

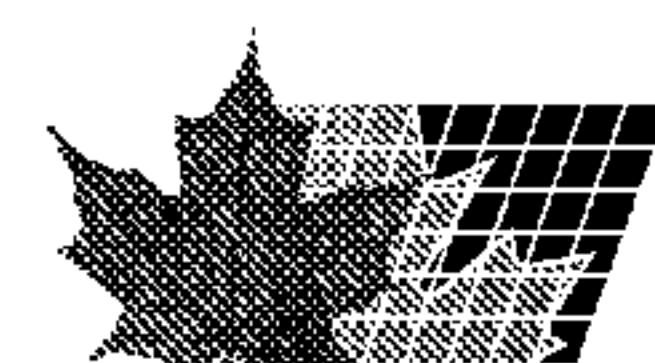


(86) Date de dépôt PCT/PCT Filing Date: 2001/06/20 (87) Date publication PCT/PCT Publication Date: 2001/12/27 (85) Entrée phase nationale/National Entry: 2002/12/20 (86) N° demande PCT/PCT Application No.: CA 2001/000913 (87) N° publication PCT/PCT Publication No.: 2001/098478 (30) Priorités/Priorities: 2000/06/22 (60/213,174) US; 2000/12/13 (60/254,932) US	(51) Cl.Int. ⁷ /Int.Cl. ⁷ C12N 15/00 (71) Demandeur/Applicant: MCGILL UNIVERSITY, CA (72) Inventeurs/Inventors: HEKIMI, SIEGFRIED, CA; BENARD, CLAIRE, CA; MCCRIGHT, BRENT, US; LAKOWSKI, BERNARD, DE; HAN, DONG, CA; LABBE, JEAN-CLAUDE, US (74) Agent: OGILVY RENAULT
---	---

(54) Titre : GENES CLK-2, CEX-7 ET COQ-4, ET UTILISATIONS ASSOCIEES
(54) Title: CLK-2, CEX-7 AND COQ-4 GENES, AND USES THEREOF

(57) **Abrégé/Abstract:**

The present invention relates to a clk-2 gene which has a function at the level of cellular physiology involved in developmental rate, telomere length and longevity, wherein clk-2 mutations cause a longer life, an altered cellular metabolism and an altered telomere length relative to the wild type, wherein clk-2 overexpression leads to telomere shortening. The present invention also relates to a clk-2 co-expressed gene which comprises a cex-7 gene having the nucleotide sequence set forth in Fig. 33 which codes for a CEX-7 protein having the amino acid sequence set forth in Fig. 34 wherein said gene is located in the clk-2 operon and said cex-7 gene is transcriptionally co-expressed with clk-2 gene present in said operon. The present invention also relates to a coq-4 gene which has a function at the level of cellular physiology involved in the regulation of developmental rate and longevity, wherein coq-4 mutations cause altered cellular metabolism and physiological relative to the wild type, wherein coq-4 gene has the identifying characteristics of nucleotide sequence set forth in Fig. 36.



(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
27 December 2001 (27.12.2001)

PCT

(10) International Publication Number
WO 01/98478 A2

(51) International Patent Classification⁷: **C12N 15/00**

(21) International Application Number: PCT/CA01/00913

(22) International Filing Date: 20 June 2001 (20.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/213,174 22 June 2000 (22.06.2000) US
60/254,932 13 December 2000 (13.12.2000) US

(71) Applicant (*for all designated States except US*): **MCGILL UNIVERSITY** [CA/CA]; 845 Sherbrooke West, Montréal, Québec H3A 2T5 (CA).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **HEKIMI, Siegfried** [CA/CA]; 4235, Avenue Powell, Montréal, Québec H4P 1E5 (CA). **BÉNARD, Claire** [CA/CA]; 3601, rue Ste-Famille, Montréal, Québec H2X 2L6 (CA). **MCCRIGHT, Brent** [US/US]; 1093, Main Street, Mount Desert, ME 04660 (US). **LAKOWSKI, Bernard** [CA/DE]; Kornwegerstr. 60, 81375 Munich (DE). **HAN, Dong** [CN/CA]; 4360, rue Henri-Julien, Montréal, Québec H2W 2K8 (CA). **LABBE, Jean-Claude** [CA/US]; 1105, Ninth Street, Durham, NC 27705 (US).

(74) Agent: **CÔTÉ, France**; Swabey Ogilvy Renault, Suite 1600, 1981 McGill College Avenue, Montréal, Québec H3A 2Y3 (CA).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: *CLK-2, CEX-7 AND COQ-4 GENES, AND USES THEREOF*

(57) Abstract: The present invention relates to a *clk-2* gene which has a function at the level of cellular physiology involved in developmental rate, telomere length and longevity, wherein *clk-2* mutations cause a longer life, an altered cellular metabolism and an altered telomere length relative to the wild type, wherein *clk-2* overexpression leads to telomere shortening. The present invention also relates to *clk-2* co-expressed gene which comprises a *cex-7* gene having the nucleotide sequence set forth in Fig. 33 which codes for a CEX-7 protein having the amino acid sequence set forth in Fig. 34 wherein said gene is located in the *clk-2* operon and said *cex-7* gene is transcriptionally co-expressed with *clk-2* gene present in said operon. The present invention also relates to a *coq-4* gene which has a function at the level of cellular physiology involved in the regulation of developmental rate and longevity, wherein *coq-4* mutations cause altered cellular metabolism and physiological relative to the wild type, wherein *coq-4* gene has the identifying characteristics of nucleotide sequence set forth in Fig. 36.



WO 01/98478 A2

CLK-2, CEX-7 AND COQ-4 GENES, AND USES THEREOFBACKGROUND OF THE INVENTION(a) Field of the Invention

5 The invention relates to the identification of three genes: the gene *clk-2*, the gene *cex-7* that is located in the same operon as *clk-2*, and the gene *coq-4*. The invention shows that these genes regulate the timing of development and behavior, and determine life span and that *clk-2* regulates the length of telomeres.

(b) Description of Prior Art

 A class of genes was identified in the nematode *Caenorhabditis elegans*, the *clk* ('clock') genes, whose activity controls how fast the worms live and die. Mutations in these genes result in an alteration of developmental and behavioral timing, including an average slow down of the animal's embryonic and post-embryonic development and of their rhythmic behaviors, as well as an increase in the animal's life span. In addition, mutations in these genes display a maternal effect, namely, homozygous mutants (*clk/clk*) derived from a heterozygous mother (*clk/+*), appear phenotypically wild-type.

 We isolated the mutations that define the genes *clk-1*, *clk-2*, *clk-3* in a screen for viable maternal-effect mutations in the nematode *Caenorhabditis elegans* (Hekimi, S. et al., *Genetics* **141**, 1351 (1995)). *gro-1* was originally identified by a spontaneous mutation isolated from a strain that had been recently established from a wild isolate (Hodgkin, J. and Doniach, T. *Genetics* **146**, 149 (1997)). Our subsequent reappraisal of this mutation revealed that it shares the characteristics of the *clk* genes (Wong, A. et al., *Genetics* **139**, 1247 (1995)).

 We have molecularly identified two of these genes, *clk-1* and *gro-1*. *clk-1* encodes a protein that is

- 2 -

highly conserved from proteobacteria to humans which is structurally similar to the yeast metabolic regulator Cat5p/Coq7p (Ewbank, J.J. et al, *Science* **275**, 980 (1997); WO98/17823). *gro-1* encodes a highly conserved
5 cellular enzyme, the dimethylallyltransferase:tRNA dimethylallyltransferase (WO99/10482).

To date, *clk-1* is the gene that has been characterized in greatest detail. In addition to the phenotypic and molecular characterization, it was found
10 that *clk-1* is ubiquitously expressed in the worm's body where it localizes to the mitochondria, the energy generating organelle of the cell (Felkai, S. et al, *EMBO Journal* **18**, 1783 (1999)). *clk-1* thus controls timing by regulating physiological rates (Branicky R,
15 C. Benard, S. Hekimi, *Bioessays* **22**, 48 (2000)).

The gene *clk-2* is defined by one allele that was isolated in a screen for viable maternal-effect mutations in *Caenorhabditis elegans* (Hekimi, S. et al., *Genetics* **141**, 1351 (1995)). The mutations in the gene
20 *clk-2* were shown to result in an alteration of the timing of several developmental and behavioral events (Hekimi, S. et al., *Genetics* **141**, 1351 (1995)) and that the activity of the gene *clk-2* controls how fast the worms live and how soon they die (Lakowski, B. and
25 Hekimi, S. *Science* **272**, 1010 (1996)). We have also noticed other phenotypes of the *clk-2* mutants such as the temperature sensitivity of the *clk-2* (*qm37*) allele. Overall, these phenotypes are similar to those of mutations in the three *clk* genes (Hekimi, S. et al.,
30 *Genetics* **141**, 1351 (1995)).

Results to date suggest that the effect of *clk* genes on the rate of aging is due to an effect on the rate of living. First, *clk-1* mutations which lead to a decrease of *clk-1* activity result in a slow down of
35 development and behavior and in an increase in life

- 3 -

span. On the other hand, overexpression of *clk-1* in transgenic animals accelerates the rate of living as revealed by the absence of a characteristic behavioral slow down with age. Second, the effect of the different *clk* genes is additive. We have shown that double *clk* mutants develop more slowly and live longer than the single *clk* mutants. Third, *clk* genes are distinct from dauer formation genes (*daf* genes) which are involved in stress resistance and also prolong life span. *Daf* genes affect life span through a separate mechanism from that of *clk*. In fact, *clk* mutants are neither dauer constitutive nor dauer defective and *daf-16* mutations cannot suppress the long life of *clk-1*, -2, -3 mutants.

The gene *coq-4* is similar to the gene *clk-1* in that both genes are required for normal ubiquinone biosynthesis in yeast and both genes have no homologues in *E. coli*. The gene *cex-7* that will be described below has been found to be a pseudoautosomal gene named XE7 in humans.

It would be highly desirable to be provided with a detailed phenotypic and molecular characterization of the gene *clk-2*, as well as with a characterization of the gene *coq-4* in an animal.

SUMMARY OF THE INVENTION

One aim of the present invention is to provide with a *clk-2* gene which has a function at the level of cellular physiology involved in developmental rate, telomere length and longevity, wherein *clk-2* mutations cause a longer life, an altered cellular metabolism and an altered telomere length relative to the wild type, wherein *clk-2* gene has the identifying characteristics of nucleotide sequence described in Fig. 1.

In accordance with the present invention there is provided the use of a *clk-2* gene to alter a function at the level of cellular physiology involved in the regulation of developmental rates, telomere length and longevity, wherein *clk-2* mutations cause a longer life, an altered cellular metabolism and physiological rates and an altered telomere length relative to the wild type, wherein *clk-2* gene has the identifying characteristics of nucleotide sequences described in Figs. 1, 4-7, 15, 16, and 20-24, or wherein the gene codes for a protein sequence as described in Figs. 2, 3, 8-14, 17-19, and 25-32 as deduced from Figs. 1, 4-7, 15, 16, 20-24.

Also is provided with the invention the use of a *clk-2* gene to alter function at the level of cellular physiology involved in the regulation of developmental rate, telomere length and longevity, wherein *clk-2* mutations cause a longer life and altered cellular metabolism and physiological rates and an altered telomere length relative to the wild type, wherein the gene codes for a protein having a sequence as set forth in Fig. 32.

In accordance with the invention there is provided a CLK-2 protein that has a function at the level of cellular physiology involved in the regulation of developmental rate, telomere length and longevity.

There is also provided with the invention a mutant CLK-2 protein which has the amino acid sequence described in Fig. 31, and the use of CLK-2 protein to alter a function at the level of cellular physiology involved in the regulation of developmental rates, telomere length and longevity, wherein the CLK-2

- 5 -

protein has the amino acid sequence as described in Figs. 2, 3, 8-14, 17-19, and 25-32.

In accordance with the invention, there is provided a *clk-2* gene which has the nucleotide sequence
5 described in Fig. 1, and the use of *clk-2* gene and homologues thereof, to manipulate the physiological rates and/or telomere biology, whereby life span of an organism is altered.

There is also provided a mouse which comprises
10 a gene knockout of the murine *clk-2* gene homologue to a *clk-2* gene.

There is provided with the present invention a method to increase the life span of multicellular organism and metazoan which comprises altering the
15 function of telomeres and mechanisms of sub-telomeric silencing.

The invention also provides the use of *clk-2* gene, CLK-2 protein, and homologues thereof, for screening drugs which decrease or increase the life
20 span of a multicellular organism, wherein the drug enhances or suppresses the expression of the *clk-2* gene or activity of the protein CLK-2, and homologues thereof.

The use of a compound is provided with the
25 invention for the manufacture of a medicament for increasing and/or decreasing physiological rates of tissues, organ, and/or whole organism of a host; wherein the compound is interfering with activity of CLK-2 protein and homologues thereof.

30 The use of a compound is also provided to promote tissue and/or organ specific reduction or increase of *clk-2* activity for the manufacture of a

medicament for the treatment of pathological conditions causing increase or decrease of physiological rate of tissue and/or organ in an individual, wherein the compound is interfering with activity of CLK-2 protein
5 and homologues thereof.

In accordance with the invention there is provided a *clk-2* co-expressed gene which comprises a *cex-7* gene having the nucleotide sequence as described in Fig. 33, and which codes for a CEX-7 protein having
10 the amino acid sequence described in Fig. 34 wherein the gene is located in the *clk-2* operon and the *cex-7* gene is transcriptionally co-expressed with *clk-2* gene present in the same operon.

A human homologue of *cex-7* gene is also
15 provided with the invention, wherein the gene codes for a protein having a sequence as described in Fig. 35.

There is provided with the invention the use of a human homologue of *cex-7* gene and homologues thereof to alter a function at the level of cellular level
20 physiology involved in the regulation of developmental rates and longevity wherein the gene codes for a protein having a sequence as described in Fig. 35.

The invention provides with a mouse which comprises a gene knock out of the murine *cex-7* gene
25 homologue of the human gene described in Fig. 35.

There is also provided with the invention the use of a compound for the manufacture of a medicament for increasing and/or decreasing physiological rates of tissues, organ, and/or whole organism of a host;
30 wherein the compound is interfering with activity of CEX-7 and homologues thereof.

- 7 -

Another aim of the invention is to provide with the use of a compound which promotes tissue and/or organ specific reduction or increase of cex-7 activity for the manufacture of a medicament for the treatment of pathological conditions causing increase or decrease of physiological rate of tissue and/or organ in an individual, wherein the compound is interfering with activity of CEX-7 and homologues thereof.

There is provided with the invention a *coq-4* gene which has a function at the level of cellular physiology involved in the regulation of developmental rate and longevity, wherein *coq-4* mutations cause altered cellular metabolism and physiological relative to the wild type, wherein *coq-4* gene has the identifying characteristics of nucleotide sequence as described in Fig. 36.

A *coq-4* gene provided with the invention has a function at the level of cellular physiology involved in the regulation of developmental rate and longevity, wherein *coq-4* mutations cause altered cellular metabolism and physiological relative to the wild type, wherein *coq-4* gene has the identifying characteristics of nucleotide sequence as described in Fig. 36, and the gene codes for a protein having a sequence as described in Fig. 37.

In accordance with the invention, there is provided with the use of *coq-4* gene to alter a function at the level of cellular physiology involved in the regulation of developmental rates, wherein *coq-4* mutations cause an altered cellular metabolism and physiological rates relative to the wild type, wherein

- 8 -

the gene codes for a protein having a sequence as described in Figs. 43 to 54 and homologues thereof.

There is also provided with the invention a mouse which comprises a gene knock out of the murine
5 *coq-4* gene as described in Fig. 47.

There is provided also with the invention the use of a compound for the manufacture of a medicament for increasing and/or decreasing physiological rates of tissues, organs and/or whole organism of a host;
10 wherein the compound is interfering with activity of *COQ-4* and homologues thereof.

A compound in accordance with the invention is provided to promote tissue and/or organ specific reduction or increase of *coq-4* activity for the
15 manufacture of a medicament for the treatment of pathological conditions causing increase or decrease of physiological rate of tissue and/or organ in an individual, wherein the compound is interfering with activity of *COQ-4* and homologues thereof.

20 Having the *clk* genes in hand can serve to manipulate the rate of development, the cell cycle, the rate of behavior and the rate of aging. Another way to look at it is that it can help to control physiological rates including for medical and industrial purposes.
25 Slowing down the rate of aging of individual organs or tissues to slow down their rate of deterioration is one medical example; accelerating the growth of farm animals or crops is an example of industrial utilization.

30 Here we describe our analysis of *clk-2* and *cex-7* and the inventions that result from this analysis, including the molecular characterization of *clk-2* and

cex-7 and the identification of homologues in several species, including humans. We also describe our identification of a new *clk* gene: *coq-4*. We have obtained a mutation in the worm *coq-4* locus and have
5 shown that the mutant animals display several of the most important characteristics of *clk* mutations.

BRIEF DESCRIPTION OF THE DRAWINGS

Figs. 1A and 1B illustrate the *Caenorhabditis*
10 *elegans clk-2* cDNA sequence.

Fig. 2 illustrates the *Caenorhabditis elegans* CLK-2 protein sequence.

Fig. 3 illustrates the *Homo sapiens* CLK-2 protein sequence (derived from clone KIAA0683).

15 Figs. 4A and 4B illustrate the *Homo sapiens clk-2* homologue nucleotide sequence (derived from AL080126).

Fig. 5 illustrates part of *Mus musculus clk-2* cDNA sequence (derived from AA671905 vl11b10.r1).

20 Fig. 6 illustrates part of *Mus musculus clk-2* cDNA sequence (derived from AA031108 mi40f03.r1).

Fig. 7 illustrates part of *Mus musculus clk-2* cDNA sequence (derived from AA230994 mw30h11.r1).

Fig. 8 illustrates part of *Mus musculus* CLK-2 protein sequence (derived from gb|AA671905.1|AA671905.)

25 Fig. 9 illustrates part of *Mus musculus* CLK-2 protein sequence (derived from gb|AA230994.1|AA230994).

Fig. 10 illustrates part of *Mus musculus* CLK-2 protein sequence (derived from gb|AA031108.1|AA031108).

30 Fig. 11 illustrates *Mus musculus* composite CLK-2 protein sequence.

Fig. 12 illustrates part of *Sus scrofa* CLK-2 protein sequence (derived from gb|AW429611.1|AW429611).

Fig. 13 illustrates the *Drosophila melanogaster* CLK-2 protein sequence.

Fig. 14 illustrates the putative *Arabidopsis thaliana* CLK-2 protein sequence (derived from
5 7630034|emb|CAB88328.1|).

Fig. 15 illustrates part of *Oryza sativa* *clk-2* cDNA sequence (derived from AU031811).

Fig. 16 illustrates part of *Oryza sativa* *clk-2* cDNA sequence (derived from D24238).

10 Fig. 17 illustrates part of *Oryza sativa* CLK-2 protein sequence (derived from dbj|D24422.1|D24422).

Fig. 18 illustrates part of *Oryza sativa* CLK-2 protein sequence (derived from
dbj|AU031811.1|AU031811).

15 Fig. 19 illustrates *Oryza sativa* composite CLK-2 protein.

Fig. 20 illustrates part of *Glycine max* *clk-2* cDNA sequence (derived from AI461201 sa76d07.y1 Gm-c1004).

20 Fig. 21 illustrates part of *Glycine max* *clk-2* cDNA sequence (derived from AW185029 se85g06.y1 Gm-c1023).

Fig. 22 illustrates part of *Glycine max* *clk-2* cDNA sequence (derived from AW350166 GM210007A10F4R Gm-
25 r1021).

Fig. 23 illustrates part of *Glycine max* *clk-2* cDNA sequence (derived from AW397826 sg68g12.y1 Gm-c1007).

Fig. 24 illustrates part of *Glycine max* *clk-2* cDNA sequence (derived from AW567713 si54a01.y1 Gm-
30 r1030).

Fig. 25 illustrates part of *Glycine max* CLK-2 protein sequence (derived from gb|AW350166.1|AW350166).

Fig. 26 illustrates part of *Glycine max* CLK-2 protein sequence (derived from gb|AI461201.1|AI461201).

