	A328W	Y332G	_	4260	13	14				
8	A199S	F227S	P285S	_	S287G	A328W	Y332G	_	2430	15	16				
9	A199S	F227S	P285G	_	S287G	A328W	Y332G	_	2480	17	18				
10	A199S	F227P	P285S	L286M	S287G	A328W	Y332G	_	2710	19	20				
11	A199S	F227A	P285S	_	S287G	A328W	_	E441D	2280	21	22				
12	A199S	F227A	P285A	_	S287G	A328W	_	E441D	1980	23	24				
13	A199S	F227P	_	L286M	S287G	A328W	Y332G	_	1140	25	26				
14	A199S	F227G	P285A	_	S287G	A328W	Y332G	_	3130	27	28				
15	A199S	F227G	P285G	_	S287G	A328W	Y332G	_	3445	29	30				
16	A199S	F227G	P285Q	_	S287G	A328W	Y332G	_	2515	31	32				
17	A199S	F227G	P285S	_	S287G	A328W	Y332G	_	3980	33	34				
18	A199S	F227A	P285E	_	S287G	A328W	Y332G	_	2180	35	36				
19	A199S	F227P	P285N	_	S287G	A328W	Y332G	_	2640	37	38				
20	A199S	F227S	P285A	_	S287G	A328W	Y332G	_	2450	39	40				
21	A199S	F227S	P285N	_	S287G	A328W	Y332G	_	2040	41	42				
22	A199S	F227S	_	L286M	S287G	A328W	Y332G	_	1774	43	44				
23	A199S	F227G	_	L286M	S287G	A328W	Y332G	_	1535	45	46				

^aCatalytic Efficiency (k_{car}/K_M) against (-)-cocaine (Approx. Fold Increase); The approximate ratio of the k_{car}/K_M value for the BChE mutant to that for the wild-type BChE against (-)-cocaine.

The terms "polypeptide", "protein", and "peptide", which are used interchangeably herein, refer to a polymer of the protein amino acids, or amino acid analogs, regardless of its size or function. Although "protein" is often used in reference to relatively large polypeptides, and "peptide" is often used in reference to small polypeptides, usage of these terms in the art overlaps and varies. The term "polypeptide" as used herein refers to peptides, polypeptides, and proteins, unless otherwise noted. The terms "protein", "polypeptide", and "peptide" are used interchangeably herein when referring to a gene product. Thus, exemplary polypeptides include gene products, naturally occurring proteins, homologs, orthologs, paralogs, fragments and other equivalents, variants, and analogs of the foregoing.

The term "variant" refers to an amino acid sequence that is different from the reference polypeptide by one or more amino acids, e.g., one or more amino acid substitutions. For example a butyrylcholinesterase (BChE) polypeptide variant differs from wild-type BChE by one or more amino acid 50 substitutions, i.e., mutations.

The terms "polypeptide fragment" or "fragment", when used in reference to a reference polypeptide, refers to a polypeptide in which amino acid residues are deleted as compared to the reference polypeptide itself, but where the 55 remaining amino acid sequence is usually identical to the corresponding positions in the reference polypeptide. Such deletions can occur at the amino-terminus, carboxy-terminus of the reference polypeptide, or alternatively both. A fragment can also be a "functional fragment," in which case the 60 fragment retains some or all of the activity of the reference polypeptide as described herein. For example, a functional fragment of a particular BChE polypeptide variant retains some or all of the cocaine hydrolysis activity, i.e., the catalytic efficiency for (-)-cocaine, of the particular BChE polypep- 65 tide variant. In this regard, the term "BChE polypeptide variant" is inclusive of functional fragments of the BChE

polypeptide variant. Such fragments are typically are at least about 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, or 550 amino acids long. One or more residues from about 1 to 67 and/or one or more residues from about 443 to 574 can be removed without substantially affecting the catalytic activity of the BChE polypeptide variant. As such, the term "BChE polypeptide variant" is inclusive of functional fragments wherein one or more residues from 1 to 67 and/or one or more residues from 443 to 574 is truncated relative to the full-length BChE polypeptide variant. In some embodiments, the fragment includes residues from 1 to 529. See Brimijoin, S. et al., *Neuropsychopharmacology* 2008, 33, 2715-2725).

The BChE polypeptide variant (e.g., SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, and 46) can be formulated in a pharmaceutical composition along with a suitable pharmaceutical carrier known to one skilled in the art.

The present BChE variant polypeptides can be used in treating a cocaine-induced condition by administering to an individual, an effective amount of a BChE variant polypeptides, (e.g., SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, and 46), to lower blood cocaine concentration. The BChE polypeptide variant can be administered in the form of a pharmaceutical composition in which the BChE polypeptide variant is included with a suitable pharmaceutical carrier. Treatment of a cocaine-induced condition using one of the aforementioned BChE polypeptide variants can be in a manner that will be understood by those skilled in the art.

The preferred dose for administration of a BChE polypeptide variant or pharmaceutical composition in accordance with the presently-described subject matter is that amount which will be effective in lowering (–)-cocaine concentration in a patient's bloodstream, and one would readily recognize that this amount will vary greatly depending on the nature of cocaine consumed, e.g., injected or inhaled, and the condition

of a patient. An "effective amount" of butyrylcholinesterase polypeptide variant or pharmaceutical composition to be used in accordance with the presently-disclosed subject matter is intended to mean a nontoxic but sufficient amount of the agent, such that the desired prophylactic or therapeutic effect is produced. Thus, the exact amount of the enzyme or a particular agent that is required will vary from subject to subject, depending on the species, age, and general condition of the subject, the severity of the condition being treated, the particular carrier or adjuvant being used and its mode of administration, and the like. Similarly, the dosing regimen should also be adjusted to suit the individual to whom the composition is administered and will once again vary with age, weight, metabolism, etc. of the individual. Accordingly, the "effective amount" of any particular butyrylcholinesterase polypeptide variant, or pharmaceutical composition thereof, will vary based on the particular circumstances, and an appropriate effective amount may be determined in each case of application by one of ordinary skill in the art using 20 only routine experimentation.

The presently-disclosed subject matter is further illustrated by the following specific but non-limiting examples. The following examples may include compilations of data that are representative of data gathered at various times during the 25 course of development and experimentation related to the presently-disclosed subject matter.

Examples

Embodiments of the BChE polypeptide variants of the presently-disclosed subject matter were made and studied using the following experimental procedure.

Site-directed mutagenesis of human BChE cDNA was performed by the QuikChange method of Braman, J.; Papworth, 35 C.; Greener, A. Methods Mol. Biol. 1996, 57, 5731, incorporated herein by this reference. Mutations were generated from wild-type human BChE in a pRc/CMV expression plasmid in accordance with Xie, W.; Altamirano, C. V.; Bartels, C. F.; 1999, 55, 83, each of which is incorporated herein by this reference. The expression plasmid pRc/CMV was kindly provided by Dr. O. Lockridge, University of Nebraska Medical Center (Omaha, Nebr.).

Using plasmid DNA as template and primers with specific 45 base-pair alterations, mutations were made by polymerase chain reaction with Pfu DNA polymerase, for replication fidelity. The PCR product was treated with Dpn I endonuclease to digest the parental DNA template. Cloned pfu DNA polymerase and Dpn I endonuclease were obtained from 50 Stratagene (La Jolla, Calif.). Modified plasmid DNA was transformed into *Escherichia coli*, amplified, and purified. The DNA sequences of the mutants were confirmed by DNA sequencing. All oligonucleotides were synthesized by the Integrated DNA Technologies, Inc. The QIAprep Spin Plas- 55 mid Miniprep Kit and Qiagen plasmid purification kit and QIAquick PCR purification kit were obtained from Qiagen (Santa Clarita, Calif.).

BChE mutants were expressed in human embryonic kidney cell line 293T/17. Cells were grown to 80-90% confluence in 6-well dishes and then transfected by Lipofectamine 2000 complexes of 4 µg plasmid DNA per each well. Cells were incubated at 37° C. in a CO_2 incubator for 24 hours and cells were moved to 60-mm culture vessel and cultured for four more days. The culture medium [10% fetal bovine serum in Dulbecco's modified Eagle's medium (DMEM)] was harvested for a BChE activity assay.

10

Human embryonic kidney 293T/17 cells were from ATCC (Manassas, Va.). Dulbecco's modified Eagle's medium (DMEM) was purchased from Fisher Scientific (Fairlawn, N.J.). Oligonucleotide primers were synthesized by the Integrated DNA Technologies and Analysis Facility of the University of Kentucky. 3,3',5,5'-Tetramethylbenzidine (TMB) was obtained from Sigma (Saint Louis, Mo.). Anti-butyrylcholinesterase (mouse monoclonal antibody, Product #HAH002-01) was purchased from Antibody Shop (Gentofte, Denmark) and Goat anti-mouse IgG HRP conjugate from Zymed (San Francisco, Calif.).

To measure cocaine and benzoic acid, the product of cocaine hydrolysis by BChE, sensitive radiometric assays based on toluene extraction of [3H]-(-)-cocaine labeled on its benzene ring were used in accordance with Zheng, F.; Yang, W.; Ko, M.-C.; Liu, J.; Cho, H.; Gao, D.; Tong, M.; Tai, H.-H.; Woods, J. H.; Zhan, C.-G. "Most Efficient Cocaine Hydrolase Designed by Virtual Screening of Transition States", J. Am. Chem. Soc. 2008, 130, 12148-12155, which is incorporated herein by this reference. ³H-(-)-cocaine (50 Ci/mmol) was purchased from PerkinElmer Life Sciences (Boston, Mass.).

In brief, to initiate reactions, 100 nCi of [³H]-(-)-cocaine was mixed with 100 µl of culture medium. Reactions proceeded at room temperature (25° C.) with varying concentrations of (-)-cocaine. Reactions were stopped by adding 300 µl of 0.02 M HCl, which neutralized the liberated benzoic acid while ensuring a positive charge on the residual cocaine. [3H]benzoic acid was extracted by 1 ml of toluene and measured by scintillation counting. Finally, the measured (-)cocaine concentration-dependent radiometric data were analyzed by using the standard Michaelis-Menten kinetics so that the catalytic efficiency (k_{cat}/K_M) was determined, along with the use of an enzyme-linked immunosorbent assay (ELISA) described by Zheng, F.; Yang, W.; Ko, M.-C.; Liu, J.; Cho, H.; Gao, D.; Tong, M.; Tai, H.-H.; Woods, J. H.; Zhan, C.-G. "Most Efficient Cocaine Hydrolase Designed by Virtual Screening of Transition States", J. Am. Chem. Soc. 2008, 130, 12148-12155.

The catalytic efficiencies (k_{cat}/K_M) of the BChE polypep-Speirs, R. J.; Cashman, J. R.; Lockridge, O. Mol. Pharmacol. 40 tide variants were found to be between about 1000 to about 4000 times the k_{cat}/K_M value $(9.11\times10^5 M^{-1} min^{-1})$ of the wild-type BChE.

Enzyme-linked immunosorbent assays (ELISA) were preformed as follows. The ELISA buffers used were the same as those described in the literature such as Brock, A.; Mortensen, V.; Loft, A. G. R.; Nergaard-Pedersen, B. J. Clin. Chem. Clin. Biochem. 1990, 28, 221-224; and Khattab, A. D.; Walker, C. H.; Johnston, G.; Siddiqui, M. K. Saphier, P. W. Environmental Toxicology and Chemistry 1994, 13, 1661-1667, both of which are incorporated herein by this reference. The coating buffer was 0.1 M sodium carbonate/bicarbonate buffer, pH 9.5. The diluent buffer (EIA buffer) was potassium phosphate monobasic/potassium phosphate monohydrate buffer, pH 7.5, containing 0.9% sodium chloride and 0.1% bovine serum albumin. The washing buffer (PBS-T) was 0.01 M potassium phosphate monobasic/potassium phosphate monohydrate buffer, pH 7.5, containing 0.05% (v/v) Tween-20. All the assays were performed in triplicate. Each well of an ELISA microtiter plate was filled with 100 µl of the mixture buffer consisting of 20 µl culture medium and 80 µl coating buffer. The plate was covered and incubated overnight at 4° C. to allow the antigen to bind to the plate. The solutions were then removed and the wells were washed four times with PBS-T. The washed wells were filled with 200 µl diluent buffer and kept shaking for 1.5 h at room temperature (25° C.). After washing with PBS-T for four times, the wells were filled with 100 μl antibody (1:8000) and were incubated for 1.5 h, fol-

lowed by washing for four times. Then, the wells were filled with 100 μl goat anti-mouse IgG HRP conjugate complex diluted to a final 1:3000 dilution, and were incubated at room temperature for 1.5 h, followed by washing for four times. The enzyme reactions were started by addition of 100 μl 5 substrate (TMB) solution. The reactions were stopped after 15 min by the addition of 100 μl of 2 M sulfuric acid, and the absorbance was read at 460 nm using a Bio-Rad ELISA plate reader.

While the terms used herein are believed to be well understood by one of ordinary skill in the art, the definitions set forth herein are provided to facilitate explanation of the presently-disclosed subject matter.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the presently-disclosed subject matter belongs. Although any methods, devices, and materials similar or equivalent to those described herein can be used in the practice or testing of the presently-disclosed subject matter, representative methods, devices, and materials are now described.

Following long-standing patent law convention, the terms "a", "an", and "the" refer to "one or more" when used in this application, including the claims. Thus, for example, reference to "a cell" includes a plurality of such cells, and so forth.

Unless otherwise indicated, all numbers expressing quantities of ingredients, properties such as reaction conditions,

and so forth used in the specification and claims are to be understood as being modified in all instances by the term "about". Accordingly, unless indicated to the contrary, the numerical parameters set forth in this specification and claims are approximations that can vary depending upon the desired properties sought to be obtained by the presently-disclosed subject matter.

12

As used herein, the term "about," when referring to a value or to an amount of mass, weight, time, volume, concentration or percentage is meant to encompass variations of in some embodiments $\pm 50\%$, in some embodiments $\pm 40\%$, in some embodiments $\pm 20\%$, in some embodiments $\pm 10\%$, in some embodiments $\pm 5\%$, in some embodiments $\pm 1\%$, in some embodiments $\pm 0.5\%$, and in some embodiments $\pm 0.1\%$ from the specified amount, as such variations are appropriate to perform the disclosed method.

As used herein, ranges can be expressed as from "about" one particular value, and/or to "about" another particular value. It is also understood that there are a number of values disclosed herein, and that each value is also herein disclosed as "about" that particular value in addition to the value itself. For example, if the value "10" is disclosed, then "about 10" is also disclosed. It is also understood that each unit between two particular units are also disclosed. For example, if 10 and 15 are disclosed, then 11, 12, 13, and 14 are also disclosed.

Throughout this document, various references are mentioned. All such references are incorporated herein by reference.

SEQUENCE LISTING

```
<160> NUMBER OF SEQ ID NOS: 46
<210> SEQ ID NO 1
<211> LENGTH: 1722
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthesized
<400> SEOUENCE: 1
gaagatgaca tcataattgc aacaaagaat ggaaaagtca gagggatgaa cttgacagtt
                                                                       60
tttggtggca cggtaacagc ctttcttgga attccctatg cacagccacc tcttggtaga
                                                                      120
cttcgattca aaaagccaca gtctctgacc aagtggtctg atatttggaa tgccacaaaa
tatgcaaatt cttgctgtca gaacatagat caaagttttc caggcttcca tggatcagag
                                                                      240
atgtggaacc caaacactga cctcagtgaa gactgtttat atctaaatgt atggattcca
                                                                      300
gcacctaaac caaaaaatgc cactgtattg atatggattt atggtggtgg ttttcaaact
                                                                      360
ggaacatcat ctttacatgt ttatgatggc aagtttctgg ctcgggttga aagagttatt
                                                                      420
qtaqtqtcaa tqaactataq qqtqqqtqcc ctaqqattct taqctttqcc aqqaaatcct
                                                                      480
gaggetecag ggaacatggg tttatttgat caacagttgg etetteagtg ggtteaaaaa
                                                                      540
aatatagcag cctttggtgg aaatcctaaa agtgtaactc tctttggaga aagttccgga
qcaqcttcaq ttaqcctqca tttqctttct cctqqaaqcc attcattqtt caccaqaqcc
                                                                      660
attotgoaaa gtggttooto caatgotoot tgggoggtaa catototta tgaagotagg
                                                                      720
                                                                      780
aacaqaacqt tqaacttaqc taaattqact qqttqctcta qaqaqaatqa qactqaaata
atcaagtgtc ttagaaataa agatccccaa gaaattcttc tgaatgaagc atttgttgtc
                                                                      840
                                                                      900
ccctatqqqa ctcctttqqq tqtaaacttt qqtccqaccq tqqatqqtqa ttttctcact
```

-continued
-continued

												COII	СТП	aca		
gaca	atgc	cag a	acata	attac	ct t	gaact	tgga	ı caa	attta	aaaa	aaac	cca	gat 1	tttg	gtgggi	t 960
gtta	aataa	aag a	atgaa	aggga	ac at	tggtt	ttta	gto	ggtg	ggtg	ctco	ctgg	ctt «	cagca	aaagat	t 1020
aaca	aataq	gta 1	tcata	aacta	ag a	aaaga	aattt	caç	ggaag	ggtt	taaa	aaata	att 1	tttt	ccagga	a 1080
gtga	agtga	agt 1	tgga	aaagg	ga at	tccat	cctt	ttt	catt	aca	caga	actg	ggt	agato	gatca	g 1140
agad	cctga	aaa a	actad	ccgt	ga g	gcctt	gggt	gat	gtt	gttg	ggga	attat	caa 1	tttca	atatgo	c 1200
cct	geett	gg a	agtto	cacca	aa ga	aagtt	ctca	ı gaa	atgg	ggaa	ataa	atgc	ctt 1	tttct	cactai	t 1260
tttç	gaaca	acc q	gatco	ctcca	aa a	cttco	gtgg	g cca	agaat	gga	tggg	gagt	gat	gcato	ggctai	t 1320
gaaa	attga	aat 1	tgto	ctttç	gg ti	ttaco	etete	g gaa	aagaa	agag	ataa	atta	cac a	aaaa	gccga	g 1380
gaaa	attt	ga 🤅	gtaga	atcca	at a	gtgaa	aacgg	j tgg	ggcaa	att	ttg	caaaa	ata 1	tggga	aatcca	a 1440
aat	gagad	ctc a	agaad	caata	ag ca	acaa	gctgc	g cct	gtct	tca	aaaq	gcact	ga a	acaa	aaatai	t 1500
ctaa	acct	ga a	ataca	agagt	c a	acaa	gaata	ato	gacga	aaac	taco	gtgct	ca a	acaat	gtcg	a 1560
ttct	ggad	cat o	cattt	ttt	cc a	aaagt	cttg	g gaa	aatga	acag	gaaa	atatt	ga 1	tgaaq	gcagaa	a 1620
tgg	gagto	gga a	aagca	aggat	t c	catco	getge	, aac	caatt	aca	tgat	gga	ctg	gaaa	aatcaa	a 1680
ttta	aacga	att a	acact	cagca	aa ga	aaaga	aaagt	. tgt	gtg	ggtc	tc					1722
<213 <213 <213 <220	L> LI 2> T: 3> OI 0> FI	ENGTI PE: RGAN: EATUI	ISM: RE:	74 Art:			Seque nthes		ā.							
< 400)> SI	EQUEI	NCE:	2												
Glu 1	Asp	Asp	Ile	Ile 5	Ile	Ala	Thr	Lys	Asn 10	Gly	ГÀЗ	Val	Arg	Gly 15	Met	
Asn	Leu	Thr	Val 20	Phe	Gly	Gly	Thr	Val 25	Thr	Ala	Phe	Leu	Gly 30	Ile	Pro	
Tyr	Ala	Gln 35	Pro	Pro	Leu	Gly	Arg 40	Leu	Arg	Phe	Lys	Lys 45	Pro	Gln	Ser	
Leu	Thr 50	Lys	Trp	Ser	Asp	Ile 55	Trp	Asn	Ala	Thr	60 Lys	Tyr	Ala	Asn	Ser	
Сув 65	Càa	Gln	Asn	Ile	Asp 70	Gln	Ser	Phe	Pro	Gly 75	Phe	His	Gly	Ser	Glu 80	
Met	Trp	Asn	Pro	Asn 85	Thr	Asp	Leu	Ser	Glu 90	Asp	Cys	Leu	Tyr	Leu 95	Asn	
Val	Trp	Ile	Pro 100	Ala	Pro	Lys	Pro	Lys 105	Asn	Ala	Thr	Val	Leu 110	Ile	Trp	
Ile	Tyr	Gly 115	Gly	Gly	Phe	Gln	Thr 120	Gly	Thr	Ser	Ser	Leu 125	His	Val	Tyr	
Asp	Gly 130	Lys	Phe	Leu	Ala	Arg 135	Val	Glu	Arg	Val	Ile 140	Val	Val	Ser	Met	
Asn 145	Tyr	Arg	Val	Gly	Ala 150	Leu	Gly	Phe	Leu	Ala 155	Leu	Pro	Gly	Asn	Pro 160	
Glu	Ala	Pro	Gly	Asn 165	Met	Gly	Leu	Phe	Asp 170	Gln	Gln	Leu	Ala	Leu 175	Gln	
Trp	Val	Gln	Lys 180	Asn	Ile	Ala	Ala	Phe 185	Gly	Gly	Asn	Pro	Lys 190	Ser	Val	
Thr	Leu	Phe 195	Gly	Glu	Ser	Ser	Gly 200	Ala	Ala	Ser	Val	Ser 205	Leu	His	Leu	

Leu Ser Pro Gly Ser His Ser Leu Phe Thr Arg Ala Ile Leu Gln Ser 210 215 220

-continued

Gly Ser Ser Asn Ala Pro Trp Ala Val Thr Ser Leu Tyr Glu Ala Arg Asn Arg Thr Leu Asn Leu Ala Lys Leu Thr Gly Cys Ser Arg Glu Asn Glu Thr Glu Ile Ile Lys Cys Leu Arg Asn Lys Asp Pro Gln Glu Ile Leu Leu Asn Glu Ala Phe Val Val Pro Tyr Gly Thr Pro Leu Gly Val Asn Phe Gly Pro Thr Val Asp Gly Asp Phe Leu Thr Asp Met Pro Asp Ile Leu Leu Glu Leu Gly Gln Phe Lys Lys Thr Gln Ile Leu Val Gly 310 Val Asn Lys Asp Glu Gly Thr Trp Phe Leu Val Gly Gly Ala Pro Gly 330 Phe Ser Lys Asp Asn Asn Ser Ile Ile Thr Arg Lys Glu Phe Gln Glu 345 Gly Leu Lys Ile Phe Phe Pro Gly Val Ser Glu Phe Gly Lys Glu Ser 360 Ile Leu Phe His Tyr Thr Asp Trp Val Asp Asp Gln Arg Pro Glu Asn Tyr Arg Glu Ala Leu Gly Asp Val Val Gly Asp Tyr Asn Phe Ile Cys 395 Pro Ala Leu Glu Phe Thr Lys Lys Phe Ser Glu Trp Gly Asn Asn Ala 410 Phe Phe Tyr Tyr Phe Glu His Arg Ser Ser Lys Leu Pro Trp Pro Glu Trp Met Gly Val Met His Gly Tyr Glu Ile Glu Phe Val Phe Gly Leu Pro Leu Glu Arg Arg Asp Asn Tyr Thr Lys Ala Glu Glu Ile Leu Ser 455 Arg Ser Ile Val Lys Arg Trp Ala Asn Phe Ala Lys Tyr Gly Asn Pro Asn Glu Thr Gln Asn Asn Ser Thr Ser Trp Pro Val Phe Lys Ser Thr Glu Gln Lys Tyr Leu Thr Leu Asn Thr Glu Ser Thr Arg Ile Met Thr Lys Leu Arg Ala Gln Gln Cys Arg Phe Trp Thr Ser Phe Phe Pro Lys 520 Val Leu Glu Met Thr Gly Asn Ile Asp Glu Ala Glu Trp Glu Trp Lys 535 540 Ala Gly Phe His Arg Trp Asn Asn Tyr Met Met Asp Trp Lys Asn Gln 550 555 Phe Asn Asp Tyr Thr Ser Lys Lys Glu Ser Cys Val Gly Leu <210> SEQ ID NO 3 <211> LENGTH: 1722 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: synthesized <400> SEQUENCE: 3 gaagatgaca tcataattgc aacaaagaat ggaaaagtca gagggatgaa cttgacagtt

60

tttqqtqqca cqqtaacaqc ctttcttqqa attccctatq cacaqccacc tcttqqtaqa

				-contir	nued			
cttcgattca a	aaaagccaca	gtctctgacc	aagtggtctg	atatttggaa	tgccacaaaa	180		
tatgcaaatt (cttgctgtca	gaacatagat	caaagttttc	caggetteca	tggatcagag	240		
atgtggaacc (caaacactga	cctcagtgaa	gactgtttat	atctaaatgt	atggattcca	300		
gcacctaaac (caaaaaatgc	cactgtattg	atatggattt	atggtggtgg	ttttcaaact	360		
ggaacatcat (ctttacatgt	ttatgatggc	aagtttctgg	ctcgggttga	aagagttatt	420		
gtagtgtcaa 1	tgaactatag	ggtgggtgcc	ctaggattct	tagctttgcc	aggaaatcct	480		
gaggctccag (ggaacatggg	tttatttgat	caacagttgg	ctcttcagtg	ggttcaaaaa	540		
aatatagcag (cctttggtgg	aaatcctaaa	agtgtaactc	tctttggaga	aagttccgga	600		
gcagcttcag 1	ttagcctgca	tttgctttct	cctggaagcc	attcattgtt	caccagagcc	660		
attctgcaaa q	gtggttccac	caatgctcct	tgggcggtaa	catctcttta	tgaagctagg	720		
aacagaacgt 1	tgaacttagc	taaattgact	ggttgctcta	gagagaatga	gactgaaata	780		
atcaagtgtc 1	ttagaaataa	agatccccaa	gaaattcttc	tgaatgaagc	atttgttgtc	840		
ccctatggga (ctcctttggg	tgtaaacttt	ggtccgaccg	tggatggtga	ttttctcact	900		
gacatgccag a	acatattact	tgaacttgga	caatttaaaa	aaacccagat	tttggtgggt	960		
gttaataaag a	atgaagggac	atggttttta	gtcggtggtg	ctcctggctt	cagcaaagat	1020		
aacaatagta 1	tcataactag	aaaagaattt	caggaaggtt	taaaaatatt	ttttccagga	1080		
gtgagtgagt 1	ttggaaagga	atccatcctt	tttcattaca	cagactgggt	agatgatcag	1140		
agacctgaaa a	actaccgtga	ggccttgggt	gatgttgttg	gggattataa	tttcatatgc	1200		
cctgccttgg a	agttcaccaa	gaagttctca	gaatggggaa	ataatgcctt	tttctactat	1260		
tttgaacacc (gatcctccaa	acttccgtgg	ccagaatgga	tgggagtgat	gcatggctat	1320		
gaaattgaat 1	ttgtctttgg	tttacctctg	gaaagaagag	ataattacac	aaaagccgag	1380		
gaaattttga 🤉	gtagatccat	agtgaaacgg	tgggcaaatt	ttgcaaaata	tgggaatcca	1440		
aatgagactc a	agaacaatag	cacaagctgg	cctgtcttca	aaagcactga	acaaaaatat	1500		
ctaaccttga a	atacagagtc	aacaagaata	atgacgaaac	tacgtgctca	acaatgtcga	1560		
ttctggacat «	cattttttcc	aaaagtcttg	gaaatgacag	gaaatattga	tgaagcagaa	1620		
tgggagtgga a	aagcaggatt	ccatcgctgg	aacaattaca	tgatggactg	gaaaaatcaa	1680		
tttaacgatt a	acactagcaa	gaaagaaagt	tgtgtgggtc	tc		1722		
<210> SEQ II <211> LENGTI <212> TYPE: <213> ORGAN: <220> FEATUI	H: 574 PRT ISM: Artifi	icial Sequer	nce					