5 Fig. 27 illustrates part of *Glycine max* CLK-2 protein sequence (derived from gb|AW185029.1|AW185029).

Fig. 28 illustrates part of *Glycine max* CLK-2 protein sequence (derived from gb|AW567713.1|AW567713).

10 Fig. 29 illustrates part of *Glycine max* CLK-2 protein sequence (derived from gb|AW397826.1|AW397826).

Fig. 30 illustrates *Glycine max* CLK-2 composite protein sequence.

Fig. 31 illustrates the *Caenorhabditis elegans* CLK-2 (QM37) mutant protein, with C to Y substitution
15 at position 772.

Fig. 32 illustrates Tel2p, the *Saccharomyces cerevisiae* CLK-2 protein.

Fig. 33 illustrates the *Caenorhabditis elegans* cex-7 cDNA sequence.

20 Fig. 34 illustrates the *Caenorhabditis elegans* CEX-7 protein sequence.

Fig. 35 illustrates the *Homo sapiens* CEX-7 protein sequence (XE7).

25 Fig. 36 illustrates the *Caenorhabditis elegans* coq-4 cDNA sequence.

Fig. 37 illustrates the *Caenorhabditis elegans* COQ-4 protein sequence.

Fig. 38A and 38B illustrate the comparison of CLK-2 eukaryotic homologues (hCLK-2: *Homo sapiens* CLK-2; *Caenorhabditis elegans* Tel2p: *Saccharomyces cerevisiae* AtCLK-2: *Arabidopsis thaliana*).

30

Fig. 39 illustrates the comparison of CLK-2 animal homologues (D.m.: *Drosophila melanogaster*, H.s.: *Homo sapiens* C.e.: *Caenorhabditis elegans*).

Fig. 40 illustrates the comparison of CLK-2
5 vertebrate homologues (H.s.: *Homo sapiens*, M.m.: *Mus musculus*, S.s.: *Sus scrofa*).

Fig. 41 illustrates the comparison of CLK-2 plant homologues (A.t.: *Arabidopsis thaliana*, G.m.: *Glycine max*, O.s.: *Oryza sativa*).

10 Fig. 42 illustrates the comparison of COQ-4 homologous proteins.

Fig. 43 illustrates *Drosophila melanogaster* COQ-4 protein (derived from gi|7293987|gb|AAF49344.1|CG3877).

15 Fig. 44 illustrates *Homo sapiens* COQ-4 protein (derived from gi|7705807|ref|NP_057119.1|CGI-92).

Fig. 45 illustrates *Schizosaccharomyces pombe* COQ-4 protein (derived from gi|7493130|pir||T37755).

20 Fig. 46 illustrates *Arabidopsis thaliana* COQ-4 protein (derived from gi|4406761|gb|AAD20072.1|).

Fig. 47A illustrates part of *Mus musculus* COQ-4 protein (derived from gb|AA274683.1|AA274683); Fig. 47B part of *Mus musculus* COQ-4 protein (derived from dbj|AU051632.1|AU051632); Fig. 47C part of *Mus musculus*
25 COQ-4 protein (derived from gb|AI157531.1|AI157531); and Fig. 47D *Mus musculus* COQ-4 consensus protein.

Fig. 48 illustrates *Glycine max* COQ-4 protein (derived from gb|AW201157.1|AW201157).

30 Fig. 49 illustrates *Bos taurus* COQ-4 protein (derived from gb|AW660771.1|AW660771).

Fig. 50 illustrates *Medicago truncatula* COQ-4 protein (derived from gb|AW696025.1|AW696025).

- 13 -

Fig. 51 illustrates *Ancylostoma caninum* COQ-4 protein (derived from gb|AW870537.1|AW870537).

Fig. 52 illustrates *Trypanosoma cruzi* COQ-4 Protein (derived from gb|AW330043.1|AW330043).

5 Fig. 53 illustrates *Rattus rattus* COQ-4 protein (derived from gb|AA800046.1|AA800046).

Fig. 54 illustrates *Gossypium hirsutum* COQ-4 protein (derived from gb|AI731097.1|AI731097).

10 Figs. 55A-C illustrate the expression pattern of *clk-2*.

Figs. 56A-E illustrate the telomere-lengthening phenotype of *clk-2(qm37)* mutants at different temperatures.

15 **DETAILED DESCRIPTION OF THE INVENTION**

For the first time, there is provided with the present invention a new method of increasing life span by modulating the biology of telomeres.

20 **The Clk phenotype of *clk-2* mutants**

We had shown previously that *clk-2* mutants have a phenotype similar to that of *clk-1* mutants, including the maternal rescue effect, their slow development and behavior, and their increased life span (Hekimi, et al., *Genetics* **141**, 1351 (1995); Lakowski, B. and Hekimi, S. *Science* **272**, 1010 (1996). We have characterized the defects of *clk-2* mutants in much further detail, the results of which follow. From 15°C to 20°C the phenotype of *clk-2* mutants is similar to
30 that of *clk-1* mutants. The average developmental, reproductive and behavioral rates are dramatically slower, and the mean and maximum life span longer, than those of the wild type as summarized in Table 1. In particular, the embryonic development of *clk-2(qm37)*

- 14 -

mutants lasts 17.0 ± 1.5 hours ($n=97$) at 20°C , while the wild type lasts 13.2 ± 0.7 hours ($n=80$). The post-embryonic development of *clk-2 (qm37)* mutants is also slower lasting 95.7 ± 1.3 hours at 20°C ($n=73$), while the wild-type worms take only 53.6 ± 8.7 hours ($n=184$).

The defecation cycles are slowed down as well, occurring every 105.7 ± 15.2 seconds in *clk-2* mutants at 20°C ($n=10$) and every 54.9 ± 0.6 seconds in the wild type ($n=70$). The pumping rate is lower, 180.9 ± 24.8 pumps per minute occurring in the *clk-2* mutants at 20°C ($n=25$), and 265.3 ± 64.4 pumps per minute in the wild type ($n=25$).

Table 1

Phenotypic characterization of *clk-2(qm37)* animals at 20°C

	Wild Type (N2)	<i>clk-2 (qm37)</i>	Maternally rescue <i>clk-2 (qm37)</i>
Embryonic Development (hours)	13.2 ± 0.7 $n=80$	17.0 ± 1.5 $n=97$	13.3 ± 1.6 $n=40$
Post-embryonic Development (hours)	53.6 ± 8.7 $n=184$	95.7 ± 1.3 $n=73$	53.9 ± 12.4 $n=98$
Self-brood Size (eggs)	302.4 ± 30.5 $n=20$	83.4 $n=10$	113.9 ± 30.3 $n=24$
Peak Egg-laying Rate (eggs per hour)	5.3 $n=10$	1.3 $n=10$	3.6 ± 0.9 $n=24$
Defecation (seconds)	54.9 ± 0.6 $n=70$	105.7 ± 15.2 $n=10$	60.3 ± 9.0 $n=8$
Pumping (pumps per minute)	265.3 ± 64.4 $n=25$	180.9 ± 24.8 $n=25$	245.2 ± 24.6 $n=11$

In addition, we have also examined the self-brood size at 20°C and found that is reduced in *clk-2*

- 15 -

mutants where it is 83.4 (n=10), while it is 302.4 ± 30.5 in the wild type (n=20). The peak egg-laying rate is 1.3 (n=10) in *clk-2* mutants at 20°C, and 5.3 (n=10) in the wild type. We have also examined the life span. *clk-2(qm37)* mutants live longer than the wild type, living on average 22.4 ± 7.4 days (n=100) at 20°C and having a maximum life span of 40 days, which is longer than the average life span of 19.3 ± 5.3 days (n=100) and maximum life span of 32 days of wild-type N2 worms.

The developmental and behavioral phenotypes are fully maternally rescued, that is to say that homozygous *clk-2/clk-2* mutants derived from a *clk-2(qm37)/+* heterozygous mother display wild-type phenotypes. In fact, the embryonic development of homozygous mutants derived from a heterozygous mother takes only 13.3 ± 1.6 hours (n=40) and their post-embryonic development lasts only 53.9 ± 12.4 hours (n=98) at 20°C. Also maternally rescued are both defecation, which occurs every 60.3 ± 9.1 seconds at 20°C (n=8) and pumping, which occurs at a rate of 245.2 ± 24.6 pumps per minute at 20°C (n=11). However, the reproductive phenotypes are only partially rescued by a wild-type copy of the gene *clk-2* in the mother. The self-brood size is 113.9 ± 30.3 at 20°C (n=24), and the peak egg-laying rate is 3.6 ± 0.9 (n=24). This indicates that the wild-type *clk-2* gene in the mother induces an epigenetic state that lasts for only one generation. Erasure of the epigenetic state in the germline prevents the animal from having a wild-type rate of reproduction. In addition, the life span of maternally rescued homozygous mutants is dramatically shortened vs. both the mutant and the wild-type life span. Indeed, homozygous mutants derived from a heterozygous mother live only 14.9 ± 4.1 days on average (n=106) and have a maximum life span of 27 days at 20°C.

Interestingly, wild-type siblings of maternally rescued *clk-2* live slightly shorter than wild-type N2 worms, 17.3±4.1 days (n=206). This observation indicates that wild-type physiological rates imposed by a maternal epigenetic setting are deleterious to animals that are partially incapable of regulating their physiological rates in response to environmental conditions.

Table 2

Life span of mutants and double mutant combinations at 20°C indicated in days

Genotype	Mean Life Span	Maximum Life Span
Wild type (N2)	19.3 ± 5.3 n=100	32
<i>clk-2(qm37)</i>	22.4 ± 7.4 n=100	40
Maternally rescued <i>clk-2(qm37)</i>	14.9 ± 4.1 n=106	27
Wild type (N2)	18.4 ± 4.6 n=260	31
<i>clk-2(qm37)</i>	22.9 ± 7.3 n=260	45
<i>daf-16(m26)</i>	18.1 ± 2.6 n=260	25
<i>daf-16(m26) clk-2(qm37)</i>	21.7 ± 5.8 n=260	41
<i>daf-2(e1370)</i>	29.3 ± 10.3 n=50	51
<i>daf-2(e1370) clk-2(qm37)</i>	54.5 ± 21.4 n=50	101
<i>eat-2(ad465)</i>	30.0 ± 7.0 n=34	42
<i>eat-2(ad465) clk-2(qm37)</i>	26.6 ± 6.3 n=50	45

We characterized the life span increase produced by *clk-2(qm37)* by comparing it to that produced by other aging genes as summarized in table 2. Among the other genes that affect life span in worms,

- 17 -

the best understood are the *daf* genes. Mutations in the *eat* genes prolong life span through caloric restriction by reducing the food intake of the animals, a process that also prolongs life span in vertebrates. Mutations in *daf* genes prolong life span by partial activation of the dauer formation pathway. The dauer stage is a dormant, long-lived, alternative developmental stage which is induced by adverse environmental conditions. The increased life span of all dauer formation mutants that have been tested is suppressed by loss of function mutations in *daf-16*.

In fact, we found that while *daf-16(m26)* lives 18.1 ± 2.6 days on average with a maximum life span of 25 days, the double mutants *daf-16(m26) clk-2(qm37)* lives an average life span of 21.7 ± 5.8 days with a maximum life span of 41 days. Furthermore, although double mutants with two long-lived dauer formation mutations do not live longer than mutants carrying only one of the component mutations, *daf-2(e1370) clk-2(qm37)* double mutants live substantially longer than *daf-2*, almost three times longer than the wild type. We have shown that while *daf-2(e1370)* lives 29.3 ± 10.3 days on average with a maximum life span of 51 days, the double mutants *daf-2(e1370) clk-2(qm37)* lives an average life span of 54.5 ± 21.4 days with a maximum life span of 101 days. In contrast to these observations, the effects of *clk-2* and *eat-2* are not additive. In fact, the double mutants live somewhat shorter than *eat-2* mutants. We have shown that *eat-2(ad465)* lives 30.0 ± 7.0 days on average with a maximum life span of 42 days, and that the double mutants *daf-2(e1370) clk-2(qm37)* live 26.6 ± 6.3 days on average with a maximum life span of 45 days. These observations are also consistent with the finding that *daf-2 eat-2* double mutants live longer than *daf-2* or *eat-2* mutants in isolation (Lakowski, B.

- 18 -

and Hekimi, S. *Science* **272**, 1010 (1996)). Together, these results show that *daf-2* and *clk-2* prolong life span by distinct mechanisms but that *clk-2* works in a way that resembles caloric restriction.

5 **The strict maternal effect of the *clk-2(qm37)* mutation**

In addition to the Clk phenotype displayed by *clk-2(qm37)* mutants, they exhibit a temperature-sensitive embryonic lethal and sterile phenotypes at 25°C. We knew that *qm37* is a temperature sensitive
10 mutation and that the mutants lay dead embryos when they are transferred to 25°C (Hekimi, S. et al., *Genetics* **141**, 1351 (1995)). These findings have now been extended, and the phenotype of *clk-2* mutants at 25°C has been examined after a number of temperature
15 shift experiments at different stages of development, from permissive to restrictive temperature and vice versa.

At the permissive temperatures (15 to 20°C), *clk-2* embryos all develop normally and grow up to
20 become long-lived adults. However, when hermaphrodites that have developed at a permissive temperature are transferred to 25°C before egg-laying begins, they produce only progeny that dies during embryogenesis at various stages of development. When these
25 hermaphrodites, that have been producing dead embryos at 25°C, are transferred back to 18°C, they lay only dead eggs at first, but start to lay live eggs that develop into adults after having been 5-6 hours at 18°C. When hermaphrodites that are kept at 18°C, and
30 that lay only live eggs, are transferred to 25°C it also takes 5-6 hours before they lay only dead eggs. Both conditions (laying live or dead progeny) are fully reversible upon temperature shift even when the animal's entire post-embryonic development was carried
35 out at a single temperature (permissive or non-

- 19 -

permissive). In addition, when larvae that developed at the permissive temperature are shifted to 25°C, some arrest development and others reach a sterile and sick adulthood. These phenotypes are fully reversible as well. Finally, all these lethality and sterility phenotypes displayed by *clk-2(qm37)* mutants at 25°C can be fully maternally rescued: heterozygous animals produce only live progeny at any temperature.

We have also found that the embryonic lethality at 25°C is a strict maternal phenotype. That is to say that despite *qm37* behaving as a recessive mutation, a wild-type allele in the genome of the embryo is not sufficient for survival if the mother was *clk-2/clk-2* homozygous mutant. When *clk-2* hermaphrodites are mated to wild-type males at 25°C they nonetheless produce only dead embryos. When shifted to 18°C at various times after mating they produce live males, indicating that the mating was successful. The strictly maternal lethal action of *clk-2* indicates a very early focus of action, before activation of the zygotic genome.

To establish how early *clk-2* acts during the development of the worm, we dissected embryos at the 2-4 cell stage from wild-type N2 and *clk-2* mutant hermaphrodites kept at either permissive (20°C) or non-permissive (25°C) temperature and transferred them to the other temperature (or not, as a control). As summarized in table 3, we found that when development up to the 2-4 cell stage proceeded at the permissive temperature, almost all eggs hatched and carried out further embryonic and post-embryonic development at 20°C {100% of dissected N2 eggs (n=35) hatched and 87% of dissected *clk-2* eggs hatched (n=91)} or 25°C {97% of dissected N2 eggs (n=36) hatched and 91% of dissected *clk-2* eggs hatched (n=93)}. In contrast, when eggs had carried out development up to the 2-4 cell stage at

- 20 -

25°C and were then transferred to 20°C, only very few *clk-2* eggs hatched and succeeded in completing development at 20°C {12% of dissected *clk-2* eggs hatched (n=136)}. As a control, when N2 eggs had carried out development up to the 2-4 cell stage at 25°C and were then transferred to 20°C, almost all hatched and succeeded in completing development at 20°C {98%, n=45}, or at 25°C {96%, n=45}. These results indicate that *clk-2* is required for viability before the 2-4 cell stage. *clk-2* is required in a narrow window between the very end of oogenesis and the initiation of embryonic development.

Table 3

Survival of eggs at the 2-4 cell stage, dissected from mothers raised at 20 or 25°C and transferred or not to another temperature

Mothers	% of eggs that hatch when developing at 20°C	% of eggs that hatch when developing at 25°C
N2 at 20°C	100 n = 35	97 n = 36
<i>clk-2</i> at 20°C	87 n = 91	91* n = 93
N2 at 25°C	98 n = 45	96 n = 45
<i>clk-2</i> at 25°C	12* n = 136	n.d.

Eggs that have reached the 2-4 cell stage at 20°C are viable even when further development is carried out at 25°C, while eggs that reached the 2-4 cell stage at 25°C die even when developing subsequently at 20°C.

*Only hatching is recorded in this table, but it should be noted that the 12% of *clk-2* eggs transferred from 25 to 20°C that succeed in hatching do not subsequently complete post-embryonic development, while the majority of the 91% *clk-2* eggs that hatch when transferred from 20 to 25°C reach adulthood.

- 21 -

Indeed, *clk-2* hermaphrodites that have spent 26 hours of adulthood at 25°C, carry on average 9.9 developing eggs in the uterus (n=125), but produce on average 10.7 dead eggs (n=133) when shifted down to permissive temperature. This observation indicates that, upon transfer from the lethal temperature, only one oocyte or embryo dies on average in addition to those that have already formed an eggshell. This corresponds to the time at which fertilization, oocyte meiosis, pronuclear formation and eggshell formation occurs. We observed early embryonic development using DIC microscopy but did not detect any obvious abnormality in the events which follow fertilization. The early embryos look invariably normal and healthy with cells and nuclei of normal size and shape. We also visualized DNA using Dapi in oocyte and early embryos and did not detect abnormal patterns of chromosome segregation or any other defects. Finally, meiosis per se is not affected as *clk-2* homozygous males can sire abundant cross-progeny at 25°C when mated to wild-type hermaphrodites.

clk-2 positional cloning, gene structure and operon

We have molecularly identified the gene *clk-2* by positional cloning. The gene was localized on the genetic map within an interval of 0.84 cM on the left cluster of linkage group III of *Caenorhabditis elegans*, between the genetic markers *sma-4* and *mab-5* (Hekimi, S. et al., *Genetics* **141**, 1351 (1995)). We refined this genetic position by a series of additional mapping experiments involving the genetic markers *sma-3*, *unc-36*, *lin-13*, and *lin-39* by multi- and two-point crosses. The following multi-point results were obtained (the genotypes whose progeny was scored is given in brackets): *dpy-17* 14 *clk-2* 18 *unc-32* (*clk-2* / *dpy-17* *unc-32*); *lon-1* 47 *clk-2* 23 *unc-36* (*clk-2* / *lon-1* *unc-*

- 22 -

36); *sma-4* 35 *clk-2* 3 *mab-5* 14 *unc-36* (*clk-2* / *sma-4*
mab-5 *unc-36*); *sma-3* 18 *clk-2* 0 *lin-13* 10 *unc-36* (*sma-3*
clk-2 *unc-36* / *lin-13*); *clk-2* 3 *lin-13* 49 *unc-32* (*lin-13*/
clk-2 *unc-32*); *sma-3* 40 *lin-39* 0 *clk-2* 33 *unc-36*
5 (*sma-3* *clk-2* *unc-36*/ *lin-39*). In addition, a two-point
cross was carried out (*clk-2* *unc-36*/ + +) and 5/630
Uncs were found to develop quickly ($p=0.4$ cM). We also
found that the deletion *nDf20* does not delete *clk-2* and
that the duplication *qDp3* does include *clk-2*. We thus
10 placed the gene *clk-2* within an interval of 0.3 cM,
between *sma-3* (at -0.9 cM on LGIII) and *lin-13* (at -0.6
cM on LGIII), and lying very close to the gene *lin-39*
(at -0.65 cM).