<400> SEQUENCE: 4

Glu Asp Asp Ile Ile Ile Ala Thr Lys Asn Gly Lys Val Arg Gly Met 1 5 10 15

Tyr Ala Gln Pro Pro Leu Gly Arg Leu Arg Phe Lys Lys Pro Gln Ser 35 40 45

Leu Thr Lys Trp Ser Asp Ile Trp Asn Ala Thr Lys Tyr Ala Asn Ser 50 $\,$

Cys Cys Gln Asn Ile Asp Gln Ser Phe Pro Gly Phe His Gly Ser Glu 65 70 75 80

<223> OTHER INFORMATION: synthesized

Met	Trp	Asn	Pro	Asn 85	Thr	Asp	Leu	Ser	Glu 90	Asp	Cys	Leu	Tyr	Leu 95	Asn
Val	Trp	Ile	Pro	Ala	Pro	Lys	Pro	Lys 105	Asn	Ala	Thr	Val	Leu 110	Ile	Trp
Ile	Tyr	Gly 115	Gly	Gly	Phe	Gln	Thr 120	Gly	Thr	Ser	Ser	Leu 125	His	Val	Tyr
Asp	Gly 130	Lys	Phe	Leu	Ala	Arg 135	Val	Glu	Arg	Val	Ile 140	Val	Val	Ser	Met
Asn 145	Tyr	Arg	Val	Gly	Ala 150	Leu	Gly	Phe	Leu	Ala 155	Leu	Pro	Gly	Asn	Pro 160
Glu	Ala	Pro	Gly	Asn 165	Met	Gly	Leu	Phe	Asp 170	Gln	Gln	Leu	Ala	Leu 175	Gln
Trp	Val	Gln	Lys 180	Asn	Ile	Ala	Ala	Phe 185	Gly	Gly	Asn	Pro	Lys 190	Ser	Val
Thr	Leu	Phe 195	Gly	Glu	Ser	Ser	Gly 200	Ala	Ala	Ser	Val	Ser 205	Leu	His	Leu
Leu	Ser 210	Pro	Gly	Ser	His	Ser 215	Leu	Phe	Thr	Arg	Ala 220	Ile	Leu	Gln	Ser
Gly 225	Ser	Thr	Asn	Ala	Pro 230	Trp	Ala	Val	Thr	Ser 235	Leu	Tyr	Glu	Ala	Arg 240
Asn	Arg	Thr	Leu	Asn 245	Leu	Ala	Lys	Leu	Thr 250	Gly	CAa	Ser	Arg	Glu 255	Asn
Glu	Thr	Glu	Ile 260	Ile	ГÀа	CÀa	Leu	Arg 265	Asn	Lys	Asp	Pro	Gln 270	Glu	Ile
Leu	Leu	Asn 275	Glu	Ala	Phe	Val	Val 280	Pro	Tyr	Gly	Thr	Pro 285	Leu	Gly	Val
Asn	Phe 290	Gly	Pro	Thr	Val	Asp 295	Gly	Asp	Phe	Leu	Thr 300	Asp	Met	Pro	Asp
Ile 305	Leu	Leu	Glu	Leu	Gly 310	Gln	Phe	ГÀЗ	ГÀЗ	Thr 315	Gln	Ile	Leu	Val	Gly 320
Val	Asn	ГÀЗ	Asp	Glu 325	Gly	Thr	Trp	Phe	Leu 330	Val	Gly	Gly	Ala	Pro 335	Gly
Phe	Ser	ГÀЗ	Asp 340	Asn	Asn	Ser	Ile	Ile 345	Thr	Arg	ГÀа	Glu	Phe 350	Gln	Glu
Gly	Leu	Lys 355	Ile	Phe	Phe	Pro	Gly 360	Val	Ser	Glu	Phe	Gly 365	ГЛЗ	Glu	Ser
Ile	Leu 370	Phe	His	Tyr	Thr	Asp 375	Trp	Val	Asp	Asp	Gln 380	Arg	Pro	Glu	Asn
Tyr 385	Arg	Glu	Ala	Leu	Gly 390	Asp	Val	Val	Gly	Asp 395	Tyr	Asn	Phe	Ile	Cys 400
Pro	Ala	Leu	Glu	Phe 405	Thr	ГÀв	Lys	Phe	Ser 410	Glu	Trp	Gly	Asn	Asn 415	Ala
Phe	Phe	Tyr	Tyr 420	Phe	Glu	His	Arg	Ser 425	Ser	ГÀа	Leu	Pro	Trp 430	Pro	Glu
Trp	Met	Gly 435	Val	Met	His	Gly	Tyr 440	Glu	Ile	Glu	Phe	Val 445	Phe	Gly	Leu
Pro	Leu 450	Glu	Arg	Arg	Asp	Asn 455	Tyr	Thr	Lys	Ala	Glu 460	Glu	Ile	Leu	Ser
Arg 465	Ser	Ile	Val	Lys	Arg 470	Trp	Ala	Asn	Phe	Ala 475	Lys	Tyr	Gly	Asn	Pro 480
Asn	Glu	Thr	Gln	Asn 485	Asn	Ser	Thr	Ser	Trp 490	Pro	Val	Phe	Lys	Ser 495	Thr
Glu	Gln	Lys	Tyr 500	Leu	Thr	Leu	Asn	Thr 505	Glu	Ser	Thr	Arg	Ile 510	Met	Thr

60

21 22

Lys Leu Arg Ala Gln Gln Cys Arg Phe Trp Thr Ser Phe Phe Pro Lys

515 520 525

Val Leu Glu Met Thr Gly Asn Ile Asp Glu Ala Glu Trp Glu Trp Lys 530 540

Ala Gly Phe His Arg Trp Asn Asn Tyr Met Met Asp Trp Lys Asn Gln 545 550 555 560

gaagatgaca tcataattgc aacaaagaat ggaaaagtca gagggatgaa cttgacagtt

Phe Asn Asp Tyr Thr Ser Lys Lys Glu Ser Cys Val Gly Leu

<210> SEQ ID NO 5

<211> LENGTH: 1722

<212> TYPE: DNA

<213 > ORGANISM: Artificial Sequence

<220> FEATURE:

<223 > OTHER INFORMATION: synthesized

<400> SEQUENCE: 5

tttggtggca cggtaacagc ctttcttgga attccctatg cacagccacc tcttggtaga 120 cttcgattca aaaagccaca gtctctgacc aagtggtctg atatttggaa tgccacaaaa 180 tatgcaaatt cttgctgtca gaacatagat caaagttttc caggcttcca tggatcagag 240 atgtggaacc caaacactga cctcagtgaa gactgtttat atctaaatgt atggattcca 300 gcacctaaac caaaaaatgc cactgtattg atatggattt atggtggtgg ttttcaaact 360 ggaacatcat ctttacatgt ttatgatggc aagtttctgg ctcgggttga aagagttatt 420 gtagtgtcaa tgaactatag ggtgggtgcc ctaggattct tagctttgcc aggaaatcct 480 gaggetecag ggaacatggg tttatttgat caacagttgg etetteagtg ggtteaaaaa 540 aatatagcag cetttggtgg aaateetaaa agtgtaacte tetttggaga aagtteegga 600 gcagcttcag ttagcctgca tttgctttct cctggaagcc attcattgtt caccagagcc 660 attotgoaaa gtggttocat gaatgotoot tgggoggtaa catototta tgaagotagg 720 aacagaacgt tgaacttagc taaattgact ggttgctcta gagagaatga gactgaaata atcaagtgtc ttagaaataa agatccccaa gaaattcttc tgaatgaagc atttgttgtc 840 ccctatggga ctcctttggg tgtaaacttt ggtccgaccg tggatggtga ttttctcact 900 gacatgccag acatattact tgaacttgga caatttaaaa aaacccagat tttggtgggt 960 1020 gttaataaag atgaagggac atggttttta gtcggtggtg ctcctggctt cagcaaagat 1080 aacaatagta tcataactag aaaagaattt caggaaggtt taaaaatatt ttttccagga qtqaqtqaqt ttqqaaaqqa atccatcctt tttcattaca caqactqqqt aqatqatcaq 1140 1200 agacctgaaa actaccgtga ggccttgggt gatgttgttg gggattataa tttcatatgc cctgccttgg agttcaccaa gaagttctca gaatggggaa ataatgcctt tttctactat 1260 tttgaacacc gatcctccaa acttccgtgg ccagaatgga tgggagtgat gcatggctat 1320 gaaattgaat ttgtctttgg tttacctctg gaaagaagag ataattacac aaaagccgag 1380 gaaattttga gtagatccat agtgaaacgg tgggcaaatt ttgcaaaata tgggaatcca 1440 aatgagactc agaacaatag cacaagctgg cctgtcttca aaagcactga acaaaaatat 1500 ctaaccttga atacagagtc aacaagaata atgacgaaac tacgtgctca acaatgtcga 1560 ttctggacat catttttcc aaaagtcttg gaaatgacag gaaatattga tgaagcagaa 1620 tgggagtgga aagcaggatt ccatcgctgg aacaattaca tgatggactg gaaaaatcaa 1680 tttaacgatt acactagcaa gaaagaaagt tgtgtgggtc tc 1722

-continued

<211 <212 <213 <220	L> LE 2> TY 3> OF 0> FE	EQ II ENGTH YPE: RGANI EATUH THER	H: 57 PRT [SM: RE:	74 Art:			=		ā.						
< 400)> SI	EQUE	ICE :	6											
Glu 1	Asp	Asp	Ile	Ile 5	Ile	Ala	Thr	Lys	Asn 10	Gly	Lys	Val	Arg	Gly 15	Met
Asn	Leu	Thr	Val 20	Phe	Gly	Gly	Thr	Val 25	Thr	Ala	Phe	Leu	Gly 30	Ile	Pro
Tyr	Ala	Gln 35	Pro	Pro	Leu	Gly	Arg 40	Leu	Arg	Phe	Lys	Lys 45	Pro	Gln	Ser
Leu	Thr 50	Lys	Trp	Ser	Asp	Ile 55	Trp	Asn	Ala	Thr	F F F	Tyr	Ala	Asn	Ser
Cys 65	Cys	Gln	Asn	Ile	Asp 70	Gln	Ser	Phe	Pro	Gly 75	Phe	His	Gly	Ser	Glu 80
Met	Trp	Asn	Pro	Asn 85	Thr	Asp	Leu	Ser	Glu 90	Asp	Cys	Leu	Tyr	Leu 95	Asn
Val	Trp	Ile	Pro 100	Ala	Pro	Lys	Pro	Lys 105	Asn	Ala	Thr	Val	Leu 110	Ile	Trp
Ile	Tyr	Gly 115	Gly	Gly	Phe	Gln	Thr 120	Gly	Thr	Ser	Ser	Leu 125	His	Val	Tyr
Asp	Gly 130	Lys	Phe	Leu	Ala	Arg 135	Val	Glu	Arg	Val	Ile 140	Val	Val	Ser	Met
Asn 145	Tyr	Arg	Val	Gly	Ala 150	Leu	Gly	Phe	Leu	Ala 155	Leu	Pro	Gly	Asn	Pro 160
Glu	Ala	Pro	Gly	Asn 165	Met	Gly	Leu	Phe	Asp 170	Gln	Gln	Leu	Ala	Leu 175	Gln
Trp	Val	Gln	Lys 180	Asn	Ile	Ala	Ala	Phe 185	Gly	Gly	Asn	Pro	Lys 190	Ser	Val
Thr	Leu	Phe 195	Gly	Glu	Ser	Ser	Gly 200	Ala	Ala	Ser	Val	Ser 205	Leu	His	Leu
Leu	Ser 210	Pro	Gly	Ser	His	Ser 215	Leu	Phe	Thr	Arg	Ala 220	Ile	Leu	Gln	Ser
Gly 225	Ser	Met	Asn	Ala	Pro 230	Trp	Ala	Val	Thr	Ser 235	Leu	Tyr	Glu	Ala	Arg 240
Asn	Arg	Thr		Asn 245					Thr 250		СЛа	Ser		Glu 255	
Glu	Thr	Glu	Ile 260	Ile	Lys	Cys	Leu	Arg 265	Asn	Lys	Asp	Pro	Gln 270	Glu	Ile
Leu	Leu	Asn 275	Glu	Ala	Phe	Val	Val 280	Pro	Tyr	Gly	Thr	Pro 285	Leu	Gly	Val
Asn	Phe 290	Gly	Pro	Thr	Val	Asp 295	Gly	Asp	Phe	Leu	Thr 300	Asp	Met	Pro	Asp
Ile 305	Leu	Leu	Glu	Leu	Gly 310	Gln	Phe	Lys	Lys	Thr 315	Gln	Ile	Leu	Val	Gly 320
Val	Asn	Lys	Asp	Glu 325	Gly	Thr	Trp	Phe	Leu 330	Val	Gly	Gly	Ala	Pro 335	Gly
Phe	Ser	Lys	Asp 340	Asn	Asn	Ser	Ile	Ile 345	Thr	Arg	Lys	Glu	Phe 350	Gln	Glu
Gly	Leu	Lys 355	Ile	Phe	Phe	Pro	Gly 360	Val	Ser	Glu	Phe	Gly 365	Lys	Glu	Ser

-continued

											_	con	tin	ued	
Ile	Leu 370	Phe	His	Tyr	Thr	Asp 375	Trp	Val	Asp	Asp	Gln 380	Arg	Pro	Glu	Asn
Tyr 385	Arg	Glu	Ala	Leu	Gly 390	Asp	Val	Val	Gly	Asp 395	Tyr	Asn	Phe	Ile	Сув 400
Pro	Ala	Leu	Glu	Phe 405	Thr	ГÀа	Lys	Phe	Ser 410	Glu	Trp	Gly	Asn	Asn 415	Ala
Phe	Phe	Tyr	Tyr 420	Phe	Glu	His	Arg	Ser 425	Ser	ГÀа	Leu	Pro	Trp 430	Pro	Glu
Trp	Met	Gly 435	Val	Met	His	Gly	Tyr 440	Glu	Ile	Glu	Phe	Val 445	Phe	Gly	Leu
Pro	Leu 450	Glu	Arg	Arg	Asp	Asn 455	Tyr	Thr	ГЛа	Ala	Glu 460	Glu	Ile	Leu	Ser
Arg 465	Ser	Ile	Val	ГÀа	Arg 470	Trp	Ala	Asn	Phe	Ala 475	ГÀа	Tyr	Gly	Asn	Pro 480
Asn	Glu	Thr	Gln	Asn 485	Asn	Ser	Thr	Ser	Trp 490	Pro	Val	Phe	Lys	Ser 495	Thr
Glu	Gln	Lys	Tyr 500	Leu	Thr	Leu	Asn	Thr 505	Glu	Ser	Thr	Arg	Ile 510	Met	Thr
ГЛа	Leu	Arg 515	Ala	Gln	Gln	Cys	Arg 520	Phe	Trp	Thr	Ser	Phe 525	Phe	Pro	Lys
Val	Leu 530	Glu	Met	Thr	Gly	Asn 535	Ile	Asp	Glu	Ala	Glu 540	Trp	Glu	Trp	Lys
Ala 545	Gly	Phe	His	Arg	Trp 550	Asn	Asn	Tyr	Met	Met 555	Asp	Trp	Lys	Asn	Gln 560
Phe	Asn	Asp	Tyr	Thr 565	Ser	Lys	ГÀа	Glu	Ser 570	CAa	Val	Gly	Leu		
<213 <213 <213 <220 <223)> FI	ENGTI PE: RGAN: EATUI THER	H: 1' DNA ISM: RE: INFO	722 Art: DRMA'	ific: FION		_		đ						
					_		-			_			_	_	acagtt

60 tttggtggca cggtaacagc ctttcttgga attccctatg cacagccacc tcttggtaga cttcgattca aaaagccaca gtctctgacc aagtggtctg atatttggaa tgccacaaaa tatgcaaatt cttgctgtca gaacatagat caaagttttc caggcttcca tggatcagag 240 atgtggaacc caaacactga cctcagtgaa gactgtttat atctaaatgt atggattcca 300 gcacctaaac caaaaaatgc cactgtattg atatggattt atggtggtgg ttttcaaact 360 ggaacatcat ctttacatgt ttatgatggc aagtttctgg ctcgggttga aagagttatt 420 gtagtgtcaa tgaactatag ggtgggtgcc ctaggattct tagctttgcc aggaaatcct 480 gaggctccag ggaacatggg tttatttgat caacagttgg ctcttcagtg ggttcaaaaa 540 aatatagcag cctttggtgg aaatcctaaa agtgtaactc tctttggaga aagttccgga 600 gcagcttcag ttagcctgca tttgctttct cctggaagcc attcattgtt caccagagcc 660 attctgcaaa gtggttcctg caatgctcct tgggcggtaa catctcttta tgaagctagg 720 780 aacagaacgt tgaacttagc taaattgact ggttgctcta gagagaatga gactgaaata atcaagtgtc ttagaaataa agatccccaa gaaattcttc tgaatgaagc atttgttgtc 840 ccctatggga ctcctttggg tgtaaacttt ggtccgaccg tggatggtga ttttctcact 900 gacatgccag acatattact tgaacttgga caatttaaaa aaacccagat tttggtgggt 960

-continued

gttaataaag atgaagggac atggttttta gtcggtggtg ctcctggctt cagcaaagat	1020
aacaatagta tcataactag aaaagaattt caggaaggtt taaaaatatt ttttccagga	1080
gtgagtgagt ttggaaagga atccatcctt tttcattaca cagactgggt agatgatcag	1140
agacctgaaa actaccgtga ggccttgggt gatgttgttg gggattataa tttcatatgc	1200
cetgeettgg agtteaceaa gaagttetea gaatggggaa ataatgeett tttetaetat	1260
tttgaacacc gatcctccaa acttccgtgg ccagaatgga tgggagtgat gcatggctat	1320
gaaattgaat ttgtctttgg tttacctctg gaaagaagag ataattacac aaaagccgag	1380
gaaattttga gtagatccat agtgaaacgg tgggcaaatt ttgcaaaata tgggaatcca	1440
aatgagactc agaacaatag cacaagctgg cctgtcttca aaagcactga acaaaaatat	1500
ctaaccttga atacagagtc aacaagaata atgacgaaac tacgtgctca acaatgtcga	1560
ttctggacat cattttttcc aaaagtcttg gaaatgacag gaaatattga tgaagcagaa	1620
tgggagtgga aagcaggatt ccatcgctgg aacaattaca tgatggactg gaaaaatcaa	1680
tttaacgatt acactagcaa gaaagaaagt tgtgtgggtc tc	1722
<pre><210> SEQ ID NO 8 <211> LENGTH: 574 <212> TYPE: PRT <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: synthesized</pre>	
<400> SEQUENCE: 8	
Glu Asp Asp Ile Ile Ile Ala Thr Lys Asn Gly Lys Val Arg Gly Met 1 10 15	
Asn Leu Thr Val Phe Gly Gly Thr Val Thr Ala Phe Leu Gly Ile Pro 20 25 30	
Tyr Ala Gln Pro Pro Leu Gly Arg Leu Arg Phe Lys Lys Pro Gln Ser 35 40 45	
Leu Thr Lys Trp Ser Asp Ile Trp Asn Ala Thr Lys Tyr Ala Asn Ser 50 55 60	
Cys Cys Gln Asn Ile Asp Gln Ser Phe Pro Gly Phe His Gly Ser Glu 65 70 75 80	
Met Trp Asn Pro Asn Thr Asp Leu Ser Glu Asp Cys Leu Tyr Leu Asn 85 90 95	
Val Trp Ile Pro Ala Pro Lys Pro Lys Asn Ala Thr Val Leu Ile Trp 100 105 110	
Ile Tyr Gly Gly Gly Phe Gln Thr Gly Thr Ser Ser Leu His Val Tyr 115 120 125	
Asp Gly Lys Phe Leu Ala Arg Val Glu Arg Val Ile Val Val Ser Met 130 135 140	
Asn Tyr Arg Val Gly Ala Leu Gly Phe Leu Ala Leu Pro Gly Asn Pro 145 150 155 160	
Glu Ala Pro Gly Asn Met Gly Leu Phe Asp Gln Gln Leu Ala Leu Gln 165 170 175	
Trp Val Gln Lys Asn Ile Ala Ala Phe Gly Gly Asn Pro Lys Ser Val	
Thr Leu Phe Gly Glu Ser Ser Gly Ala Ala Ser Val Ser Leu His Leu 195 200 205	
Leu Ser Pro Gly Ser His Ser Leu Phe Thr Arg Ala Ile Leu Gln Ser	

Gly Ser Cys Asn Ala Pro Trp Ala Val Thr Ser Leu Tyr Glu Ala Arg

-continued

												con	tin	ued		
225					230					235					240	
Asn	Arg	Thr	Leu	Asn 245	Leu	Ala	ГÀа	Leu	Thr 250	Gly	CAa	Ser	Arg	Glu 255	Asn	
Glu	Thr	Glu	Ile 260	Ile	Lys	CAa	Leu	Arg 265	Asn	Lys	Asp	Pro	Gln 270	Glu	Ile	
Leu	Leu	Asn 275	Glu	Ala	Phe	Val	Val 280	Pro	Tyr	Gly	Thr	Pro 285	Leu	Gly	Val	
Asn	Phe 290	Gly	Pro	Thr	Val	Asp 295	Gly	Asp	Phe	Leu	Thr 300	Asp	Met	Pro	Asp	
Ile 305	Leu	Leu	Glu	Leu	Gly 310	Gln	Phe	ГЛа	ГÀа	Thr 315	Gln	Ile	Leu	Val	Gly 320	
Val	Asn	ГÀа	Asp	Glu 325	Gly	Thr	Trp	Phe	Leu 330	Val	Gly	Gly	Ala	Pro 335	Gly	
Phe	Ser	Lys	Asp 340	Asn	Asn	Ser	Ile	Ile 345	Thr	Arg	Lys	Glu	Phe 350	Gln	Glu	
Gly	Leu	Lys 355	Ile	Phe	Phe	Pro	Gly 360	Val	Ser	Glu	Phe	Gly 365	Lys	Glu	Ser	
Ile	Leu 370	Phe	His	Tyr	Thr	Asp 375	Trp	Val	Asp	Asp	Gln 380	Arg	Pro	Glu	Asn	
Tyr 385	Arg	Glu	Ala	Leu	Gly 390	Asp	Val	Val	Gly	Asp 395	Tyr	Asn	Phe	Ile	Cys 400	
Pro	Ala	Leu	Glu	Phe 405	Thr	Lys	Lys	Phe	Ser 410	Glu	Trp	Gly	Asn	Asn 415	Ala	
Phe	Phe	Tyr	Tyr 420	Phe	Glu	His	Arg	Ser 425	Ser	Lys	Leu	Pro	Trp 430	Pro	Glu	
Trp	Met	Gly 435	Val	Met	His	Gly	Tyr 440	Glu	Ile	Glu	Phe	Val 445	Phe	Gly	Leu	
Pro	Leu 450	Glu	Arg	Arg	Asp	Asn 455	Tyr	Thr	ГÀз	Ala	Glu 460	Glu	Ile	Leu	Ser	
Arg 465	Ser	Ile	Val	Lys	Arg 470	Trp	Ala	Asn	Phe	Ala 475	rys	Tyr	Gly	Asn	Pro 480	
Asn	Glu	Thr	Gln	Asn 485	Asn	Ser	Thr	Ser	Trp 490	Pro	Val	Phe	ГÀз	Ser 495	Thr	
Glu	Gln	Lys	Tyr 500	Leu	Thr	Leu	Asn	Thr 505	Glu	Ser	Thr	Arg	Ile 510	Met	Thr	
Lys	Leu	Arg 515	Ala	Gln	Gln	CAa	Arg 520	Phe	Trp	Thr	Ser	Phe 525	Phe	Pro	Lys	
Val	Leu 530	Glu	Met	Thr	Gly	Asn 535	Ile	Asp	Glu	Ala	Glu 540	Trp	Glu	Trp	ГÀа	
Ala 545	Gly	Phe	His	Arg	Trp 550	Asn	Asn	Tyr	Met	Met 555	Asp	Trp	Lys	Asn	Gln 560	
Phe	Asn	Asp	Tyr	Thr 565	Ser	Lys	Lys	Glu	Ser 570	Cys	Val	Gly	Leu			
<213 <213 <213 <220 <223	D> SI 1> LI 2> TY 3> OF D> FF 0> SI D> SI	ENGTI PE: RGAN EATUI THER	H: 1' DNA ISM: RE: INFO	722 Art: DRMA			-		Ē							
gaa	gatga	aca t	cata	aatto	gc aa	acaa	agaat	gga	aaaq	gtca	gag	ggat	gaa (cttg	acagtt 60	
ttt	ggtgg	gca d	eggta	aacaq	gc ct	ttc	ttgga	a att	ccct	atg	caca	agcca	acc 1	cctt	ggtaga 120	

cttcgattca aaaagccaca gtctctgacc aagtggtctg atatttggaa tgccacaaaa

-continued

tatgcaaatt	cttgctgtca	gaacatagat	caaagttttc	caggetteca	tggatcagag	240
atgtggaacc	caaacactga	cctcagtgaa	gactgtttat	atctaaatgt	atggattcca	300
gcacctaaac	caaaaaatgc	cactgtattg	atatggattt	atggtggtgg	ttttcaaact	360
ggaacatcat	ctttacatgt	ttatgatggc	aagtttctgg	ctcgggttga	aagagttatt	420
gtagtgtcaa	tgaactatag	ggtgggtgcc	ctaggattct	tagctttgcc	aggaaatcct	480
gaggctccag	ggaacatggg	tttatttgat	caacagttgg	ctcttcagtg	ggttcaaaaa	540
aatatagcag	cctttggtgg	aaatcctaaa	agtgtaactc	tctttggaga	aagttccgga	600
gcagcttcag	ttagcctgca	tttgctttct	cctggaagcc	attcattgtt	caccagagcc	660
attctgcaaa	gtggttccgc	taatgctcct	tgggcggtaa	catctcttta	tgaagctagg	720
aacagaacgt	tgaacttagc	taaattgact	ggttgctcta	gagagaatga	gactgaaata	780
atcaagtgtc	ttagaaataa	agatececaa	gaaattcttc	tgaatgaagc	atttgttgtc	840
ccctatggga	ctaacttggg	tgtaaacttt	ggtccgaccg	tggatggtga	ttttctcact	900
gacatgccag	acatattact	tgaacttgga	caatttaaaa	aaacccagat	tttggtgggt	960
gttaataaag	atgaagggac	atggtttta	gtcggtggtg	ctcctggctt	cagcaaagat	1020
aacaatagta	tcataactag	aaaagaattt	caggaaggtt	taaaaatatt	ttttccagga	1080
gtgagtgagt	ttggaaagga	atccatcctt	tttcattaca	cagactgggt	agatgatcag	1140
agacctgaaa	actaccgtga	ggccttgggt	gatgttgttg	gggattataa	tttcatatgc	1200
cctgccttgg	agttcaccaa	gaagttctca	gaatggggaa	ataatgcctt	tttctactat	1260
tttgaacacc	gatcctccaa	acttccgtgg	ccagaatgga	tgggagtgat	gcatggctat	1320
gaaattgaat	ttgtctttgg	tttacctctg	gaaagaagag	ataattacac	aaaagccgag	1380
gaaattttga	gtagatccat	agtgaaacgg	tgggcaaatt	ttgcaaaata	tgggaatcca	1440
aatgagactc	agaacaatag	cacaagctgg	cctgtcttca	aaagcactga	acaaaaatat	1500
ctaaccttga	atacagagtc	aacaagaata	atgacgaaac	tacgtgctca	acaatgtcga	1560
ttctggacat	catttttcc	aaaagtcttg	gaaatgacag	gaaatattga	tgaagcagaa	1620
tgggagtgga	aagcaggatt	ccatcgctgg	aacaattaca	tgatggactg	gaaaaatcaa	1680
tttaacgatt	acactagcaa	gaaagaaagt	tgtgtgggtc	tc		1722
<210> SEQ : <211> LENG' <212> TYPE	TH: 574					