By aligning the genetic and physical maps, we
15 predicted the physical region which likely would
contain the *clk-2* gene. Groups of cosmids from this
region were tested for their ability to rescue the *clk-2*
mutant by DNA microinjection. *clk-2* was rescued by a
pool of 4 cosmids (*H14A12*, *K07D8*, *C34A5*, *C07H6*).
20 Individual injection of cosmids *C07H6* and *C34A5* also
rescued the *clk-2* phenotype, narrowing the physical
position of *clk-2* to within approximately 15kb.
Fragments of cosmid *C07H6* (obtained by restriction
digests from base pair 31,528 to base pair 36,545 of
25 cosmid *C07H6* [Accession: AC006605]) were then tested
for rescue and a short region of approximately 5kb was
shown to fully rescue the phenotype, indicating that
this 5kb fragment contains the *clk-2* gene.

The identity of the gene was further confirmed
30 by phenocopying the *clk-2* phenotype with RNA
interference (RNAi) experiments, that is the injection
of double stranded RNA corresponding to the coding mRNA
sequence of a gene of interest to fully abolish the

- 23 -

function of this gene. Double stranded RNA was produced by *in vitro* transcription from a cDNA (EST 447b4, gift of Y. Kohara) that mapped to this region, and injected into wild-type as well as into *clk-2(qm37)* worms. All wild-type and *clk-2* animals injected with *clk-2* dsRNA initially produced embryos that hatched and developed into worms phenotypically resembling *clk-2(qm37)*, that is, slow development, slow defecation and sterility. After 24 hours, the injected animals started laying only dead eggs. These results confirmed the identity of *clk-2*. The observation that RNAi-treated mothers produce dead eggs, a phenotype more severe than the weak embryonic lethality normally present in the *clk-2(qm37)* strain, indicated that *qm37* is a partial loss-of-function mutation that displays the null phenotype only at 25°C. We further confirmed the identity of the gene by characterizing the molecular lesion underlying the *clk-2* mutation. Genomic DNA from the *clk-2(qm37)* strain was isolated and the nucleotide sequence of the *clk-2* region determined. The *qm37* mutation is a G→A transition at in base 2321 of the cDNA.

The structure of the gene was established experimentally by determining the nucleotide sequence of the EST yk447b4 cDNA, thus defining the actual intron/exon boundaries *in vivo* and allowing to predict the encoded protein. The gene *clk-2* is SL2 transpliced. We have further established the gene structure by RT-PCR experiments, which not only showed that *clk-2* is SL2 transpliced, but also that the gene just upstream to *clk-2*, which we called *cex-7*, is expressed and is SL1 transpliced. The transplicing by SL1 of a gene

- 24 -

placed upstream, and by SL2 of a gene downstream constitutes a hallmark of genes which are in an operon, and are transcriptionally co-expressed. Therefore, *clk-2* and *cex-7* are transcriptionally co-expressed, and thus play functionally related roles. The cDNA (yk215f6) that corresponds to *cex-7* was also sequenced. The gene *cex-7* encodes a predicted protein of 481 amino acid residues in length (Fig. 34), that is similar to a human polypeptide of 550 amino acids (Fig. 35).

clk-2 encodes a predicted protein of 877 amino acids and the *clk-2(qm37)* mutation is a cysteine to tyrosine substitution at residue 772 of the predicted protein. We have been able to detect the expressed protein by western blot analysis of protein extracted from both mutant and wild-type worms at different temperatures. CLK-2 is similar to unique predicted proteins in human (Fig. 3), *Drosophila* (Fig. 13), rice (Fig. 19), soybean (Fig. 26-30) and to *Saccharomyces cerevisiae* Tel2p (Fig. 32) and in other species (Figs. 7-12, 14, 17-19). The structural conservation among these proteins is illustrated by the alignment presented in Figs. 38, 39, 40 and 41. No homologue of Tel2p had previously been recognized because aligning multiple sequences is necessary to reveal the homology. Tel2p has been shown to bind yeast telomeric DNA in a sequence-specific manner (Kota, R.S. Runge, K.W. *Chromosoma* 108, 278 (1999); Kota, R.S., Runge, K.W. *Nucleic Acids Research* 26, 1528 (1998)) and to affect the length of telomeres.

30 Expression pattern of *clk-2*

We determined the spatial and temporal expression pattern of the gene *clk-2* by analyzing

- 25 -

transcript and protein levels (Fig. 55) and by examining transgenic worms carrying reporter fusions. Panel A of Fig. 55 illustrates Northern and Western (37) analyses of *clk-2* at all developmental stages. The level of *clk-2* mRNA appears uniform throughout pre-adult development (E, embryos; L1-L4, larval stages; A, adult; *glp-4*, adult *glp-4(bn2ts)* mutants at 25°C). The low level of *clk-2* expression in L4 larvae and in *glp-4* mutants that lack a germline at 25°C suggest that most *clk-2* RNA in adults is located in gametes. In contrast to the finding with mRNA, the level of CLK-2 protein is similar at all stages including adults (lower panel of A). Panel B of Fig. 55, *clk-2* mRNA and protein levels (lower panel) in mutant backgrounds (*glp-4(bn2ts)*, *fem-3(q20ts)*, which produces only sperm at 25°C, and *fem-2(b245ts)*, which produces only oocytes at 25°C). The mRNA and protein levels of *clk-2* expression are similar to the wild type in *fem-3* and elevated in *fem-2* mutants. *glp-4* mutants have wild type protein levels but reduced mRNA levels. *clk-2* mRNA appears strongly elevated in *clk-2* mutants. Panel C of Fig. 55, CLK-2 protein levels in wild type and *clk-2* mutants at three temperatures. *clk-2(qm37)* is a missense (C772Y) and temperature-sensitive mutation. The level of CLK-2 is greatly reduced in the mutant, but does not change as a function of temperature in either the wild type or the mutant. Worms were raised at 20°C except when specified otherwise.

We grew populations of worms synchronized at different developmental stages and extracted total or polyA+ selected RNA from them. The highest level of *clk-2* mRNA is detected in young adults. We used several mutants to determine the origin of the transcript level in young adults. Since *clk-2* mRNA level is highly reduced in *glp-4(bn2ts)* mutants that do not develop a

- 26 -

germline at the non-permissive temperature, most of the RNA present in wild-type young adults is in the germline. Given the low abundance of RNA in L4 larvae which possess an already large germline but only a few
5 male gametes, most of the *clk-2* mRNA in wild-type adults is localized to meiotic gametes, in particular to oocytes.

We have analyzed the CLK-2 protein level in different genetic backgrounds and in worms grown at
10 different temperatures. We immunodetected CLK-2 protein on western blots by using two different polyclonal antibodies, MG19 and MG20. We obtained these antibodies by injecting rabbits with a bacterially expressed His₁₀-CLK-2 protein. We found that the content of CLK-2
15 protein is uniform across developmental stages in wild type and in *clk-2* animals. Furthermore, the concentration of CLK-2 is not different from the wild type in *glp-4* mutants which have no germline, nor in *fem-3* and *fem-2* mutants that contain only sperm and
20 only oocytes, respectively. Taken together these results indicate that gametes specifically accumulate high levels of *clk-2* mRNA, presumably as a store to be used by the embryo. Finally, we observed that in *qm37* mutants, while the level of *clk-2* mRNA appears slightly
25 elevated, the level of CLK-2 protein is greatly reduced.

We constructed three reporter constructs of the *clk-2* gene that comprised different upstream promoter regions and /or the coding region of the *clk-2* gene
30 fused to the green fluorescent protein. Two of the constructs are transcriptional fusions, one containing bases 36932 to 37319 and the other containing bases

- 27 -

36932 to 40010 of cosmid C07H6 [Accession: AC006605]. A third reporter construct (pMQ251) is a translational fusion that contains bases 30501 to 37319, except bases 35078 to 36545 which are part of the gene *cex-7*. We microinjected these reporter genes into wild type and *clk-2(qm37)* mutant worms, and analyzed numerous worms from several transgenic lines carrying these reporters. We observed that the *clk-2* promoter region directs expression in all somatic tissues, including hypodermis, muscles, neurons, excretory system, gut, pharynx, somatic gonad, vulva, and presumably all cells. No expression was visible in the germline, despite the use of both standard and complex array mixes. This is commonly the case for transgenes in *C. elegans* and does not indicate an absence of expression in the germline tissue. A full length fusion protein between CLK-2 and GFP (encoded by the construct pMQ251) that complements the mutant phenotype for development, behavior and viability at 25°C, is localized exclusively into the cytoplasm, which is consistent with the absence of an obvious nuclear localization signal in the predicted protein. The pattern observed is not a consequence of overexpression as very small transgene concentrations have been used in complex arrays (Kelly et al., *Genetics* 146:227-238, 1997). However, although the nucleus appears dark in the fluorescent images, it still may contain very small amounts of the fusion protein. This analysis of expression indicates that CLK-2 protein is indeed produced in the nematode, as shown by western analysis on total *C. elegans* extracts using anti-CLK-2 antibodies.

- 28 -

Yeast Tel2p has been found to bind telomeric repeats *in vitro*, and thus is expected to be nuclear *in vivo*. However, it was found that CLK-2::GFP is excluded from the nucleus. Subtelomeric silencing and telomere length regulation can also be affected by events in the cytosol. For example, Hst2p, a cytosolic NAD⁺-dependent deacetylase homologous to Sir2p, can modulate nucleolar and telomeric silencing in yeast Perrod et al., *EMBO J.*, **20**(Nos 1 & 2), 197-209, 2001), and the nonsense-mediated mRNA decay pathway appears to affect both telomeric silencing and telomere length regulation (Lew et al., *Molecular and Cellular Biology*, **18**(10):6121-6130, 1998). Other proteins that affect telomere length, like tankyrase Smith, S. and De Lange, Titia, *J. of cell Science*, **112**:3649-3656, 1999), are mostly extranuclear Chi, N.-W., and Lodish, H.F., *J. of Biological Chemistry*, **275**(49):38437-38444, 2000), with only a very small amount of protein localized to the telomeres Smith et al., *Science* **282**:1484-1487, 1998).

20 The role of *clk-2*

Telomere function has been found to affect replicative life span in yeast and in vertebrate cells. It also has also been shown to affect the immortality of the germline in *C. elegans*. However, an involvement of telomere function in determining the life span of multicellular organisms has not been established prior to this work. Here we have shown that the maternal-effect *clk-2* gene of *C. elegans* regulates telomere length, and prolongs life span by a mechanism that is distinct from the regulation of dauer formation but resembles caloric restriction, and encodes a protein

- 29 -

that is similar to the yeast telomere binding protein Tel2p.

The timing of the lethal action of *clk-2(qm37)* indicates a function for *clk-2* during the events that immediately follow fertilization, including oocyte meiosis, pronuclei formation and karyogamy, and this would be consistent with the known importance of telomeres in meiosis. However, our examination of the morphology of chromosomes in oocytes and early embryos did not reveal any abnormalities. Similarly, although telomere function appears linked to double strand break repair and chromosome stability, including in worms, *clk-2* mutants appear only moderately sensitive to ionizing radiation and do not display signs of chromosome instability. In fact, we examined the response of *clk-2(qm37)* mutants to gamma-radiation and found that among the progeny of irradiated animals, the proportion of dead eggs and larvae was about 10 times higher than among the progeny of irradiated wild-type animals. There is also no report of a function of Tel2p in the response to ionizing radiation in yeast.

The null phenotype of *tel2* is lethal but a hypomorphic mutation of *tel2* results in short telomeres and slow growth (Runge, K.W. and Zakian, V.A. *Molecular & Cellular Biology* **16**, 3094 (1996)). Tel2p has been shown to be involved in telomere position effect (TPE) and thus contributes to silencing of sub-telomeric regions (Runge, K.W. and Zakian, V.A. *Molecular & Cellular Biology* **16**, 3094 (1996)), one of the best studied examples of epigenesis. Mutations in other genes, such as *tell1*, that also result in telomere shortening do not result in abnormal TPE, indicating

- 30 -

that the TPE defect in *tel2* mutants is not a simple consequence of short telomeres. Furthermore, the rapid death and abnormal cellular morphology of cells fully lacking Tel2p suggests that Tel2p, like Rap1p and the Sir proteins, also functions at non-telomeric sites (Zakian, V.A. *Ann. Rev. Genet.* **30**, 141 (1996)). In light of this, the absolute requirement for maternal *clk-2* in embryogenesis suggests a function for CLK-2 in silencing genes that are needed during some part of the worm's life cycle but that are deleterious when expressed during early development. The study of the *mes* genes which are required for the specification of the germline in *C. elegans* and can confer maternal-effect sterile phenotype has shown that mechanisms of silencing are part of the normal development of worms. Indeed, some of the *mes* genes have been found to encode proteins that resemble Polycomb group proteins and appear generally to be involved in the regulation of chromatin structure.

Mutations in *clk-1* and *clk-2(qm37)* at the permissive temperature confer a similar CLK phenotype and in particular an increase of life span of similar magnitude (Lakowski, B. and Hekimi, S. *Science* **272**, 1010 (1996) and show similar pattern of interactions with other aging genes (Lakowski, B. Hekimi, S. *Proc. Nat. Acad. Sci. US* **95**, 13091 (1998)). CLK-1 is a mitochondrial protein of unknown function (Felkai, S. *et al*, *EMBO Journal* **18**, 1783 (1999)). In an attempt to explain many puzzling features of the *clk-1* phenotype, including the maternal effect, we have suggested that the action of CLK-1 is to indirectly, but specifically, regulate nuclear gene expression

- 31 -

(Branicky R, C. Benard, S. Hekimi, *Bioessays* 22:48, 2000). One possibility might be that CLK-2 might be one of the molecules that implements changes in gene expression in response to alteration of CLK-1 activity.

5 *clk-1 clk-2* double mutants have a phenotype that is more severe than either of the single mutants (Lakowski, B. and Hekimi, S. *Science* 272, 1010 (1996). However, the phenotype of a double mutants containing the null allele *clk-1(qm30)* is not more

10 severe than a double mutant containing the much weaker allele *clk-1(e2519)*, in contrast to the situation with *clk-3*, for which double mutants with *clk-1(qm30)* are much more severe than with *clk-1(e2519)* (Lakowski, B. and Hekimi, S. *Science* 272, 1010 (1996). These

15 observations indicate that at least part of the activity of *clk-1* requires *clk-2*. Furthermore, *clk-1 clk-2* double mutant embryos resemble *clk-1* mutant in that the interphases of the embryonic cell cycles are slowed down, but mitoses appear unaltered. This

20 indicates that *clk-2* as well as *clk-1* is involved in determining the rate of cellular multiplication, and thus affects mechanisms which are known to lead to cancer when deregulated.

Telomere function has also been implicated in

25 the replicative life span of yeast, where Sir proteins mediate silencing at the telomeres and the HM loci. When displaced from the telomeres by mutation or by shortage of telomeric DNA, part of the Sir complex can move to the nucleolus where its action appears to

30 prolong replicative life span. These and other studies indicate that telomeres are a reserve compartment for silencing factors and participate in regulating

- 32 -

silencing in other parts of the genome. It has been suggested that the effect on cellular senescence of expressing telomerase in cultured human cells might be mediated by an effect on silencing rather than by preventing chromosome erosion. Therefore, *clk-2* must be involved in determining cellular senescence, including in vertebrates, and affect in this manner aging and diseases linked to cellular senescence such as cancer.

As mentioned earlier, CLK-2 is similar to predicted proteins in vertebrates and plants as well as to *Saccharomyces cerevisiae* Tel2p. Tel2p has been shown to bind yeast telomeric DNA in a sequence-specific manner, and to affect the length of telomeres. We found that *clk-2* also affected the length of telomeres in worms (Fig. 56). In worms, genomic DNA hybridization to telomeric probes after restriction digestion with *HinfI* reveals the end fragments of the chromosomes carrying the telomeres, which appear as smears, as well as fragments carrying tracts of telomeric repeats that are internal to the chromosome, which appear as discrete bands. The regions where the telomeric smears are the most intense are indicated by stippled lines. Two lanes are shown for each genotype and each temperature.

The length of telomeres in wild-type and *clk-2* mutants was examined by Southern blotting at three temperatures, including the lethal temperature. For 18 and 20°C, worms were grown for numerous generations at each temperature before DNA extraction. Since *clk-2(qm37)* is lethal at 25°C, mixed stage worms from 20°C were transferred to and grown at 25°C for 3-4 days. Genomic DNA was prepared, *HinfI* digested and separated

- 33 -

on a 0.6% agarose gel at 1.2Vcm^{-1} . Southern blots were hybridized with gamma ^{32}P dATP end-labelled TTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGG oligonucleotide. Use of a second type of probe, made by
5 direct incorporation of alpha ^{32}P dATP during PCR amplification of telomeric repeats from the plasmid cTel55X with primers T7 and SHP1617 (GAATAATGAGAATTTTCAGGC), gave identical results. The extrachromosomal array in MQ691 *clk-2(qm37); qmEx159*
10 contains a clone with the entire coding sequence of *clk-2* as well as the promoter of the operon but excluding *cux-7* (bases 37319 to 31528 of cosmid C07H6, except bases 36544 to 35077) and rescues *clk-2* mutant phenotypes. In *clk-2* mutants, telomeres are two to
15 three times longer than in the wild type on average (Figure 56). However, the chromosomes are of wild-type length in strain MQ691, which carries an extrachromosomal array expressing wild-type CLK-2 in a *clk-2(qm37)* chromosomal background (Fig. 56) indicating
20 that the alteration of telomere length *clk-2(qm37)* mutants is indeed due to abnormal function of *clk-2* in these mutants.

The length of terminal telomeric fragments in the animals of the strain MQ691, which carries an
25 extrachromosomal array (*qmEx159*) containing functional wild-type CLK-2 that rescues development and behavior at 25°C in a *clk-2(qm37)* chromosomal background, was further analyzed. A similar clone containing the *qm37* mutation fails to rescue the *Clk-2* phenotypes. In MQ691
30 animals, the length of terminal telomeric fragments appear very similar to the wild-type, and even shorter, indicating that the lengthened telomere phenotype of

- 34 -

qm37 mutants is rescued by the expression of *clk-2(+)*. The telomere length of non-transgenic animals of the strain MQ931, derived from MQ691, which have lost the extrachromosomal array and thus again lack *clk-2(+)* has
5 been further examined. The terminal telomeric repeats in this strain are long again. Thus, the lengthened telomere phenotype of *clk-2(qm37)* can be rescued by *clk-2(+)* and reverses back to mutant length after the loss of the transgene.

10 In *C. elegans*, tracks of numerous TTAGGC telomeric repeats are present at the ends of the 6 chromosomes (Wicky C., et al., *Proc. Natl. Acad. Sci. USA*, **93**:8983-8988, 1996). In addition, numerous
15 interstitial blocks of perfect and degenerate telomeric repeats are located more internally to the chromosomes (*C. elegans* II. Edited by Riddell D et al. Published by Plainview, N.Y.: Cold Spring Harbor Laboratory Press (1997), pp 56-59, Chapter 3). Analysis of genomic DNA
20 after restriction digestion with a frequent cutter that does not cleave within the telomeric repeats (HinfI), electrophoresis, and hybridization to telomeric probes, reveals the telomere-carrying end fragments of the
25 chromosomes (Wicky C., et al., *Proc. Natl. Acad. Sci. USA*, **93**:8983-8988, 1996). Telomeres, and thus the restriction fragments containing them, are heterogeneous in size and appear as smears. On the other hand, restriction fragments carrying tracts of
30 internal telomeric repeats are of fixed size and appear as discrete bands in the 0.5-3 kb range (Ahmed S, Hodgkin J. *Nature*, **403**(6766):159-64, 2000; and Wicky C., et al., *Proc. Natl. Acad. Sci. USA*, **93**:8983-8988, 1996). The quality of visualization of the length of

- 35 -

telomeres in *C. elegans* with a hybridization probe that detects telomeric repeats is marred by the numerous internal repeats that also hybridize to the probe. In particular, they can mask the detection of the telomeres of chromosomes that have small HinfI terminal telomeric fragments. To further describe the telomere phenotype of *clk-2(qm37)* mutants, the length of individual telomeres has been characterized. The subtelomeric regions just adjacent to the terminal telomeric repeats share no sequence homology among the chromosomes (Wicky C., et al., *Proc. Natl. Acad. Sci. USA*, **93**:8983-8988, 1996). Taking advantage of this sequence diversity, probes specific to particular telomeres were designed. The size of a given HinfI terminal fragment is related to the fixed distance between the most exterior HinfI site of the chromosome and the beginning of the telomeric repeats, and by the variable number of terminal telomeric repeats. Upon genomic DNA digestion with HinfI and Southern blotting with a probe specific to a particular telomere, the terminal fragments, which are heterogeneous in size, again appear as a smear. Detailed results obtained for two individual telomeres are illustrated in Fig. 56.

The length of the terminal fragment of the left telomere of chromosome X is ~ 1 kb longer in *qm37* than in the wild type, ranging from 2.4 to 4.2 kb and from 1.7 to 2.8 kb, respectively. This telomere is of wild-type length in MQ691, which carries the rescuing transgene, and lengthens again to the *clk-2(qm37)* values in the non-rescued MQ931 strain. The length of another terminal fragment (left telomere of chromosome IV) is also ~1 kb longer in *qm37* than in the wild type,

- 36 -

ranging from 2.2 to 3.9 kb and from 1.8 to 2.8 kb respectively. This telomere becomes shorter than the wild type in MQ691, ranging from 1.3 to 2 kb only. This telomere acquires the mutant length again after loss of
5 the transgene in MQ931. Thus, the overexpression of *clk-2* can shorten the tracks of telomeric repeats, but not at each telomere.

Identity of the gene *coq-4* and the *coq-4* (*qm143*) mutants

10 The gene *COQ7/CAT5* of the yeast *S. cerevisiae* is the homologous gene to *clk-1* (Ewbank, J.J. et al, *Science* **275**, 980 (1997); PCT/CA97/00768). While *Coq7p* does not structurally resemble an enzyme, it is required for ubiquinone biosynthesis in yeast. A second
15 gene, *COQ4* (Marbois, B.N. and Clark, C. *J Biol Chem*, **271**, 2995 (1996) (Accession: NP_010490), that is also required for ubiquinone biosynthesis in yeast, does not code for an enzyme, and like *COQ7*, has no homologue in bacteria. We have generated a deletion mutant worm to
20 describe the role of the gene *coq-4* and its functional relationships with the *clk* genes we have identified and described, including *clk-1*, *clk-2*, and *gro-1*.

The gene *coq-4* in *C. elegans* largely corresponds to the predicted gene T03F1.2 of the cosmid
25 T03F1 (Accession U88169). It is localized on LGI, between *unc-73* and *unc-11*. *coq-4* is less than 100kb away from the characterized gene, *unc-73*, and less than 40kb away from the other characterized gene, *unc-11*. *coq-4* is 843 bp long and has four exons. We
30 experimentally established the structure of the gene *coq-4* by sequencing a cDNA clone, *ykl40a2*. A second gene, T03F1.3, which is highly similar to phosphoglycerate kinase (PGK), is 264 bp upstream of *coq-4* and, as we have shown, forms an operon with *coq-4*
35 and is thus transcriptionally co-expressed. We showed

- 37 -

that *coq-4* is in the same operon as T03F1.3 by RT-PCR, that *coq-4* is SL2 trans-spliced and that T03F1.3 is SL1 trans-spliced.

We have generated a *coq-4* (*qm143*) deletion mutant by carrying out PCR-based mutant screen following a large scale EMS mutagenesis wild-type worm. *coq-4* (*qm143*) has a 1469 bp deletion, which starts from 44 bp downstream of T03F1.3, and ends 406 bp downstream of *coq-4*. The predicted gene downstream is 1521 bp away from *coq-4* and 1115 bp away from the deletion. Therefore, *coq-4* (*qm143*) is a null mutant and it does not affect the coding sequence of any gene other than *coq-4*.

The phenotype of *coq-4* mutants

coq-4 (*qm143*) is a non-strict maternal-effect lethal mutation. Most of the progeny, from a homozygous *coq-4* hermaphrodite, dies during embryogenesis. Very few eggs hatch, and those which do hatch fail to complete development and die as young larvae. We have also shown that maternal *coq-4* product is sufficient for homozygous *coq-4* to develop normally until adulthood. However, homozygous *coq-4* adult worms from a heterozygous hermaphrodite (*coq-4*/+) are paralytic and are defective in egg-laying. Moreover, *coq-4* homozygous mutants can be mated by N2 males and produce progeny, which grow normally. Taken together, these results indicate that either maternal or zygotic *coq-4* product is sufficient for *coq-4* mutant to go through embryonic and post-embryonic development. *coq-4* deletion (*qm143*) is kept as a balanced strain, *coq-4* (*qm143*)/*unc-73*(*e936*). We demonstrated that the phenotypes of the *coq-4* mutants, in particular the sterility, can be rescued by an extrachromosomal wild-type copy of *coq-4* DNA fragment.

Expression pattern of *coq-4*

The spatial expression pattern of *coq-4* was determined by using translational reporter fusion to the green fluorescence protein, containing 2.2 kb of upstream promoter region. These constructs were injected into both N2 and heterozygous *coq-4* (*coq-4/unc-73*), and animals of several transgenic lines were examined. We found that a functional *coq-4::gfp* is expressed in the hypodermis, muscles, the gut, the excretory canal and embryos. In addition, we detected that the reporter fusion localizes to the mitochondria, in particular, in muscle cells.

While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and this application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure as come within known or customary practice within the art to which the invention pertains and as may be applied to the essential features hereinbefore set forth, and as follows in the scope of the appended claims.

- 1/40 -

SEQUENCE LISTING

<110> McGill University
 HEKIMI, Siegfried
 BÉNARD, Claire
 McCRIGHT, Brent
 LAKOWSKI, Bernard
 HAN, Dong
 LABBE, Jean-Claude

<120> CLK-2, CEX-7 AND COQ-4 GENES, AND USES
 THEREOF

<130> 1770-267PCT

<150> us 60/254,932

<151> 2000-12-13

<150> us 60/213,174

<151> 2000-06-22

<160> 52

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 2814

<212> DNA

<213> C. elegans clk-2 cDNA

<400> 1

ctcaagatga	atttacgaag	tcgcctggta	aatgccacgg	aacgtgctgt	gctttttcaa	60
attttcaaag	atgtgcagaa	tgatccggaa	aagtacgaca	atgcagttga	ggcgatctgt	120
gaatcaatcg	actattttgg	caaatttttg	accgatagcg	agtatcttac	acaaatcaag	180
ccgattcttg	atacacagt	cccaacaaag	tcgataattt	gcttctcgaa	atgtttgaca	240
aaagtgagca	cagatataaa	tactaccaca	tttcgagatg	tgatcaccat	gctcgactgg	300
ttgaagtatg	tcgttgaaaa	atcgctgaca	agtgctat	gtagcagtct	gaaagttaaa	360
gaaactgatg	tcagtgcagt	tcagttgtat	cgagaattcg	catcagcatg	ttcaaata	420
ccggagaaag	tttcgaattg	ttgtgcaaag	gcattgtctg	gcgagcatgt	caaataatc	480
aacacgggta	agtggatatt	caaatgaat	ctgggtgcaag	gaattcaaaa	ggctatgctt	540
cttgctcacg	acgacattgt	aactgctgcc	ccgttcactt	cattctacgg	atccgggtgg	600
ccttatatga	agactgtcgc	agaaattatt	tcattctggaa	gaaacataga	tatcaccaac	660
aaggatgggc	ttctagttca	aatgattgaa	tggattgggt	cactaaacaa	ttttgattct	720
caatggcgtc	ggatgatgtt	tctcatcttc	caagagccca	catatcaggg	aattcaagtt	780
catgaatcac	tactgacaac	attgttccta	atttcgaaaa	gtgaccaa	cttgaaacga	840
tgtatcgaag	ccactgatct	gactggaaca	ctgaagcgtg	tagtgatgg	taagctcccg	900
tttcagcgag	ttctcaaacg	aaagaccatc	gagattctta	tcaattttgt	ttatcgaact	960
aaggaacaat	ttgccatcca	gctattagag	acttctgtga	aaatctggag	tgatctcaat	1020
tacgcaaaaa	gtgctccaga	atcacaagaa	cgacacatag	tcagaatgat	attatacttg	1080
gttcatcttt	tcagaacatg	ttcttcaatc	gattgggagt	cactcttcct	gaactctatg	1140
gatggagttc	attgtcgaat	gagcatgttg	cctatgtacg	tccaaagtgg	tatttttggt	1200
aatcaagcac	tgtgcaagca	agcgacaaag	catcgatcga	aaacgcacgg	atcagatgag	1260
caacctccag	agactctaga	agaaaacaaa	ttcgtttcaa	gtgaagtggg	aaaaatatgg	1320
tttgaagaga	tgacgtcaat	tttggaaacat	ggatttaatt	cttctacagt	gaaagattct	1380

- 2/40 -

```

gagcgcgagttc gagaaaccgc caacgaaata accaaagacg attcgggtga agaatttgaa 1440
gaaaccaatg cacagcgtct tcaaaacaac aaagattcgg cagcaatcac atcgaaaaac 1500
aatctacggt tagattctga tgatgacgaa gactttcctg actatcaagt taatgaatca 1560
gaaaagatct tcaagaattt agaaattgga gaagaaccga aaaataaagt gacacctcca 1620
gcatacattg cagatgcttt cgaaatgcta ttagagaaag aaaaatacga ggtttttgaa 1680
gcagctttct ttaatattac gaatttgatc aatcgccggc caattggatt tccacaaatc 1740
gctgagaagt tgttcatccg aatcctccat cttcaaaaca attttggaac gcctaaattc 1800
aaggaaactg ttgatgaaat tgcagttgca tgtatcactc agcgtccaga aattgtccca 1860
tctgtagtgc gtctgatcat tgcaccaggt caaggtttca gtatcaaaca acgtcttctt 1920
cattacattc acaatgctgc tgatggaatg ggtgcattgg ataagaaact tgaagagtgt 1980
gtaatggcgc aacaattgag aattgggtgg ccaacgttaa gtattattct tcatcgaact 2040
ataaacactg attatgacga tgaggatgaa gatccccaca gacttttagt tcctgaatgg 2100
cgacgaatgg tggatgctcg cattgctgca aataccagaa gaattggaac gacgcgagag 2160
ccaccaagag ccggagtgtg caatcgtctc gcacaagctg ccaaataatat gttttatcct 2220
ttgttggttt tgccacgtgg tgagaacgca agtcttttg gcaaggactc cgatctactc 2280
gcctcactca tcatggttgc atcgatggtt tatgtgagat gtggcgtatg tcctcaaatt 2340
catcgaatgt caagtgcgct tatatcatat gcaacgcctc atcgattctc tgaaaacgca 2400
aaactacgga ctgcctgcat cattgccccat ctgaatgtga cgactttgtt gcctggagat 2460
cttatggatg aactgtttga tgtaccggca cttattggat ggtttgattg ggccaattct 2520
gtactggtta atgcatcttc atcacaattg gaaaaggata tgactcgcca gtttggtcat 2580
agtgttacaa aacaccttca acgttatcat ccagctgtac tgcaacacca agacgtttaa 2640
atagttacta ttcacttggt ttccttcttt tcaagtactg tatcatttct tactatcttg 2700
ccaacacttt gatctctacc tcgttcactt cttgctttgc caccggttga tatcacctgt 2760
ctcattcatt tatcagcatg ttcataatat caaaaataaa atcttatcaa atgt 2814

```

<210> 2

<211> 876

<212> PRT

<213> C. elegans

<220>

<223> clk-2 protein

<400> 2

```

Met Asn Leu Arg Ser Arg Leu Val Asn Ala Thr Glu Arg Ala Val Leu
 1           5           10           15
Phe Gln Ile Phe Lys Asp Val Gln Asn Asp Pro Glu Lys Tyr Asp Asn
           20           25           30
Ala Val Glu Ala Ile Cys Glu Ser Ile Asp Tyr Phe Gly Lys Phe Leu
           35           40           45
Thr Asp Ser Glu Tyr Leu Thr Gln Ile Lys Pro Ile Leu Asp Thr Gln
           50           55           60
Cys Pro Thr Lys Ser Ile Ile Cys Phe Ser Lys Cys Leu Thr Lys Val
65           70           75           80
Ser Thr Asp Ile Asn Thr Thr Thr Phe Arg Asp Val Ile Thr Met Leu
           85           90           95
Asp Trp Leu Lys Tyr Val Val Glu Lys Ser Leu Thr Ser Ala Ile Cys
           100          105          110
Ser Ser Leu Lys Val Lys Glu Thr Asp Val Ser Ala Val Gln Leu Tyr
           115          120          125
Arg Glu Phe Ala Ser Ala Cys Ser Asn Ile Pro Glu Lys Val Ser Asn
           130          135          140
Cys Cys Ala Lys Ala Leu Ser Gly Glu His Val Lys Tyr Ile Asn Thr
145          150          155          160

```

- 3/40 -

Val	Lys	Trp	Ile	Phe	Lys	Met	Asn	Leu	Val	Gln	Gly	Ile	Gln	Lys	Ala	
				165					170					175		
Met	Leu	Leu	Ala	His	Asp	Asp	Ile	Val	Thr	Ala	Ala	Pro	Phe	Thr	Ser	
			180					185					190			
Phe	Tyr	Gly	Ser	Gly	Gly	Pro	Tyr	Met	Lys	Thr	Val	Ala	Glu	Ile	Ile	
		195					200					205				
Ser	Ser	Gly	Arg	Ile	Asp	Ile	Thr	Asn	Lys	Asp	Gly	Leu	Leu	Val	Gln	
	210					215					220					
Met	Ile	Glu	Trp	Ile	Gly	Ser	Leu	Asn	Asn	Phe	Asp	Ser	Gln	Trp	Arg	
225					230					235					240	
Arg	Met	Met	Phe	Leu	Ile	Phe	Gln	Glu	Pro	Thr	Tyr	Gln	Gly	Ile	Gln	
			245						250					255		
Val	His	Glu	Ser	Leu	Leu	Thr	Thr	Leu	Phe	Leu	Ile	Ser	Lys	Ser	Asp	
			260					265					270			
Gln	Ile	Leu	Lys	Arg	Cys	Ile	Glu	Ala	Thr	Asp	Leu	Thr	Gly	Thr	Leu	
		275					280					285				
Lys	Arg	Val	Val	Met	Val	Lys	Leu	Pro	Phe	Gln	Arg	Val	Leu	Lys	Arg	
	290					295					300					
Lys	Thr	Ile	Glu	Ile	Leu	Ile	Asn	Phe	Val	Tyr	Arg	Thr	Lys	Glu	Gln	
305					310					315					320	
Phe	Ala	Ile	Gln	Leu	Leu	Glu	Thr	Ser	Val	Lys	Ile	Trp	Ser	Asp	Leu	
			325						330					335		
Asn	Tyr	Ala	Lys	Ser	Ala	Pro	Glu	Ser	Gln	Glu	Arg	His	Ile	Val	Arg	
		340					345						350			
Met	Ile	Leu	Tyr	Leu	Val	His	Leu	Phe	Arg	Thr	Cys	Ser	Ser	Ile	Asp	
		355					360					365				
Trp	Glu	Ser	Leu	Phe	Leu	Asn	Ser	Met	Asp	Gly	Val	His	Cys	Arg	Met	
	370					375					380					
Ser	Met	Leu	Pro	Met	Tyr	Val	Gln	Ser	Gly	Ile	Phe	Val	Asn	Gln	Ala	
385					390					395					400	
Leu	Cys	Lys	Gln	Ala	Thr	Lys	His	Arg	Ser	Lys	Thr	His	Gly	Ser	Asp	
			405						410					415		
Glu	Gln	Pro	Pro	Glu	Thr	Leu	Glu	Glu	Asn	Lys	Phe	Val	Ser	Ser	Glu	
			420					425					430			
Val	Gly	Lys	Ile	Trp	Phe	Glu	Glu	Met	Thr	Ser	Ile	Leu	Glu	His	Gly	
		435					440					445				
Phe	Asn	Ser	Ser	Thr	Val	Lys	Asp	Ser	Glu	Arg	Val	Arg	Glu	Thr	Ala	
	450					455					460					
Asn	Glu	Ile	Thr	Lys	Asp	Asp	Ser	Gly	Glu	Glu	Phe	Glu	Glu	Thr	Asn	
465					470					475					480	
Ala	Gln	Arg	Leu	Gln	Asn	Asn	Lys	Asp	Ser	Ala	Ala	Ile	Thr	Ser	Lys	
			485						490					495		
Asn	Asn	Leu	Arg	Leu	Asp	Ser	Asp	Asp	Asp	Glu	Asp	Phe	Pro	Asp	Tyr	
		500						505					510			
Gln	Val	Asn	Glu	Ser	Glu	Lys	Ile	Phe	Lys	Asn	Leu	Glu	Ile	Gly	Glu	
		515					520					525				
Glu	Pro	Lys	Asn	Lys	Val	Thr	Pro	Pro	Ala	Tyr	Ile	Ala	Asp	Ala	Phe	
	530					535					540					
Glu	Met	Leu	Leu	Glu	Lys	Glu	Lys	Tyr	Glu	Val	Phe	Glu	Ala	Ala	Phe	
545					550					555					560	
Phe	Asn	Ile	Thr	Asn	Leu	Ile	Asn	Arg	Arg	Pro	Ile	Gly	Phe	Pro	Gln	
			565						570					575		
Ile	Ala	Glu	Lys	Leu	Phe	Ile	Arg	Ile	Leu	His	Leu	Gln	Asn	Asn	Phe	
			580					585					590			

- 4/40 -

Gly	Thr	Pro	Lys	Phe	Lys	Glu	Thr	Val	Asp	Glu	Ile	Ala	Val	Ala	Cys
		595						600				605			
Ile	Thr	Gln	Arg	Pro	Glu	Ile	Val	Pro	Ser	Val	Val	Arg	Leu	Ile	Ile
	610					615					620				
Ala	Pro	Gly	Gln	Gly	Phe	Ser	Ile	Lys	Gln	Arg	Leu	Leu	His	Tyr	Ile
625					630					635					640
His	Asn	Ala	Ala	Asp	Gly	Met	Gly	Ala	Leu	Asp	Lys	Lys	Leu	Glu	Glu
				645					650					655	
Cys	Val	Met	Ala	Gln	Gln	Leu	Arg	Ile	Gly	Gly	Pro	Thr	Leu	Ser	Ile
			660					665					670		
Ile	Leu	His	Arg	Thr	Ile	Asn	Thr	Asp	Tyr	Asp	Asp	Glu	Asp	Glu	Asp
		675					680					685			
Pro	His	Arg	Leu	Leu	Val	Pro	Glu	Trp	Arg	Arg	Met	Val	Asp	Ala	Arg
	690					695					700				
Ile	Ala	Ala	Asn	Thr	Arg	Arg	Ile	Gly	Thr	Thr	Arg	Glu	Pro	Pro	Arg
705					710					715					720
Ala	Gly	Val	Val	Asn	Arg	Leu	Ala	Gln	Ala	Ala	Lys	Tyr	Met	Phe	Tyr
				725					730					735	
Pro	Leu	Leu	Val	Leu	Pro	Arg	Gly	Glu	Asn	Ala	Ser	Leu	Leu	Gly	Lys
			740					745				750			
Asp	Ser	Asp	Leu	Leu	Ala	Ser	Leu	Ile	Met	Val	Ala	Ser	Met	Val	Tyr
		755					760				765				
Val	Arg	Cys	Gly	Val	Cys	Pro	Gln	Ile	His	Arg	Met	Ser	Ser	Glu	Leu
	770					775					780				
Ile	Ser	Tyr	Ala	Thr	Pro	His	Arg	Phe	Ser	Glu	Asn	Ala	Lys	Leu	Arg
785					790					795					800
Thr	Ala	Cys	Ile	Ile	Ala	His	Leu	Asn	Val	Thr	Thr	Leu	Leu	Pro	Gly
				805					810					815	
Asp	Leu	Met	Asp	Glu	Leu	Phe	Asp	Val	Pro	Ala	Leu	Ile	Gly	Trp	Phe
			820					825					830		
Asp	Trp	Ala	Asn	Ser	Val	Leu	Val	Asn	Ala	Ser	Ser	Ser	Gln	Leu	Glu
	835						840					845			
Lys	Asp	Met	Thr	Arg	Gln	Phe	Gly	His	Ser	Val	Thr	Lys	His	Leu	Gln
	850					855					860				
Arg	His	His	Pro	Ala	Val	Leu	Gln	His	Gln	Asp	Val				
865					870					875					