<400> SEQUENCE: 10

Glu Asp Asp Ile Ile Ile Ala Thr Lys Asn Gly Lys Val Arg Gly Met 1 $$ 10 $$ 15

Asn Leu Thr Val Phe Gly Gly Thr Val Thr Ala Phe Leu Gly Ile Pro $20 \ \ \,$ 25 $\ \ \,$ 30

Tyr Ala Gln Pro Pro Leu Gly Arg Leu Arg Phe Lys Lys Pro Gln Ser 35 40 45

Leu Thr Lys Trp Ser Asp Ile Trp Asn Ala Thr Lys Tyr Ala Asn Ser 50 $\,$

Cys Cys Gln Asn Ile Asp Gln Ser Phe Pro Gly Phe His Gly Ser Glu 65 70 75 80

Met Trp Asn Pro Asn Thr Asp Leu Ser Glu Asp Cys Leu Tyr Leu Asn

<213 > ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthesized

-continued

Val	Trp	Ile	Pro 100	Ala	Pro	Lys	Pro	Lys 105	Asn	Ala	Thr	Val	Leu 110	Ile	Trp
Ile	Tyr	Gly 115	Gly	Gly	Phe	Gln	Thr 120	Gly	Thr	Ser	Ser	Leu 125	His	Val	Tyr
Asp	Gly 130	Lys	Phe	Leu	Ala	Arg 135	Val	Glu	Arg	Val	Ile 140	Val	Val	Ser	Met
Asn 145	Tyr	Arg	Val	Gly	Ala 150	Leu	Gly	Phe	Leu	Ala 155	Leu	Pro	Gly	Asn	Pro 160
Glu	Ala	Pro	Gly	Asn 165	Met	Gly	Leu	Phe	Asp 170	Gln	Gln	Leu	Ala	Leu 175	Gln
Trp	Val	Gln	Lys 180	Asn	Ile	Ala	Ala	Phe 185	Gly	Gly	Asn	Pro	Lys 190	Ser	Val
Thr	Leu	Phe 195	Gly	Glu	Ser	Ser	Gly 200	Ala	Ala	Ser	Val	Ser 205	Leu	His	Leu
Leu	Ser 210	Pro	Gly	Ser	His	Ser 215	Leu	Phe	Thr	Arg	Ala 220	Ile	Leu	Gln	Ser
Gly 225	Ser	Ala	Asn	Ala	Pro 230	Trp	Ala	Val	Thr	Ser 235	Leu	Tyr	Glu	Ala	Arg 240
Asn	Arg	Thr	Leu	Asn 245	Leu	Ala	Lys	Leu	Thr 250	Gly	Cys	Ser	Arg	Glu 255	Asn
Glu	Thr	Glu	Ile 260	Ile	Lys	Cys	Leu	Arg 265	Asn	Lys	Asp	Pro	Gln 270	Glu	Ile
Leu	Leu	Asn 275	Glu	Ala	Phe	Val	Val 280	Pro	Tyr	Gly	Thr	Asn 285	Leu	Gly	Val
Asn	Phe 290	Gly	Pro	Thr	Val	Asp 295	Gly	Asp	Phe	Leu	Thr 300	Asp	Met	Pro	Asp
Ile 305	Leu	Leu	Glu	Leu	Gly 310	Gln	Phe	Lys	Lys	Thr 315	Gln	Ile	Leu	Val	Gly 320
Val	Asn	Lys	Asp	Glu 325	Gly	Thr	Trp	Phe	Leu 330	Val	Gly	Gly	Ala	Pro 335	Gly
Phe	Ser	Lys	Asp 340	Asn	Asn	Ser	Ile	Ile 345	Thr	Arg	Lys	Glu	Phe 350	Gln	Glu
Gly	Leu	355 Lys	Ile	Phe	Phe	Pro	Gly 360	Val	Ser	Glu	Phe	Gly 365	Lys	Glu	Ser
Ile	Leu 370	Phe	His	Tyr	Thr	Asp 375	Trp	Val	Asp	Asp	Gln 380	Arg	Pro	Glu	Asn
Tyr 385	Arg	Glu	Ala	Leu	Gly 390	Asp	Val	Val	Gly	Asp 395	Tyr	Asn	Phe	Ile	Суз 400
Pro	Ala	Leu	Glu	Phe 405	Thr	Lys	Lys	Phe	Ser 410	Glu	Trp	Gly	Asn	Asn 415	Ala
Phe	Phe	Tyr	Tyr 420	Phe	Glu	His	Arg	Ser 425	Ser	ГЛа	Leu	Pro	Trp 430	Pro	Glu
Trp	Met	Gly 435	Val	Met	His	Gly	Tyr 440	Glu	Ile	Glu	Phe	Val 445	Phe	Gly	Leu
Pro	Leu 450	Glu	Arg	Arg	Asp	Asn 455	Tyr	Thr	Lys	Ala	Glu 460	Glu	Ile	Leu	Ser
Arg 465	Ser	Ile	Val	Lys	Arg 470	Trp	Ala	Asn	Phe	Ala 475	Lys	Tyr	Gly	Asn	Pro 480
Asn	Glu	Thr	Gln	Asn 485	Asn	Ser	Thr	Ser	Trp 490	Pro	Val	Phe	Lys	Ser 495	Thr
Glu	Gln	Lys	Tyr 500	Leu	Thr	Leu	Asn	Thr 505	Glu	Ser	Thr	Arg	Ile 510	Met	Thr
Lys	Leu	Arg	Ala	Gln	Gln	CAa	Arg	Phe	Trp	Thr	Ser	Phe	Phe	Pro	Lys

525

35 36

Val Leu Glu Met Thr Gly Asn Ile Asp Glu Ala Glu Trp Glu Trp Lys

520

Ala Gly Phe His Arg Trp Asn Asn Tyr Met Met Asp Trp Lys Asn Gln

Phe Asn Asp Tyr Thr Ser Lys Lys Glu Ser Cys Val Gly Leu

<210> SEQ ID NO 11

515

- <211> LENGTH: 1722
- <212> TYPE: DNA
- <213> ORGANISM: Artificial Sequence
- <220> FEATURE:
- <223> OTHER INFORMATION: synthesized

<400> SEQUENCE: 11

gaagatgaca tcataattgc aacaaagaat ggaaaagtca gagggatgaa cttgacagtt tttggtggca cggtaacagc ctttcttgga attccctatg cacagccacc tcttggtaga 120 cttcqattca aaaaqccaca qtctctqacc aaqtqqtctq atatttqqaa tqccacaaaa 180 tatgcaaatt cttgctgtca gaacatagat caaagttttc caggcttcca tggatcagag 240 atgtggaacc caaacactga cctcagtgaa gactgtttat atctaaatgt atggattcca 300 gcacctaaac caaaaaatgc cactgtattg atatggattt atggtggtgg ttttcaaact 360 ggaacatcat ctttacatgt ttatgatggc aagtttctgg ctcgggttga aagagttatt 420 gtagtgtcaa tgaactatag ggtgggtgcc ctaggattct tagctttgcc aggaaatcct 480 gaggetecag ggaacatggg tttatttgat caacagttgg etetteagtg ggtteaaaaa 540 aatatagcag cctttggtgg aaatcctaaa agtgtaactc tctttggaga aagttccgga 600 gcagcttcag ttagcctgca tttgctttct cctggaagcc attcattgtt caccagagcc 660 attotgoaaa gtggttocoo gaatgotoot tgggeggtaa catototta tgaagotagg 720 aacagaacgt tgaacttagc taaattgact ggttgctcta gagagaatga gactgaaata 780 atcaagtgtc ttagaaataa agatccccaa gaaattcttc tgaatgaagc atttgttgtc ccctatggga ctgctttggg tgtaaacttt ggtccgaccg tggatggtga ttttctcact 900 gacatgccag acatattact tgaacttgga caatttaaaa aaacccagat tttggtgggt 960 gttaataaag atgaagggac atggttttta gtcggtggtg ctcctggctt cagcaaagat 1020 1080 aacaatagta tcataactag aaaagaattt caggaaggtt taaaaatatt ttttccagga qtqaqtqaqt ttqqaaaqqa atccatcctt tttcattaca caqactqqqt aqatqatcaq 1140 agacctgaaa actaccgtga ggccttgggt gatgttgttg gggattataa tttcatatgc 1200 cctqccttqq aqttcaccaa qaaqttctca qaatqqqqaa ataatqcctt tttctactat 1260 tttgaacacc gatcctccaa acttccgtgg ccagaatgga tgggagtgat gcatggctat 1320 qaaattqaat ttgtctttqq tttacctctq qaaaqaaqaq ataattacac aaaaqccqaq 1380 gaaattttga gtagatccat agtgaaacgg tgggcaaatt ttgcaaaata tgggaatcca 1440 aatgagactc agaacaatag cacaagctgg cctgtcttca aaagcactga acaaaaatat 1500 ctaaccttga atacagagtc aacaagaata atgacgaaac tacgtgctca acaatgtcga 1560 ttctggacat catttttcc aaaagtcttg gaaatgacag gaaatattga tgaagcagaa 1620 tgggagtgga aagcaggatt ccatcgctgg aacaattaca tgatggactg gaaaaatcaa 1680 tttaacgatt acactagcaa gaaagaaagt tgtgtgggtc tc 1722

37

<210>	SEO	TD	MO	12

<211> LENGTH: 574

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthesized

<400> SEQUENCE: 12

Glu Asp Asp Ile Ile Ile Ala Thr Lys Asn Gly Lys Val Arg Gly Met 1 $$ 5 $$ 10 $$ 15

Asn Leu Thr Val Phe Gly Gly Thr Val Thr Ala Phe Leu Gly Ile Pro $20 \\ 25 \\ 30$

Tyr Ala Gln Pro Pro Leu Gly Arg Leu Arg Phe Lys Lys Pro Gln Ser 35 40 45

Leu Thr Lys Trp Ser Asp Ile Trp Asn Ala Thr Lys Tyr Ala Asn Ser 50 60

Cys Cys Gln Asn Ile Asp Gln Ser Phe Pro Gly Phe His Gly Ser Glu 65 70 75 80

Met Trp Asn Pro Asn Thr Asp Leu Ser Glu Asp Cys Leu Tyr Leu Asn 85 90 95

Val Trp Ile Pro Ala Pro Lys Pro Lys Asn Ala Thr Val Leu Ile Trp 100 105 110

Ile Tyr Gly Gly Gly Phe Gln Thr Gly Thr Ser Ser Leu His Val Tyr \$115\$ \$120\$ \$125\$

Asp Gly Lys Phe Leu Ala Arg Val Glu Arg Val Ile Val Val Ser Met 130 135 140

Asn Tyr Arg Val Gly Ala Leu Gly Phe Leu Ala Leu Pro Gly Asn Pro 145 150 155 160

Glu Ala Pro Gly Asn Met Gly Leu Phe Asp Gln Gln Leu Ala Leu Gln 165 170 175

Trp Val Gln Lys Asn Ile Ala Ala Phe Gly Gly Asn Pro Lys Ser Val \$180\$ \$180\$ \$185\$ \$190\$

Thr Leu Phe Gly Glu Ser Ser Gly Ala Ala Ser Val Ser Leu His Leu 195 200205

Leu Ser Pro Gly Ser His Ser Leu Phe Thr Arg Ala Ile Leu Gln Ser 210 215 220

Gly Ser Pro Asn Ala Pro Trp Ala Val Thr Ser Leu Tyr Glu Ala Arg 225 230 235 240

Asn Arg Thr Leu Asn Leu Ala Lys Leu Thr Gly Cys Ser Arg Glu Asn 245 250 255

Leu Leu Asn Glu Ala Phe Val Val Pro Tyr Gly Thr Ala Leu Gly Val \$275\$ \$280\$ \$285\$

Asn Phe Gly Pro Thr Val Asp Gly Asp Phe Leu Thr Asp Met Pro Asp 290 295 300

Ile Leu Leu Glu Leu Gly Gln Phe Lys Lys Thr Gln Ile Leu Val Gly 305 \$310\$ 315 320

Val Asn Lys Asp Glu Gly Thr Trp Phe Leu Val Gly Gly Ala Pro Gly 325 330 335

Phe Ser Lys Asp Asn Asn Ser Ile Ile Thr Arg Lys Glu Phe Gln Glu \$340\$ \$345\$ \$350

Gly Leu Lys Ile Phe Phe Pro Gly Val Ser Glu Phe Gly Lys Glu Ser 355 \$360\$ 365

Ile Leu Phe His Tyr Thr Asp Trp Val Asp Asp Gln Arg Pro Glu Asn $370 \hspace{1.5cm} 375 \hspace{1.5cm} 380 \hspace{1.5cm}$

-continued

Tyr Arg Glu Ala Leu Gly Asp Val Val Gly Asp Tyr Asn Phe Ile Cys Pro Ala Leu Glu Phe Thr Lys Lys Phe Ser Glu Trp Gly Asn Asn Ala Phe Phe Tyr Tyr Phe Glu His Arg Ser Ser Lys Leu Pro Trp Pro Glu Trp Met Gly Val Met His Gly Tyr Glu Ile Glu Phe Val Phe Gly Leu Pro Leu Glu Arg Arg Asp Asn Tyr Thr Lys Ala Glu Glu Ile Leu Ser Arg Ser Ile Val Lys Arg Trp Ala Asn Phe Ala Lys Tyr Gly Asn Pro 470 Asn Glu Thr Gln Asn Asn Ser Thr Ser Trp Pro Val Phe Lys Ser Thr Glu Gln Lys Tyr Leu Thr Leu Asn Thr Glu Ser Thr Arg Ile Met Thr 505 Lys Leu Arg Ala Gln Gln Cys Arg Phe Trp Thr Ser Phe Phe Pro Lys 520 Val Leu Glu Met Thr Gly Asn Ile Asp Glu Ala Glu Trp Glu Trp Lys 530 535 Ala Gly Phe His Arg Trp Asn Asn Tyr Met Met Asp Trp Lys Asn Gln 545 Phe Asn Asp Tyr Thr Ser Lys Lys Glu Ser Cys Val Gly Leu 565 <210> SEQ ID NO 13 <211> LENGTH: 1722 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223 > OTHER INFORMATION: synthesized <400> SEQUENCE: 13 gaagatgaca tcataattgc aacaaagaat ggaaaagtca gagggatgaa cttgacagtt tttggtggca cggtaacagc ctttcttgga attccctatg cacagccacc tcttggtaga

cttcqattca aaaaqccaca qtctctqacc aaqtqqtctq atatttqqaa tqccacaaaa 180 tatgcaaatt cttgctgtca gaacatagat caaagttttc caggcttcca tggatcagag 240 atgtggaacc caaacactga cctcagtgaa gactgtttat atctaaatgt atggattcca 300 gcacctaaac caaaaaatgc cactgtattg atatggattt atggtggtgg ttttcaaact 360 qqaacatcat ctttacatqt ttatqatqqc aaqtttctqq ctcqqqttqa aaqaqttatt 420 qtaqtqtcaa tqaactataq qqtqqqtqcc ctaqqattct taqctttqcc aqqaaatcct 480 gaggetecag ggaacatggg tttatttgat caacagttgg etetteagtg ggtteaaaaa 540 aatatagcag cctttggtgg aaatcctaaa agtgtaactc tctttggaga aagttccgga 600 geagetteag ttageetgea tttgetttet eetggaagee atteattgtt eaceagagee 660 attetgeaaa gtggtteete caatgeteet tgggeggtaa catetetta tgaagetagg 720 aacagaacgt tgaacttagc taaattgact ggttgctcta gagagaatga gactgaaata 780 atcaagtgtc ttagaaataa agatccccaa gaaattcttc tgaatgaagc atttgttgtc 840 ccctatggga ctcagttggg tgtaaacttt ggtccgaccg tggatggtga ttttctcact 900 gacatgccag acatattact tgaacttgga caatttaaaa aaacccagat tttggtgggt 960

gttaataaag atgaagggac atggttttta gtcggtggtg ctcctggctt cagcaaagat

-continued

aaca	aataq	gta 1	cata	aacta	ag aa	aaaga	aattt	caç	ggaaq	ggtt	taaa	aata	att 1	ttttc	ccagga	10	80
gtga	agtga	agt 1	tgga	aaag	ga at	ccat	cctt	ttt	catt	aca	caga	ctg	ggt	agato	gatcag	11	40
agad	cctga	aaa a	actad	ccgt	ga go	gcctt	gggt	gat	gtt	gttg	ggga	attat	caa 1	tttca	atatgc	12	00
cct	gcctt	gg a	agtto	cacca	aa ga	aagtt	cctca	a gaa	atggg	ggaa	ataa	atgc	ctt 1	tttct	actat	12	60
tttç	gaaca	acc q	gatco	eteca	aa a	cttc	gtgg	g cca	agaat	gga	tggg	gagt	gat	gcato	ggctat	13	20
gaaa	attga	aat 1	tgt	cttt	gg ti	taco	ctctç	g gaa	aagaa	agag	ataa	atta	cac a	aaaag	gccgag	13	80
gaaa	attt	ga 🤅	gtaga	atcca	at aç	gtgaa	aacgg	g tgg	ggcaa	att	ttg	caaaa	ata 1	tggga	aatcca	14	40
aat	gaga	ctc a	agaad	caata	ag ca	acaa	gatgo	g cct	gtct	tca	aaag	gcact	ga a	acaaa	aaatat	15	00
ctaa	acctt	ga a	ataca	agagt	c aa	acaaç	gaata	a ato	gacga	aaac	taco	gtgct	ca a	acaat	gtcga	15	60
ttct	ggad	cat o	cattt	ttt	cc aa	aaagt	ctto	g gaa	aatga	acag	gaaa	atati	ga 1	tgaag	gcagaa	16	20
tggg	gagto	gga a	aagca	aggat	t co	catco	gctgg	gaad	caatt	aca	tgat	gga	ctg (gaaaa	aatcaa	16	80
ttta	aacga	att a	acact	agca	aa ga	aaaga	aaagt	tgt	gtg	ggtc	tc					17	22
<213 <213 <213 <220 <223	L> LI 2> T: 3> OI 0> FI 3> O:	ENGTI PE: RGAN: EATUI THER	ISM:	74 Art: ORMA:			-		Ē								
Glu 1	Asp	Asp	Ile	Ile 5	Ile	Ala	Thr	Lys	Asn 10	Gly	Lys	Val	Arg	Gly 15	Met		
Asn	Leu	Thr	Val 20	Phe	Gly	Gly	Thr	Val 25	Thr	Ala	Phe	Leu	Gly 30	Ile	Pro		
Tyr	Ala	Gln 35	Pro	Pro	Leu	Gly	Arg 40	Leu	Arg	Phe	Lys	Lys 45	Pro	Gln	Ser		
Leu	Thr 50	Lys	Trp	Ser	Asp	Ile 55	Trp	Asn	Ala	Thr	60 Lys	Tyr	Ala	Asn	Ser		
Cys	CAa	Gln	Asn	Ile	Asp 70	Gln	Ser	Phe	Pro	Gly 75	Phe	His	Gly	Ser	Glu 80		
Met	Trp	Asn	Pro	Asn 85	Thr	Asp	Leu	Ser	Glu 90	Asp	CAa	Leu	Tyr	Leu 95	Asn		
Val	Trp	Ile	Pro 100	Ala	Pro	Lys	Pro	Lys 105	Asn	Ala	Thr	Val	Leu 110	Ile	Trp		
Ile	Tyr	Gly 115	Gly	Gly	Phe	Gln	Thr 120	Gly	Thr	Ser	Ser	Leu 125	His	Val	Tyr		
Asp	Gly 130	ГЛа	Phe	Leu	Ala	Arg 135	Val	Glu	Arg	Val	Ile 140	Val	Val	Ser	Met		
Asn 145	Tyr	Arg	Val	Gly	Ala 150	Leu	Gly	Phe	Leu	Ala 155	Leu	Pro	Gly	Asn	Pro 160		
Glu	Ala	Pro	Gly	Asn 165	Met	Gly	Leu	Phe	Asp 170	Gln	Gln	Leu	Ala	Leu 175	Gln		
Trp	Val	Gln	Lys 180	Asn	Ile	Ala	Ala	Phe 185	Gly	Gly	Asn	Pro	Lys 190	Ser	Val		
Thr	Leu	Phe	Gly	Glu	Ser	Ser	Gly 200	Ala	Ala	Ser	Val	Ser 205	Leu	His	Leu		
Leu	Ser 210	Pro	Gly	Ser	His	Ser 215	Leu	Phe	Thr	Arg	Ala 220	Ile	Leu	Gln	Ser		

Gly Ser Ser Asn Ala Pro Trp Ala Val Thr Ser Leu Tyr Glu Ala Arg 225 230 230 235 240

-continued

											-	con	tın	ued		
Asn	Arg	Thr	Leu	Asn 245	Leu	Ala	Lys	Leu	Thr 250	Gly	CAa	Ser	Arg	Glu 255	Asn	
Glu	Thr	Glu	Ile 260	Ile	Lys	CÀa	Leu	Arg 265	Asn	Lys	Asp	Pro	Gln 270	Glu	Ile	
Leu	Leu	Asn 275	Glu	Ala	Phe	Val	Val 280	Pro	Tyr	Gly	Thr	Gln 285	Leu	Gly	Val	
Asn	Phe 290	Gly	Pro	Thr	Val	Asp 295	Gly	Asp	Phe	Leu	Thr 300	Asp	Met	Pro	Asp	
Ile 305	Leu	Leu	Glu	Leu	Gly 310	Gln	Phe	Lys	Lys	Thr 315	Gln	Ile	Leu	Val	Gly 320	
Val	Asn	Lys	Asp	Glu 325	Gly	Thr	Trp	Phe	Leu 330	Val	Gly	Gly	Ala	Pro 335	Gly	
Phe	Ser	Lys	Asp 340	Asn	Asn	Ser	Ile	Ile 345	Thr	Arg	Lys	Glu	Phe 350	Gln	Glu	
Gly	Leu	Lys 355	Ile	Phe	Phe	Pro	Gly 360	Val	Ser	Glu	Phe	Gly 365	Lys	Glu	Ser	
Ile	Leu 370	Phe	His	Tyr	Thr	Asp 375	Trp	Val	Asp	Asp	Gln 380	Arg	Pro	Glu	Asn	
Tyr 385	Arg	Glu	Ala	Leu	Gly 390	Asp	Val	Val	Gly	Asp 395	Tyr	Asn	Phe	Ile	Cys 400	
Pro	Ala	Leu	Glu	Phe 405	Thr	Lys	Lys	Phe	Ser 410	Glu	Trp	Gly	Asn	Asn 415	Ala	
Phe	Phe	Tyr	Tyr 420	Phe	Glu	His	Arg	Ser 425	Ser	Lys	Leu	Pro	Trp 430	Pro	Glu	
Trp	Met	Gly 435	Val	Met	His	Gly	Tyr 440	Glu	Ile	Glu	Phe	Val 445	Phe	Gly	Leu	
Pro	Leu 450	Glu	Arg	Arg	Asp	Asn 455	Tyr	Thr	Lys	Ala	Glu 460	Glu	Ile	Leu	Ser	
Arg 465	Ser	Ile	Val	Lys	Arg 470	Trp	Ala	Asn	Phe	Ala 475	Lys	Tyr	Gly	Asn	Pro 480	
Asn	Glu	Thr	Gln	Asn 485	Asn	Ser	Thr	Ser	Trp 490	Pro	Val	Phe	Lys	Ser 495	Thr	
Glu	Gln	Lys	Tyr 500	Leu	Thr	Leu	Asn	Thr 505	Glu	Ser	Thr	Arg	Ile 510	Met	Thr	
Lys	Leu	Arg 515	Ala	Gln	Gln	CAa	Arg 520	Phe	Trp	Thr	Ser	Phe 525	Phe	Pro	Lys	
Val	Leu 530	Glu	Met	Thr	Gly	Asn 535	Ile	Asp	Glu	Ala	Glu 540	Trp	Glu	Trp	Lys	
Ala 545	Gly	Phe	His	Arg	Trp 550	Asn	Asn	Tyr	Met	Met 555	Asp	Trp	Lys	Asn	Gln 560	
Phe	Asn	Asp	Tyr	Thr 565	Ser	Lys	Lys	Glu	Ser 570	Cys	Val	Gly	Leu			
<213 <213 <213 <220 <223	0> FE 3> O	ENGTI (PE : RGAN: EATUI THER	H: 1' DNA ISM: RE: INFO	722 Art: DRMA'	ific: TION		_		Ē							
< 400	D> SI	EQUEI	ICE:	15												
	_				=		_			=			=	_	acagtt	
					-					_		_		-	ggtaga	
ctt	cgatt	ca a	aaaa	gcca	ca gt	ctct	gac	c aag	gtggt	ctg	atai	ttg	gaa 1	gcca	acaaaa	180

tatgcaaatt cttgctgtca gaacatagat caaagttttc caggcttcca tggatcagag

atgtggaacc caaacactga cctcagtgaa gactgtttat atctaaatgt atggattcca	300
gcacctaaac caaaaaatgc cactgtattg atatggattt atggtggtgg ttttcaaact	360
ggaacatcat ctttacatgt ttatgatggc aagtttctgg ctcgggttga aagagttatt	420
gtagtgtcaa tgaactatag ggtgggtgcc ctaggattct tagctttgcc aggaaatcct	480
gaggetecag ggaacatggg tttatttgat caacagttgg etetteagtg ggtteaaaaa	540
aatatagcag cctttggtgg aaatcctaaa agtgtaactc tctttggaga aagttccgga	600
gcagcttcag ttagcctgca tttgctttct cctggaagcc attcattgtt caccagagcc	660
attctgcaaa gtggttcctc caatgctcct tgggcggtaa catctcttta tgaagctagg	720
aacagaacgt tgaacttagc taaattgact ggttgctcta gagagaatga gactgaaata	780
atcaagtgtc ttagaaataa agatccccaa gaaattcttc tgaatgaagc atttgttgtc	840
ccctatggga cttccttggg tgtaaacttt ggtccgaccg tggatggtga ttttctcact	900
gacatgccag acatattact tgaacttgga caatttaaaa aaacccagat tttggtgggt	960
gttaataaag atgaagggac atggttttta gtcggtggtg ctcctggctt cagcaaagat	1020
aacaatagta tcataactag aaaagaattt caggaaggtt taaaaatatt ttttccagga	1080
gtgagtgagt ttggaaagga atccatcctt tttcattaca cagactgggt agatgatcag	1140
agacctgaaa actaccgtga ggccttgggt gatgttgttg gggattataa tttcatatgc	1200
cctgccttgg agttcaccaa gaagttctca gaatggggaa ataatgcctt tttctactat	1260
tttgaacacc gatcctccaa acttccgtgg ccagaatgga tgggagtgat gcatggctat	1320
gaaattgaat ttgtctttgg tttacctctg gaaagaagag ataattacac aaaagccgag	1380
gaaattttga gtagatccat agtgaaacgg tgggcaaatt ttgcaaaata tgggaatcca	1440
aatgagactc agaacaatag cacaagctgg cctgtcttca aaagcactga acaaaaatat	1500
ctaaccttga atacagagtc aacaagaata atgacgaaac tacgtgctca acaatgtcga	1560
ttctggacat cattttttcc aaaagtcttg gaaatgacag gaaatattga tgaagcagaa	1620
tgggagtgga aagcaggatt ccatcgctgg aacaattaca tgatggactg gaaaaatcaa	1680
tttaacgatt acactagcaa gaaagaaagt tgtgtgggtc tc	1722
<210> SEQ ID NO 16 <211> LENGTH: 574 <212> TYPE: PRT <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: synthesized	
<400> SEQUENCE: 16	
Glu Asp Asp Ile Ile Ile Ala Thr Lys Asn Gly Lys Val Arg Gly Met 1 10 15	
Asn Leu Thr Val Phe Gly Gly Thr Val Thr Ala Phe Leu Gly Ile Pro 20 25 30	
Tyr Ala Gln Pro Pro Leu Gly Arg Leu Arg Phe Lys Lys Pro Gln Ser 35 40 45	
Leu Thr Lys Trp Ser Asp Ile Trp Asn Ala Thr Lys Tyr Ala Asn Ser 50 60	
Cys Cys Gln Asn Ile Asp Gln Ser Phe Pro Gly Phe His Gly Ser Glu 65 70 75 80	
Met Trp Asn Pro Asn Thr Asp Leu Ser Glu Asp Cys Leu Tyr Leu Asn 85 90 95	