<210> 3

<211> 836

<212> PRT

<213> Homo sapiens

<220>

<223> clk-2 protein

<400> 3

Met	Glu	Pro	Ala	Pro	Ser	Glu	Val	Arg	Leu	Ala	Val	Arg	Glu	Ala	Ile
1				5					10					15	
His	Ala	Leu	Ser	Ser	Ser	Glu	Asp	Gly	Gly	His	Ile	Phe	Cys	Thr	Leu
			20					25					30		
Glu	Ser	Leu	Lys	Arg	Tyr	Leu	Gly	Glu	Met	Glu	Pro	Pro	Ala	Leu	Pro
	35					40					45				
Arg	Glu	Lys	Glu	Glu	Phe	Ala	Ser	Ala	His	Phe	Ser	Pro	Val	Leu	Arg
	50					55					60				

- 5/40 -

Cys	Leu	Ala	Ser	Arg	Leu	Ser	Pro	Ala	Trp	Leu	Glu	Leu	Leu	Pro	His
65					70					75				80	
Gly	Arg	Leu	Glu	Glu	Leu	Trp	Ala	Ser	Phe	Phe	Leu	Glu	Gly	Pro	Ala
				85					90					95	
Asp	Gln	Ala	Phe	Leu	Val	Leu	Met	Glu	Thr	Ile	Glu	Gly	Ala	Ala	Gly
			100					105					110		
Pro	Ser	Phe	Arg	Leu	Met	Lys	Met	Ala	Arg	Leu	Leu	Ala	Arg	Phe	Leu
		115					120					125			
Arg	Glu	Gly	Arg	Leu	Ala	Val	Leu	Met	Glu	Ala	Gln	Cys	Arg	Gln	Gln
	130					135					140				
Thr	Gln	Pro	Gly	Phe	Ile	Leu	Leu	Arg	Glu	Thr	Leu	Leu	Gly	Lys	Val
145					150					155					160
Val	Ala	Leu	Pro	Asp	His	Leu	Gly	Asn	Arg	Leu	Gln	Gln	Glu	Asn	Leu
				165					170					175	
Ala	Glu	Phe	Phe	Pro	Gln	Asn	Tyr	Phe	Arg	Leu	Leu	Gly	Glu	Glu	Val
			180					185					190		
Val	Arg	Val	Leu	Gln	Ala	Val	Val	Asp	Ser	Leu	Gln	Gly	Gly	Leu	Asp
		195					200					205			
Ser	Ser	Val	Ser	Phe	Val	Ser	Gln	Val	Leu	Gly	Lys	Ala	Cys	Val	His
	210					215					220				
Gly	Arg	Gln	Gln	Glu	Ile	Leu	Gly	Val	Leu	Val	Pro	Arg	Leu	Ala	Ala
225					230					235					240
Leu	Thr	Gln	Gly	Ser	Tyr	Leu	His	Gln	Arg	Val	Cys	Trp	Arg	Leu	Val
				245					250					255	
Glu	Gln	Val	Pro	Asp	Arg	Ala	Met	Glu	Ala	Val	Leu	Thr	Gly	Leu	Val
			260					265					270		
Glu	Ala	Ala	Leu	Gly	Pro	Glu	Val	Leu	Ser	Arg	Leu	Leu	Gly	Asn	Leu
		275					280					285			
Val	Val	Lys	Asn	Lys	Lys	Ala	Gln	Phe	Val	Met	Thr	Gln	Lys	Leu	Leu
	290					295					300				
Phe	Leu	Gln	Ser	Arg	Leu	Thr	Thr	Pro	Met	Leu	Gln	Ser	Leu	Leu	Gly
305					310					315					320
His	Leu	Ala	Met	Asp	Ser	Gln	Arg	Arg	Pro	Leu	Leu	Leu	Gln	Val	Leu
				325					330					335	
Lys	Glu	Leu	Leu	Glu	Thr	Trp	Gly	Ser	Ser	Ser	Ala	Ile	Arg	His	Thr
			340					345					350		
Pro	Leu	Pro	Gln	Gln	Arg	His	Val	Ser	Lys	Ala	Val	Leu	Ile	Cys	Leu
		355					360						365		
Ala	Gln	Leu	Gly	Glu	Pro	Glu	Leu	Arg	Asp	Ser	Arg	Asp	Glu	Leu	Leu
						375					380				
Ala	Ser	Met	Met	Ala	Gly	Val	Lys	Cys	Arg	Leu	Asp	Ser	Ser	Leu	Pro
385					390					395					400
Pro	Val	Arg	Arg	Leu	Gly	Met	Ile	Val	Ala	Glu	Val	Val	Ser	Ala	Arg
				405					410					415	
Ile	His	Pro	Glu	Gly	Pro	Pro	Leu	Lys	Phe	Gln	Tyr	Glu	Glu	Asp	Glu
			420					425					430		
Leu	Ser	Leu	Glu	Leu	Leu	Ala	Leu	Ala	Ser	Pro	Gln	Pro	Ala	Gly	Asp
		435					440					445			
Gly	Ala	Ser	Glu	Ala	Gly	Thr	Ser	Leu	Val	Pro	Ala	Thr	Ala	Glu	Pro
	450					455					460				
Pro	Ala	Glu	Thr	Pro	Ala	Glu	Ile	Val	Asp	Gly	Gly	Val	Pro	Gln	Ala
465					470					475					480
Gln	Leu	Ala	Gly	Ser	Asp	Ser	Asp	Leu	Asp	Ser	Asp	Asp	Glu	Phe	Val
				485					490					495	

- 6/40 -

Pro Tyr Asp Met Ser Gly Asp Arg Glu Leu Lys Ser Ser Lys Ala Pro
 500 505 510
 Ala Tyr Val Arg Asp Cys Val Glu Ala Leu Thr Thr Ser Glu Asp Ile
 515 520 525
 Glu Arg Trp Glu Ala Ala Leu Arg Ala Leu Glu Gly Leu Val Tyr Arg
 530 535 540
 Ser Pro Thr Ala Thr Arg Glu Val Ser Val Glu Leu Ala Lys Val Leu
 545 550 555 560
 Leu His Leu Glu Glu Lys Thr Cys Val Val Gly Phe Ala Gly Leu Arg
 565 570 575
 Gln Arg Ala Leu Val Ala Val Thr Val Thr Asp Pro Ala Pro Val Ala
 580 585 590
 Asp Tyr Leu Thr Ser Gln Phe Tyr Ala Leu Asn Tyr Ser Leu Arg Gln
 595 600 605
 Arg Met Asp Ile Leu Asp Val Leu Thr Leu Ala Ala Gln Glu Leu Ser
 610 615 620
 Arg Pro Gly Cys Leu Gly Arg Thr Pro Gln Pro Gly Ser Pro Ser Pro
 625 630 635 640
 Asn Thr Pro Cys Leu Pro Glu Ala Ala Val Ser Gln Pro Gly Ser Ala
 645 650 655
 Val Ala Ser Asp Trp Arg Val Val Val Glu Glu Arg Ile Arg Ser Lys
 660 665 670
 Thr Gln Arg Leu Ser Lys Gly Gly Pro Arg Gln Gly Pro Ala Gly Ser
 675 680 685
 Pro Ser Arg Phe Asn Ser Val Ala Gly His Phe Phe Phe Pro Leu Leu
 690 695 700
 Gln Arg Phe Asp Arg Pro Leu Val Thr Phe Asp Leu Leu Gly Glu Asp
 705 710 715 720
 Gln Leu Val Leu Gly Arg Leu Ala His Thr Leu Gly Ala Leu Met Cys
 725 730 735
 Leu Ala Val Asn Thr Thr Val Ala Val Ala Met Gly Lys Ala Leu Leu
 740 745 750
 Glu Phe Val Trp Ala Leu Arg Phe His Ile Asp Ala Tyr Val Arg Gln
 755 760 765
 Gly Leu Leu Ser Ala Val Ser Ser Val Leu Leu Ser Leu Pro Ala Ala
 770 775 780
 Leu Leu Glu Asp Leu Met Asp Glu Leu Leu Glu Ala Arg Ser Trp Leu
 785 790 795 800
 Ala Asp Val Ala Glu Lys Asp Pro Asp Glu Asp Cys Arg Thr Leu Ala
 805 810 815
 Leu Arg Ala Leu Leu Leu Leu Gln Arg Leu Lys Asn Arg Leu Leu Pro
 820 825 830
 Pro Ala Ser Pro
 835

<210> 4

<211> 3320

<212> DNA

<213> Homo sapiens clk-2

<400> 4

gcgccccaga	ggctcaagaa	aacccgcggg	agcctcgccc	ggacccagga	actcgtgctc	60
ggggccaacc	ggctggggccg	cgatcgcggtt	tcgtccgggg	ccgcggcggc	cgtgggggaat	120
cggctgcagc	gaatcgggtg	cgcgcggcgc	ctgagcgcgc	tgcagtcacc	cgggagccgg	180
gtccagggtcg	tcttcccgtg	acgccagat	ctgtcctgca	ggatggagcc	agcaccctca	240

- 7/40 -

gaggttcgac	tcgccgtccg	ggaagccatt	catgccctct	cgtcttcgga	ggatggcggc	300
cacatcttct	gcaccctgga	gtccctgaag	cggatatctc	gtgagatgga	gcctccagcg	360
ctcccgaggg	agaaggagga	gtttgcctcg	gccacttctt	cgcctgtcct	cagatgtctt	420
gccagcaggc	tgagcccagc	ctggctggag	ctgctgcccc	atggccgcct	ggaggagctg	480
tgggccagct	tcttcctgga	gggcccggcg	gaccaagcct	tcctgggtgt	gatggagacc	540
atcgaggggtg	ctgcggggccc	cagcttccgg	ctgatgaaga	tggcgcggct	gctggccaga	600
ttcctgcgcg	agggccggct	ggcagtgtct	atggaggcgc	agtgtcggca	gcagacgcgg	660
cccggcttca	tcctgctccg	ggagacgctg	ctgggcaagg	tgggtggccct	gcccgatcac	720
ctgggcaacc	gcctgcagca	ggagaacttg	gccgagttct	ttccccagaa	ctacttccgc	780
ctgctcggcg	aggaggtcgt	ccgggtgtct	caggcgggtg	tggactctct	ccaaggtggc	840
ctggattcct	ccgtgtcctt	cgtgtctcag	gtccttggga	aagcctgtgt	ccacgggagg	900
cagcaggaga	tcctgggctg	gctggtacct	cggctggcag	cgctcaccga	gggcagctac	960
ctgcaccagc	gcgtctgtct	gcgcctgggt	gagcaagtgc	cggaccgggc	catggaggct	1020
gtgctgaccg	ggctgggtgga	ggccgcactg	gggcctgagg	tcctttcgag	actgctgggg	1080
aacctgggtg	tgaagaacaa	gaaggcccag	tttgtgatga	cccagaagct	tctgttctta	1140
cagtcccggc	tcacgacgcc	catgctgcag	agcctgctgg	gccatctggc	catggacagc	1200
cagcggcgcc	cgctcctgct	gcaggtgtct	aaggagctgt	tggagacgtg	gggcagcagc	1260
agtgccatcc	gccacactcc	cctgccgcag	cagcgccacg	tcagcaaggc	tgtcctcatc	1320
tgcttggcgc	aactcgggga	gccggaactg	cgggacagcc	gggatgaact	gctggccagc	1380
atgatggcgg	gcgtgaagtg	ccgcctggac	agtagcctgc	cccccgctgc	acgcctgggc	1440
atgatcgtgg	cagaggtcgt	tagtgcccgg	atccaccccg	aggggcctcc	cctgaaattc	1500
cagtacgaag	aggatgaact	gagcctcgag	ctgctggcct	tggcctcccc	ccagcctgcg	1560
ggtgacggcg	cctcggaggc	gggcacgtcc	ctcgttccag	ccacggcaga	gccccctgca	1620
gagacccccg	cagagatcgt	ggatggcggc	gtcccccaag	cacagctggc	gggctctgac	1680
tcggacctgg	acagcgatga	tgagtttgtc	ccctacgaca	tgtcggggga	cagagagctg	1740
aagagcagca	aggctcctgc	ctacgtccgg	gactgcgtgg	aagccctgac	cacgtctgag	1800
gacatagagc	gctgggaggc	agccctgcgg	gcccttgagg	gcctgggtcta	caggagcccc	1860
acagccactc	gggaggtgag	cgtggagctg	gccaaggtgc	ttctgcatct	ggaggagaag	1920
acctgtgtgg	tgggatttgc	agggctgcgc	cagagagccc	tgggtggcct	cacggtcaca	1980
gacccggccc	cgggtggccga	ctatctgacc	tcacagttct	atgccctcaa	ctacagcctc	2040
cggcagcgca	tggacatcct	ggatgtgtct	actctggctg	cccaggagct	gtctaggcct	2100
gggtgcctcg	ggaggactcc	ccaacctggc	ttcccaagtc	ccaacacccc	gtgcctgcca	2160
gaggcagccg	tctctcagcc	tggcagtgcc	gtggcgtctg	actggcgggt	ggtggtggag	2220
gagcggatca	gaagcaagac	ccagcggctc	tccaaggggtg	gcccagaggca	gggcccggca	2280
ggcagcccca	gcagattcaa	ctccgtggcc	ggccacttct	tcttccccct	ccttcagcgc	2340
tttgacaggg	ctctgggtgac	cttcgacctc	ttgggagaag	accagctggg	tctcggaagg	2400
ctggcgcaca	ccttaggggc	cctgatgtgc	ctggctgtta	acaccacggg	ggctgtggcc	2460
atgggcaagg	ccctgctgga	attcgtgtgg	gcccttcgct	ttcacatcga	tgcctacgtg	2520
cgccaggggc	tgttgctcggc	cgtctcctcc	gtcctgctca	gcctgcctgc	tgcgcgcctg	2580
ctggaggacc	tgatggacga	gctgctggaa	gcccggctct	ggctggcggg	cgtggctgag	2640
aaagacccgg	acgaggactg	caggacgctg	gcactgaggg	ccctgctgct	tctgcagaga	2700
ctcaagaaca	ggctcctccc	acccgcgtct	ccctagtccc	tggaggcctc	cccaggacca	2760
ccctcgccga	cagcaaggca	ggcggctgag	cagcggcctg	gagcagcaga	gccaggcttt	2820
gtagcagagg	caggtcttct	gccgcattcc	gtacggagag	tgcagatgca	ggaaggcccg	2880
gcctgccgct	atttatagtg	cagccagtcc	gctaaaaata	cactgggcct	gggcactgcc	2940
cgccgggaca	tggcagcctg	gacgtggggc	tggggctgtg	ggcgtgtgtg	gcgggggttg	3000
ctcttccagt	gagggcagaa	ccaggctggc	aggagggggg	gacgggtgtac	ctgctgtctc	3060
gagcccccaa	ggctctcctc	tgagagccac	caagcaggac	agagcagctc	ttgtcccagg	3120
tcctcggggc	tgagcgccgt	gtcaccagga	gaatagtgtc	cacagcccag	gcagggtgtg	3180
tggctcctgg	atgggctcgt	ggggcgggat	gggacagggc	acgggctctc	agaaaataaa	3240
ctgctttatt	ggaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	3300
aaaaaaaaaa	aaaaaaaaaa					3320

<210> 5

<211> 466

- 8/40 -

<212> DNA

<213> Mus musculus clk-2 cDNA

<400> 5

gtgaccccgca	agctcctgct	cctgcagtac	cagcacacga	cacccatggt	gcagagcctg	60
ttgggggtact	tggctctaga	cagtcagcgg	cggccactcc	tcatacaggt	gcttaaggag	120
ctgctggaga	cctggggctg	cagcagtgct	gtccgacaca	cacccttgga	gcagcagtgt	180
tacatcagca	aggccatcct	tgtctgcctg	gcacacctcg	gggagccgga	gctgcaggac	240
atccgggatg	aattgctggc	cagcatgatg	gcaggcgtga	agtgccgcct	ggatagcagc	300
ctgccccctg	tgcgccgctt	gggcatgatt	gtggccgagg	tcatcagctc	caggatccac	360
cctgaggggc	ctctcctgaa	attccaatat	gaagatgacg	agatgagccg	tgagttgctg	420
gccttggcta	ccccagagcc	tgcgggtgac	tgctcctcgg	tgtcac		466

<210> 6

<211> 267

<212> DNA

<213> Artificial Sequence

<220>

<223> part of Mus musculus cDNA

<400> 6

tgtcccatgt	gctgacttct	gaggtggctg	tgctagtggg	taaggccctg	ctggagtttg	60
tatggggcct	tgccttccac	gttgacattt	acgtgcgccg	gggcttgctg	tctgctgtgt	120
cctctgtcct	cctcagtgtg	cccacagagc	ggctgctggg	ggacctgcca	gatgagctac	180
tggaagccag	atcctggttg	gcagatgtgg	ctgagaagga	tgtggacgag	gactgtaggg	240
agctggcagt	aagggtctct	ctgcttc				267

<210> 7

<211> 337

<212> DNA

<213> Artificial Sequence

<220>

<223> part of Mus musculus cDNA

<400> 7

tcacagttct	atggcctaaa	ctatagcctc	cgccagcgca	tggacatcct	ggacgtcctt	60
gttctggctg	cccaggcact	gtctcggcca	aagagcctgc	agagacgttc	ccagcacggg	120
ccccccgttc	ctggcaccat	gtgttcacca	gcactagccg	tttctcagac	tggcaatgtc	180
gctgctcctg	actggcaggt	ggttgtggag	gagcggatca	gaagcaagac	ccggagggtc	240
tcgaagggct	gtcctcagcg	ggacgtgtca	ggcgtcccca	atgaattcag	ctctgtggct	300
ggctacttct	tcttccccct	ccttcagcac	tttgaca			337

<210> 8

<211> 153

<212> PRT

<213> unknown

<220>

<223> part of Mus musculus clk-2 protein

<400> 8

Val	Thr	Arg	Lys	Leu	Leu	Leu	Leu	Gln	Tyr	Gln	His	Thr	Thr	Pro	Met
1				5				10						15	

- 9/40 -

Val	Gln	Ser	Leu	Leu	Gly	Tyr	Leu	Ala	Leu	Asp	Ser	Gln	Arg	Arg	Pro
			20					25					30		
Leu	Leu	Ile	Gln	Val	Leu	Lys	Glu	Leu	Leu	Glu	Thr	Trp	Gly	Cys	Ser
		35					40					45			
Ser	Ala	Val	Arg	His	Thr	Pro	Leu	Glu	Gln	Gln	Cys	Tyr	Ile	Ser	Lys
	50					55					60				
Ala	Ile	Leu	Val	Cys	Leu	Ala	His	Leu	Gly	Glu	Pro	Glu	Leu	Gln	Asp
65					70					75					80
Ile	Arg	Asp	Glu	Leu	Leu	Ala	Ser	Met	Met	Ala	Gly	Val	Lys	Cys	Arg
				85					90					95	
Leu	Asp	Ser	Ser	Leu	Pro	Pro	Val	Arg	Arg	Leu	Gly	Met	Ile	Val	Ala
			100					105					110		
Glu	Val	Ile	Ser	Ser	Arg	Ile	His	Pro	Glu	Gly	Pro	Leu	Leu	Lys	Phe
		115					120					125			
Gln	Tyr	Glu	Asp	Asp	Glu	Met	Ser	Arg	Glu	Leu	Leu	Ala	Leu	Ala	Thr
	130					135					140				
Pro	Glu	Pro	Ala	Gly	Asp	Cys	Ser	Ser							
145					150										

<210> 9
 <211> 112
 <212> PRT
 <213> unknown

<220>
 <223> part of Mus musculus clk-2 protein

Ser	Gln	Phe	Tyr	Gly	Leu	Asn	Tyr	Ser	Leu	Arg	Gln	Arg	Met	Asp	Ile
1				5					10					15	
Leu	Asp	Val	Leu	Val	Leu	Ala	Ala	Gln	Ala	Leu	Ser	Arg	Pro	Lys	Ser
		20						25					30		
Leu	Gln	Arg	Arg	Ser	Gln	His	Gly	Pro	Pro	Val	Pro	Gly	Thr	Met	Cys
		35					40					45			
Ser	Pro	Ala	Leu	Ala	Val	Ser	Gln	Thr	Gly	Asn	Val	Ala	Ala	Pro	Asp
	50					55					60				
Trp	Gln	Val	Val	Val	Glu	Glu	Arg	Ile	Arg	Ser	Lys	Thr	Arg	Arg	Phe
65					70				75						80
Ser	Lys	Gly	Cys	Pro	Gln	Arg	Asp	Val	Ser	Gly	Val	Pro	Asn	Glu	Phe
			85					90					95		
Ser	Ser	Val	Ala	Gly	Tyr	Phe	Phe	Phe	Pro	Leu	Leu	Gln	His	Phe	Asp
		100						105					110		

<210> 10
 <211> 85
 <212> PRT
 <213> unknown

<220>
 <223> part of Mus musculus clk-2 protein

Leu	Thr	Ser	Glu	Val	Ala	Val	Leu	Val	Gly	Lys	Ala	Leu	Leu	Glu	Phe
1				5					10					15	

- 10/40 -

Val	Trp	Ala	Leu	Arg	Phe	His	Val	Asp	Ile	Tyr	Val	Arg	Arg	Gly	Leu
			20					25					30		
Leu	Ser	Ala	Val	Ser	Ser	Val	Leu	Leu	Ser	Val	Pro	Thr	Glu	Arg	Leu
		35					40					45			
Leu	Gly	Asp	Leu	Pro	Asp	Glu	Leu	Leu	Glu	Ala	Arg	Ser	Trp	Leu	Ala
	50					55					60				
Asp	Val	Ala	Glu	Lys	Asp	Val	Asp	Glu	Asp	Cys	Arg	Glu	Leu	Ala	Val
65					70					75					80
Arg	Ala	Leu	Leu	Leu											
				85											

<210> 11

<211> 350

<212> PRT

<213> unknown

<220>

<223> composite protein sequence of Mus musculus clk-2

<400> 11

Val	Thr	Arg	Lys	Leu	Leu	Leu	Leu	Gln	Tyr	Gln	His	Thr	Thr	Pro	Met
1				5				10						15	
Val	Gln	Ser	Leu	Leu	Gly	Tyr	Leu	Ala	Leu	Asp	Ser	Gln	Arg	Arg	Pro
			20					25					30		
Leu	Leu	Ile	Gln	Val	Leu	Lys	Glu	Leu	Leu	Glu	Thr	Trp	Gly	Cys	Ser
		35					40					45			
Ser	Ala	Val	Arg	His	Thr	Pro	Leu	Glu	Gln	Gln	Cys	Tyr	Ile	Ser	Lys
	50					55					60				
Ala	Ile	Leu	Val	Cys	Leu	Ala	His	Leu	Gly	Glu	Pro	Glu	Leu	Gln	Asp
65					70					75					80
Ile	Arg	Asp	Glu	Leu	Leu	Ala	Ser	Met	Met	Ala	Gly	Val	Lys	Cys	Arg
				85				90					95		
Leu	Asp	Ser	Ser	Leu	Pro	Pro	Val	Arg	Arg	Leu	Gly	Met	Ile	Val	Ala
			100					105					110		
Glu	Val	Ile	Ser	Ser	Arg	Ile	His	Pro	Glu	Gly	Pro	Leu	Leu	Lys	Phe
		115					120					125			
Gln	Tyr	Glu	Asp	Asp	Glu	Met	Ser	Arg	Glu	Leu	Leu	Ala	Leu	Ala	Thr
	130					135						140			
Pro	Glu	Pro	Ala	Gly	Asp	Cys	Ser	Ser	Ser	Gln	Phe	Tyr	Gly	Leu	Asn
145					150					155					160
Tyr	Ser	Leu	Arg	Gln	Arg	Met	Asp	Ile	Leu	Asp	Val	Leu	Val	Leu	Ala
				165				170						175	
Ala	Gln	Ala	Leu	Ser	Arg	Pro	Lys	Ser	Leu	Gln	Arg	Arg	Ser	Gln	His
			180					185					190		
Gly	Pro	Pro	Val	Pro	Gly	Thr	Met	Cys	Ser	Pro	Ala	Leu	Ala	Val	Ser
		195					200					205			
Gln	Thr	Gly	Asn	Val	Ala	Ala	Pro	Asp	Trp	Gln	Val	Val	Val	Glu	Glu
	210					215					220				
Arg	Ile	Arg	Ser	Lys	Thr	Arg	Arg	Phe	Ser	Lys	Gly	Cys	Pro	Gln	Arg
225					230					235					240
Asp	Val	Ser	Gly	Val	Pro	Asn	Glu	Phe	Ser	Ser	Val	Ala	Gly	Tyr	Phe
				245				250						255	
Phe	Phe	Pro	Leu	Leu	Gln	His	Phe	Asp	Leu	Thr	Ser	Glu	Val	Ala	Val
			260					265					270		

- 11/40 -

Leu	Val	Gly	Lys	Ala	Leu	Leu	Glu	Phe	Val	Trp	Ala	Leu	Arg	Phe	His
		275					280					285			
Val	Asp	Ile	Tyr	Val	Arg	Arg	Gly	Leu	Leu	Ser	Ala	Val	Ser	Ser	Val
	290					295					300				
Leu	Leu	Ser	Val	Pro	Thr	Glu	Arg	Leu	Leu	Gly	Asp	Leu	Pro	Asp	Glu
305					310					315					320
Leu	Leu	Glu	Ala	Arg	Ser	Trp	Leu	Ala	Asp	Val	Ala	Glu	Lys	Asp	Val
				325					330					335	
Asp	Glu	Asp	Cys	Arg	Glu	Leu	Ala	Val	Arg	Ala	Leu	Leu	Leu		
			340					345						350	

<210> 12
 <211> 122
 <212> PRT
 <213> unknown

<220>
 <223> part of Sus Scrofa clk-2 protein

Lys	Ala	Pro	Val	Tyr	Val	Arg	Asp	Cys	Val	Glu	Ala	Leu	Thr	Ala	Ser
1				5					10					15	
Glu	Asp	Trp	Glu	Arg	Trp	Glu	Ala	Ala	Leu	Arg	Ala	Leu	Glu	Gly	Leu
			20				25						30		
Val	Phe	Arg	Ser	Pro	Ala	Ala	Ala	Arg	Glu	Val	Ser	Val	Glu	Leu	Ala
		35				40						45			
Lys	Val	Leu	Leu	His	Leu	Glu	Glu	Lys	Thr	Ala	Val	Ala	Gly	Phe	Glu
50					55					60					
Gly	Leu	Arg	Gln	Arg	Ala	Leu	Val	Ala	Val	Thr	Val	Thr	Asp	Pro	Ala
65					70				75					80	
Arg	Val	Ala	Glu	Tyr	Leu	Thr	Ala	Gln	Phe	Tyr	Ala	Leu	Asn	Tyr	Ser
				85				90					95		
Leu	Arg	Gln	Arg	Met	Asp	Ile	Leu	Asp	Val	Leu	Thr	Leu	Ala	Ala	Gln
			100					105					110		
Glu	Leu	Ser	Arg	Pro	Gly	Arg	Leu	Gly	Arg						
		115					120								

<210> 13
 <211> 554
 <212> PRT
 <213> D. melanogaster clk-2

Leu	Ile	Ser	Leu	Pro	Ala	Gln	Val	Ala	Asn	Arg	Leu	Gly	Arg	Arg	Leu
1				5					10					15	
Pro	Glu	Thr	Phe	Ala	Pro	Val	Ser	Tyr	Gln	Lys	Leu	Leu	Leu	Arg	Gln
			20					25					30		
Trp	Leu	Lys	Ser	Leu	His	Phe	Val	Leu	Gln	Cys	Asp	Asp	Asn	Arg	Glu
		35				40						45			
Tyr	Phe	Asp	Leu	Glu	Pro	Tyr	Ser	Trp	Leu	Leu	Ser	Gln	Ala	Ile	Asn
50					55					60					
Leu	Ile	Tyr	Asp	Val	Ser	Thr	Leu	Glu	Ser	Leu	Leu	Arg	Val	Leu	Lys
65					70				75					80	
Asp	Tyr	Ala	Val	Ala	Pro	Arg	Gly	Arg	Lys	Val	Val	His	Thr	Ile	Leu
				85					90					95	

- 12/40 -

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

- 13/40 -

Leu Gln Arg Trp Leu Asn Asp Phe Ile Arg Ser Pro Leu Val Gly Gly
 530 535 540
 Glu Thr Ser Glu Glu Cys Arg Glu Leu Ala
 545 550

<210> 14
 <211> .1017
 <212> PRT
 <213> unknown

<220>
 <223> A. thaliana clk-2 putative protein

<400> 14
 Met Ala Glu Gly Thr Lys Gln Glu Arg Thr Leu Glu Asn Asn Leu Leu
 1 5 10 15
 His Lys Val Gly Glu Ala Val Ser Ala Ile Ser Asp Ala Lys His Val
 20 25 30
 Asp Gln Val Ile Ser Ala Ile His Ser Val Ala Val Leu Leu Phe Pro
 35 40 45
 Val Asp Pro Ser Leu Phe Ser Gly Asn Phe Glu Met Leu His Ile Val
 50 55 60
 Arg Gly Ser Gly Thr Phe Gly Leu Leu Met Ile Leu Tyr Leu Gly Ser
 65 70 75 80
 Ala Gln Val Cys Ser Ser Val Val Pro Ser Ala Asp Glu Arg Asn Glu
 85 90 95
 Trp Leu Glu Thr Phe Tyr Arg Gly Val Ala Phe Pro Thr Phe Ala Arg
 100 105 110
 Val Leu Leu Leu Asp Val Ala Ser Asp Trp Leu Ser Cys Phe Pro Ile
 115 120 125
 Ser Val Gln Lys His Leu Tyr Asp Lys Phe Phe Leu Asp Gly Ser Val
 130 135 140
 Ile Glu Val Val Gln Val Leu Val Pro Phe Leu His His Val Gly Asp
 145 150 155 160
 Gly Gly Val Asn Ala Asn Ser Val Gln Thr Asn Val Glu Arg Leu Leu
 165 170 175
 Ile Leu Cys Leu Leu Glu Asn Asp Gly Val Leu Lys Ile Thr Lys Glu
 180 185 190
 Ile Gly Asn Ile Tyr Gln Gly His Asn Ser Ser Asn Gly Ser Leu Lys
 195 200 205
 Pro Leu Leu Ser Arg Leu Ser Gln Ile Leu Thr Ser Ile Pro Asp Lys
 210 215 220
 Ala Arg Ala Ser Cys Thr Glu Ala Asn Cys Thr Val Ile Val Leu Ser
 225 230 235 240
 Phe Val Gly Glu Val Phe Ser Arg Ile Cys Arg Arg Gly Leu Ser Asp
 245 250 255
 Leu Leu Leu Ser Glu Val Thr Pro His Val Leu Ala Gln Val Arg Arg
 260 265 270
 Leu Leu Asn Ser Lys Ile Gly Ala Ile Glu Val Asp Thr Phe Gln Leu
 275 280 285
 Asp Pro Thr Thr Arg Ile Trp Ser Lys Thr Met Glu Ala Val Thr Asp
 290 295 300
 Pro Tyr Ala Val Glu Lys Met Ala Glu Gln Leu Leu His Gln Leu Tyr
 305 310 315 320

- 14/40 -

Ala	Glu	His	Pro	Ser	Asp	Val	Glu	Ala	Phe	Trp	Thr	Ile	Trp	Thr	Leu
				325					330					335	
Phe	His	Arg	Asn	Val	Ile	His	Gln	Ala	Ser	Val	Arg	Gln	Ala	Lys	Cys
			340					345					350		
Phe	Leu	Trp	Gln	Leu	Asp	Ser	Phe	Phe	Arg	Tyr	Pro	Phe	Phe	Phe	Phe
		355					360					365			
His	Phe	His	Pro	Asn	Ala	Val	Lys	Gln	Cys	Val	Leu	Glu	Cys	Pro	Pro
	370					375					380				
Val	Thr	Asn	Thr	Leu	Ala	Lys	Gly	Asp	Val	Thr	Gln	Gly	Leu	Leu	Glu
385					390					395					400
Thr	Thr	Gln	Arg	Leu	Ala	Ser	Val	Trp	Ser	Lys	Arg	Glu	Phe	Leu	Gln
			405					410						415	
Ser	Val	Gln	Leu	Glu	Gln	Gln	Ala	Tyr	Leu	Gln	Phe	Leu	Phe	Pro	Val
			420					425					430		
Thr	Asp	Ile	Ser	Asp	Ile	Thr	Ala	Ala	Leu	Gly	Leu	Cys	Leu	Glu	Asn
	435					440						445			
Met	Ser	Arg	Glu	Glu	Leu	Asp	Arg	Thr	Lys	Asp	Val	Met	His	Ser	Ile
	450					455					460				
Leu	Gln	Gly	Val	Ser	Cys	Arg	Leu	Glu	Asn	Pro	Gly	Asp	Leu	Val	Arg
465					470					475					480
Lys	Met	Ala	Ser	Ser	Ile	Ala	Phe	Met	Phe	Ser	Lys	Val	Ile	Asp	Pro
				485					490					495	
Lys	Asn	Pro	Leu	Tyr	Leu	Asp	Asp	Ser	Ile	Thr	Asp	Asn	Ala	Ile	Asp
			500					505					510		
Trp	Glu	Phe	Gly	Leu	Gln	Thr	Ala	Ser	Ile	Thr	Asn	Thr	Met	Glu	Asn
	515						520						525		
Gly	Asp	Gly	Glu	Asn	Lys	Arg	Ser	Ala	Ser	Leu	Thr	Glu	Val	Asn	Glu
	530					535					540				
Ser	Ser	Arg	Arg	Asn	Lys	Gln	Lys	Glu	Asn	Arg	Lys	Ser	Lys	Asn	Ile
545					550					555					560
Ser	Ala	Phe	Val	Leu	Ala	Asp	Pro	Asn	Glu	Ile	Val	Asp	Leu	Ala	Thr
			565						570					575	
Leu	Asn	Cys	Asp	Thr	Glu	Ser	Asp	Lys	Asp	Asp	Gly	Asp	Asp	Asp	Ala
			580					585					590		
Ser	Val	Ser	Ser	Asp	Asn	Ser	Ser	Val	Thr	Ser	Leu	Glu	Pro	Tyr	Asp
	595						600					605			
Leu	Met	Asp	Asp	Asp	Lys	Asp	Leu	Gly	Lys	Gln	Phe	Thr	His	Leu	Val
	610					615					620				
Asp	Val	Val	Gly	Ala	Leu	Arg	Lys	Thr	Asp	Asp	Ala	Asp	Gly	Val	Glu
625					630					635					640
Lys	Ala	Ile	Tyr	Val	Ala	Glu	Lys	Leu	Val	Arg	Ala	Ser	Pro	Asp	Glu
				645					650					655	
Leu	Thr	His	Ile	Ala	Gly	Asp	Leu	Ala	Arg	Thr	Leu	Val	Gln	Val	Arg
			660					665					670		
Cys	Ser	Asp	Ile	Ala	Ile	Glu	Gly	Glu	Glu	Asp	Ser	Ala	Glu	Glu	Lys
	675						680					685			
Arg	Gln	Arg	Ala	Leu	Ile	Ala	Leu	Leu	Val	Thr	Arg	Pro	Phe	Glu	Ser
	690					695					700				
Leu	Glu	Thr	Leu	Asn	Asn	Ile	Leu	Tyr	Ser	Pro	Asn	Val	Asp	Val	Ser
705					710					715					720
Gln	Arg	Ile	Met	Ile	Leu	Asp	Val	Met	Ala	Glu	Ala	Ala	Arg	Glu	Leu
			725						730					735	
Ala	Asn	Ser	Lys	Thr	Leu	Lys	Pro	Lys	His	Glu	Ala	Arg	Gly	Pro	Leu
			740					745					750		

- 15/40 -

```

Ile Ser Asn Ile Ser Asp Pro Gln Pro Trp Tyr Leu Pro Ser Asn Ala
      755                      760                      765
Ser Thr Pro Trp Lys Lys Val Ser Glu Thr Gly Ser Phe His Leu Asn
      770                      775                      780
Trp Ala Asn Arg Phe Glu Arg Glu Leu Gln Ser Lys Pro Gly Gln Thr
785                      790                      795                      800
Lys Lys Gly Lys Ser Arg Arg Trp Ser Leu Lys Ser Ala Asp Arg Asp
      805                      810                      815
Gln Asn Ser Thr Asp Trp Ser Gln Asn Arg Phe Pro Leu Tyr Ala Ala
      820                      825                      830
Ala Phe Met Leu Pro Ala Met Lys Glu Phe Asp Lys Lys Arg His Gly
      835                      840                      845
Val Asp Leu Leu Gly Arg Asp Phe Val Val Leu Gly Lys Leu Val His
      850                      855                      860
Met Leu Gly Val Cys Met Gln Cys Ala Ser Met His Pro Glu Ala Ser
865                      870                      875                      880
Ala Leu Ala Ile Ser Leu Leu Asp Met Leu Gln Arg Arg Glu Val Cys
      885                      890                      895
Asn His Pro Glu Ala Tyr Val Arg Arg Ala Val Leu Phe Ala Ala Ser
      900                      905                      910
Ser Val Leu Val Ser Leu His Pro Ser Tyr Ile Val Ser Thr Leu Val
      915                      920                      925
Glu Gly Asn Leu Asp Leu Ser Arg Ala Leu Glu Trp Ile Arg Thr Trp
      930                      935                      940
Ala Leu Gln Ile Ala Asp Ser Asp Ile Asp Arg Asp Cys Tyr Thr Met
945                      950                      955                      960
Ala Leu Ser Cys Leu Gln Leu His Ala Glu Met Ala Leu Gln Thr Ser
      965                      970                      975
Arg Ala Leu Glu Ser Thr Gly Gly Ser Ser Ser Ser Ser Ser Ile Arg
      980                      985                      990
Pro Met Asn Ile Ser Leu Pro Ser Gly Ile Ser Lys Leu Thr Ser Ile
      995                      1000                      1005
Lys Leu Pro Ser Ser Asn Val His Leu
      1010                      1015

```

<210> 15

<211> 469

<212> DNA

<213> Artificial Sequence

<220>

<223> part of Oryza sativa clk-2 cDNA

<400> 15

```

gtttctactg ggtttggaat ggatacgcac atgggctctt ccatgttgca gaaacagatc      60
ctgatacaga atgcacatca atggctatga ccctccctgc ggctttcatt ctgagatggc      120
ccttcagaca tcacgagcac tggaatcggc ggatcacagc aaggccagca gcagcagcag      180
caggtcgcta ccttctaagc ttgataacat catcatacca ttgccaaca tgatgtgatc      240
gatgtactag attaagttgt aaataagcat atattatctt gccatgtaat attagaaatt      300
ccaggtcagc taattctgat taggcttgat attgtcatat tacagggttt tttaatctgg      360
gatttgaaaa tgctagaatt tttgtcgatt cgtgtggaga ttgtatggaa gttgtcctga      420
tccagattaa agagaattga tgaaaattgt acccaaaaaa aaaaaaaaaa      469

```

<210> 16

<211> 376

- 16/40 -

<212> DNA

<213> Artificial Sequence

<220>

<223> part of Oryza sativa clk-2 cDNA

<221> misc_feature

<222> (1)...(376)

<223> n = A,T,C or G

<400> 16

```

agttcctagt gatcagggac ctgctggtgc aggtctntgg agggaagtct cagaatcagg      60
aacacttctg aattgggtcac accggtatga aagagaagtt ccatctagat ctggtcaagt    120
taaatcagga aaatctcgta aatgggggtct tggaaaagct aaagatttgc agacagagtg    180
gtcaaaaaac agatttcctt tatatgctgc tgcttttatg ctccctgtta tgcaaggata    240
tgataaaaaga tcacatggtg ttgacttgct caatcgggac tttgttgtcc taggtaaatt    300
gatatacatg cttggtgtct gtatgaagtg catggcaatg catccagaag catcagctgt    360
tgccccagct cttctt                                     376