Val Trp Ile Pro Ala Pro Lys Pro Lys Asn Ala Thr Val Leu Ile Trp

-continued
-continued

												COII	LIII	uea	
			100					105					110		
Ile	Tyr	Gly 115	Gly	Gly	Phe	Gln	Thr 120	Gly	Thr	Ser	Ser	Leu 125	His	Val	Tyr
Asp	Gly 130	Lys	Phe	Leu	Ala	Arg 135	Val	Glu	Arg	Val	Ile 140	Val	Val	Ser	Met
Asn 145	Tyr	Arg	Val	Gly	Ala 150	Leu	Gly	Phe	Leu	Ala 155	Leu	Pro	Gly	Asn	Pro 160
Glu	Ala	Pro	Gly	Asn 165	Met	Gly	Leu	Phe	Asp 170	Gln	Gln	Leu	Ala	Leu 175	Gln
Trp	Val	Gln	Lys 180	Asn	Ile	Ala	Ala	Phe 185	Gly	Gly	Asn	Pro	Lys 190	Ser	Val
Thr	Leu	Phe 195	Gly	Glu	Ser	Ser	Gly 200	Ala	Ala	Ser	Val	Ser 205	Leu	His	Leu
Leu	Ser 210	Pro	Gly	Ser	His	Ser 215	Leu	Phe	Thr	Arg	Ala 220	Ile	Leu	Gln	Ser
Gly 225	Ser	Ser	Asn	Ala	Pro 230	Trp	Ala	Val	Thr	Ser 235	Leu	Tyr	Glu	Ala	Arg 240
Asn	Arg	Thr	Leu	Asn 245	Leu	Ala	ГÀа	Leu	Thr 250	Gly	CÀa	Ser	Arg	Glu 255	Asn
Glu	Thr	Glu	Ile 260	Ile	ГÀа	CÀa	Leu	Arg 265	Asn	ГЛа	Asp	Pro	Gln 270	Glu	Ile
Leu	Leu	Asn 275	Glu	Ala	Phe	Val	Val 280	Pro	Tyr	Gly	Thr	Ser 285	Leu	Gly	Val
Asn	Phe 290	Gly	Pro	Thr	Val	Asp 295	Gly	Asp	Phe	Leu	Thr 300	Asp	Met	Pro	Asp
Ile 305	Leu	Leu	Glu	Leu	Gly 310	Gln	Phe	Lys	Lys	Thr 315	Gln	Ile	Leu	Val	Gly 320
Val	Asn	Lys	Asp	Glu 325	Gly	Thr	Trp	Phe	Leu 330	Val	Gly	Gly	Ala	Pro 335	Gly
Phe	Ser	Lys	Asp 340	Asn	Asn	Ser	Ile	Ile 345	Thr	Arg	ràa	Glu	Phe 350	Gln	Glu
Gly	Leu	Lys 355	Ile	Phe	Phe	Pro	Gly 360	Val	Ser	Glu	Phe	Gly 365	ГЛЗ	Glu	Ser
Ile	Leu 370	Phe	His	Tyr	Thr	Asp 375	Trp	Val	Asp	Asp	Gln 380	Arg	Pro	Glu	Asn
Tyr 385	Arg	Glu	Ala	Leu	Gly 390	Asp	Val	Val	Gly	Asp 395	Tyr	Asn	Phe	Ile	Сув 400
Pro	Ala	Leu	Glu	Phe 405	Thr	Lys	Lys	Phe	Ser 410	Glu	Trp	Gly	Asn	Asn 415	Ala
Phe	Phe	Tyr	Tyr 420	Phe	Glu	His	Arg	Ser 425	Ser	Lys	Leu	Pro	Trp 430	Pro	Glu
Trp	Met	Gly 435	Val	Met	His	Gly	Tyr 440	Glu	Ile	Glu	Phe	Val 445	Phe	Gly	Leu
Pro	Leu 450	Glu	Arg	Arg	Asp	Asn 455	Tyr	Thr	Lys	Ala	Glu 460	Glu	Ile	Leu	Ser
Arg 465	Ser	Ile	Val	Lys	Arg 470	Trp	Ala	Asn	Phe	Ala 475	Lys	Tyr	Gly	Asn	Pro 480
Asn	Glu	Thr	Gln	Asn 485	Asn	Ser	Thr	Ser	Trp 490	Pro	Val	Phe	Lys	Ser 495	Thr
Glu	Gln	Lys	Tyr 500	Leu	Thr	Leu	Asn	Thr 505	Glu	Ser	Thr	Arg	Ile 510	Met	Thr
ГÀз	Leu	Arg 515	Ala	Gln	Gln	Сув	Arg 520	Phe	Trp	Thr	Ser	Phe 525	Phe	Pro	Lys

60

49 50

Val Leu Glu Met Thr Gly Asn Ile Asp Glu Ala Glu Trp Glu Trp Lys 535

Ala Gly Phe His Arg Trp Asn Asn Tyr Met Met Asp Trp Lys Asn Gln

Phe Asn Asp Tyr Thr Ser Lys Lys Glu Ser Cys Val Gly Leu

<210> SEQ ID NO 17

<211> LENGTH: 1722

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthesized

<400> SEQUENCE: 17

gaagatgaca tcataattgc aacaaagaat ggaaaagtca gagggatgaa cttgacagtt tttqqtqqca cqqtaacaqc ctttcttqqa attccctatq cacaqccacc tcttqqtaqa 120 cttcgattca aaaagccaca gtctctgacc aagtggtctg atatttggaa tgccacaaaa 180 tatqcaaatt cttqctqtca qaacataqat caaaqttttc caqqcttcca tqqatcaqaq 240 atgtggaacc caaacactga cctcagtgaa gactgtttat atctaaatgt atggattcca 300 gcacctaaac caaaaaatgc cactgtattg atatggattt atggtggtgg ttttcaaact 360 ggaacatcat ctttacatgt ttatgatggc aagtttctgg ctcgggttga aagagttatt 420 gtagtgtcaa tgaactatag ggtgggtgcc ctaggattct tagctttgcc aggaaatcct 480 gaggetecag ggaacatggg tttatttgat caacagttgg etetteagtg ggtteaaaaa 540 aatatagcag cetttggtgg aaateetaaa agtgtaacte tetttggaga aagtteegga 600 gcagcttcag ttagcctgca tttgctttct cctggaagcc attcattgtt caccagagcc 660 attctgcaaa gtggttcctc caatgctcct tgggcggtaa catctcttta tgaagctagg 720 aacagaacgt tgaacttagc taaattgact ggttgctcta gagagaatga gactgaaata 780 atcaagtgtc ttagaaataa agatccccaa gaaattcttc tgaatgaagc atttgttgtc 840 ccctatggga ctggtttggg tgtaaacttt ggtccgaccg tggatggtga ttttctcact gacatgccag acatattact tgaacttgga caatttaaaa aaacccagat tttggtgggt 960 gttaataaag atgaagggac atggttttta gtcggtggtg ctcctggctt cagcaaagat 1020 aacaatagta tcataactag aaaagaattt caggaaggtt taaaaaatatt ttttccagga 1080 qtqaqtqaqt ttqqaaaqqa atccatcctt tttcattaca caqactqqqt aqatqatcaq 1140 1200 agacctgaaa actaccgtga ggccttgggt gatgttgttg gggattataa tttcatatgc cctqccttqq aqttcaccaa qaaqttctca qaatqqqqaa ataatqcctt tttctactat 1260 1320 tttqaacacc qatcctccaa acttccqtqq ccaqaatqqa tqqqaqtqat qcatqqctat gaaattgaat ttgtctttgg tttacctctg gaaagaagag ataattacac aaaagccgag 1380 gaaattttga gtagatccat agtgaaacgg tgggcaaatt ttgcaaaata tgggaatcca 1440 aatgagactc agaacaatag cacaagctgg cctgtcttca aaagcactga acaaaaatat 1500 1560 ctaaccttga atacagagtc aacaagaata atgacgaaac tacgtgctca acaatgtcga ttctggacat catttttcc aaaagtcttg gaaatgacag gaaatattga tgaagcagaa 1620 tgggagtgga aagcaggatt ccatcgctgg aacaattaca tgatggactg gaaaaatcaa 1680 tttaacgatt acactagcaa gaaagaaagt tgtgtgggtc tc 1722

<210> SEQ ID NO 18 <211> LENGTH: 574

51 52

<212>	TYPE:	PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223 > OTHER INFORMATION: synthesized

<400> SEQUENCE: 18

Glu Asp Asp Ile Ile Ile Ala Thr Lys Asn Gly Lys Val Arg Gly Met
1 5 10 15

Asn Leu Thr Val Phe Gly Gly Thr Val Thr Ala Phe Leu Gly Ile Pro

Tyr Ala Gln Pro Pro Leu Gly Arg Leu Arg Phe Lys Lys Pro Gln Ser 35 40 45

Leu Thr Lys Trp Ser Asp Ile Trp Asn Ala Thr Lys Tyr Ala Asn Ser

Cys Cys Gln Asn Ile Asp Gln Ser Phe Pro Gly Phe His Gly Ser Glu 65 70 75 80

Met Trp Asn Pro Asn Thr Asp Leu Ser Glu Asp Cys Leu Tyr Leu Asn 85 90 95

Val Trp Ile Pro Ala Pro Lys Pro Lys Asn Ala Thr Val Leu Ile Trp 100 105 110

Ile Tyr Gly Gly Gly Phe Gln Thr Gly Thr Ser Ser Leu His Val Tyr \$115\$ \$120\$ \$125\$

Asp Gly Lys Phe Leu Ala Arg Val Glu Arg Val Ile Val Val Ser Met 130 135 140

Asn Tyr Arg Val Gly Ala Leu Gly Phe Leu Ala Leu Pro Gly Asn Pro 145 150 155 160

Glu Ala Pro Gly Asn Met Gly Leu Phe Asp Gln Gln Leu Ala Leu Gln 165 \$170\$

Trp Val Gln Lys Asn Ile Ala Ala Phe Gly Gly Asn Pro Lys Ser Val

Thr Leu Phe Gly Glu Ser Ser Gly Ala Ala Ser Val Ser Leu His Leu 195 200205

Leu Ser Pro Gly Ser His Ser Leu Phe Thr Arg Ala Ile Leu Gln Ser 210 \$215\$

Gly Ser Ser Asn Ala Pro Trp Ala Val Thr Ser Leu Tyr Glu Ala Arg 225 230 230 235

Asn Arg Thr Leu Asn Leu Ala Lys Leu Thr Gly Cys Ser Arg Glu Asn 245 250 255

Glu Thr Glu Ile Ile Lys Cys Leu Arg Asn Lys Asp Pro Gln Glu Ile 260 265 270

Leu Leu Asn Glu Ala Phe Val Val Pro Tyr Gly Thr Gly Leu Gly Val \$275\$ \$280\$ \$285\$

Asn Phe Gly Pro Thr Val Asp Gly Asp Phe Leu Thr Asp Met Pro Asp 290 295 300

Ile Leu Leu Glu Leu Gly Gln Phe Lys Lys Thr Gln Ile Leu Val Gly 305 310 315 320

Val Asn Lys Asp Glu Gly Thr Trp Phe Leu Val Gly Gly Ala Pro Gly 325 330 335

Phe Ser Lys Asp Asn Asn Ser Ile Ile Thr Arg Lys Glu Phe Glu Glu \$340\$ \$345\$ \$350

Gly Leu Lys Ile Phe Phe Pro Gly Val Ser Glu Phe Gly Lys Glu Ser 355 \$360\$

Ile Leu Phe His Tyr Thr Asp Trp Val Asp Asp Gln Arg Pro Glu Asn 370 375 380

Tyr Arg Glu Ala Leu Gly Asp Val Val Gly Asp Tyr Asn Phe Ile Cys

-continued

385					390					395					400
Pro	Ala	Leu	Glu	Phe 405	Thr	Lys	Lys	Phe	Ser 410	Glu	Trp	Gly	Asn	Asn 415	Ala
Phe	Phe	Tyr	Tyr 420	Phe	Glu	His	Arg	Ser 425	Ser	Lys	Leu	Pro	Trp 430	Pro	Glu
Trp	Met	Gly 435	Val	Met	His	Gly	Tyr 440	Glu	Ile	Glu	Phe	Val 445	Phe	Gly	Leu
Pro	Leu 450	Glu	Arg	Arg	Asp	Asn 455	Tyr	Thr	Lys	Ala	Glu 460	Glu	Ile	Leu	Ser
Arg 465	Ser	Ile	Val	Lys	Arg 470	Trp	Ala	Asn	Phe	Ala 475	Lys	Tyr	Gly	Asn	Pro 480
Asn	Glu	Thr	Gln	Asn 485	Asn	Ser	Thr	Ser	Trp 490	Pro	Val	Phe	Lys	Ser 495	Thr
Glu	Gln	Lys	Tyr 500	Leu	Thr	Leu	Asn	Thr 505	Glu	Ser	Thr	Arg	Ile 510	Met	Thr
Lys	Leu	Arg 515	Ala	Gln	Gln	Cys	Arg 520	Phe	Trp	Thr	Ser	Phe 525	Phe	Pro	Lys
Val	Leu 530	Glu	Met	Thr	Gly	Asn 535	Ile	Asp	Glu	Ala	Glu 540	Trp	Glu	Trp	Lys
Ala 545	Gly	Phe	His	Arg	Trp 550	Asn	Asn	Tyr	Met	Met 555	Asp	Trp	Lys	Asn	Gln 560
Phe	Asn	Asp	Tyr	Thr 565	Ser	Lys	Lys	Glu	Ser 570	Cya	Val	Gly	Leu		

<210> SEQ ID NO 19

<211> LENGTH: 1722

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthesized

<400> SEQUENCE: 19

gaagatgaca tcataattgc aacaaagaat ggaaaagtca gagggatgaa cttgacagtt 60 tttggtggca cggtaacagc ctttcttgga attccctatg cacagccacc tcttggtaga cttcgattca aaaagccaca gtctctgacc aagtggtctg atatttggaa tgccacaaaa tatgcaaatt cttgctgtca gaacatagat caaagttttc caggcttcca tggatcagag 240 atgtggaacc caaacactga cctcagtgaa gactgtttat atctaaatgt atggattcca 300 gcacctaaac caaaaaatgc cactgtattg atatggattt atggtggtgg ttttcaaact 360 ggaacatcat ctttacatgt ttatgatggc aagtttctgg ctcgggttga aagagttatt 420 gtagtgtcaa tgaactatag ggtgggtgcc ctaggattct tagctttgcc aggaaatcct 480 gaggetecag ggaacatggg tttatttgat caacagttgg etetteagtg ggtteaaaaa 540 aatatagcag cetttggtgg aaateetaaa agtgtaacte tetttggaga aagtteegga 600 gcagetteag ttageetgea tttgetttet eetggaagee atteattgtt eaceagagee 660 attctgcaaa gtggttcccc gaatgctcct tgggcggtaa catctcttta tgaagctagg 720 aacagaacgt tgaacttagc taaattgact ggttgctcta gagagaatga gactgaaata 780 atcaagtgtc ttagaaataa agatccccaa gaaattcttc tgaatgaagc atttgttgtc 840 ccctatggga cttccatggg tgtaaacttt ggtccgaccg tggatggtga ttttctcact 900 gacatgccag acatattact tgaacttgga caatttaaaa aaacccagat tttggtgggt 960 gttaataaag atgaagggac atggttttta gtcggtggtg ctcctggctt cagcaaagat 1020 aacaatagta tcataactag aaaagaattt caggaaggtt taaaaaatatt ttttccagga 1080

-continued

gtga	agtga	agt 1	tgga	aaag	ga at	ccat	cctt	: ttt	tcati	aca	caga	actg	ggt a	agato	gatcag	1140
agad	cctga	aaa a	actac	ccgt	ga g	gaatt	gggt	gat	tgtt	gttg	ggg	attat	taa 1	tttca	atatgc	1200
ccts	geett	gg a	agtto	cacca	aa ga	aagtt	ctca	a gaa	atgg	ggaa	ataa	atgc	ett 1	tttct	cactat	1260
tttç	gaaca	acc q	gatco	eteca	aa a	ette	gtgg	g cca	agaat	gga	tgg	gagto	gat q	gcato	ggctat	1320
gaaa	attga	aat 1	tgt	cttt	gg tt	taco	ctctç	g gaa	aagaa	agag	ataa	atta	cac a	aaaag	gccgag	1380
gaaa	attt	ga	gtaga	atcca	at aç	gtgaa	aacgg	g tgg	ggcaa	aatt	ttg	caaaa	ata 1	tggga	aatcca	1440
aato	gagad	etc a	agaad	caata	ag ca	acaaq	gctg	g cct	tgtc	tca	aaaq	gcact	tga a	acaaa	aaatat	1500
ctaa	acctt	ga a	ataca	agagt	tc aa	acaa	gaata	a ato	gacga	aaac	taco	gtgct	tca a	acaat	gtcga	1560
ttct	ggad	cat o	catt	ttt	cc aa	aaagt	ctto	g gaa	aatga	acag	gaaa	atati	tga 1	tgaag	gcagaa	1620
tggg	gagto	gga a	aagca	aggat	tt co	catco	gatg	g aad	caati	caca	tgat	gga	ctg q	gaaaa	aatcaa	1680
ttta	aacga	att a	acact	agca	aa ga	aaaga	aaagt	t tgt	tgtg	ggtc	tc					1722
<211 <212 <213 <220	L> LI 2> T 3> OI 0> FI	ENGTI (PE : RGAN: EATUI	ISM:	74 Art:			_		đ							
< 400)> SI	EQUEI	ICE:	20												
Glu 1	Asp	Asp	Ile	Ile 5	Ile	Ala	Thr	Lys	Asn 10	Gly	ГÀв	Val	Arg	Gly 15	Met	
Asn	Leu	Thr	Val 20	Phe	Gly	Gly	Thr	Val 25	Thr	Ala	Phe	Leu	Gly 30	Ile	Pro	
Tyr	Ala	Gln 35	Pro	Pro	Leu	Gly	Arg 40	Leu	Arg	Phe	rys	Lys 45	Pro	Gln	Ser	
Leu	Thr 50	Lys	Trp	Ser	Asp	Ile 55	Trp	Asn	Ala	Thr	Lys 60	Tyr	Ala	Asn	Ser	
Cys 65	Cys	Gln	Asn	Ile	Asp 70	Gln	Ser	Phe	Pro	Gly 75	Phe	His	Gly	Ser	Glu 80	
Met	Trp	Asn	Pro	Asn 85	Thr	Asp	Leu	Ser	Glu 90	Asp	CAa	Leu	Tyr	Leu 95	Asn	
Val	Trp	Ile	Pro	Ala	Pro	Lys	Pro	Lys 105	Asn	Ala	Thr	Val	Leu 110	Ile	Trp	
Ile	Tyr	Gly 115		Gly	Phe		Thr 120		Thr	Ser		Leu 125	His	Val	Tyr	
Asp	Gly 130	Lys	Phe	Leu	Ala	Arg 135	Val	Glu	Arg	Val	Ile 140	Val	Val	Ser	Met	
Asn 145	Tyr	Arg	Val	Gly	Ala 150	Leu	Gly	Phe	Leu	Ala 155	Leu	Pro	Gly	Asn	Pro 160	
Glu	Ala	Pro	Gly	Asn 165	Met	Gly	Leu	Phe	Asp 170	Gln	Gln	Leu	Ala	Leu 175	Gln	
Trp	Val	Gln	Lys 180	Asn	Ile	Ala	Ala	Phe 185	Gly	Gly	Asn	Pro	Lys 190	Ser	Val	
Thr	Leu	Phe 195	Gly	Glu	Ser	Ser	Gly 200	Ala	Ala	Ser	Val	Ser 205	Leu	His	Leu	
Leu	Ser 210	Pro	Gly	Ser	His	Ser 215	Leu	Phe	Thr	Arg	Ala 220	Ile	Leu	Gln	Ser	
Gly 225	Ser	Pro	Asn	Ala	Pro 230	Trp	Ala	Val	Thr	Ser 235	Leu	Tyr	Glu	Ala	Arg 240	

-continued

Glu Thr Glu Ile Ile Lys Cys Leu Arg Asn Lys Asp Pro Gln Glu Ile

Leu Leu Asn Glu Ala Phe Val Val Pro Tyr Gly Thr Ser Met Gly Val Asn Phe Gly Pro Thr Val Asp Gly Asp Phe Leu Thr Asp Met Pro Asp Ile Leu Leu Glu Leu Gly Gln Phe Lys Lys Thr Gln Ile Leu Val Gly Val Asn Lys Asp Glu Gly Thr Trp Phe Leu Val Gly Gly Ala Pro Gly 330 Phe Ser Lys Asp Asn Asn Ser Ile Ile Thr Arg Lys Glu Phe Gln Glu 345 Gly Leu Lys Ile Phe Phe Pro Gly Val Ser Glu Phe Gly Lys Glu Ser Ile Leu Phe His Tyr Thr Asp Trp Val Asp Asp Gln Arg Pro Glu Asn Tyr Arg Glu Ala Leu Gly Asp Val Val Gly Asp Tyr Asn Phe Ile Cys 390 Pro Ala Leu Glu Phe Thr Lys Lys Phe Ser Glu Trp Gly Asn Asn Ala 405 410 Phe Phe Tyr Tyr Phe Glu His Arg Ser Ser Lys Leu Pro Trp Pro Glu Trp Met Gly Val Met His Gly Tyr Glu Ile Glu Phe Val Phe Gly Leu 440 Pro Leu Glu Arg Arg Asp Asn Tyr Thr Lys Ala Glu Glu Ile Leu Ser 455 Arg Ser Ile Val Lys Arg Trp Ala Asn Phe Ala Lys Tyr Gly Asn Pro Asn Glu Thr Gln Asn Asn Ser Thr Ser Trp Pro Val Phe Lys Ser Thr Glu Gln Lys Tyr Leu Thr Leu Asn Thr Glu Ser Thr Arg Ile Met Thr Lys Leu Arg Ala Gln Gln Cys Arg Phe Trp Thr Ser Phe Phe Pro Lys Val Leu Glu Met Thr Gly Asn Ile Asp Glu Ala Glu Trp Glu Trp Lys Ala Gly Phe His Arg Trp Asn Asn Tyr Met Met Asp Trp Lys Asn Gln Phe Asn Asp Tyr Thr Ser Lys Lys Glu Ser Cys Val Gly Leu 565 <210> SEQ ID NO 21 <211> LENGTH: 1722 <212> TYPE: DNA <213 > ORGANISM: Artificial Sequence <220> FEATURE: <223 > OTHER INFORMATION: synthesized <400> SEQUENCE: 21 gaagatgaca tcataattgc aacaaagaat ggaaaagtca gagggatgaa cttgacagtt tttggtggca cggtaacagc ctttcttgga attccctatg cacagccacc tcttggtaga 120 cttcgattca aaaagccaca gtctctgacc aagtggtctg atatttggaa tgccacaaaa tatgcaaatt cttgctgtca gaacatagat caaagttttc caggcttcca tggatcagag atgtggaacc caaacactga cctcagtgaa gactgtttat atctaaatgt atggattcca

gcacctaaac	caaaaaatgc	cactgtattg	atatggattt	atggtggtgg	ttttcaaact	360
ggaacatcat	ctttacatgt	ttatgatggc	aagtttctgg	ctcgggttga	aagagttatt	420
gtagtgtcaa	tgaactatag	ggtgggtgcc	ctaggattct	tagctttgcc	aggaaatcct	480
gaggctccag	ggaacatggg	tttatttgat	caacagttgg	ctcttcagtg	ggttcaaaaa	540
aatatagcag	cctttggtgg	aaatcctaaa	agtgtaactc	tctttggaga	aagttccgga	600
gcagcttcag	ttagcctgca	tttgctttct	cctggaagcc	attcattgtt	caccagagcc	660
attctgcaaa	gtggttccgc	taatgctcct	tgggcggtaa	catctcttta	tgaagctagg	720
aacagaacgt	tgaacttagc	taaattgact	ggttgctcta	gagagaatga	gactgaaata	780
atcaagtgtc	ttagaaataa	agatccccaa	gaaattcttc	tgaatgaagc	atttgttgtc	840
ccctatggga	cttccttggg	tgtaaacttt	ggtccgaccg	tggatggtga	ttttctcact	900
gacatgccag	acatattact	tgaacttgga	caatttaaaa	aaacccagat	tttggtgggt	960
gttaataaag	atgaagggac	atggtttta	gtctacggtg	ctcctggctt	cagcaaagat	1020
aacaatagta	tcataactag	aaaagaattt	caggaaggtt	taaaaatatt	ttttccagga	1080
gtgagtgagt	ttggaaagga	atccatcctt	tttcattaca	cagactgggt	agatgatcag	1140
agacctgaaa	actaccgtga	ggccttgggt	gatgttgttg	gggattataa	tttcatatgc	1200
cctgccttgg	agttcaccaa	gaagttctca	gaatggggaa	ataatgcctt	tttctactat	1260
tttgaacacc	gatcctccaa	acttccgtgg	ccagaatgga	tgggagtgat	gcatggctat	1320
gacattgaat	ttgtctttgg	tttacctctg	gaaagaagag	ataattacac	aaaagccgag	1380
gaaattttga	gtagatccat	agtgaaacgg	tgggcaaatt	ttgcaaaata	tgggaatcca	1440
aatgagactc	agaacaatag	cacaagctgg	cctgtcttca	aaagcactga	acaaaaatat	1500
ctaaccttga	atacagagtc	aacaagaata	atgacgaaac	tacgtgctca	acaatgtcga	1560
ttctggacat	catttttcc	aaaagtcttg	gaaatgacag	gaaatattga	tgaagcagaa	1620
tgggagtgga	aagcaggatt	ccatcgctgg	aacaattaca	tgatggactg	gaaaaatcaa	1680
tttaacgatt	acactagcaa	gaaagaaagt	tgtgtgggtc	tc		1722
010. GEO	rp. No. 00					

```
<210> SEQ ID NO 22
```

<400> SEQUENCE: 22

Glu Asp Asp Ile Ile Ile Ala Thr Lys Asn Gly Lys Val Arg Gly Met 1 5 10 15

Asn Leu Thr Val Phe Gly Gly Thr Val Thr Ala Phe Leu Gly Ile Pro 20 25 30

Tyr Ala Gln Pro Pro Leu Gly Arg Leu Arg Phe Lys Lys Pro Gln Ser $35 \ \ \,$ 40 $\ \ \,$ 45

Leu Thr Lys Trp Ser Asp Ile Trp Asn Ala Thr Lys Tyr Ala Asn Ser 50 $\,$ 60 $\,$

Cys Cys Gln Asn Ile Asp Gln Ser Phe Pro Gly Phe His Gly Ser Glu 70 75

Met Trp Asn Pro Asn Thr Asp Leu Ser Glu Asp Cys Leu Tyr Leu Asn 85 90

Val Trp Ile Pro Ala Pro Lys Pro Lys Asn Ala Thr Val Leu Ile Trp 105

<211> LENGTH: 574 <212> TYPE: PRT

<213 > ORGANISM: Artificial Sequence

<220> FEATURE:

<223 > OTHER INFORMATION: synthesized

-continued

Ile	Tyr	Gly 115	Gly	Gly	Phe	Gln	Thr 120	Gly	Thr	Ser	Ser	Leu 125	His	Val	Tyr
Asp	Gly 130	Lys	Phe	Leu	Ala	Arg 135	Val	Glu	Arg	Val	Ile 140	Val	Val	Ser	Met
Asn 145	Tyr	Arg	Val	Gly	Ala 150	Leu	Gly	Phe	Leu	Ala 155	Leu	Pro	Gly	Asn	Pro 160
Glu	Ala	Pro	Gly	Asn 165	Met	Gly	Leu	Phe	Asp 170	Gln	Gln	Leu	Ala	Leu 175	Gln
Trp	Val	Gln	Lys 180	Asn	Ile	Ala	Ala	Phe 185	Gly	Gly	Asn	Pro	Lys 190	Ser	Val
Thr	Leu	Phe 195	Gly	Glu	Ser	Ser	Gly 200	Ala	Ala	Ser	Val	Ser 205	Leu	His	Leu
Leu	Ser 210	Pro	Gly	Ser	His	Ser 215	Leu	Phe	Thr	Arg	Ala 220	Ile	Leu	Gln	Ser
Gly 225	Ser	Ala	Asn	Ala	Pro 230	Trp	Ala	Val	Thr	Ser 235	Leu	Tyr	Glu	Ala	Arg 240
Asn	Arg	Thr	Leu	Asn 245	Leu	Ala	Lys	Leu	Thr 250	Gly	CÀa	Ser	Arg	Glu 255	Asn
Glu	Thr	Glu	Ile 260	Ile	ГÀв	Cys	Leu	Arg 265	Asn	Lys	Asp	Pro	Gln 270	Glu	Ile
Leu	Leu	Asn 275	Glu	Ala	Phe	Val	Val 280	Pro	Tyr	Gly	Thr	Ser 285	Leu	Gly	Val
Asn	Phe 290	Gly	Pro	Thr	Val	Asp 295	Gly	Asp	Phe	Leu	Thr 300	Asp	Met	Pro	Asp
Ile 305	Leu	Leu	Glu	Leu	Gly 310	Gln	Phe	Lys	Lys	Thr 315	Gln	Ile	Leu	Val	Gly 320
Val	Asn	Lys	Asp	Glu 325	Gly	Thr	Trp	Phe	Leu 330	Val	Tyr	Gly	Ala	Pro 335	Gly
Phe	Ser	Lys	Asp 340	Asn	Asn	Ser	Ile	Ile 345	Thr	Arg	Lys	Glu	Phe 350	Gln	Glu
Gly	Leu	Lys 355	Ile	Phe	Phe	Pro	Gly 360	Val	Ser	Glu	Phe	Gly 365	Lys	Glu	Ser
Ile	Leu 370	Phe	His	Tyr	Thr	Asp 375	Trp	Val	Asp	Asp	Gln 380	Arg	Pro	Glu	Asn
Tyr 385	Arg	Glu	Ala	Leu	Gly 390	Asp	Val	Val	Gly	Asp 395	Tyr	Asn	Phe	Ile	Cys 400
Pro	Ala	Leu	Glu	Phe 405	Thr	ГЛа	ГЛа	Phe	Ser 410	Glu	Trp	Gly	Asn	Asn 415	Ala
Phe	Phe	Tyr	Tyr 420	Phe	Glu	His	Arg	Ser 425	Ser	Lys	Leu	Pro	Trp 430	Pro	Glu
Trp	Met	Gly 435	Val	Met	His	Gly	Tyr 440	Asp	Ile	Glu	Phe	Val 445	Phe	Gly	Leu
Pro	Leu 450	Glu	Arg	Arg	Asp	Asn 455	Tyr	Thr	Lys	Ala	Glu 460	Glu	Ile	Leu	Ser
Arg 465	Ser	Ile	Val	ГÀа	Arg 470	Trp	Ala	Asn	Phe	Ala 475	ГÀа	Tyr	Gly	Asn	Pro 480
Asn	Glu	Thr	Gln	Asn 485	Asn	Ser	Thr	Ser	Trp 490	Pro	Val	Phe	ГЛа	Ser 495	Thr
Glu	Gln	Lys	Tyr 500	Leu	Thr	Leu	Asn	Thr 505	Glu	Ser	Thr	Arg	Ile 510	Met	Thr
Lys	Leu	Arg 515	Ala	Gln	Gln	Cys	Arg 520	Phe	Trp	Thr	Ser	Phe 525	Phe	Pro	Lys
Val	Leu 530	Glu	Met	Thr	Gly	Asn 535	Ile	Asp	Glu	Ala	Glu 540	Trp	Glu	Trp	Lys

-continued

Ala Gly Phe His Arg Trp Asn Asn Tyr Met Met Asp Trp Lys Asn Gln 545 550 555 560

Phe Asn Asp Tyr Thr Ser Lys Lys Glu Ser Cys Val Gly Leu \$565\$

<210> SEQ ID NO 23

<211> LENGTH: 1722

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthesized

<400> SEQUENCE: 23

gaagatgaca tcataattgc aacaaagaat ggaaaagtca gagggatgaa cttgacagtt 60 tttggtggca cggtaacagc ctttcttgga attccctatg cacagccacc tcttggtaga 120 cttcgattca aaaagccaca gtctctgacc aagtggtctg atatttggaa tgccacaaaa 180 tatgcaaatt cttgctgtca gaacatagat caaagttttc caggcttcca tggatcagag 240 atqtqqaacc caaacactqa cctcaqtqaa qactqtttat atctaaatqt atqqattcca 300 gcacctaaac caaaaaatgc cactgtattg atatggattt atggtggtgg ttttcaaact 360 ggaacatcat ctttacatgt ttatgatggc aagtttctgg ctcgggttga aagagttatt 420 gtagtgtcaa tgaactatag ggtgggtgcc ctaggattct tagctttgcc aggaaatcct 480 gaggetecag ggaacatggg tttatttgat caacagttgg etetteagtg ggtteaaaaa 540 aatatagcag cetttggtgg aaateetaaa agtgtaacte tetttggaga aagtteegga 600 gcagcttcag ttagcctgca tttgctttct cctggaagcc attcattgtt caccagagcc 660 attctgcaaa gtggttccgc taatgctcct tgggcggtaa catctcttta tgaagctagg 720 aacagaacgt tgaacttagc taaattgact ggttgctcta gagagaatga gactgaaata 780 atcaagtgtc ttagaaataa agatccccaa gaaattcttc tgaatgaagc atttgttgtc 840 900 ccctatggga ctgctttggg tgtaaacttt ggtccgaccg tggatggtga ttttctcact gacatgccag acatattact tgaacttgga caatttaaaa aaacccagat tttggtgggt gttaataaag atgaagggac atggttttta gtctacggtg ctcctggctt cagcaaagat 1020 aacaatagta tcataactag aaaagaattt caggaaggtt taaaaaatatt ttttccagga 1080 gtgagtgagt ttggaaagga atccatcctt tttcattaca cagactgggt agatgatcag 1140 1200 agacctgaaa actaccgtga ggccttgggt gatgttgttg gggattataa tttcatatgc cctgccttgg agttcaccaa gaagttctca gaatggggaa ataatgcctt tttctactat 1260 tttqaacacc qatcctccaa acttccqtqq ccaqaatqqa tqqqaqtqat qcatqqctat 1320 1380 qacattqaat ttqtctttqq tttacctctq qaaaqaaqaq ataattacac aaaaqccqaq gaaattttga gtagatccat agtgaaacgg tgggcaaatt ttgcaaaata tgggaatcca 1440 aatgagactc agaacaatag cacaagctgg cctgtcttca aaagcactga acaaaaatat 1500 1560 ctaaccttga atacagagtc aacaagaata atgacgaaac tacgtgctca acaatgtcga 1620 ttctggacat catttttcc aaaagtcttg gaaatgacag gaaatattga tgaagcagaa tgggagtgga aagcaggatt ccatcgctgg aacaattaca tgatggactg gaaaaatcaa 1680 1722 tttaacgatt acactagcaa gaaagaaagt tgtgtgggtc tc

<210> SEQ ID NO 24

<211> LENGTH: 574

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE: <223> OTHER INFORMATION: synthesized <400> SEQUENCE: 24 Glu Asp Asp Ile Ile Ile Ala Thr Lys Asn Gly Lys Val Arg Gly Met Asn Leu Thr Val Phe Gly Gly Thr Val Thr Ala Phe Leu Gly Ile Pro Tyr Ala Gln Pro Pro Leu Gly Arg Leu Arg Phe Lys Lys Pro Gln Ser $35 \ \ \,$ 45 Leu Thr Lys Trp Ser Asp Ile Trp Asn Ala Thr Lys Tyr Ala Asn Ser 50 $\,$ 60 Cys Cys Gln Asn Ile Asp Gln Ser Phe Pro Gly Phe His Gly Ser Glu 65 70 75 80Met Trp Asn Pro Asn Thr Asp Leu Ser Glu Asp Cys Leu Tyr Leu Asn 90 Val Trp Ile Pro Ala Pro Lys Pro Lys Asn Ala Thr Val Leu Ile Trp Ile Tyr Gly Gly Gly Phe Gln Thr Gly Thr Ser Ser Leu His Val Tyr 120 Asp Gly Lys Phe Leu Ala Arg Val Glu Arg Val Ile Val Val Ser Met 135 Asn Tyr Arg Val Gly Ala Leu Gly Phe Leu Ala Leu Pro Gly Asn Pro 150 Glu Ala Pro Gly Asn Met Gly Leu Phe Asp Gln Gln Leu Ala Leu Gln Trp Val Gln Lys Asn Ile Ala Ala Phe Gly Gly Asn Pro Lys Ser Val 185 Thr Leu Phe Gly Glu Ser Ser Gly Ala Ala Ser Val Ser Leu His Leu Leu Ser Pro Gly Ser His Ser Leu Phe Thr Arg Ala Ile Leu Gln Ser Gly Ser Ala Asn Ala Pro Trp Ala Val Thr Ser Leu Tyr Glu Ala Arg Asn Arg Thr Leu Asn Leu Ala Lys Leu Thr Gly Cys Ser Arg Glu Asn Glu Thr Glu Ile Ile Lys Cys Leu Arg Asn Lys Asp Pro Gln Glu Ile Leu Leu Asn Glu Ala Phe Val Val Pro Tyr Gly Thr Ala Leu Gly Val 280 Asn Phe Gly Pro Thr Val Asp Gly Asp Phe Leu Thr Asp Met Pro Asp Ile Leu Leu Glu Leu Gly Gln Phe Lys Lys Thr Gln Ile Leu Val Gly 310 315 Val Asn Lys Asp Glu Gly Thr Trp Phe Leu Val Tyr Gly Ala Pro Gly 325 330 Phe Ser Lys Asp Asn Asn Ser Ile Ile Thr Arg Lys Glu Phe Gln Glu Gly Leu Lys Ile Phe Phe Pro Gly Val Ser Glu Phe Gly Lys Glu Ser 360 Ile Leu Phe His Tyr Thr Asp Trp Val Asp Asp Gln Arg Pro Glu Asn 375 Tyr Arg Glu Ala Leu Gly Asp Val Val Gly Asp Tyr Asn Phe Ile Cys 395

-continued

Pro Ala Leu Glu Phe Thr Lys Lys Phe Ser Glu Trp Gly Asn Asn Ala 405 Phe Phe Tyr Tyr Phe Glu His Arg Ser Ser Lys Leu Pro Trp Pro Glu Trp Met Gly Val Met His Gly Tyr Asp Ile Glu Phe Val Phe Gly Leu Pro Leu Glu Arg Arg Asp Asn Tyr Thr Lys Ala Glu Glu Ile Leu Ser Arg Ser Ile Val Lys Arg Trp Ala Asn Phe Ala Lys Tyr Gly Asn Pro 470 Asn Glu Thr Gln Asn Asn Ser Thr Ser Trp Pro Val Phe Lys Ser Thr 485 Glu Gln Lys Tyr Leu Thr Leu Asn Thr Glu Ser Thr Arg Ile Met Thr 505 Lys Leu Arg Ala Gln Gln Cys Arg Phe Trp Thr Ser Phe Phe Pro Lys 520 Val Leu Glu Met Thr Gly Asn Ile Asp Glu Ala Glu Trp Glu Trp Lys 535 540 Ala Gly Phe His Arg Trp Asn Asn Tyr Met Met Asp Trp Lys Asn Gln 550

Phe Asn Asp Tyr Thr Ser Lys Lys Glu Ser Cys Val Gly Leu 565 570

<210> SEQ ID NO 25

<211> LENGTH: 1722

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthesized

<400> SEQUENCE: 25

gaagatgaca tcataattgc aacaaagaat ggaaaagtca gagggatgaa cttgacagtt 60 tttggtggca cggtaacagc ctttcttgga attccctatg cacagccacc tcttggtaga cttcgattca aaaagccaca gtctctgacc aagtggtctg atatttggaa tgccacaaaa tatgcaaatt cttgctgtca gaacatagat caaagttttc caggcttcca tggatcagag 240 atgtggaacc caaacactga cctcagtgaa gactgtttat atctaaatgt atggattcca 300 gcacctaaac caaaaaatgc cactgtattg atatggattt atggtggtgg ttttcaaact 360 ggaacatcat ctttacatgt ttatgatggc aagtttctgg ctcgggttga aagagttatt 420 gtagtgtcaa tgaactatag ggtgggtgcc ctaggattct tagctttgcc aggaaatcct 480 qaqqctccaq qqaacatqqq tttatttqat caacaqttqq ctcttcaqtq qqttcaaaaa 540 aatataqcaq cctttqqtqq aaatcctaaa aqtqtaactc tctttqqaqa aaqttccqqa 600 gcagcttcag ttagcctgca tttgctttct cctggaagcc attcattgtt caccagagcc 660 attetgeaaa gtggtteece gaatgeteet tgggeggtaa catetetta tgaagetagg 720 aacagaacgt tgaacttagc taaattgact ggttgctcta gagagaatga gactgaaata 780 atcaagtgtc ttagaaataa agatccccaa gaaattcttc tgaatgaagc atttgttgtc 840 ccctatggga ctcctatggg tgtaaacttt ggtccgaccg tggatggtga ttttctcact 900 gacatgccag acatattact tgaacttgga caatttaaaa aaacccagat tttggtgggt 960 gttaataaag atgaagggac atggttttta gtcggtggtg ctcctggctt cagcaaagat 1020 aacaatagta tcataactag aaaagaattt caggaaggtt taaaaaatatt ttttccagga 1080 qtqaqtqaqt ttqqaaaqqa atccatcctt tttcattaca caqactqqqt aqatqatcaq 1140

agad	ctga	aa a	actac	ccgto	ga go	gaatt	gggt	gat	gttg	gttg	ggga	attat	aa t	ttca	atatgc	1200
ccts	geett	gg a	agtto	cacca	aa ga	agtt	ctca	a gaa	tggg	ggaa	ataa	atgco	ett t	ttct	actat	1260
tttç	gaaca	icc ç	gatco	ctcca	aa ac	ettec	gtgg	g cca	igaat	gga	tggg	gagto	gat q	gcato	ggctat	1320
gaaa	ittga	at t	tgto	ettte	gg tt	taco	etete	g gaa	agaa	ıgag	ataa	attad	cac a	aaaag	gccgag	1380
gaaa	atttt	ga g	gtaga	atcca	at aç	gtgaa	acgo	g tgg	gcaa	att	ttgo	caaaa	ata t	ggga	atcca	1440
aatg	gagac	etc a	agaad	caata	ag ca	caaç	gatgo	g cct	gtct	tca	aaag	gcact	ga a	acaaa	aatat	1500
ctaa	ecctt	ga a	ataca	agagt	c aa	caaç	gaata	ato	gacga	aac	taco	gtgct	ca a	acaat	gtcga	1560
ttct	ggad	at o	attt	tttc	cc aa	aagt	ctte	g gaa	atga	cag	gaaa	atatt	ga t	gaag	gcagaa	1620
tggg	gagto	gga a	agca	aggat	t co	catco	getge	g aac	aatt	aca	tgat	ggad	ctg ç	gaaaa	atcaa	1680
ttta	acga	att a	acact	agca	aa ga	aaga	aagt	tgt:	gtgg	ggtc	tc					1722
<211 <212 <213 <220)> SE L> LE 2> T\ 3> OF 0> FE 3> OT	ENGTH PE: RGANI EATUR	H: 57 PRT ISM: RE:	74 Arti			-		ı							
< 400)> SE	EQUE1	ICE :	26												
Glu 1	Asp	Asp	Ile	Ile 5	Ile	Ala	Thr	ГЛа	Asn 10	Gly	Lys	Val	Arg	Gly 15	Met	
Asn	Leu	Thr	Val 20	Phe	Gly	Gly	Thr	Val 25	Thr	Ala	Phe	Leu	Gly 30	Ile	Pro	
Tyr	Ala	Gln 35	Pro	Pro	Leu	Gly	Arg 40	Leu	Arg	Phe	Lys	Lys 45	Pro	Gln	Ser	
Leu	Thr 50	Lys	Trp	Ser	Asp	Ile 55	Trp	Asn	Ala	Thr	60 Fåa	Tyr	Ala	Asn	Ser	
65 Cys	Cys	Gln	Asn	Ile	Asp 70	Gln	Ser	Phe	Pro	Gly 75	Phe	His	Gly	Ser	Glu 80	
Met	Trp	Asn	Pro	Asn 85	Thr	Asp	Leu	Ser	Glu 90	Asp	Cys	Leu	Tyr	Leu 95	Asn	
Val	Trp	Ile	Pro 100	Ala	Pro	Lys	Pro	Lys 105	Asn	Ala	Thr	Val	Leu 110	Ile	Trp	
Ile	Tyr	Gly 115	Gly	Gly	Phe	Gln	Thr 120	Gly	Thr	Ser	Ser	Leu 125	His	Val	Tyr	
Asp	Gly 130	Lys	Phe	Leu	Ala	Arg 135	Val	Glu	Arg	Val	Ile 140	Val	Val	Ser	Met	
Asn 145	Tyr	Arg	Val	Gly	Ala 150	Leu	Gly	Phe	Leu	Ala 155	Leu	Pro	Gly	Asn	Pro 160	
Glu	Ala	Pro	Gly	Asn 165	Met	Gly	Leu	Phe	Asp 170	Gln	Gln	Leu	Ala	Leu 175	Gln	
Trp	Val	Gln	Lys 180	Asn	Ile	Ala	Ala	Phe 185	Gly	Gly	Asn	Pro	Lys 190	Ser	Val	
Thr	Leu	Phe 195	Gly	Glu	Ser	Ser	Gly 200	Ala	Ala	Ser	Val	Ser 205	Leu	His	Leu	
Leu	Ser 210	Pro	Gly	Ser	His	Ser 215	Leu	Phe	Thr	Arg	Ala 220	Ile	Leu	Gln	Ser	
Gly 225	Ser	Pro	Asn	Ala	Pro 230	Trp	Ala	Val	Thr	Ser 235	Leu	Tyr	Glu	Ala	Arg 240	
Asn	Arg	Thr	Leu	Asn 245	Leu	Ala	Lys	Leu	Thr 250	Gly	Cya	Ser	Arg	Glu 255	Asn	

Glu Thr Glu Ile Ile Lys Cys Leu Arg Asn Lys Asp Pro Gln Glu Ile

												COII	L IIII	uea			
			260					265					270				
Leu	Leu	Asn 275	Glu	Ala	Phe	Val	Val 280	Pro	Tyr	Gly	Thr	Pro 285	Met	Gly	Val		
Asn	Phe 290	Gly	Pro	Thr	Val	Asp 295	Gly	Asp	Phe	Leu	Thr 300	Asp	Met	Pro	Asp		
Ile 305	Leu	Leu	Glu	Leu	Gly 310	Gln	Phe	Lys	Lys	Thr 315	Gln	Ile	Leu	Val	Gly 320		
Val	Asn	Lys	Asp	Glu 325	Gly	Thr	Trp	Phe	Leu 330	Val	Gly	Gly	Ala	Pro 335	Gly		
Phe	Ser	Lys	Asp 340	Asn	Asn	Ser	Ile	Ile 345	Thr	Arg	Lys	Glu	Phe 350	Gln	Glu		
Gly	Leu	355 Lys	Ile	Phe	Phe	Pro	Gly 360	Val	Ser	Glu	Phe	Gly 365	Lys	Glu	Ser		
Ile	Leu 370	Phe	His	Tyr	Thr	Asp 375	Trp	Val	Asp	Asp	Gln 380	Arg	Pro	Glu	Asn		
Tyr 385	Arg	Glu	Ala	Leu	Gly 390	Asp	Val	Val	Gly	Asp 395	Tyr	Asn	Phe	Ile	Cys 400		
Pro	Ala	Leu	Glu	Phe 405	Thr	Lys	Lys	Phe	Ser 410	Glu	Trp	Gly	Asn	Asn 415	Ala		
Phe	Phe	Tyr	Tyr 420	Phe	Glu	His	Arg	Ser 425	Ser	Lys	Leu	Pro	Trp 430	Pro	Glu		
Trp	Met	Gly 435	Val	Met	His	Gly	Tyr 440	Glu	Ile	Glu	Phe	Val 445	Phe	Gly	Leu		
Pro	Leu 450	Glu	Arg	Arg	Asp	Asn 455	Tyr	Thr	Lys	Ala	Glu 460	Glu	Ile	Leu	Ser		
Arg 465	Ser	Ile	Val	ГЛа	Arg 470	Trp	Ala	Asn	Phe	Ala 475	rys	Tyr	Gly	Asn	Pro 480		
Asn	Glu	Thr	Gln	Asn 485	Asn	Ser	Thr	Ser	Trp 490	Pro	Val	Phe	Lys	Ser 495	Thr		
Glu	Gln	ГÀЗ	Tyr 500	Leu	Thr	Leu	Asn	Thr 505	Glu	Ser	Thr	Arg	Ile 510	Met	Thr		
Lys	Leu	Arg 515	Ala	Gln	Gln	Cys	Arg 520	Phe	Trp	Thr	Ser	Phe 525	Phe	Pro	Lys		
Val	Leu 530	Glu	Met	Thr	Gly	Asn 535	Ile	Asp	Glu	Ala	Glu 540	Trp	Glu	Trp	Lys		
Ala 545	Gly	Phe	His	Arg	Trp 550	Asn	Asn	Tyr	Met	Met 555	Asp	Trp	Lys	Asn	Gln 560		
Phe	Asn	Asp	Tyr	Thr 565	Ser	ГÀа	ГЛа	Glu	Ser 570	CÀa	Val	Gly	Leu				
<211 <212 <213 <220	-> LF 2> TY 3> OF 0> FF	EQ II ENGTH YPE: RGANI EATUH	H: 17 DNA [SM: RE:	722 Art:			_		1								
< 400)> SI	EQUE1	ICE :	27													
gaaç	gatga	aca t	cata	aatto	gc aa	acaaa	agaat	gga	aaaq	gtca	gag	ggato	gaa d	cttga	acagtt	60	
tttç	gtgg	gca d	ggta	aacaç	gc ct	ttct	tgga	a att	ccct	atg	caca	agcca	acc t	ctt	ggtaga	120	
ctto	gatt	ca a	aaaq	gccad	ca gt	ctct	gaco	c aaq	gtggt	ctg	atat	ttg	gaa t	gcca	acaaaa	180	
tato	gcaaa	att o	ettge	ctgto	ca ga	acat	agat	caa	agtt	ttc	cago	gette	cca t	ggat	cagag	240	
atgt	ggaa	acc c	caaac	cacto	ga co	ctcaç	gtgaa	a gad	etgtt	tat	atct	caaat	gt a	atgga	attcca	300	

gcacctaaac caaaaaatgc cactgtattg atatggattt atggtggtgg ttttcaaact

-continued

ggaacatcat ctttacatgt ttatgatggc aagtttctgg ctcgggttga aagagttatt	420
gtagtgtcaa tgaactatag ggtgggtgcc ctaggattct tagctttgcc aggaaatcct	480
gaggetecag ggaacatggg tttatttgat caacagttgg etetteagtg ggtteaaaaa	540
aatatagcag cetttggtgg aaateetaaa agtgtaaete tetttggaga aagtteegga	600
gcagettcag ttageetgea tttgetttet eetggaagee atteattgtt caccagagee	660
attotgoaaa gtggttoogg taatgotoot tgggoggtaa catotottta tgaagotagg	720
aacagaacgt tgaacttagc taaattgact ggttgctcta gagagaatga gactgaaata	780
atcaagtgtc ttagaaataa agatccccaa gaaattcttc tgaatgaagc atttgttgtc	840
ccctatggga ctgctttggg tgtaaacttt ggtccgaccg tggatggtga ttttctcact	900
gacatgccag acatattact tgaacttgga caatttaaaa aaacccagat tttggtgggt	960
gttaataaag atgaagggac atggttttta gtcggtggtg ctcctggctt cagcaaagat	1020
aacaatagta tcataactag aaaagaattt caggaaggtt taaaaatatt ttttccagga	1080
gtgagtgagt ttggaaagga atccatcctt tttcattaca cagactgggt agatgatcag	1140
agacctgaaa actaccgtga ggccttgggt gatgttgttg gggattataa tttcatatgc	1200
cctgccttgg agttcaccaa gaagttctca gaatggggaa ataatgcctt tttctactat	1260
tttgaacacc gatcctccaa acttccgtgg ccagaatgga tgggagtgat gcatggctat	1320
gaaattgaat ttgtctttgg tttacctctg gaaagaagag ataattacac aaaagccgag	1380
gaaattttga gtagatccat agtgaaacgg tgggcaaatt ttgcaaaata tgggaatcca	1440
aatgagactc agaacaatag cacaagctgg cctgtcttca aaagcactga acaaaaatat	1500
ctaacettga atacagagte aacaagaata atgacgaaac tacgtgetea acaatgtega	1560
ttctggacat cattttttcc aaaagtcttg gaaatgacag gaaatattga tgaagcagaa	1620
tgggagtgga aagcaggatt ccatcgctgg aacaattaca tgatggactg gaaaaatcaa	1680
tttaacgatt acactagcaa gaaagaaagt tgtgtgggtc tc	1722
<210> SEQ ID NO 28 <211> LENGTH: 574 <212> TYPE: PRT <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: synthesized	
<400> SEQUENCE: 28	
Glu Asp Asp Ile Ile Ile Ala Thr Lys Asn Gly Lys Val Arg Gly Met 1 10 15	
Asn Leu Thr Val Phe Gly Gly Thr Val Thr Ala Phe Leu Gly Ile Pro 20 25 30	
Tyr Ala Gln Pro Pro Leu Gly Arg Leu Arg Phe Lys Lys Pro Gln Ser 35 40 45	
Leu Thr Lys Trp Ser Asp Ile Trp Asn Ala Thr Lys Tyr Ala Asn Ser 50 55 60	
Cys Cys Gln Asn Ile Asp Gln Ser Phe Pro Gly Phe His Gly Ser Glu 65 70 75 80	
Met Trp Asn Pro Asn Thr Asp Leu Ser Glu Asp Cys Leu Tyr Leu Asn 85 90 95	

Val Trp Ile Pro Ala Pro Lys Pro Lys Asn Ala Thr Val Leu Ile Trp 100 105 110

Ile Tyr Gly Gly Gly Phe Gln Thr Gly Thr Ser Ser Leu His Val Tyr \$115\$ \$120\$ \$125\$