```

<210> 17

<211> 122

<212> PRT

<213> unknown

<220>

<223> part of Oryza sativa clk-2 protein

<400> 17

```

Pro Trp Arg Glu Val Ser Glu Ser Gly Thr Leu Leu Asn Trp Ser His
 1              5              10              15
Arg Tyr Glu Arg Glu Val Pro Ser Arg Ser Gly Gln Val Lys Ser Gly
      20              25              30
Lys Ser Arg Lys Trp Gly Leu Gly Lys Ala Lys Asp Leu Gln Thr Glu
      35              40              45
Trp Ser Lys Asn Arg Phe Pro Leu Tyr Ala Ala Ala Phe Met Leu Pro
      50              55              60
Val Met Gln Gly Tyr Asp Lys Arg Ser His Gly Val Asp Leu Leu Asn
      65              70              75              80
Arg Asp Phe Val Val Leu Gly Lys Leu Ile Tyr Met Leu Gly Val Cys
      85              90              95
Met Lys Cys Met Ala Met His Pro Glu Ala Ser Ala Val Ala Pro Ala
      100              105              110
Leu Leu Asp Met Ile Arg Ser Arg Ala Val
      115              120

```

<210> 18

<211> 70

<212> PRT

<213> unknown

<220>

<223> part of Oryza sativa clk-2 protein

<400> 18

- 17/40 -

```

Leu Glu Trp Ile Arg Thr Trp Ala Leu Pro Cys Cys Arg Asn Tyr Arg
 1           5           10           15
Met His Ile Asn Gly Tyr Asp Pro Pro Cys Gly Phe His Ser Glu Met
          20           25           30
Ala Leu Gln Thr Ser Arg Ala Leu Glu Ser Ala Asp His Ser Lys Ala
          35           40           45
Ser Ser Ser Ser Ser Arg Ser Leu Pro Ser Lys Leu Asp Asn Ile Ile
          50           55           60
Ile Pro Phe Ala Asn Met
.65           70

```

<210> 19
 <211> 192
 <212> PRT
 <213> unknown

<220>

<223> composite protein sequence of Oryza sativa clk-2 protein

<400> 19

```

Pro Trp Arg Glu Val Ser Glu Ser Gly Thr Leu Leu Asn Trp Ser His
 1           5           10           15
Arg Tyr Glu Arg Glu Val Pro Ser Arg Ser Gly Gln Val Lys Ser Gly
          20           25           30
Lys Ser Arg Lys Trp Gly Leu Gly Lys Ala Lys Asp Leu Gln Thr Glu
          35           40           45
Trp Ser Lys Asn Arg Phe Pro Leu Tyr Ala Ala Ala Phe Met Leu Pro
          50           55           60
Val Met Gln Gly Tyr Asp Lys Arg Ser His Gly Val Asp Leu Leu Asn
65           70           75           80
Arg Asp Phe Val Val Leu Gly Lys Leu Ile Tyr Met Leu Gly Val Cys
          85           90           95
Met Lys Cys Met Ala Met His Pro Glu Ala Ser Ala Val Ala Pro Ala
          100          105          110
Leu Leu Asp Met Ile Arg Ser Arg Ala Val Leu Glu Trp Ile Arg Thr
          115          120          125
Trp Ala Leu Pro Cys Cys Arg Asn Tyr Arg Met His Ile Asn Gly Tyr
          130          135          140
Asp Pro Pro Cys Gly Phe His Ser Glu Met Ala Leu Gln Thr Ser Arg
145           150           155           160
Ala Leu Glu Ser Ala Asp His Ser Lys Ala Ser Ser Ser Ser Ser Arg
          165           170           175
Ser Leu Pro Ser Lys Leu Asp Asn Ile Ile Ile Pro Phe Ala Asn Met
          180          185          190

```

<210> 20
 <211> 464
 <212> DNA
 <213> Artificial Sequence

<220>

<223> part of Glycine max clk-2 cDNA

<400> 20

- 18/40 -

```

tttttaggaat tgggtcaaata gctatgagag ggaacttccc ccaaaaccta atcaggtcaa      60
gaaagggaaa acacgccggt ggagcctaca atctcccaca caacaaaacc agatggagta      120
ttctcataat aagttacca tgtatgctgc tgcattcatg cttcctgcca tggagggata      180
tgataaaaaa aggcaaggtg ttgacttgct tggaagagat ttatttgtct tggggaaact      240
catttatatg cttgggggtct gtatgaaatc tgtagccatg catccagaag cttctatgct      300
ggctccttcc ctcctaaata tgttaagatc caggagggtg tgccatcacc aggaagcata      360
tgtgagaaga gccgtgcttt ttgcagctgc atgtgtattg gttgcccttc atcctactta      420
catttcatcc accttactcg aaggaaatgc tgaaatttcg actg                          464

```

<210> 21

<211> 521

<212> DNA

<213> Artificial Sequence

<220>

<223> part of Glycine max clk-2 cDNA

<400> 21

```

tccggagaag gactctgatt ccccttccaa taaagagaaa agtatttggt taaagggtaa      60
aaagaagtta ttggacttta atgcgcttga tccagatgag attattgatc cagcatcact      120
gaatcttgaa tcagacgata gcgatgagga tgctgacgat ggtgctagtg agaattcata      180
ttcttcaagt gattcatctt tacggccata tgatttgtca gatgatgact cagatttgaa      240
aagaaaaatt tcacagttgg ctgatgtagt tgcagctctt agaaaatcca atgatgccga      300
tgggggtggaa agggctattg atgtagctga aaagctcata agagcatccc ccgatgaact      360
aaaacatgca gcaagggata tgaccagaac tcttgttcag gttcgggtgct ctgatatagc      420
tttagaaggt gcagaagaat caactgaaga caaaagacaa agatcattag ttgccttagt      480
agttacctgc ccatttgaat cacttgagtc actaaacaac c                          521

```

<210> 22

<211> 741

<212> DNA

<213> Artificial Sequence

<220>

<223> part of Glycine max clk-2 cDNA

<221> misc_feature

<222> (1)...(741)

<223> n = A,T,C or G

<400> 22

```

atnnatnnnn nntagcagaa ataattagca gtttagcacg tttagaattt atattctggg      60
ctagaagcta ttcacaatag ctagaagcat cactaatccc catttttagta gggaatttta      120
atggtcactt tggatgcac agaaggaaga acaggacctg ccttgagtga acttcttact      180
gactccaatg ctcgggaagt ttgaagagcc atctcggcat ggagttgtat acatgtcata      240
gccatcgat agcattcttt atctgtgtct gactcggcta tgtcaagtgc ccatgtgcga      300
atccattcaa ggccagtcga aatttcagca ttctcttcga gtaagggtgga tgaaatgtaa      360
gtaggatgaa gggcaaccaa tacacatgca gctgcaaaaa gcacggctct tctcacatat      420
gcttcctggg gatggcatac ctccctggat cttaacatat ttaggaggga aggagccagc      480
atagaagctt ctggatgcat ggctacagat ttcatacaga cccaagcat ataaatgagt      540
ttccccaaga caataaaatc tcttccaagc aagtcaacac cttgcctttt tttatcatat      600
ccctccatgg caggaagcat gaatgcagca gcatacatgg gtaacttatt atgagaatac      660
tccatctggg tttgttgtgt gggnnnnngt nnnncnccacn nncgtgtttt ccctttcttg      720
acctgattng nttttgnnng a                          741

```

- 19/40 -

<210> 23
 <211> 456
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> part of Glycine max clk-2 cDNA

<400> 23
 ttgaaatttc aactggcctt gaatggattc gcacatgggc acttgatgta gccgagtcgg 60
 acacagataa agaatgctat acgatggcta tgacatgtat acagctccat gttgagatgg 120
 ctcttcaaac ttcccagagca ttggagtcag taagaaattc actcaaggca ggtcctgttc 180
 ttccttctga tgcattccaa gtgaccatta aaattcccca cttaaattggg gattagtgat 240
 gcttctagct attgtggatg gattctagcc cagaatacaa attctatatc atgctaaact 300
 gctaattatt tctgctactc tattaattta atgttttatg gggtccttgt tttcgtaaga 360
 aattgtatat tttaggattg tatatatctt cacaagggga cgaggttgat gtcaatgctc 420
 ccctacacac ccacccttga tgtagcacgt gttact 456

<210> 24
 <211> 455
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> part of Glycine max clk-2 cDNA

<400> 24
 aaactcgcgg cggaaggcag cagagtggcg ccgatggagg agggttttaga gaagagagaa 60
 ttggaaggcg aagtagttac caaggttgtc gaagtgggtt cggtataaaa gaatgcgaag 120
 cacgtcgatc aagtcattcg cgcgcttcat tccttagtca cccttctttt cccctttgac 180
 tcttcactcc tatcagatag cattgaccag agttaccgag accagggtcg agttccttct 240
 gcagaaaaac gacatgcttg gtggcgtgcg ttttatcgag gagctgcttt tcctacactg 300
 gctaggtttt tattacttga tgttgctcgc aactgggttg gttgttttcc ctttatggcg 360
 cagaagtaca tttatgatgt tttctttgtt cgtggattgg tcaactgacgt tctgcagatt 420
 ctggttcctt ttcttcagct gagtgcgagt gatgg 455

<210> 25
 <211> 208
 <212> PRT
 <213> unknown

<220>
 <223> part of Glycine max clk-2 protein

<221> VARIANT
 <222> (1)...(208)
 <223> Xaa = Any Amino Acid

<400> 25
 Lys Xaa Asn Gln Val Lys Lys Gly Lys Thr Arg Xaa Trp Xaa Xaa Xaa
 1 5 10 15
 Xaa Pro Thr Gln Gln Asn Gln Met Glu Tyr Ser His Asn Lys Leu Pro
 20 25 30
 Met Tyr Ala Ala Ala Phe Met Leu Pro Ala Met Glu Gly Tyr Asp Lys
 35 40 45

- 20/40 -

Lys Arg Gln Gly Val Asp Leu Leu Gly Arg Asp Phe Ile Val Leu Gly
 50 55 60
 Lys Leu Ile Tyr Met Leu Gly Val Cys Met Lys Ser Val Ala Met His
 65 70 75 80
 Pro Glu Ala Ser Met Leu Ala Pro Ser Leu Leu Asn Met Leu Arg Ser
 85 90 95
 Arg Glu Val Cys His His Gln Glu Ala Tyr Val Arg Arg Ala Val Leu
 100 105 110
 Phe Ala Ala Ala Cys Val Leu Val Ala Leu His Pro Thr Tyr Ile Ser
 115 120 125
 Ser Thr Leu Leu Glu Gly Asn Ala Glu Ile Ser Thr Gly Leu Glu Trp
 130 135 140
 Ile Arg Thr Trp Ala Leu Asp Ile Ala Glu Ser Asp Thr Asp Lys Glu
 145 150 155 160
 Cys Tyr Thr Met Ala Met Thr Cys Ile Gln Leu His Ala Glu Met Ala
 165 170 175
 Leu Gln Thr Ser Arg Ala Leu Glu Ser Val Arg Ser Ser Leu Lys Ala
 180 185 190
 Gly Pro Val Leu Pro Ser Asp Ala Ser Lys Val Thr Ile Lys Ile Pro
 195 200 205

<210> 26

<211> 151

<212> PRT

<213> unknown

<220>

<223> part of Glycine max clk-2 protein

<400> 26

Asn Trp Ser Asn Ser Tyr Glu Arg Glu Leu Pro Pro Lys Pro Asn Gln
 1 5 10 15
 Val Lys Lys Gly Lys Thr Arg Arg Trp Ser Leu Gln Ser Pro Thr Gln
 20 25 30
 Gln Asn Gln Met Glu Tyr Ser His Asn Lys Leu Pro Met Tyr Ala Ala
 35 40 45
 Ala Phe Met Leu Pro Ala Met Glu Gly Tyr Asp Lys Lys Arg Gln Gly
 50 55 60
 Val Asp Leu Leu Gly Arg Asp Phe Ile Val Leu Gly Lys Leu Ile Tyr
 65 70 75 80
 Met Leu Gly Val Cys Met Lys Ser Val Ala Met His Pro Glu Ala Ser
 85 90 95
 Met Leu Ala Pro Ser Leu Leu Asn Met Leu Arg Ser Arg Glu Val Cys
 100 105 110
 His His Gln Glu Ala Tyr Val Arg Arg Ala Val Leu Phe Ala Ala Ala
 115 120 125
 Cys Val Leu Val Ala Leu His Pro Thr Tyr Ile Ser Ser Thr Leu Leu
 130 135 140
 Glu Gly Asn Ala Glu Ile Ser
 145 150

<210> 27

<211> 172

<212> PRT

<213> unknown

- 21/40 -

<22.0>

<223> part of Glycine max clk-2 protein

<400> 27

Glu 1	Lys	Asp	Ser	Asp 5	Ser	Pro	Ser	Asn	Lys 10	Glu	Lys	Ser	Ile	Cys 15	Leu
Lys	Gly	Lys	Lys	Lys	Leu	Leu	Asp	Phe	Asn	Ala	Leu	Asp	Pro	Asp	Glu
			20					25					30		
Ile	Ile	Asp	Pro	Ala	Ser	Leu	Asn	Leu	Glu	Ser	Asp	Asp	Ser	Asp	Glu
		35					40					45			
Asp	Ala	Asp	Asp	Gly	Ala	Ser	Glu	Asn	Ser	Tyr	Ser	Ser	Ser	Asp	Ser
	50					55					60				
Ser	Leu	Arg	Pro	Tyr	Asp	Leu	Ser	Asp	Asp	Asp	Ser	Asp	Leu	Lys	Arg
65					70					75					80
Lys	Ile	Ser	Gln	Leu	Ala	Asp	Val	Val	Ala	Ala	Leu	Arg	Lys	Ser	Asn
				85					90					95	
Asp	Ala	Asp	Gly	Val	Glu	Arg	Ala	Ile	Asp	Val	Ala	Glu	Lys	Leu	Ile
			100					105					110		
Arg	Ala	Ser	Pro	Asp	Glu	Leu	Lys	His	Ala	Ala	Arg	Asp	Met	Thr	Arg
		115					120					125			
Thr	Leu	Val	Gln	Val	Arg	Cys	Ser	Asp	Ile	Ala	Leu	Glu	Gly	Ala	Glu
	130					135					140				
Glu	Ser	Thr	Glu	Asp	Lys	Arg	Gln	Arg	Ser	Leu	Val	Ala	Leu	Val	Val
145					150					155					160
Thr	Cys	Pro	Phe	Glu	Ser	Leu	Glu	Ser	Leu	Asn	Asn				
				165					170						

<210> 28

<211> 134

<212> PRT

<213> unknown

<220>

<223> part of Glycine max clk-2 protein

<400> 28

[illegible]

- 22/40 -

<210> 29
 <211> 75
 <212> PRT
 <213> unknown

<220>
 <223> part of Glycine max clk-2 protein

<400> 29
 Glu Ile Ser Thr Gly Leu Glu Trp Ile Arg Thr Trp Ala Leu Asp Val
 1 5 10 15
 Ala Glu Ser Asp Thr Asp Lys Glu Cys Tyr Thr Met Ala Met Thr Cys
 20 25 30
 Ile Gln Leu His Val Glu Met Ala Leu Gln Thr Ser Arg Ala Leu Glu
 35 40 45
 Ser Val Arg Asn Ser Leu Lys Ala Gly Pro Val Leu Pro Ser Asp Ala
 50 55 60
 Ser Lys Val Thr Ile Lys Ile Pro His Leu Asn
 65 70 75

<210> 30
 <211> 528
 <212> PRT
 <213> unknown

<220>
 <223> composite of Glycine max clk-2 protein

<400> 30
 Met Glu Glu Gly Leu Glu Lys Arg Glu Leu Glu Gly Glu Val Val Thr
 1 5 10 15
 Lys Val Val Glu Val Val Ser Ala Ile Lys Asn Ala Lys His Val Asp
 20 25 30
 Gln Val Ile Arg Ala Leu His Ser Leu Val Thr Leu Leu Phe Pro Phe
 35 40 45
 Asp Ser Ser Leu Leu Ser Asp Ser Ile Asp Gln Ser Tyr Arg Asp Gln
 50 55 60
 Val Glu Val Pro Ser Ala Glu Lys Arg His Ala Trp Trp Arg Ala Phe
 65 70 75 80
 Tyr Arg Gly Ala Ala Phe Pro Thr Leu Ala Arg Phe Leu Leu Leu Asp
 85 90 95
 Val Ala Ser Asn Trp Leu Gly Cys Phe Pro Phe Met Ala Gln Lys Tyr
 100 105 110
 Ile Tyr Asp Val Phe Phe Val Arg Gly Leu Val Thr Asp Val Leu Gln
 115 120 125
 Ile Leu Val Pro Phe Leu Glu Lys Asp Ser Asp Ser Pro Ser Asn Lys
 130 135 140
 Glu Lys Ser Ile Cys Leu Lys Gly Lys Lys Lys Leu Leu Asp Phe Asn
 145 150 155 160
 Ala Leu Asp Pro Asp Glu Ile Ile Asp Pro Ala Ser Leu Asn Leu Glu
 165 170 175
 Ser Asp Asp Ser Asp Glu Asp Ala Asp Asp Gly Ala Ser Glu Asn Ser
 180 185 190

- 23/40 -

```

Tyr Ser Ser Ser Asp Ser Ser Leu Arg Pro Tyr Asp Leu Ser Asp Asp
      195                200                205
Asp Ser Asp Leu Lys Arg Lys Ile Ser Gln Leu Ala Asp Val Val Ala
      210                215                220
Ala Leu Arg Lys Ser Asn Asp Ala Asp Gly Val Glu Arg Ala Ile Asp
      225                230                235                240
Val Ala Glu Lys Leu Ile Arg Ala Ser Pro Asp Glu Leu Lys His Ala
      245                250                255
Ala Arg Asp Met Thr Arg Thr Leu Val Gln Val Arg Cys Ser Asp Ile
      260                265                270
Ala Leu Glu Gly Ala Glu Glu Ser Thr Glu Asp Lys Arg Gln Arg Ser
      275                280                285
Leu Val Ala Leu Val Val Thr Cys Pro Phe Glu Ser Leu Glu Ser Leu
      290                295                300
Asn Asn Glu Ile Ser Thr Gly Leu Glu Trp Ile Arg Thr Trp Ala Leu
      305                310                315                320
Asp Val Ala Glu Ser Asp Thr Asp Lys Glu Cys Tyr Thr Met Ala Met
      325                330                335
Thr Cys Ile Gln Leu His Val Glu Met Ala Leu Gln Thr Ser Arg Ala
      340                345                350
Leu Glu Ser Val Arg Asn Ser Leu Lys Ala Gly Pro Val Leu Pro Ser
      355                360                365
Asp Ala Ser Lys Val Thr Ile Lys Ile Pro His Leu Asn Asn Trp Ser
      370                375                380
Asn Ser Tyr Glu Arg Glu Leu Pro Pro Lys Pro Asn Gln Val Lys Lys
      385                390                395                400
Gly Lys Thr Arg Arg Trp Ser Leu Gln Ser Pro Thr Gln Gln Asn Gln
      405                410                415
Met Glu Tyr Ser His Asn Lys Leu Pro Met Tyr Ala Ala Ala Phe Met
      420                425                430
Leu Pro Ala Met Glu Gly Tyr Asp Lys Lys Arg Gln Gly Val Asp Leu
      435                440                445
Leu Gly Arg Asp Phe Ile Val Leu Gly Lys Leu Ile Tyr Met Leu Gly
      450                455                460
Val Cys Met Lys Ser Val Ala Met His Pro Glu Ala Ser Met Leu Ala
      465                470                475                480
Pro Ser Leu Leu Asn Met Leu Arg Ser Arg Glu Val Cys His His Gln
      485                490                495
Glu Ala Tyr Val Arg Arg Ala Val Leu Phe Ala Ala Ala Cys Val Leu
      500                505                510
Val Ala Leu His Pro Thr Tyr Ile Ser Ser Thr Leu Leu Glu Gly Asn
      515                520                525