Asp Gly L	ys Phe	Leu Al	a Arg		Glu	Arg	Val	Ile 140	Val	Val	Ser	Met
Asn Tyr A 145	rg Val	Gly Al		. Gly	Phe	Leu	Ala 155	Leu	Pro	Gly	Asn	Pro 160
Glu Ala P	ro Gly	Asn Me	t Gly	Leu	Phe	Asp 170	Gln	Gln	Leu	Ala	Leu 175	Gln
Trp Val G	ln Lys 180	Asn II	e Ala	Ala	Phe 185	Gly	Gly	Asn	Pro	Lys 190	Ser	Val
Thr Leu P	he Gly 95	Glu Se	r Ser	Gly 200	Ala	Ala	Ser	Val	Ser 205	Leu	His	Leu
Leu Ser P 210	ro Gly	Ser H	s Ser 215		Phe	Thr	Arg	Ala 220	Ile	Leu	Gln	Ser
Gly Ser G 225	ly Asn	Ala Pi 23		Ala	Val	Thr	Ser 235	Leu	Tyr	Glu	Ala	Arg 240
Asn Arg T	hr Leu	Asn Le 245	u Ala	. Lys	Leu	Thr 250	Gly	Cys	Ser	Arg	Glu 255	Asn
Glu Thr G	lu Ile 260	Ile Ly	a Cya	Leu	Arg 265	Asn	Lys	Asp	Pro	Gln 270	Glu	Ile
	75			280					285			
Asn Phe G 290			295					300	_			_
Ile Leu L 305	eu Glu	Leu Gi		Phe	ГÀа	Lys	Thr 315	Gln	Ile	Leu	Val	Gly 320
Val Asn L	ya Aap	Glu G 325	y Thr	Trp	Phe	Leu 330	Val	Gly	Gly	Ala	Pro 335	Gly
Phe Ser L	ys Asp 340	Asn As	n Ser	Ile	Ile 345	Thr	Arg	ГÀв	Glu	Phe 350	Gln	Glu
Gly Leu L 3	ys Ile 55	Phe Pl	e Pro	360	Val	Ser	Glu	Phe	Gly 365	Lys	Glu	Ser
Ile Leu P 370			375					380				
Tyr Arg G 385		39	0			-	395	-				400
Pro Ala L		405	_	_		410		_	_		415	
Phe Phe T	yr Tyr 420		u His		Ser 425			Leu		Trp 430		Glu
	35		_	440					445		_	
Pro Leu G 450	lu Arg	Arg As	p Asn 455	-	Thr	Lys	Ala	Glu 460	Glu	Ile	Leu	Ser
Arg Ser I 465	le Val	Lys Ai		Ala	Asn	Phe	Ala 475	ГÀа	Tyr	Gly	Asn	Pro 480
Asn Glu T	hr Gln	Asn As 485	n Ser	Thr	Ser	Trp 490	Pro	Val	Phe	ГÀв	Ser 495	Thr
Glu Gln L	500				505					510		
Lys Leu A 5	rg Ala 15	Gln G	n Cys	Arg 520	Phe	Trp	Thr	Ser	Phe 525	Phe	Pro	ГЛа
Val Leu G 530	lu Met	Thr G	y Asn 535		Asp	Glu	Ala	Glu 540	Trp	Glu	Trp	Lys
Ala Gly P	he His	Arg Ti	p Asn	Asn	Tyr	Met	Met	Asp	Trp	Lys	Asn	Gln

555

-continued

Phe Asn Asp Tyr Thr Ser Lys Lys Glu Ser Cys Val Gly Leu

<210> SEQ ID NO 29

<211> LENGTH: 1722

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

545

<223 > OTHER INFORMATION: synthesized

<400> SEQUENCE: 29

gaagatgaca tcataattgc aacaaagaat ggaaaagtca gagggatgaa cttgacagtt 60 tttggtggca cggtaacagc ctttcttgga attccctatg cacagccacc tcttggtaga 120 cttcgattca aaaagccaca gtctctgacc aagtggtctg atatttggaa tgccacaaaa 180 tatgcaaatt cttgctgtca gaacatagat caaagttttc caggcttcca tggatcagag 240 atgtggaacc caaacactga cctcagtgaa gactgtttat atctaaatgt atggattcca 300 qcacctaaac caaaaaatqc cactqtattq atatqqattt atqqtqqtqq ttttcaaact 360 ggaacatcat ctttacatgt ttatgatggc aagtttctgg ctcgggttga aagagttatt 420 gtagtgtcaa tgaactatag ggtgggtgcc ctaggattct tagctttgcc aggaaatcct 480 gaggetecag ggaacatggg tttatttgat caacagttgg etetteagtg ggtteaaaaa 540 aatatagcag cetttggtgg aaateetaaa agtgtaacte tetttggaga aagtteegga 600 gcagcttcag ttagcctgca tttgctttct cctggaagcc attcattgtt caccagagcc 660 attctgcaaa gtggttccgg taatgctcct tgggcggtaa catctcttta tgaagctagg 720 aacagaacgt tgaacttagc taaattgact ggttgctcta gagagaatga gactgaaata 780 atcaagtgtc ttagaaataa agatccccaa gaaattcttc tgaatgaagc atttgttgtc 840 ccctatggga ctggtttggg tgtaaacttt ggtccgaccg tggatggtga ttttctcact 900 gacatgccag acatattact tgaacttgga caatttaaaa aaacccagat tttggtgggt 960 gttaataaag atgaagggac atggttttta gtcggtggtg ctcctggctt cagcaaagat aacaatagta tcataactag aaaagaattt caggaaggtt taaaaatatt ttttccagga 1080 gtgagtgagt ttggaaagga atccatcctt tttcattaca cagactgggt agatgatcag 1140 agacctgaaa actaccgtga ggccttgggt gatgttgttg gggattataa tttcatatgc 1200 cctgccttgg agttcaccaa gaagttctca gaatggggaa ataatgcctt tttctactat 1260 1320 tttgaacacc gatcctccaa acttccgtgg ccagaatgga tgggagtgat gcatggctat qaaattqaat ttqtctttqq tttacctctq qaaaqaaqaq ataattacac aaaaqccqaq 1380 gaaattttga gtagatccat agtgaaacgg tgggcaaatt ttgcaaaata tgggaatcca 1440 aatgagactc agaacaatag cacaagctgg cctgtcttca aaagcactga acaaaaatat 1500 ctaaccttga atacagagtc aacaagaata atgacgaaac tacgtgctca acaatgtcga 1560 ttctggacat catttttcc aaaagtcttg gaaatgacag gaaatattga tgaagcagaa 1620 1680 tgggagtgga aagcaggatt ccatcgctgg aacaattaca tgatggactg gaaaaatcaa 1722 tttaacgatt acactagcaa gaaagaaagt tgtgtgggtc tc

<210> SEQ ID NO 30

<211> LENGTH: 574

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223 > OTHER INFORMATION: synthesized

<400> SEQUENCE:	30					
Glu Asp Asp Ile 1	Ile Ile Ala 5	Thr Lys	Asn Gly 10	Lys Val	Arg Gly 15	Met
Asn Leu Thr Val	Phe Gly Gly	Thr Val 25	Thr Ala	Phe Leu	Gly Ile 30	Pro
Tyr Ala Gln Pro 35	Pro Leu Gly	Arg Leu 40	Arg Phe	Lys Lys 45	Pro Gln	Ser
Leu Thr Lys Trp 50	Ser Asp Ile 55	Trp Asn	Ala Thr	Lys Tyr 60	Ala Asn	Ser
Cys Cys Gln Asn 65	Ile Asp Gln 70	Ser Phe	Pro Gly 75	Phe His	Gly Ser	Glu 80
Met Trp Asn Pro	Asn Thr Asp 85	Leu Ser	90 Glu Asp	Cys Leu	Tyr Leu 95	Asn
Val Trp Ile Pro 100	Ala Pro Lys	Pro Lys 105	Asn Ala	Thr Val	Leu Ile 110	Trp
Ile Tyr Gly Gly 115	Gly Phe Gln	Thr Gly 120	Thr Ser	Ser Leu 125		Tyr
Asp Gly Lys Phe 130	Leu Ala Arg 135		Arg Val	Ile Val 140	Val Ser	Met
Asn Tyr Arg Val 145	Gly Ala Leu 150	. Gly Phe	Leu Ala 155	Leu Pro	Gly Asn	Pro 160
Glu Ala Pro Gly	Asn Met Gly 165	Leu Phe	Asp Gln 170	Gln Leu	Ala Leu 175	Gln
Trp Val Gln Lys 180	Asn Ile Ala	Ala Phe 185	Gly Gly	Asn Pro	Lys Ser 190	Val
Thr Leu Phe Gly 195	Glu Ser Ser	Gly Ala 200	Ala Ser	Val Ser 205		Leu
Leu Ser Pro Gly 210	Ser His Ser 215		Thr Arg	Ala Ile 220	Leu Gln	Ser
Gly Ser Gly Asn 225	Ala Pro Trp 230	Ala Val	Thr Ser 235	Leu Tyr	Glu Ala	Arg 240
Asn Arg Thr Leu	Asn Leu Ala 245	. Lys Leu	Thr Gly 250	Cys Ser	Arg Glu 255	Asn
Glu Thr Glu Ile 260	Ile Lys Cys	Leu Arg 265	Asn Lys	Asp Pro	Gln Glu 270	Ile
Leu Leu Asn Glu 275	Ala Phe Val	Val Pro 280	Tyr Gly	Thr Gly 285		Val
Asn Phe Gly Pro 290	Thr Val Asp 295		Phe Leu	Thr Asp	Met Pro	Asp
Ile Leu Leu Glu 305	Leu Gly Gln 310	Phe Lys	Lys Thr 315	Gln Ile	Leu Val	Gly 320
Val Asn Lys Asp	Glu Gly Thr 325	Trp Phe	Leu Val 330	Gly Gly	Ala Pro 335	Gly
Phe Ser Lys Asp 340	Asn Asn Ser	Ile Ile 345	Thr Arg	Lys Glu	Phe Gln 350	Glu
Gly Leu Lys Ile 355	Phe Phe Pro	Gly Val 360	Ser Glu	Phe Gly 365	-	Ser
Ile Leu Phe His 370	Tyr Thr Asp		Asp Asp	Gln Arg 380	Pro Glu	Asn
Tyr Arg Glu Ala 385	Leu Gly Asp 390	Val Val	Gly Asp 395	Tyr Asn	Phe Ile	Cys 400
Pro Ala Leu Glu	Phe Thr Lys 405	Lys Phe	Ser Glu 410	Trp Gly	Asn Asn 415	Ala

-continued

 Phe
 Phe
 Tyr 420
 Phe 61u His 420
 <t

<210> SEQ ID NO 31

<211> LENGTH: 1722

<211> LENGIH: 1/2 <212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthesized

<400> SEQUENCE: 31

gaagatgaca tcataattgc aacaaagaat ggaaaagtca gagggatgaa cttgacagtt 60 tttggtggca cggtaacagc ctttcttgga attccctatg cacagccacc tcttggtaga 120 cttcgattca aaaagccaca gtctctgacc aagtggtctg atatttggaa tgccacaaaa tatgcaaatt cttgctgtca gaacatagat caaagttttc caggcttcca tggatcagag atgtggaacc caaacactga cctcagtgaa gactgtttat atctaaatgt atggattcca 300 gcacctaaac caaaaaatgc cactgtattg atatggattt atggtggtgg ttttcaaact 360 ggaacatcat ctttacatgt ttatgatggc aagtttctgg ctcgggttga aagagttatt 420 gtagtgtcaa tgaactatag ggtgggtgcc ctaggattct tagctttgcc aggaaatcct 480 gaggetecag ggaacatggg tttatttgat caacagttgg etetteagtg ggtteaaaaa 540 aatataqcaq cctttqqtqq aaatcctaaa aqtqtaactc tctttqqaqa aaqttccqqa 600 gcagcttcag ttagcctgca tttgctttct cctggaagcc attcattgtt caccagagcc 660 attctgcaaa gtggttccgg taatgctcct tgggcggtaa catctcttta tgaagctagg 720 aacagaacgt tgaacttagc taaattgact ggttgctcta gagagaatga gactgaaata 780 atcaagtgtc ttagaaataa agatccccaa gaaattcttc tgaatgaagc atttgttgtc 840 ccctatggga ctcagttggg tgtaaacttt ggtccgaccg tggatggtga ttttctcact 900 gacatgccag acatattact tgaacttgga caatttaaaa aaacccagat tttggtgggt 960 gttaataaag atgaagggac atggttttta gtcggtggtg ctcctggctt cagcaaagat 1020 aacaatagta tcataactag aaaagaattt caggaaggtt taaaaatatt ttttccagga 1080 gtgagtgagt ttggaaagga atccatcctt tttcattaca cagactgggt agatgatcag 1140 agacctgaaa actaccgtga ggccttgggt gatgttgttg gggattataa tttcatatgc 1200

-continued

cctgccttgg agttcaccaa gaagttctca gaatggggaa ataatgcctt tttctactat tttgaacacc gatcctccaa acttccgtgg ccagaatgga tgggagtgat gcatggctat gaaattgaat ttgtctttgg tttacctctg gaaagaagag ataattacac aaaagccgag gaaattttga gtagatccat agtgaaacgg tgggcaaatt ttgcaaaata tgggaatcca aatqaqactc aqaacaataq cacaaqctqq cctqtcttca aaaqcactqa acaaaaatat 1500 ctaaccttqa atacaqaqtc aacaaqaata atqacqaaac tacqtqctca acaatqtcqa ttctggacat catttttcc aaaagtcttg gaaatgacag gaaatattga tgaagcagaa 1620 tgggagtgga aagcaggatt ccatcgctgg aacaattaca tgatggactg gaaaaatcaa 1680 1722 tttaacgatt acactagcaa gaaagaaagt tgtgtgggtc tc <210> SEQ ID NO 32 <211> LENGTH: 574 <212> TYPE: PRT <213 > ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: synthesized <400> SEQUENCE: 32 Glu Asp Asp Ile Ile Ile Ala Thr Lys Asn Gly Lys Val Arg Gly Met 10 Asn Leu Thr Val Phe Gly Gly Thr Val Thr Ala Phe Leu Gly Ile Pro 25 Tyr Ala Gln Pro Pro Leu Gly Arg Leu Arg Phe Lys Lys Pro Gln Ser Leu Thr Lys Trp Ser Asp Ile Trp Asn Ala Thr Lys Tyr Ala Asn Ser Cys Cys Gln Asn Ile Asp Gln Ser Phe Pro Gly Phe His Gly Ser Glu Met Trp Asn Pro Asn Thr Asp Leu Ser Glu Asp Cys Leu Tyr Leu Asn Val Trp Ile Pro Ala Pro Lys Pro Lys Asn Ala Thr Val Leu Ile Trp Ile Tyr Gly Gly Gly Phe Gln Thr Gly Thr Ser Ser Leu His Val Tyr Asp Gly Lys Phe Leu Ala Arg Val Glu Arg Val Ile Val Val Ser Met Asn Tyr Arg Val Gly Ala Leu Gly Phe Leu Ala Leu Pro Gly Asn Pro Glu Ala Pro Gly Asn Met Gly Leu Phe Asp Gln Gln Leu Ala Leu Gln Trp Val Gln Lys Asn Ile Ala Ala Phe Gly Gly Asn Pro Lys Ser Val 185 Thr Leu Phe Gly Glu Ser Ser Gly Ala Ala Ser Val Ser Leu His Leu 200 Leu Ser Pro Gly Ser His Ser Leu Phe Thr Arg Ala Ile Leu Gln Ser Gly Ser Gly Asn Ala Pro Trp Ala Val Thr Ser Leu Tyr Glu Ala Arg 230 235 Asn Arg Thr Leu Asn Leu Ala Lys Leu Thr Gly Cys Ser Arg Glu Asn 245 250

Glu Thr Glu Ile Ile Lys Cys Leu Arg Asn Lys Asp Pro Gln Glu Ile 265

-continued

Leu Leu Asn Glu Ala Phe Val Val Pro Tyr Gly Thr Gln Leu Gly Val 275 280 285

Asn	Phe 290	Gly	Pro	Thr	Val	Asp 295	Gly	Asp	Phe	Leu	Thr 300	Asp	Met	Pro	Asp	
Ile 305	Leu	Leu	Glu	Leu	Gly 310	Gln	Phe	Lys	Lys	Thr 315	Gln	Ile	Leu	Val	Gly 320	
Val	Asn	Lys	Asp	Glu 325	Gly	Thr	Trp	Phe	Leu 330	Val	Gly	Gly	Ala	Pro 335	Gly	
Phe	Ser	Lys	Asp 340	Asn	Asn	Ser	Ile	Ile 345	Thr	Arg	Lys	Glu	Phe 350	Gln	Glu	
Gly	Leu	Lys 355	Ile	Phe	Phe	Pro	Gly 360	Val	Ser	Glu	Phe	Gly 365	Lys	Glu	Ser	
	Leu 370	Phe	His	Tyr	Thr	Asp 375	Trp	Val	Asp	Asp	Gln 380	Arg	Pro	Glu	Asn	
Tyr 385	Arg	Glu	Ala	Leu	Gly 390	Asp	Val	Val	Gly	Asp 395	Tyr	Asn	Phe	Ile	Cys 400	
Pro	Ala	Leu	Glu	Phe 405	Thr	Lys	Lys	Phe	Ser 410	Glu	Trp	Gly	Asn	Asn 415	Ala	
Phe	Phe	Tyr	Tyr 420	Phe	Glu	His	Arg	Ser 425	Ser	Lys	Leu	Pro	Trp 430	Pro	Glu	
Trp	Met	Gly 435	Val	Met	His	Gly	Tyr 440	Glu	Ile	Glu	Phe	Val 445	Phe	Gly	Leu	
	Leu 450	Glu	Arg	Arg	Asp	Asn 455	Tyr	Thr	Lys	Ala	Glu 460	Glu	Ile	Leu	Ser	
Arg 465	Ser	Ile	Val	Lys	Arg 470	Trp	Ala	Asn	Phe	Ala 475	Lys	Tyr	Gly	Asn	Pro 480	
Asn	Glu	Thr	Gln	Asn 485	Asn	Ser	Thr	Ser	Trp 490	Pro	Val	Phe	Lys	Ser 495	Thr	
Glu	Gln	Lys	Tyr 500	Leu	Thr	Leu	Asn	Thr 505	Glu	Ser	Thr	Arg	Ile 510	Met	Thr	
Lys	Leu	Arg 515	Ala	Gln	Gln	Càa	Arg 520	Phe	Trp	Thr	Ser	Phe 525	Phe	Pro	Lys	
Val	Leu 530	Glu	Met	Thr	Gly	Asn 535	Ile	Asp	Glu	Ala	Glu 540	Trp	Glu	Trp	Lys	
Ala 545	Gly	Phe	His	Arg	Trp 550	Asn	Asn	Tyr	Met	Met 555	Asp	Trp	Lys	Asn	Gln 560	
Phe	Asn	Asp	Tyr	Thr 565	Ser	ГЛа	Lys	Glu	Ser 570	CÀa	Val	Gly	Leu			
<211 <212 <213 <220	> LH > TY > OH > FH > OT	ENGTI (PE : RGANI EATUI THER	DNA ISM: RE: INFO	722 Art: DRMA'	ific: rion		-		ī							
gaag	atga	aca t	tcata	aatt	gc aa	acaaa	agaat	gga	aaaa	gtca	gag	ggat	gaa (cttga	acagtt	60
															ggtaga	120
															acaaaa	180 240
															cagag attcca	300
															caaact	360
															gttatt	420

-continued

gtagtgtcaa	tgaactatag	ggtgggtgcc	ctaggattct	tagctttgcc	aggaaatcct	480				
gaggctccag	ggaacatggg	tttatttgat	caacagttgg	ctcttcagtg	ggttcaaaaa	540				
aatatagcag	cctttggtgg	aaatcctaaa	agtgtaactc	tctttggaga	aagttccgga	600				
gcagcttcag	ttagcctgca	tttgctttct	cctggaagcc	attcattgtt	caccagagcc	660				
attctgcaaa	gtggttccgg	taatgctcct	tgggcggtaa	catctcttta	tgaagctagg	720				
aacagaacgt	tgaacttagc	taaattgact	ggttgctcta	gagagaatga	gactgaaata	780				
atcaagtgtc	ttagaaataa	agatececaa	gaaattcttc	tgaatgaagc	atttgttgtc	840				
ccctatggga	cttccttggg	tgtaaacttt	ggtccgaccg	tggatggtga	ttttctcact	900				
gacatgccag	acatattact	tgaacttgga	caatttaaaa	aaacccagat	tttggtgggt	960				
gttaataaag	atgaagggac	atggttttta	gtcggtggtg	ctcctggctt	cagcaaagat	1020				
aacaatagta	tcataactag	aaaagaattt	caggaaggtt	taaaaatatt	ttttccagga	1080				
gtgagtgagt	ttggaaagga	atccatcctt	tttcattaca	cagactgggt	agatgatcag	1140				
agacctgaaa	actaccgtga	ggccttgggt	gatgttgttg	gggattataa	tttcatatgc	1200				
cctgccttgg	agttcaccaa	gaagttctca	gaatggggaa	ataatgcctt	tttctactat	1260				
tttgaacacc	gatcctccaa	acttccgtgg	ccagaatgga	tgggagtgat	gcatggctat	1320				
gaaattgaat	ttgtctttgg	tttacctctg	gaaagaagag	ataattacac	aaaagccgag	1380				
gaaattttga	gtagatccat	agtgaaacgg	tgggcaaatt	ttgcaaaata	tgggaatcca	1440				
aatgagactc	agaacaatag	cacaagctgg	cctgtcttca	aaagcactga	acaaaaatat	1500				
ctaaccttga	atacagagtc	aacaagaata	atgacgaaac	tacgtgctca	acaatgtcga	1560				
ttctggacat	catttttcc	aaaagtcttg	gaaatgacag	gaaatattga	tgaagcagaa	1620				
tgggagtgga	aagcaggatt	ccatcgctgg	aacaattaca	tgatggactg	gaaaaatcaa	1680				
tttaacgatt	acactagcaa	gaaagaaagt	tgtgtgggtc	tc		1722				
<pre><210> SEQ ID NO 34 <211> LENGTH: 574 <212> TYPE: PRT <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: synthesized <400> SEQUENCE: 34 Glu Asp Asp Ile Ile Ile Ala Thr Lys Asn Gly Lys Val Arg Gly Met 1 5 10 15</pre>										

Asn Leu Thr Val Phe Gly Gly Thr Val Thr Ala Phe Leu Gly Ile Pro 20

Tyr Ala Gln Pro Pro Leu Gly Arg Leu Arg Phe Lys Lys Pro Gln Ser $35 \ \ \,$ 40 $\ \ \,$ 45

Cys Cys Gln Asn Ile Asp Gln Ser Phe Pro Gly Phe His Gly Ser Glu 65 70 75 80

Met Trp Asn Pro Asn Thr Asp Leu Ser Glu Asp Cys Leu Tyr Leu Asn 85 90 95

Val Trp Ile Pro Ala Pro Lys Pro Lys Asn Ala Thr Val Leu Ile Trp \$100\$ \$105 \$110

Ile Tyr Gly Gly Gly Phe Gln Thr Gly Thr Ser Ser Leu His Val Tyr \$115\$ \$120\$ \$125\$

Asp Gly Lys Phe Leu Ala Arg Val Glu Arg Val Ile Val Val Ser Met

												COII	CIII	aca	
	130					135					140				
Asn 145	Tyr	Arg	Val	Gly	Ala 150	Leu	Gly	Phe	Leu	Ala 155	Leu	Pro	Gly	Asn	Pro 160
Glu	Ala	Pro	Gly	Asn 165	Met	Gly	Leu	Phe	Asp 170	Gln	Gln	Leu	Ala	Leu 175	Gln
Trp	Val	Gln	180	Asn	Ile	Ala	Ala	Phe 185	Gly	Gly	Asn	Pro	Lys 190	Ser	Val
Thr	Leu	Phe 195	Gly	Glu	Ser	Ser	Gly 200	Ala	Ala	Ser	Val	Ser 205	Leu	His	Leu
Leu	Ser 210	Pro	Gly	Ser	His	Ser 215	Leu	Phe	Thr	Arg	Ala 220	Ile	Leu	Gln	Ser
Gly 225	Ser	Gly	Asn	Ala	Pro 230	Trp	Ala	Val	Thr	Ser 235	Leu	Tyr	Glu	Ala	Arg 240
Asn	Arg	Thr	Leu	Asn 245	Leu	Ala	Lys	Leu	Thr 250	Gly	Cys	Ser	Arg	Glu 255	Asn
Glu	Thr	Glu	Ile 260	Ile	Lys	CÀa	Leu	Arg 265	Asn	Lys	Asp	Pro	Gln 270	Glu	Ile
Leu	Leu	Asn 275	Glu	Ala	Phe	Val	Val 280	Pro	Tyr	Gly	Thr	Ser 285	Leu	Gly	Val
Asn	Phe 290	Gly	Pro	Thr	Val	Asp 295	Gly	Asp	Phe	Leu	Thr 300	Asp	Met	Pro	Asp
Ile 305	Leu	Leu	Glu	Leu	Gly 310	Gln	Phe	Lys	Lys	Thr 315	Gln	Ile	Leu	Val	Gly 320
Val	Asn	Lys	Asp	Glu 325	Gly	Thr	Trp	Phe	Leu 330	Val	Gly	Gly	Ala	Pro 335	Gly
Phe	Ser	Lys	Asp 340	Asn	Asn	Ser	Ile	Ile 345	Thr	Arg	Lys	Glu	Phe 350	Gln	Glu
Gly	Leu	Lys 355	Ile	Phe	Phe	Pro	Gly 360	Val	Ser	Glu	Phe	Gly 365	ГÀа	Glu	Ser
Ile	Leu 370	Phe	His	Tyr	Thr	Asp 375	Trp	Val	Asp	Asp	Gln 380	Arg	Pro	Glu	Asn
Tyr 385	Arg	Glu	Ala	Leu	Gly 390	Asp	Val	Val	Gly	Asp 395	Tyr	Asn	Phe	Ile	Сув 400
Pro	Ala	Leu	Glu	Phe 405	Thr	Lys	Lys	Phe	Ser 410	Glu	Trp	Gly	Asn	Asn 415	Ala
Phe	Phe	Tyr	Tyr 420	Phe	Glu	His	Arg	Ser 425	Ser	ГÀа	Leu	Pro	Trp 430	Pro	Glu
Trp	Met	Gly 435	Val	Met	His	Gly	Tyr 440	Glu	Ile	Glu	Phe	Val 445	Phe	Gly	Leu
Pro	Leu 450	Glu	Arg	Arg	Asp	Asn 455	Tyr	Thr	ГÀа	Ala	Glu 460	Glu	Ile	Leu	Ser
Arg 465	Ser	Ile	Val	Lys	Arg 470	Trp	Ala	Asn	Phe	Ala 475	ГÀа	Tyr	Gly	Asn	Pro 480
Asn	Glu	Thr	Gln	Asn 485	Asn	Ser	Thr	Ser	Trp 490	Pro	Val	Phe	Lys	Ser 495	Thr
Glu	Gln	Lys	Tyr 500	Leu	Thr	Leu	Asn	Thr 505	Glu	Ser	Thr	Arg	Ile 510	Met	Thr
ГÀа	Leu	Arg 515	Ala	Gln	Gln	Cys	Arg 520	Phe	Trp	Thr	Ser	Phe 525	Phe	Pro	Lys
Val	Leu 530	Glu	Met	Thr	Gly	Asn 535	Ile	Asp	Glu	Ala	Glu 540	Trp	Glu	Trp	Lys
Ala 545	Gly	Phe	His	Arg	Trp 550	Asn	Asn	Tyr	Met	Met 555	Asp	Trp	Lys	Asn	Gln 560

-continued

Phe Asn Asp Tyr Thr Ser Lys Lys Glu Ser Cys Val Gly Leu

```
565
<210> SEQ ID NO 35
<211> LENGTH: 1722
<212> TYPE: DNA
<213 > ORGANISM: Artificial Sequence
<220> FEATURE:
<223 > OTHER INFORMATION: synthesized
<400> SEQUENCE: 35
gaagatgaca tcataattgc aacaaagaat ggaaaagtca gagggatgaa cttgacagtt
                                                                      60
tttqqtqqca cqqtaacaqc ctttcttqqa attccctatq cacaqccacc tcttqqtaqa
                                                                     120
cttcgattca aaaagccaca gtctctgacc aagtggtctg atatttggaa tgccacaaaa
                                                                     180
tatgcaaatt cttgctgtca gaacatagat caaagttttc caggcttcca tggatcagag
                                                                     240
atgtggaacc caaacactga cctcagtgaa gactgtttat atctaaatgt atggattcca
                                                                     300
gcacctaaac caaaaaatgc cactgtattg atatggattt atggtggtgg ttttcaaact
                                                                     360
ggaacatcat ctttacatgt ttatgatggc aagtttctgg ctcgggttga aagagttatt
                                                                     420
gtagtgtcaa tgaactatag ggtgggtgcc ctaggattct tagctttgcc aggaaatcct
                                                                     480
gaggetecag ggaacatggg tttatttgat caacagttgg etetteagtg ggtteaaaaa
                                                                     540
aatatagcag cetttggtgg aaateetaaa agtgtaacte tetttggaga aagtteegga
                                                                     600
gcagcttcag ttagcctgca tttgctttct cctggaagcc attcattgtt caccagagcc
                                                                     660
attctgcaaa gtggttccgc taatgctcct tgggcggtaa catctcttta tgaagctagg
                                                                     720
aacagaacgt tgaacttagc taaattgact ggttgctcta gagagaatga gactgaaata
                                                                     780
atcaagtgtc ttagaaataa agatccccaa gaaattcttc tgaatgaagc atttgttgtc
                                                                     840
ccctatggga ctgaattggg tgtaaacttt ggtccgaccg tggatggtga ttttctcact
                                                                     900
gacatgccag acatattact tgaacttgga caatttaaaa aaacccagat tttggtgggt
                                                                     960
gttaataaag atgaagggac atggttttta gtcggtggtg ctcctggctt cagcaaagat
                                                                    1020
aacaatagta tcataactag aaaagaattt caggaaggtt taaaaatatt ttttccagga
                                                                    1080
gtgagtgagt ttggaaagga atccatcctt tttcattaca cagactgggt agatgatcag
                                                                    1140
agacctgaaa actaccgtga ggccttgggt gatgttgttg gggattataa tttcatatgc
                                                                    1200
cctgccttgg agttcaccaa gaagttctca gaatggggaa ataatgcctt tttctactat
                                                                    1260
                                                                    1320
tttgaacacc gatcctccaa acttccgtgg ccagaatgga tgggagtgat gcatggctat
                                                                    1380
gaaattgaat ttgtctttgg tttacctctg gaaagaagag ataattacac aaaagccgag
qaaattttqa qtaqatccat aqtqaaacqq tqqqcaaatt ttqcaaaata tqqqaatcca
                                                                    1440
aatgagactc agaacaatag cacaagctgg cctgtcttca aaagcactga acaaaaatat
                                                                    1500
ctaaccttga atacagagtc aacaagaata atgacgaaac tacgtgctca acaatgtcga
                                                                    1560
ttctggacat catttttcc aaaagtcttg gaaatgacag gaaatattga tgaagcagaa
                                                                    1620
tgggagtgga aagcaggatt ccatcgctgg aacaattaca tgatggactg gaaaaatcaa
                                                                    1680
tttaacgatt acactagcaa gaaagaaagt tgtgtgggtc tc
                                                                    1722
```

<210> SEQ ID NO 36

<211> LENGTH: 574

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthesized

Glu Asp Asp 1	Ile Ile 5	Ile A	la Thr	ГЛа	Asn 10	Gly	ГЛа	Val	Arg	Gly 15	Met
Asn Leu Thr	Val Phe 20	Gly G	ly Thr	Val 25	Thr	Ala	Phe	Leu	Gly 30	Ile	Pro
Tyr Ala Gln 35	Pro Pro	Leu G	ly Arg 40	Leu	Arg	Phe	Lys	Lys 45	Pro	Gln	Ser
Leu Thr Lys 50	Trp Ser	Asp I 5		Asn	Ala	Thr	Lys 60	Tyr	Ala	Asn	Ser
Cys Cys Gln 65	Asn Ile	Asp G 70	ln Ser	Phe	Pro	Gly 75	Phe	His	Gly	Ser	Glu 80
Met Trp Asn	Pro Asn 85	Thr A	sp Leu	Ser	Glu 90	Asp	CAa	Leu	Tyr	Leu 95	Asn
Val Trp Ile	Pro Ala 100	Pro L	ys Pro	Lys 105	Asn	Ala	Thr	Val	Leu 110	Ile	Trp
Ile Tyr Gly 115		Phe G	ln Thr 120	Gly	Thr	Ser	Ser	Leu 125	His	Val	Tyr
Asp Gly Lys 130	Phe Leu		rg Val 35	Glu	Arg	Val	Ile 140	Val	Val	Ser	Met
Asn Tyr Arg 145	Val Gly	Ala L 150	eu Gly	Phe	Leu	Ala 155	Leu	Pro	Gly	Asn	Pro 160
Glu Ala Pro	Gly Asn 165	Met G	ly Leu	Phe	Asp 170	Gln	Gln	Leu	Ala	Leu 175	Gln
Trp Val Gln	Lys Asn 180	Ile A	la Ala	Phe 185	Gly	Gly	Asn	Pro	Lys 190	Ser	Val
Thr Leu Phe 195		Ser S	er Gly 200	Ala	Ala	Ser	Val	Ser 205	Leu	His	Leu
Leu Ser Pro 210	Gly Ser		er Leu 15	Phe	Thr	Arg	Ala 220	Ile	Leu	Gln	Ser
Gly Ser Ala 225	Asn Ala	Pro T 230	rp Ala	Val	Thr	Ser 235	Leu	Tyr	Glu	Ala	Arg 240
Asn Arg Thr	Leu Asn 245	Leu A	la Lys	Leu	Thr 250	Gly	Cys	Ser	Arg	Glu 255	Asn
Glu Thr Glu	Ile Ile 260	Lys C	ys Leu	Arg 265	Asn	ГЛа	Asp	Pro	Gln 270	Glu	Ile
Leu Leu Asn 275	Glu Ala	Phe V	al Val 280	Pro	Tyr	Gly	Thr	Glu 285	Leu	Gly	Val
Asn Phe Gly 290	Pro Thr		ap Gly 95	Asp	Phe	Leu	Thr 300	Asp	Met	Pro	Asp
Ile Leu Leu 305	Glu Leu	Gly G 310	ln Phe	ГÀа	Lys	Thr 315	Gln	Ile	Leu	Val	Gly 320
Val Asn Lys	Asp Glu 325	Gly T	hr Trp	Phe	Leu 330	Val	Gly	Gly	Ala	Pro 335	Gly
Phe Ser Lys	Asp Asn 340	Asn S	er Ile	Ile 345	Thr	Arg	Lys	Glu	Phe 350	Gln	Glu
Gly Leu Lys 355		Phe P	ro Gly 360	Val	Ser	Glu	Phe	Gly 365	Lys	Glu	Ser
Ile Leu Phe 370	His Tyr		ap Trp 75	Val	Asp	Asp	Gln 380	Arg	Pro	Glu	Asn
Tyr Arg Glu 385	Ala Leu	Gly A 390	sp Val	Val	Gly	Asp 395	Tyr	Asn	Phe	Ile	Сув 400
Pro Ala Leu	Glu Phe 405	Thr L	Às PÀs	Phe	Ser 410	Glu	Trp	Gly	Asn	Asn 415	Ala
Phe Phe Tyr	Tyr Phe	Glu H	is Arg	Ser	Ser	Lys	Leu	Pro	Trp	Pro	Glu

95

-continued

420 425 430 Trp Met Gly Val Met His Gly Tyr Glu Ile Glu Phe Val Phe Gly Leu Pro Leu Glu Arg Arg Asp Asn Tyr Thr Lys Ala Glu Glu Ile Leu Ser Arg Ser Ile Val Lys Arg Trp Ala Asn Phe Ala Lys Tyr Gly Asn Pro Asn Glu Thr Gln Asn Asn Ser Thr Ser Trp Pro Val Phe Lys Ser Thr Glu Gln Lys Tyr Leu Thr Leu Asn Thr Glu Ser Thr Arg Ile Met Thr 505 Lys Leu Arg Ala Gln Gln Cys Arg Phe Trp Thr Ser Phe Phe Pro Lys 520 Val Leu Glu Met Thr Gly Asn Ile Asp Glu Ala Glu Trp Glu Trp Lys 530 535 540 Ala Gly Phe His Arg Trp Asn Asn Tyr Met Met Asp Trp Lys Asn Gln 545 550 Phe Asn Asp Tyr Thr Ser Lys Lys Glu Ser Cys Val Gly Leu 565 570

gaagatgaca tcataattgc aacaaagaat ggaaaagtca gagggatgaa cttgacagtt

<210> SEQ ID NO 37

<211> LENGTH: 1722

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthesized

<400> SEQUENCE: 37

tttggtggca cggtaacagc ctttcttgga attccctatg cacagccacc tcttggtaga 120 cttcgattca aaaagccaca gtctctgacc aagtggtctg atatttggaa tgccacaaaa tatgcaaatt cttgctgtca gaacatagat caaagttttc caggcttcca tggatcagag atgtggaacc caaacactga cctcagtgaa gactgtttat atctaaatgt atggattcca gcacctaaac caaaaaatgc cactgtattg atatggattt atggtggtgg ttttcaaact 360 ggaacatcat ctttacatgt ttatgatggc aagtttctgg ctcgggttga aagagttatt 420 gtagtgtcaa tgaactatag ggtgggtgcc ctaggattct tagctttgcc aggaaatcct 480 540 gaggetecag ggaacatggg tttatttgat caacagttgg etetteagtg ggtteaaaaa aatatagcag cetttggtgg aaateetaaa agtgtaacte tetttggaga aagtteegga 600 qcaqcttcaq ttaqcctqca tttqctttct cctqqaaqcc attcattqtt caccaqaqcc 660 attetqeaaa qtqqtteecc qaatqeteet tqqqeqqtaa catetetta tqaaqetaqq 720 aacagaacgt tgaacttagc taaattgact ggttgctcta gagagaatga gactgaaata 780 atcaagtqtc ttagaaataa agatccccaa gaaattcttc tgaatgaagc atttgttgtc 840 ccctatggga ctaacttggg tgtaaacttt ggtccgaccg tggatggtga ttttctcact 900 gacatgccag acatattact tgaacttgga caatttaaaa aaacccagat tttggtgggt 960 gttaataaag atgaagggac atggttttta gtcggtggtg ctcctggctt cagcaaagat 1020 aacaatagta tcataactag aaaagaattt caggaaggtt taaaaatatt ttttccagga 1080 gtgagtgagt ttggaaagga atccatcctt tttcattaca cagactgggt agatgatcag 1140 agacctgaaa actaccgtga ggccttgggt gatgttgttg gggattataa tttcatatgc 1200 cctqccttqq aqttcaccaa qaaqttctca qaatqqqqaa ataatqcctt tttctactat 1260

-continued

_																
tttg	gaaca	acc ç	gatco	ctcca	aa a	cttco	gtgg	g cca	agaat	gga	tgg	gagt	gat ç	gcato	ggctat	1320
gaaa	ittga	aat t	tgto	ettte	gg ti	taco	etete	g gaa	aagaa	agag	ataa	atta	cac a	aaaa	gccgag	1380
gaaa	tttt	ga g	gtaga	atcca	at aç	gtgaa	acgo	g tgg	ggcaa	aatt	ttg	caaa	ata t	ggga	aatcca	1440
aato	gagad	etc a	agaad	caata	ag ca	acaaç	gctgg	g cct	gtct	tca	aaaq	gcact	cga a	acaaa	aaatat	1500
ctaa	ecct	ga a	ataca	agagt	cc aa	acaaç	gaata	a atç	gacga	aaac	tac	gtgct	ca a	acaat	gtcga	1560
ttct	ggad	cat o	cattt	ttt	cc aa	aaagt	cttg	g gaa	aatga	acag	gaaa	atatt	ga t	gaag	gcagaa	1620
tggg	gagto	gga a	aagca	aggat	t c	catco	gctgg	gaad	caatt	caca	tgat	gga	ctg g	gaaaa	aatcaa	1680
ttta	acga	att a	acact	agca	aa ga	aaaga	aaagt	tgt	gtgg	ggtc	tc					1722
<211 <212 <213 <220 <223	L> LE 2> TY 3> OF 0> FE 3> OT	ENGTI (PE : RGANI EATUI THER	ISM: RE:	74 Art: DRMA:		ial :			Ē							
		_			Tla	Ala	Thr	Lare	Δen	Gl ₃₇	Lare	Val.	Δra	Glv	Mat	
1	vob	App	116	5	116	AIA	1111	цув	10	GIY	пуъ	vai	Arg	15	Mec	
Asn	Leu	Thr	Val 20	Phe	Gly	Gly	Thr	Val 25	Thr	Ala	Phe	Leu	Gly 30	Ile	Pro	
Tyr	Ala	Gln 35	Pro	Pro	Leu	Gly	Arg 40	Leu	Arg	Phe	Lys	Lys 45	Pro	Gln	Ser	
Leu	Thr 50	Lys	Trp	Ser	Asp	Ile 55	Trp	Asn	Ala	Thr	60 Lys	Tyr	Ala	Asn	Ser	
Cys	Cys	Gln	Asn	Ile	Asp 70	Gln	Ser	Phe	Pro	Gly 75	Phe	His	Gly	Ser	Glu 80	
Met	Trp	Asn	Pro	Asn 85	Thr	Asp	Leu	Ser	Glu 90	Asp	Cys	Leu	Tyr	Leu 95	Asn	
Val	Trp	Ile	Pro 100	Ala	Pro	Lys	Pro	Lys 105	Asn	Ala	Thr	Val	Leu 110	Ile	Trp	
Ile	Tyr	Gly 115	Gly	Gly	Phe	Gln	Thr 120	Gly	Thr	Ser	Ser	Leu 125	His	Val	Tyr	
Asp	Gly 130	Lys	Phe	Leu	Ala	Arg 135	Val	Glu	Arg	Val	Ile 140	Val	Val	Ser	Met	
Asn 145	Tyr	Arg	Val	Gly	Ala 150	Leu	Gly	Phe	Leu	Ala 155	Leu	Pro	Gly	Asn	Pro 160	
Glu	Ala	Pro	Gly	Asn 165	Met	Gly	Leu	Phe	Asp 170	Gln	Gln	Leu	Ala	Leu 175	Gln	
Trp	Val	Gln	Lys 180	Asn	Ile	Ala	Ala	Phe 185	Gly	Gly	Asn	Pro	Lys 190	Ser	Val	
Thr	Leu	Phe 195	Gly	Glu	Ser	Ser	Gly 200	Ala	Ala	Ser	Val	Ser 205	Leu	His	Leu	
Leu	Ser 210	Pro	Gly	Ser	His	Ser 215	Leu	Phe	Thr	Arg	Ala 220	Ile	Leu	Gln	Ser	
Gly 225	Ser	Pro	Asn	Ala	Pro 230	Trp	Ala	Val	Thr	Ser 235	Leu	Tyr	Glu	Ala	Arg 240	
Asn	Arg	Thr	Leu	Asn 245	Leu	Ala	Lys	Leu	Thr 250	Gly	Cys	Ser	Arg	Glu 255	Asn	
Glu	Thr	Glu	Ile 260	Ile	Lys	СЛа	Leu	Arg 265	Asn	ГЛа	Asp	Pro	Gln 270	Glu	Ile	

Leu Leu Asn Glu Ala Phe Val Val Pro Tyr Gly Thr Asn Leu Gly Val 275 280 280

Asn Phe Gly Pro Thr Val Asp Gly Asp Phe Leu Thr Asp Met Pro Asp	
290 295 300 Ile Leu Leu Glu Leu Gly Gln Phe Lys Lys Thr Gln Ile Leu Val Gly	
305 310 315 320	
Val Asn Lys Asp Glu Gly Thr Trp Phe Leu Val Gly Gly Ala Pro Gly 325 330 335	
Phe Ser Lys Asp Asn Asn Ser Ile Ile Thr Arg Lys Glu Phe Gln Glu 340 345 350	
Gly Leu Lys Ile Phe Phe Pro Gly Val Ser Glu Phe Gly Lys Glu Ser 355 360 365	
Ile Leu Phe His Tyr Thr Asp Trp Val Asp Asp Gln Arg Pro Glu Asn 370 375 380	
Tyr Arg Glu Ala Leu Gly Asp Val Val Gly Asp Tyr Asn Phe Ile Cys 385 390 395 400	
Pro Ala Leu Glu Phe Thr Lys Lys Phe Ser Glu Trp Gly Asn Asn Ala 405 410 415	
Phe Phe Tyr Tyr Phe Glu His Arg Ser Ser Lys Leu Pro Trp Pro Glu 420 425 430	
Trp Met Gly Val Met His Gly Tyr Glu Ile Glu Phe Val Phe Gly Leu 435 440 445	
Pro Leu Glu Arg Arg Asp Asn Tyr Thr Lys Ala Glu Glu Ile Leu Ser 450 455 460	
Arg Ser Ile Val Lys Arg Trp Ala Asn Phe Ala Lys Tyr Gly Asn Pro 465 470 475 480	
Asn Glu Thr Gln Asn Asn Ser Thr Ser Trp Pro Val Phe Lys Ser Thr 485 490 495	
Glu Gln Lys Tyr Leu Thr Leu Asn Thr Glu Ser Thr Arg Ile Met Thr 500 505 510	
Lys Leu Arg Ala Gln Gln Cys Arg Phe Trp Thr Ser Phe Phe Pro Lys 515 520 525	
Val Leu Glu Met Thr Gly Asn Ile Asp Glu Ala Glu Trp Glu Trp Lys 530 535 540	
Ala Gly Phe His Arg Trp Asn Asn Tyr Met Met Asp Trp Lys Asn Gln 545 550 555 560	
Phe Asn Asp Tyr Thr Ser Lys Lys Glu Ser Cys Val Gly Leu 565 570	
<210> SEQ ID NO 39 <211> LENGTH: 1722 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: synthesized <400> SEQUENCE: 39	
gaagatgaca tcataattgc aacaaagaat ggaaaagtca gagggatgaa cttgacagtt	60
tttggtggca cggtaacagc ctttcttgga attccctatg cacagccacc tcttggtaga	120
cttcgattca aaaagccaca gtctctgacc aagtggtctg atatttggaa tgccacaaaa	180
tatgcaaatt cttgctgtca gaacatagat caaagttttc caggcttcca tggatcagag	240
atgtggaacc caaacactga cctcagtgaa gactgtttat atctaaatgt atggattcca	300
gcacctaaac caaaaaatgc cactgtattg atatggattt atggtggtgg ttttcaaact	360
ggaacatcat ctttacatgt ttatgatggc aagtttctgg ctcgggttga aagagttatt	420
gtagtgtcaa tgaactatag ggtgggtgcc ctaggattct tagctttgcc aggaaatcct	480

-continued

gaggctccag ggaacatgg	g tttatttgat ca	acagttgg ct	cttcagtg ggttcaaaaa	540
aatatagcag cctttggtg	g aaatcctaaa ag	tgtaactc tc	tttggaga aagttccgga	600
gcagcttcag ttagcctgc	a tttgctttct cc	tggaagcc at	tcattgtt caccagagcc	660
attetgeaaa gtggtteet	c caatgeteet tg	ggcggtaa ca	tctcttta tgaagctagg	720
aacagaacgt tgaacttag	c taaattgact gg	ttgctcta ga	gagaatga gactgaaata	780
atcaagtgtc ttagaaata	a agatececaa ga	aattette tg	aatgaagc atttgttgtc	840
ccctatggga ctgctttgg	g tgtaaacttt gg	teegaeeg tg	gatggtga ttttctcact	900
gacatgccag acatattac	t tgaacttgga ca	atttaaaa aa	acccagat tttggtgggt	960
gttaataaag atgaaggga	c atggttttta gt	cggtggtg ct	cctggctt cagcaaagat	1020
aacaatagta tcataacta	g aaaagaattt ca	ggaaggtt ta	aaaatatt ttttccagga	1080
gtgagtgagt ttggaaagg	a atccatcctt tt	tcattaca ca	gactgggt agatgatcag	1140
agacctgaaa actaccgtg	a ggccttgggt ga	tgttgttg gg	gattataa tttcatatgo	1200
cctgccttgg agttcacca	a gaagttetea ga	atggggaa at	aatgeett tttetaetat	1260
tttgaacacc gatcctcca	a actteegtgg ee	agaatgga tg	ggagtgat gcatggctat	1320
gaaattgaat ttgtctttg	g tttacctctg ga	aagaagag at	aattacac aaaagccgag	1380
gaaattttga gtagatcca	t agtgaaacgg tg	ggcaaatt tt	gcaaaata tgggaatcca	1440
aatgagactc agaacaata	g cacaagctgg co	tgtcttca aa	agcactga acaaaaatat	1500
ctaaccttga atacagagt	c aacaagaata at	gacgaaac ta	cgtgctca acaatgtcga	1560
ttctggacat catttttc	c aaaagtcttg ga	aatgacag ga	aatattga tgaagcagaa	1620
tgggagtgga aagcaggat	t ccatcgctgg aa	caattaca tg	atggactg gaaaaatcaa	1680
tttaacgatt acactagca	a gaaagaaagt tg	tgtgggtc tc		1722
<210> SEQ ID NO 40 <211> LENGTH: 574 <212> TYPE: PRT <213> ORGANISM: Arti <220> FEATURE: <223> OTHER INFORMAT	-			
<400> SEQUENCE: 40				
Glu Asp Asp Ile Ile 1 5	Ile Ala Thr Lys	Asn Gly Ly	s Val Arg Gly Met 15	
Asn Leu Thr Val Phe 20	Gly Gly Thr Val 25	Thr Ala Ph	e Leu Gly Ile Pro 30	
Tyr Ala Gln Pro Pro 35	Leu Gly Arg Leu 40	Arg Phe Ly	rs Lys Pro Gln Ser 45	
Leu Thr Lys Trp Ser 50	Asp Ile Trp Asn 55	Ala Thr Ly		
Cys Cys Gln Asn Ile	Asp Gln Ser Phe 70	Pro Gly Pho 75	e His Gly Ser Glu 80	
Met Trp Asn Pro Asn 85	Thr Asp Leu Ser	Glu Asp Cy 90	s Leu Tyr Leu Asn 95	
Val Trp Ile Pro Ala 100	Pro Lys Pro Lys 105		r Val Leu Ile Trp 110	
Ile Tyr Gly Gly Gly 115	Phe Gln Thr Gly 120	Thr Ser Se	r Leu His Val Tyr 125	

Asp Gly Lys Phe Leu Ala Arg Val Glu Arg Val Ile Val Val Ser Met 130 135 140

												COII	CIII	aca	
Asn 145	Tyr	Arg	Val	Gly	Ala 150	Leu	Gly	Phe	Leu	Ala 155	Leu	Pro	Gly	Asn	Pro 160
Glu	Ala	Pro	Gly	Asn 165	Met	Gly	Leu	Phe	Asp 170	Gln	Gln	Leu	Ala	Leu 175	Gln
Trp	Val	Gln	Lys 180	Asn	Ile	Ala	Ala	Phe 185	Gly	Gly	Asn	Pro	Lys 190	Ser	Val
Thr	Leu	Phe 195	Gly	Glu	Ser	Ser	Gly 200	Ala	Ala	Ser	Val	Ser 205	Leu	His	Leu
Leu	Ser 210	Pro	Gly	Ser	His	Ser 215	Leu	Phe	Thr	Arg	Ala 220	Ile	Leu	Gln	Ser
Gly 225	Ser	Ser	Asn	Ala	Pro 230	Trp	Ala	Val	Thr	Ser 235	Leu	Tyr	Glu	Ala	Arg 240
Asn	Arg	Thr	Leu	Asn 245	Leu	Ala	ГЛа	Leu	Thr 250	Gly	CAa	Ser	Arg	Glu 255	Asn
Glu	Thr	Glu	Ile 260	Ile	Lys	CÀa	Leu	Arg 265	Asn	ГÀа	Asp	Pro	Gln 270	Glu	Ile
Leu	Leu	Asn 275	Glu	Ala	Phe	Val	Val 280	Pro	Tyr	Gly	Thr	Ala 285	Leu	Gly	Val
Asn	Phe 290	Gly	Pro	Thr	Val	Asp 295	Gly	Asp	Phe	Leu	Thr 300	Asp	Met	Pro	Asp
Ile 305	Leu	Leu	Glu	Leu	Gly 310	Gln	Phe	ГÀв	Lys	Thr 315	Gln	Ile	Leu	Val	Gly 320
Val	Asn	Lys	Asp	Glu 325	Gly	Thr	Trp	Phe	Leu 330	Val	Gly	Gly	Ala	Pro 335	Gly
Phe	Ser	Lys	Asp 340	Asn	Asn	Ser	Ile	Ile 345	Thr	Arg	Lys	Glu	Phe 350	Gln	Glu
Gly	Leu	Lys 355	Ile	Phe	Phe	Pro	Gly 360	Val	Ser	Glu	Phe	Gly 365	Lys	Glu	Ser
Ile	Leu 370	Phe	His	Tyr	Thr	Asp 375	Trp	Val	Asp	Asp	Gln 380	Arg	Pro	Glu	Asn
Tyr 385	Arg	Glu	Ala	Leu	Gly 390	Asp	Val	Val	Gly	Asp 395	Tyr	Asn	Phe	Ile	Сув 400
Pro	Ala	Leu	Glu	Phe 405	Thr	Lys	Lys	Phe	Ser 410	Glu	Trp	Gly	Asn	Asn 415	Ala
Phe	Phe	Tyr	Tyr 420	Phe	Glu	His	Arg	Ser 425	Ser	Lys	Leu	Pro	Trp 430	Pro	Glu
Trp	Met	Gly 435	Val	Met	His	Gly	Tyr 440	Glu	Ile	Glu	Phe	Val 445	Phe	Gly	Leu
Pro	Leu 450	Glu	Arg	Arg	Asp	Asn 455	Tyr	Thr	Lys	Ala	Glu 460	Glu	Ile	Leu	Ser
Arg 465	Ser	Ile	Val	Lys	Arg 470	Trp	Ala	Asn	Phe	Ala 475	Lys	Tyr	Gly	Asn	Pro 480
Asn	Glu	Thr	Gln	Asn 485	Asn	Ser	Thr	Ser	Trp 490	Pro	Val	Phe	Lys	Ser 495	Thr
Glu	Gln	Lys	Tyr 500	Leu	Thr	Leu	Asn	Thr 505	Glu	Ser	Thr	Arg	Ile 510	Met	Thr
ГÀа	Leu	Arg 515	Ala	Gln	Gln	CAa	Arg 520	Phe	Trp	Thr	Ser	Phe 525	Phe	Pro	Lys
Val	Leu 530	Glu	Met	Thr	Gly	Asn 535	Ile	Asp	Glu	Ala	Glu 540	Trp	Glu	Trp	Lys
Ala 545	Gly	Phe	His	Arg	Trp 550	Asn	Asn	Tyr	Met	Met 555	Asp	Trp	Lys	Asn	Gln 560
Phe	Asn	Asp	Tyr	Thr 565	Ser	ГÀа	Lys	Glu	Ser 570	Cys	Val	Gly	Leu		

-continued

```
<210> SEQ ID NO 41
<211> LENGTH: 1722
<212> TYPE: DNA
<213 > ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthesized
<400> SEQUENCE: 41
gaagatgaca tcataattgc aacaaagaat ggaaaagtca gagggatgaa cttgacagtt
                                                                      60
tttggtggca cggtaacagc ctttcttgga attccctatg cacagccacc tcttggtaga
cttcgattca aaaagccaca gtctctgacc aagtggtctg atatttggaa tgccacaaaa
                                                                     180
tatgcaaatt cttgctgtca gaacatagat caaagttttc caggcttcca tggatcagag
                                                                     240
atqtqqaacc caaacactqa cctcaqtqaa qactqtttat atctaaatqt atqqattcca
                                                                     300
gcacctaaac caaaaaatgc cactgtattg atatggattt atggtggtgg ttttcaaact
                                                                     360
ggaacatcat ctttacatgt ttatgatggc aagtttctgg ctcgggttga aagagttatt
                                                                     420
gtagtgtcaa tgaactatag ggtgggtgcc ctaggattct tagctttgcc aggaaatcct
                                                                     480
gaggetecag ggaacatggg tttatttgat caacagttgg etetteagtg ggtteaaaaa
                                                                     540
aatatagcag cctttggtgg aaatcctaaa agtgtaactc tctttggaga aagttccgga
                                                                     600
gcagetteag ttageetgea tttgetttet eetggaagee atteattgtt eaceagagee
                                                                     660
attctgcaaa gtggttcctc caatgctcct tgggcggtaa catctcttta tgaagctagg
                                                                     720
aacagaacgt tgaacttagc taaattgact ggttgctcta gagagaatga gactgaaata
                                                                     780
atcaagtgtc ttagaaataa agatccccaa gaaattcttc tgaatgaagc atttgttgtc
                                                                     840
ccctatggga ctaacttggg tgtaaacttt ggtccgaccg tggatggtga ttttctcact
                                                                     900
gacatgccag acatattact tgaacttgga caatttaaaa aaacccagat tttggtgggt
                                                                     960
gttaataaag atgaagggac atggttttta gtcggtggtg ctcctggctt cagcaaagat
                                                                    1020
aacaatagta tcataactag aaaagaattt caggaaggtt taaaaatatt ttttccagga
                                                                    1080
gtgagtgagt ttggaaagga atccatcctt tttcattaca cagactgggt agatgatcag
                                                                    1140
agacctgaaa actaccgtga ggccttgggt gatgttgttg gggattataa tttcatatgc
cctgccttgg agttcaccaa gaagttctca gaatggggaa ataatgcctt tttctactat
                                                                    1260
tttgaacacc gatcctccaa acttccgtgg ccagaatgga tgggagtgat gcatggctat
                                                                    1320
gaaattgaat ttgtctttgg tttacctctg gaaagaagag ataattacac aaaagccgag
                                                                    1380
gaaattttga gtagatccat agtgaaacgg tgggcaaatt ttgcaaaata tgggaatcca
                                                                    1440
                                                                    1500
aatgagactc agaacaatag cacaagctgg cctgtcttca aaagcactga acaaaaatat
ctaaccttqa atacaqaqtc aacaaqaata atqacqaaac tacqtqctca acaatqtcqa
                                                                    1560
ttctggacat catttttcc aaaagtcttg gaaatgacag gaaatattga tgaagcagaa
                                                                    1620
tgggagtgga aagcaggatt ccatcgctgg aacaattaca tgatggactg gaaaaatcaa
                                                                    1680
                                                                    1722
tttaacgatt acactagcaa gaaagaaagt tgtgtgggtc tc
<210> SEQ ID NO 42
<211> LENGTH: 574
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223 > OTHER INFORMATION: synthesized
<400> SEQUENCE: 42
```

Glu Asp Asp Ile Ile Ile Ala Thr Lys Asn Gly Lys Val Arg Gly Met 1 $$ 5 $$ 10 $$ 15

Asn	Leu	Thr	Val 20	Phe	Gly	Gly	Thr	Val 25	Thr	Ala	Phe	Leu	Gly 30	Ile	Pro
Tyr	Ala	Gln 35	Pro	Pro	Leu	Gly	Arg 40	Leu	Arg	Phe	Lys	Lys 45	Pro	Gln	Ser
Leu	Thr 50	Lys	Trp	Ser	Asp	Ile 55	Trp	Asn	Ala	Thr	Lys 60	Tyr	Ala	Asn	Ser
Сув 65	Сув	Gln	Asn	Ile	Asp 70	Gln	Ser	Phe	Pro	Gly 75	Phe	His	Gly	Ser	Glu 80
Met	Trp	Asn	Pro	Asn 85	Thr	Asp	Leu	Ser	Glu 90	Asp	САв	Leu	Tyr	Leu 95	Asn
Val	Trp	Ile	Pro 100	Ala	Pro	Lys	Pro	Lys 105	Asn	Ala	Thr	Val	Leu 110	Ile	Trp
Ile	Tyr	Gly 115	Gly	Gly	Phe	Gln	Thr 120	Gly	Thr	Ser	Ser	Leu 125	His	Val	Tyr
Asp	Gly 130	ГЛа	Phe	Leu	Ala	Arg 135	Val	Glu	Arg	Val	Ile 140	Val	Val	Ser	Met
Asn 145	Tyr	Arg	Val	Gly	Ala 150	Leu	Gly	Phe	Leu	Ala 155	Leu	Pro	Gly	Asn	Pro 160
Glu	Ala	Pro	Gly	Asn 165	Met	Gly	Leu	Phe	Asp 170	Gln	Gln	Leu	Ala	Leu 175	Gln
Trp	Val	Gln	Lys 180	Asn	Ile	Ala	Ala	Phe 185	Gly	Gly	Asn	Pro	Lys 190	Ser	Val
Thr	Leu	Phe 195	Gly	Glu	Ser	Ser	Gly 200	Ala	Ala	Ser	Val	Ser 205	Leu	His	Leu
Leu	Ser 210	Pro	Gly	Ser	His	Ser 215	Leu	Phe	Thr	Arg	Ala 220	Ile	Leu	Gln	Ser
Gly 225	Ser	Ser	Asn	Ala	Pro 230	Trp	Ala	Val	Thr	Ser 235	Leu	Tyr	Glu	Ala	Arg 240
Asn	Arg	Thr	Leu	Asn 245	Leu	Ala	Lys	Leu	Thr 250	Gly	Cys	Ser	Arg	Glu 255	Asn
Glu	Thr	Glu	Ile 260	Ile	Lys	Cys	Leu	Arg 265	Asn	Lys	Asp	Pro	Gln 270	Glu	Ile
Leu	Leu	Asn 275	Glu	Ala	Phe	Val	Val 280	Pro	Tyr	Gly	Thr	Asn 285	Leu	Gly	Val
Asn	Phe 290	Gly	Pro	Thr	Val	Asp 295	Gly	Asp	Phe	Leu	Thr 300	Asp	Met	Pro	Asp
Ile 305	Leu	Leu	Glu	Leu	Gly 310	Gln	Phe	Lys	Lys	Thr 315	Gln	Ile	Leu	Val	Gly 320
Val	Asn	Lys	Asp	Glu 325	Gly	Thr	Trp	Phe	Leu 330	Val	Gly	Gly	Ala	Pro 335	Gly
Phe	Ser	Lys	Asp 340	Asn	Asn	Ser	Ile	Ile 345	Thr	Arg	Lys	Glu	Phe 350	Gln	Glu
Gly	Leu	Lys 355	Ile	Phe	Phe	Pro	Gly 360	Val	Ser	Glu	Phe	Gly 365	Lys	Glu	Ser
Ile	Leu 370	Phe	His	Tyr	Thr	Asp 375	Trp	Val	Asp	Asp	Gln 380	Arg	Pro	Glu	Asn
Tyr 385	Arg	Glu	Ala	Leu	Gly 390	Asp	Val	Val	Gly	Asp 395	Tyr	Asn	Phe	Ile	Cys 400
Pro	Ala	Leu	Glu	Phe 405	Thr	Lys	ГÀа	Phe	Ser 410	Glu	Trp	Gly	Asn	Asn 415	Ala
Phe	Phe	Tyr	Tyr 420	Phe	Glu	His	Arg	Ser 425	Ser	ГÀв	Leu	Pro	Trp 430	Pro	Glu
Trp	Met	Gly 435	Val	Met	His	Gly	Tyr 440	Glu	Ile	Glu	Phe	Val 445	Phe	Gly	Leu

-continued

Pro Leu Glu Arg Arg Asp Asn Tyr Thr Lys Ala Glu Glu Ile Leu Ser 455 Arg Ser Ile Val Lys Arg Trp Ala Asn Phe Ala Lys Tyr Gly Asn Pro 470 Asn Glu Thr Gln Asn Asn Ser Thr Ser Trp Pro Val Phe Lys Ser Thr Glu Gln Lys Tyr Leu Thr Leu Asn Thr Glu Ser Thr Arg Ile Met Thr Lys Leu Arg Ala Gln Gln Cys Arg Phe Trp Thr Ser Phe Phe Pro Lys 520 Val Leu Glu Met Thr Gly Asn Ile Asp Glu Ala Glu Trp Glu Trp Lys 535 530 Ala Gly Phe His Arg Trp Asn Asn Tyr Met Met Asp Trp Lys Asn Gln 545 550 555 Phe Asn Asp Tyr Thr Ser Lys Lys Glu Ser Cys Val Gly Leu 565 570 <210> SEO ID NO 43 <211> LENGTH: 1722 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: synthesized <400> SEOUENCE: 43 gaagatgaca tcataattgc aacaaagaat ggaaaagtca gagggatgaa cttgacagtt 60 tttggtggca cggtaacagc ctttcttgga attccctatg cacagccacc tcttggtaga 120 cttcgattca aaaagccaca gtctctgacc aagtggtctg atatttggaa tgccacaaaa 180 tatgcaaatt cttgctgtca gaacatagat caaagttttc caggcttcca tggatcagag 240 atgtggaacc caaacactga cctcagtgaa gactgtttat atctaaatgt atggattcca 300 gcacctaaac caaaaaatgc cactgtattg atatggattt atggtggtgg ttttcaaact ggaacatcat ctttacatgt ttatgatggc aagtttctgg ctcgggttga aagagttatt gtagtgtcaa tgaactatag ggtgggtgcc ctaggattct tagctttgcc aggaaatcct gaggetecag ggaacatggg tttatttgat caacagttgg etetteagtg ggtteaaaaa aatatagcag cctttggtgg aaatcctaaa agtgtaactc tctttggaga aagttccgga 600 gcagetteag ttageetgea tttgetttet eetggaagee atteattgtt caccagagee 660 attetqeaaa qtqqtteete caatqeteet tqqqeqqtaa catetetta tqaaqetaqq 720 aacagaacgt tgaacttagc taaattgact ggttgctcta gagagaatga gactgaaata 780 atcaagtgtc ttagaaataa agatccccaa gaaattcttc tgaatgaagc atttgttgtc 840 ccctatggga ctcctatggg tgtaaacttt ggtccgaccg tggatggtga ttttctcact 900 qacatqccaq acatattact tqaacttqqa caatttaaaa aaacccaqat tttqqtqqqt 960 gttaataaag atgaagggac atggttttta gtcggtggtg ctcctggctt cagcaaagat 1020 aacaatagta tcataactag aaaagaattt caggaaggtt taaaaatatt ttttccagga gtgagtgagt ttggaaagga atccatcctt tttcattaca cagactgggt agatgatcag 1140 agacctgaaa actaccgtga ggccttgggt gatgttgttg gggattataa tttcatatgc 1200

cctqccttqq aqttcaccaa qaaqttctca qaatqqqqaa ataatqcctt tttctactat

tttgaacacc gatcctccaa acttccgtgg ccagaatgga tgggagtgat gcatggctat

qaaattqaat ttqtctttqq tttacctctq qaaaqaaqaq ataattacac aaaaqccqaq

1260

1320

- (70	m	+	٦	n	110	\mathcal{C}

					-con	ıtinu	ed				
gaaattttga	gtagatcca	ıt agtga	aacgg t	gggcaaatt	ttgcaaa	ata to	gggaatcca	1440			
aatgagactc	agaacaata	g cacaa	gctgg c	ctgtcttca	aaagcac	tga ac	caaaaatat	1500			
ctaaccttga	atacagagt	c aacaa	gaata a	tgacgaaad	: tacgtgc	tca ac	caatgtcga	1560			
ttctggacat	cattttttc	c aaaag	tcttg g	aaatgacaq	gaaatat	tga tg	gaagcagaa	1620			
tgggagtgga	aagcaggat	t ccatc	gctgg a	acaattaca	ı tgatgga	ctg ga	aaaatcaa	1680			
tttaacgatt	acactagca	ıa gaaag	aaagt t	gtgtgggtd	tc tc			1722			
<210> SEQ ID NO 44 <211> LENGTH: 574 <212> TYPE: PRT <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: synthesized <400> SEQUENCE: 44											
Glu Asp Asp 1	Ile Ile 5	Ile Ala	Thr Ly	s Asn Gly 10	v Lys Val	_	Gly Met 15				
Asn Leu Thr	Val Phe 20	Gly Gly	Thr Va 25	l Thr Ala	Phe Leu	Gly 1 30	Ile Pro				
Tyr Ala Gln 35	Pro Pro	Leu Gly	Arg Le	u Arg Phe	Lys Lys 45	Pro C	Gln Ser				
Leu Thr Lys 50	Trp Ser	Asp Ile 55	Trp As:	n Ala Thi	Lys Tyr	Ala A	Asn Ser				
Cys Cys Gln 65	Asn Ile	Asp Gln 70	Ser Ph	e Pro Gly 75	Phe His	Gly S	Ser Glu 80				
Met Trp Asn	Pro Asn 85	Thr Asp	Leu Se	r Glu Ası 90	Cys Leu		Leu Asn 95				
Val Trp Ile	Pro Ala 100	Pro Lys	Pro Ly		Thr Val	Leu 1	Ile Trp				
Ile Tyr Gly 115	Gly Gly	Phe Gln	Thr Gl	y Thr Sei	Ser Leu 125		/al Tyr				
Asp Gly Lys	Phe Leu	Ala Arg 135	Val Gl	u Arg Val	. Ile Val	Val S	Ser Met				
Asn Tyr Arg 145	Val Gly	Ala Leu 150	Gly Ph	e Leu Ala 159		Gly A	Asn Pro 160				
Glu Ala Pro	Gly Asn 165	Met Gly	Leu Ph	e Asp Glr 170	n Gln Leu	-	Leu Gln 175				
Trp Val Gln	Lys Asn 180	Ile Ala	Ala Ph		Asn Pro	Lys S	Ser Val				
Thr Leu Phe	Gly Glu	Ser Ser	Gly Al	a Ala Sei	Val Ser 205		His Leu				
Leu Ser Pro 210	Gly Ser	His Ser 215		e Thr Arç	, Ala Ile 220	Leu (Gln Ser				
Gly Ser Ser	Asn Ala		Ala Va	l Thr Sei 235	Leu Tyr	Glu A	Ala Arg 240				
Asn Arg Thr			Lys Le	u Thr Gly		_	Glu Asn				
Glu Thr Glu	245 Ile Ile	Lya Cya	Leu Ar	250 g Asn Lys	s Asp Pro		255 Glu Ile				
Leu Leu Asn	260 Glu Ala	Phe Val	26 Val Pr		Thr Pro	270 Met 0	Gly Val				
275			280		285						
Asn Phe Gly 290	Pro Thr	Val Asp 295	GIA Ya	p rne Lei	Thr Asp 300	Met I	ro Asp				

Ile Le	eu 1	Leu	Glu	Leu	Gly 310	Gln	Phe	Lys	Lys	Thr 315	Gln	Ile	Leu	Val	Gly 320	
Val A	sn 1	Lys	Asp	Glu 325	Gly	Thr	Trp	Phe	Leu 330	Val	Gly	Gly	Ala	Pro 335	Gly	
Phe S	er 1	Lys	Asp 340	Asn	Asn	Ser	Ile	Ile 345		Arg	Lys	Glu	Phe 350	Gln	Glu	
Gly L		Lys 355	Ile	Phe	Phe	Pro	Gly 360	Val	Ser	Glu	Phe	Gly 365	Lys	Glu	Ser	
Ile Le	eu 1	Phe	His	Tyr	Thr	Asp 375	Trp	Val	Asp	Asp	Gln 380	Arg	Pro	Glu	Asn	
Tyr A:	rg (Glu	Ala	Leu	Gly 390	Asp	Val	Val	Gly	Asp 395		Asn	Phe	Ile	Cys 400	
Pro A	la 1	Leu	Glu	Phe 405	Thr	Lys	Lys	Phe	Ser 410	Glu	Trp	Gly	Asn	Asn 415	Ala	
Phe Pl	he '	Tyr	Tyr 420	Phe	Glu	His	Arg	Ser 425	Ser	ГЛа	Leu	Pro	Trp 430	Pro	Glu	
Trp M		Gly 435	Val	Met	His	Gly	Tyr 440	Glu	Ile	Glu	Phe	Val 445	Phe	Gly	Leu	
Pro Le	eu (Glu	Arg	Arg	Asp	Asn 455	Tyr	Thr	Lys	Ala	Glu 460	Glu	Ile	Leu	Ser	
Arg Se	er :	Ile	Val	Lys	Arg 470	Trp	Ala	Asn	Phe	Ala 475	Lys	Tyr	Gly	Asn	Pro 480	
Asn G	lu '	Thr	Gln	Asn 485	Asn	Ser	Thr	Ser	Trp 490	Pro	Val	Phe	Lys	Ser 495	Thr	
Glu G	ln 1	Lys	Tyr 500	Leu	Thr	Leu	Asn	Thr 505	Glu	Ser	Thr	Arg	Ile 510	Met	Thr	
ràa r		Arg 515	Ala	Gln	Gln	Cys	Arg 520	Phe	Trp	Thr	Ser	Phe 525	Phe	Pro	Lys	
Val Le 5	eu (30	Glu	Met	Thr	Gly	Asn 535	Ile	Asp	Glu	Ala	Glu 540	Trp	Glu	Trp	Lys	
Ala G	ly 1	Phe	His	Arg	Trp 550	Asn	Asn	Tyr	Met	Met 555	Asp	Trp	Lys	Asn	Gln 560	
Phe A	sn i	Asp	Tyr	Thr 565	Ser	Lys	Lys	Glu	Ser 570	Сув	Val	Gly	Leu			
<210> SEQ ID NO 45 <211> LENGTH: 1722 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: synthesized																
<400>																
gaagat	-						_							_	_	
cttcg															ggtaga	
															cagag	
															attcca	
gcacc	taa	ac c	aaaa	aaat	ge ea	actgt	atto	g ata	atgga	attt	atg	gtggt	gg 1	tttt	caaact	360
ggaac	atc	at c	ttta	acato	gt tt	atga	atggo	c aaq	gttt	ctgg	ctc	gggti	ga	aagaq	gttatt	420
gtagt	gtc	aa t	gaad	ctata	ag gg	gtggg	gtgc	c cta	aggat	tct	tago	cttt	gee a	aggaa	atcct	480
gagge	tcc	ag g	gaad	catg	gg tt	tatt	tgat	caa	acagt	tgg	ctct	tcaç	gtg (ggtto	caaaaa	540
aatat	agc	ag c	cttt	ggt	gg aa	aatco	ctaaa	a agt	gtaa	actc	tctt	tgga	aga a	aagtt	ccgga	600

	113			
			-continued	
gcagcttcag ttago	cctgca tttgctt	ttct cctggaagcc	attcattgtt caccagagcc	660
attctgcaaa gtggt	tccgg taatgct	teet tgggeggtaa	catctcttta tgaagctagg	720
aacagaacgt tgaac	cttagc taaattg	gact ggttgctcta	gagagaatga gactgaaata	780
atcaagtgtc ttaga	aaataa agatccc	ccaa gaaattcttc	tgaatgaagc atttgttgtc	840
ccctatggga ctcct	atggg tgtaaac	cttt ggtccgaccg	tggatggtga ttttctcact	900
gacatgccag acata	attact tgaactt	tgga caatttaaaa	aaacccagat tttggtgggt	960
gttaataaag atgaa	agggac atggttt	ttta gtcggtggtg	ctcctggctt cagcaaagat	1020
aacaatagta tcata	aactag aaaagaa	attt caggaaggtt	taaaaatatt ttttccagga	1080
gtgagtgagt ttgga	aaagga atccatc	cctt tttcattaca	cagactgggt agatgatcag	1140
agacctgaaa actac	cegtga ggeette	gggt gatgttgttg	gggattataa tttcatatgc	1200
cctgccttgg agtto	caccaa gaagtto	ctca gaatggggaa	ataatgcctt tttctactat	1260
tttgaacacc gatco	ctccaa acttccg	gtgg ccagaatgga	tgggagtgat gcatggctat	1320
gaaattgaat ttgto	ctttgg tttacct	tctg gaaagaagag	ataattacac aaaagccgag	1380
gaaattttga gtaga	atccat agtgaaa	acgg tgggcaaatt	ttgcaaaata tgggaatcca	1440
aatgagactc agaac	caatag cacaago	ctgg cctgtcttca	aaagcactga acaaaaatat	1500
ctaaccttga ataca	agagtc aacaaga	aata atgacgaaac	tacgtgctca acaatgtcga	1560
ttctggacat cattt	tttcc aaaagtc	cttg gaaatgacag	gaaatattga tgaagcagaa	1620
tgggagtgga aagca	aggatt ccatcgo	ctgg aacaattaca	tgatggactg gaaaaatcaa	1680
tttaacgatt acact	agcaa gaaagaa	aagt tgtgtgggtc	tc	1722
<pre><210> SEQ ID NO <211> LENGTH: 57 <212> TYPE: PRT <213> ORGANISM: <220> FEATURE: <223> OTHER INFO <400> SEQUENCE:</pre>	74 Artificial Se DRMATION: synt	_		
-		Thr Lve Aen Clv	Lys Val Arq Gly Met	
1	5	10	15	
Asn Leu Thr Val 20	Phe Gly Gly T	Thr Val Thr Ala 25	Phe Leu Gly Ile Pro 30	
Tyr Ala Gln Pro 35	_	Arg Leu Arg Phe 40	Lys Lys Pro Gln Ser 45	
Leu Thr Lys Trp 50	Ser Asp Ile T 55	_	Lys Tyr Ala Asn Ser 60	
Cys Cys Gln Asn 65	Ile Asp Gln S 70	Ser Phe Pro Gly 75	Phe His Gly Ser Glu 80	
Met Trp Asn Pro	Asn Thr Asp I 85	Leu Ser Glu Asp 90	Cys Leu Tyr Leu Asn 95	
Val Trp Ile Pro	Ala Pro Lys F	Pro Lys Asn Ala 105	Thr Val Leu Ile Trp 110	
Ile Tyr Gly Gly 115	-	Thr Gly Thr Ser 120	Ser Leu His Val Tyr 125	
Asp Gly Lys Phe 130	Leu Ala Arg V 135	-	Ile Val Val Ser Met 140	
Asn Tyr Arg Val 145	Gly Ala Leu G	Gly Phe Leu Ala 155	Leu Pro Gly Asn Pro 160	
Glu Ala Pro Gly	Asn Met Gly I 165	Leu Phe Asp Gln 170	Gln Leu Ala Leu Gln 175	

												con	tln.	ued	_
Trp	Val	Gln	Lys 180	Asn	Ile	Ala	Ala	Phe 185	Gly	Gly	Asn	Pro	Lys 190	Ser	V
Thr	Leu	Phe	Gly	Glu	Ser	Ser	Gly 200	Ala	Ala	Ser	Val	Ser 205	Leu	His	Le
Leu	Ser 210	Pro	Gly	Ser	His	Ser 215	Leu	Phe	Thr	Arg	Ala 220	Ile	Leu	Gln	Ser
Gly 225	Ser	Gly	Asn	Ala	Pro 230	Trp	Ala	Val	Thr	Ser 235	Leu	Tyr	Glu	Ala	Arg 240
Asn	Arg	Thr	Leu	Asn 245		Ala	ГÀа	Leu	Thr 250	Gly	Cys	Ser	Arg	Glu 255	Asn
Glu	Thr	Glu	Ile 260			Cys	Leu	Arg 265		ГÀа	Asp	Pro	Gln 270		Ile
Leu	Leu	Asn 275		Ala	Phe	Val	Val 280		Tyr	Gly	Thr	Pro 285		Gly	Val
Asn	Phe		Pro	Thr	Val	Asp 295		Asp	Phe	Leu	Thr		Met	Pro	Asp
	Leu	Leu	Glu	Leu	_		Phe	Lys	Lys			Ile	Leu	Val	_
305 Val	Asn	Lys	Asp		_	Thr	Trp	Phe		315 Val	Gly	Gly	Ala		320 Gly
Phe	Ser	Lys	_	325 Asn		Ser	Ile		330 Thr	Arg	Lys	Glu		335 Gln	Glu
Gly	Leu	_	340 Ile	Phe	Phe	Pro	Gly	345 Val	Ser	Glu	Phe	Gly	TAa 320	Glu	Ser
Ile	Leu	355 Phe	His	Tyr	Thr	Asp	360 Trp	Val	Asp	Asp	Gln	365 Arg	Pro	Glu	Asn
	370 Arg					375					380				
385	_				390	_			-	395	-				400
	Ala			405		-	-		410		_			415	
Phe	Phe	Tyr	Tyr 420	Phe	Glu	His	Arg	Ser 425	Ser	Lys	Leu	Pro	Trp 430	Pro	Glu
Trp	Met	Gly 435	Val	Met	His	Gly	Tyr 440	Glu	Ile	Glu	Phe	Val 445	Phe	Gly	Leu
Pro	Leu 450	Glu	Arg	Arg	Asp	Asn 455	Tyr	Thr	Lys	Ala	Glu 460	Glu	Ile	Leu	Ser
Arg 465	Ser	Ile	Val	Lys	Arg 470	Trp	Ala	Asn	Phe	Ala 475	Lys	Tyr	Gly	Asn	Pro 480
Asn	Glu	Thr	Gln	Asn 485		Ser	Thr	Ser	Trp 490	Pro	Val	Phe	Lys	Ser 495	Thr
Glu	Gln	Lys	Tyr 500	Leu	Thr	Leu	Asn	Thr 505	Glu	Ser	Thr	Arg	Ile 510	Met	Thr
Lys	Leu	Arg 515	Ala	Gln	Gln	Cys	Arg 520	Phe	Trp	Thr	Ser	Phe 525	Phe	Pro	Lys
Val	Leu 530	Glu	Met	Thr	Gly	Asn 535	Ile	Asp	Glu	Ala	Glu 540	Trp	Glu	Trp	Lys
Ala 545	Gly	Phe	His	Arg	Trp 550		Asn	Tyr	Met	Met 555		Trp	Lys	Asn	Gln 560
	Asn	Asp	Tyr			Lys	Lys	Glu			Val	Gly	Leu		300
_				565					570						

What is claimed is:

- 1. An isolated butyrylcholinesterase polypeptide variant comprising the amino acid sequence of SEQ ID NO:14.
 - 2. A pharmaceutical composition comprising: a butyrylcholinesterase polypeptide variant of claim 1; and 5 a suitable pharmaceutical carrier.
- 3. A method of treating a cocaine-induced condition comprising administering to an individual an effective amount of a pharmaceutical composition of claim 2 to lower blood cocaine concentration.
- **4**. A method of treating a cocaine-induced condition comprising administering to an individual an effective amount of a butyrylcholinesterase polypeptide variant of claim **1** to lower blood cocaine concentration.
- **5**. The method of claim **4**, wherein said butyrylcholinest- 15 erase polypeptide variant exhibits a one-hundred-fold or more increase in cocaine hydrolysis catalytic efficiency compared to butyrylcholinesterase.

* * * * *