```

<210> 31
 <211> 876
 <212> PRT
 <213> unknown

<220>

<223> C. elegans clk-2 (QM37) mutant protein, with C to
 Y substitution at position 772

<400> 31

```

Met Asn Leu Arg Ser Arg Leu Val Asn Ala Thr Glu Arg Ala Val Leu
  1              5              10              15

```

- 24/40 -

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

- 25/40 -

Phe	Asn	Ser	Ser	Thr	Val	Lys	Asp	Ser	Glu	Arg	Val	Arg	Glu	Thr	Ala
450						455					460				
Asn	Glu	Ile	Thr	Lys	Asp	Asp	Ser	Gly	Glu	Glu	Phe	Glu	Glu	Thr	Asn
465					470					475					480
Ala	Gln	Arg	Leu	Gln	Asn	Asn	Lys	Asp	Ser	Ala	Ala	Ile	Thr	Ser	Lys
				485					490					495	
Asn	Asn	Leu	Arg	Leu	Asp	Ser	Asp	Asp	Asp	Glu	Asp	Phe	Pro	Asp	Tyr
		500					505						510		
Gln	Val	Asn	Glu	Ser	Glu	Lys	Ile	Phe	Lys	Asn	Leu	Glu	Ile	Gly	Glu
	515					520					525				
Glu	Pro	Lys	Asn	Lys	Val	Thr	Pro	Pro	Ala	Tyr	Ile	Ala	Asp	Ala	Phe
530						535					540				
Glu	Met	Leu	Leu	Glu	Lys	Glu	Lys	Tyr	Glu	Val	Phe	Glu	Ala	Ala	Phe
545					550					555					560
Phe	Asn	Ile	Thr	Asn	Leu	Ile	Asn	Arg	Arg	Pro	Ile	Gly	Phe	Pro	Gln
				565				570						575	
Ile	Ala	Glu	Lys	Leu	Phe	Ile	Arg	Ile	Leu	His	Leu	Gln	Asn	Asn	Phe
			580				585						590		
Gly	Thr	Pro	Lys	Phe	Lys	Glu	Thr	Val	Asp	Glu	Ile	Ala	Val	Ala	Cys
	595						600					605			
Ile	Thr	Gln	Arg	Pro	Glu	Ile	Val	Pro	Ser	Val	Val	Arg	Leu	Ile	Ile
610						615					620				
Ala	Pro	Gly	Gln	Gly	Phe	Ser	Ile	Lys	Gln	Arg	Leu	Leu	His	Tyr	Ile
625					630					635					640
His	Asn	Ala	Ala	Asp	Gly	Met	Gly	Ala	Leu	Asp	Lys	Lys	Leu	Glu	Glu
				645					650					655	
Cys	Val	Met	Ala	Gln	Gln	Leu	Arg	Ile	Gly	Gly	Pro	Thr	Leu	Ser	Ile
			660				665						670		
Ile	Leu	His	Arg	Thr	Ile	Asn	Thr	Asp	Tyr	Asp	Asp	Glu	Asp	Glu	Asp
	675					680						685			
Pro	His	Arg	Leu	Leu	Val	Pro	Glu	Trp	Arg	Arg	Met	Val	Asp	Ala	Arg
690						695					700				
Ile	Ala	Ala	Asn	Thr	Arg	Arg	Ile	Gly	Thr	Thr	Arg	Glu	Pro	Pro	Arg
705					710					715					720
Ala	Gly	Val	Val	Asn	Arg	Leu	Ala	Gln	Ala	Ala	Lys	Tyr	Met	Phe	Tyr
				725					730					735	
Pro	Leu	Leu	Val	Leu	Pro	Arg	Gly	Glu	Asn	Ala	Ser	Leu	Leu	Gly	Lys
			740					745					750		
Asp	Ser	Asp	Leu	Leu	Ala	Ser	Leu	Ile	Met	Val	Ala	Ser	Met	Val	Tyr
	755						760					765			
Val	Arg	Tyr	Gly	Val	Cys	Pro	Gln	Ile	His	Arg	Met	Ser	Ser	Glu	Leu
	770					775					780				
Ile	Ser	Tyr	Ala	Thr	Pro	His	Arg	Phe	Ser	Glu	Asn	Ala	Lys	Leu	Arg
785					790					795					800
Thr	Ala	Cys	Ile	Ile	Ala	His	Leu	Asn	Val	Thr	Thr	Leu	Leu	Pro	Gly
				805					810					815	
Asp	Leu	Met	Asp	Glu	Leu	Phe	Asp	Val	Pro	Ala	Leu	Ile	Gly	Trp	Phe
			820				825						830		
Asp	Trp	Ala	Asn	Ser	Val	Leu	Val	Asn	Ala	Ser	Ser	Ser	Gln	Leu	Glu
	835						840						845		
Lys	Asp	Met	Thr	Arg	Gln	Phe	Gly	His	Ser	Val	Thr	Lys	His	Leu	Gln
850						855					860				
Arg	His	His	Pro	Ala	Val	Leu	Gln	His	Gln	Asp	Val				
865					870					875					

- 26/40 -

<210> 32
 <211> 688
 <212> PRT
 <213> unknown

<220>
 <223> Tel2p, *S. cerevisiae* clk-2 protein

<400> 32
 Met Val Leu Glu Thr Leu Lys Gln Gly Leu Asp Ser Ser Gln Ile His
 1 5 10 15
 Glu Ala Leu Ile Gln Leu Asp Ser Tyr Pro Arg Glu Pro Val Asp Leu
 20 25 30
 Asp Ala Ser Met Val Leu Ile Lys Phe Val Ile Pro Val Tyr Pro Ser
 35 40 45
 Leu Pro Glu Arg Ser Lys Val Ile Leu Arg Arg Leu Ala Ser Lys Ser
 50 55 60
 Phe Thr Phe Leu Cys Gln Ile Val Thr Phe Ser Arg Thr Ile Ser Gly
 65 70 75 80
 Arg Asp Gly Leu Gln Glu Ile Arg Ile Tyr Gln Glu Ile Leu Glu Asp
 85 90 95
 Ile Ile Ser Phe Glu Pro Gly Cys Leu Thr Phe Tyr Leu Lys Ala Ser
 100 105 110
 Thr Thr Ser Lys Ala Asp Arg Asp Ser Ile Lys Ala Leu Phe Phe Gly
 115 120 125
 Ser Lys Leu Phe Asn Val Leu Ala Asn Arg Ile Asp Met Ala Lys Tyr
 130 135 140
 Leu Gly Tyr Leu Arg Leu Gln Trp Lys Phe Leu Leu Glu Ser Asn Glu
 145 150 155 160
 Thr Asp Pro Pro Gly Phe Leu Gly Glu Trp Leu Val Ser Ser Phe Leu
 165 170 175
 Leu Asn Pro Val Leu Ala Ala Asp Met Leu Leu Gly Glu Leu Phe Leu
 180 185 190
 Leu Lys Glu Ser Tyr Phe Phe Ser Phe Gln Lys Ile Ile Ser Ala Ser
 195 200 205
 Ser Leu Ile Asp Gln Lys Arg Leu Ile Ala Lys Phe Leu Leu Pro Tyr
 210 215 220
 Ile Gln Val Ile Val Thr Leu Glu Asn Leu Asn Asp Val Arg Lys Ile
 225 230 235 240

Leu Arg Arg Phe Asp Leu Asp Lys Ile Ile Ser Leu Ser Val Leu Phe
 245 250 255
 Glu Ile Gln Ser Leu Pro Leu Lys Glu Val Ile Val Arg Leu Met Ser
 260 265 270
 Asn His Ser Ser Thr Lys Phe Val Ser Ala Leu Val Ser Lys Phe Ala
 275 280 285
 Asp Phe Thr Asp Glu Glu Val Asp Thr Lys Thr Cys Glu Leu Leu Val
 290 295 300
 Leu Phe Ala Val His Asn Leu Asn His Ser Gln Arg Glu Glu Ile Ala
 305 310 315 320
 His Asp Glu Arg Phe Leu Asn Gly Val Thr Lys His Leu Gly Ser Asn
 325 330 335
 Glu Arg Glu Ala Arg Glu Arg Ala Met Phe Ile Ala Lys Leu Leu Ser
 340 345 350

- 27/40 -

Gly	Gly	His	Leu	Lys	Tyr	Glu	Ser	Asp	Phe	Lys	Ile	Asn	Ile	Pro	Asn		
		355					360					365					
Val	Lys	Phe	Glu	Ser	Asn	Ser	Asp	Asp	Lys	Ile	Ile	Asp	Phe	Gln	Ser		
	370					375					380						
Leu	Lys	Asn	Pro	Ser	Ile	Cys	Asn	Thr	Gln	Thr	Asp	Val	Gly	Lys	Asp		
385					390					395					400		
Lys	Ile	Thr	Glu	Val	Ser	Gly	His	Val	Gln	Ser	Leu	Thr	Leu	Asp	Cys		
			405						410					415			
Ser	Asp	Ser	Asp	Asp	Glu	Asp	Glu	Asn	Asp	Glu	Arg	Glu	Ile	Val	Lys		
			420					425					430				
Arg	Ile	Val	Phe	Leu	Lys	Asp	Leu	Met	Lys	Glu	Tyr	Glu	Lys	Thr	Gly		
	435					440						445					
Glu	Ser	Arg	Lys	Ala	Pro	Leu	Ile	Pro	Leu	Leu	Lys	Gln	Thr	Val	Lys		
	450					455					460						
Leu	Ile	Arg	Gln	Lys	Ala	Asp	Phe	Gln	Leu	Glu	Val	Gly	Tyr	Tyr	Ala		
465					470				475						480		
Gln	Gly	Ile	Leu	Ser	Ser	Ile	Val	Cys	Leu	Asn	Asn	Glu	Phe	Asp	Glu		
			485					490						495			
Pro	Leu	Phe	Glu	Gln	Trp	Arg	Ile	Asn	Ala	Leu	Thr	Ser	Ile	Leu	Val		
		500						505					510				
Val	Leu	Pro	Glu	Lys	Val	Asn	Gly	Ala	Ile	Asn	Ile	Leu	Phe	Asn	Ser		
	515					520						525					
Glu	Leu	Ser	Leu	Gln	Gln	Arg	Met	Ser	Leu	Leu	Ser	Ala	Leu	Gly	Leu		
	530					535					540						
Ser	Ala	Arg	Glu	Leu	Arg	Gly	Leu	Asp	Asp	Pro	Thr	Ile	Val	Lys	Pro		
545					550					555					560		
Lys	Phe	Asp	Phe	Pro	Thr	Asn	Arg	Leu	Pro	Trp	Asp	Asp	Gln	Ser	His		
			565					570						575			
His	Asn	Ser	Arg	Leu	Val	Glu	Val	Gln	Glu	Ser	Thr	Ser	Met	Ile	Lys		
		580						585					590				
Lys	Thr	Lys	Thr	Val	Trp	Lys	Ser	Arg	Lys	Leu	Gly	Lys	Asp	Arg	Glu		
	595					600						605					
Lys	Gly	Thr	Gln	Asn	Arg	Phe	Arg	Lys	Tyr	Ala	Gly	Leu	Phe	Phe	Tyr		
	610					615					620						
Pro	Leu	Ala	His	Gly	Trp	Leu	Asn	Gly	Ile	Asp	Val	Gly	Thr	Tyr	Asn		
625					630					635					640		
Gln	Leu	Phe	Lys	Ser	His	Tyr	Leu	Thr	Thr	Leu	Arg	Ile	Ile	Tyr	Ser		
			645						650					655			
Cys	Ala	Asn	Pro	Val	His	Asp	Phe	Glu	Ser	Met	Thr	Glu	Leu	Met	Asn		
		660						665					670				
His	Ile	Ile	Ser	Ser	Ala	Ile	Glu	Glu	Gly	Ile	Ser	Leu	Asn	Lys	Gly		
	675						680					685					

<210> 33

<211> 1520

<212> DNA

<213> C. elegans cex-7 cDNA

<400> 33

tcgaacttca	atztatatcg	atgtttcttc	tatttttagtg	caaatttttaa	tcaaaaaaat	60
taaacctttt	gtcaacatgg	cgcgaaattt	tcccactgct	ctcgattacc	tgggatcaga	120
agctgaagat	tttaacaagg	cgcaacactt	gtatctgaaa	ccgatggctg	ttataaaaat	180
taccgtagtg	ctgcctcgga	tgacgatccc	gggacaatca	atatctaatt	gggatctcat	240
ggaaagacta	aaacgtgcaa	ttgatccaat	tcaaatggat	tcttgcaaag	ttcgtgagag	300
caatatcgac	agtgttattt	ttgaagcgga	acttctttcg	ctaggaatca	tgcagaaaac	360

- 28/40 -

```

gatgaagatt ctcgatggat tctctatgaa agtgtctgga tttgctgagc ccttgaaagt 420
taaaacaaaa gaggcaaagc ttgattttcc aagtcgtcat gattgggatg attggtttat 480
gaaacataag atggacgaga tgaaaccccg agaacgtcct gatacagttt atttggcaag 540
aattccagtg aaatggttct gcgatgggta caatgatctt ctttctgaac gacgtcttcg 600
cgttgcaatg gaagcgttcg gatctgttag agttgtcgac attccaattt gtgacccact 660
ccgatctcga atgaattcaa aaatctctgg tattcagcaa aaaggtttcg gattagggca 720
ggatgtattht tttgaagcat atgtgcaatt tatggagtat aaaggtttcg ccactgccat 780
ggattccctg agaaatcgta aatgggcaaa acgaattgat ggacggttct ttcaggctaa 840
tgtcaagggtg gatttcgatc gatcacgtca tctcagcgaa gttcaaattg caaagcgagc 900
ggaagaacga cgtcaaattg aaacggaacg actgcggcaa gaagaagagg agttgaatat 960
cagacgtcaa gaagaactga aagttaaaca agaactcgat gataaagata gaagacgaga 1020
agatagagaa cggaaacgcc gcgaaaagag ggaactcgaa cgaatggctg aagaagagaa 1080
aaaacgcctt gaaaaggaac gacttgaagc agagcaacga tcacgagcca ctgccgtht 1140
gcaagggtgtt cgtcttttga agtttctgtt tgaaaaaatt gaagcacgag aagagagacg 1200
aaagagaaaa gaagaagaaa agttgaaaga tgagctgagc aaaatcaaag aacttgaaga 1260
acaacctgta gaacaagaag acgcattgag gcaagctcta cttcagcaga gggaaattcg 1320
aatgcgagaa cgcttgaaag agaagatgaa agcttcagga gcagagaagg ataagaagag 1380
agataaaaaat aagacttcta gaagaagacg aactaataga catgacactt catcctcttc 1440
atccgaagat tctgattctc catctgattc tccccactct cgtcgacaaa gaaaacgtca 1500
aagtgaagga gatgatttcc 1520

```

<210> 34

<211> 478

<212> PRT

<213> unknown

<220>

<223> C. elegans cex-7 protein

<400> 34

```

Met Ala Arg Asn Phe Pro Leu Tyr Leu Gly Ser Glu Ala Glu Asp Phe
 1           5           10          15
Asn Lys Ala Gln His Leu Tyr Leu Lys Pro Met Ala Val Ile Lys Ile
          20          25          30
Thr Val Val Leu Pro Arg Met Thr Ile Pro Gly Gln Ser Ile Ser Asn
          35          40          45
Trp Asp Leu Met Glu Arg Leu Lys Arg Ala Ile Asp Pro Ile Gln Met
          50          55          60
Asp Ser Cys Lys Val Arg Glu Ser Asn Ile Asp Ser Val Ile Phe Glu
          65          70          75          80
Ala Glu Leu Leu Ser Leu Gly Ile Met Gln Lys Thr Met Lys Ile Leu
          85          90          95
Asp Gly Phe Ser Met Lys Val Ser Gly Phe Ala Glu Pro Leu Lys Val
          100         105         110
Lys Thr Lys Glu Ala Lys Leu Asp Phe Pro Ser Arg His Asp Trp Asp
          115         120         125
Asp Trp Phe Met Lys His Lys Met Asp Glu Met Lys Pro Gly Glu Arg
          130         135         140
Pro Asp Thr Val Tyr Leu Ala Arg Ile Pro Val Lys Trp Phe Cys Asp
          145         150         155         160
Gly Tyr Asn Asp Leu Pro Ser Glu Arg Arg Leu Arg Val Ala Met Glu
          165         170         175
Ala Phe Gly Ser Val Arg Val Val Asp Ile Pro Ile Cys Asp Pro Leu
          180         185         190

```

- 29/40 -

Arg Ser Arg Met Asn Ser Lys Ile Ser Gly Ile Gln Gln Lys Gly Phe
 195 200 205
 Gly Leu Gly Gln Asp Val Phe Phe Glu Ala Tyr Val Gln Phe Met Glu
 210 215 220
 Tyr Lys Gly Phe Ala Thr Ala Met Asp Ser Leu Arg Asn Arg Lys Trp
 225 230 235 240
 Ala Lys Arg Ile Asp Gly Arg Phe Phe Gln Ala Asn Val Lys Val Asp
 245 250 255
 Phe Asp Arg Ser Arg His Leu Ser Glu Val Gln Ile Ala Lys Arg Ala
 260 265 270
 Glu Glu Arg Arg Gln Ile Glu Thr Glu Arg Leu Arg Gln Glu Glu Glu
 275 280 285
 Glu Leu Asn Ile Arg Arg Gln Glu Glu Leu Lys Val Lys Gln Glu Leu
 290 295 300
 Asp Asp Lys Asp Arg Arg Arg Glu Asp Arg Glu Arg Lys Arg Arg Glu

305 310 315 320
 Lys Arg Glu Leu Glu Arg Met Ala Glu Glu Glu Lys Lys Arg Leu Glu
 325 330 335
 Lys Glu Arg Leu Glu Ala Glu Gln Arg Ser Arg Ala Thr Arg Arg Leu

340 345 350
 Gln Gly Val Arg Leu Leu Lys Phe Leu Phe Glu Lys Ile Glu Ala Arg
 355 360 365
 Glu Glu Arg Arg Lys Arg Lys Glu Glu Glu Lys Leu Lys Asp Glu Leu
 370 375 380
 Ser Lys Ile Lys Glu Leu Glu Glu Gln Pro Val Glu Gln Glu Asp Ala
 385 390 395 400
 Leu Arg Gln Ala Leu Leu Gln Gln Arg Glu Ile Arg Met Arg Glu Arg
 405 410 415
 Leu Lys Glu Lys Met Lys Ala Ser Gly Ala Glu Lys Asp Lys Lys Arg
 420 425 430
 Asp Lys Asn Lys Thr Ser Arg Arg Arg Arg Thr Asn Arg His Asp Thr
 435 440 445
 Ser Ser Ser Ser Ser Glu Asp Ser Asp Ser Pro Ser Asp Ser Pro His
 450 455 460
 Ser Arg Arg Gln Arg Lys Arg Gln Ser Glu Gly Asp Asp Phe
 465 470 475

<210> 35

<211> 550

<212> PRT

<213> unknown

<220>

<223> XE7, Homo sapiens cex-7 protein

<400> 35

Met Ala Ala Ala Thr Ile Val His Asp Thr Ser Glu Ala Val Glu Leu
 1 5 10 15
 Cys Pro Ala Tyr Gly Leu Tyr Leu Lys Pro Ile Thr Lys Met Thr Ile
 20 25 30

- 30/40 -

Ser	Val	Ala	Leu	Pro	Gln	Leu	Lys	Gln	Pro	Gly	Lys	Ser	Ile	Ser	Asn
		35					40					45			
Trp	Glu	Val	Met	Glu	Arg	Leu	Lys	Gly	Met	Val	Gln	Asn	His	Gln	Phe
	50					55					60				
Ser	Thr	Leu	Arg	Ile	Ser	Lys	Ser	Thr	Met	Asp	Phe	Ile	Arg	Phe	Glu
65					70					75					80
Gly	Glu	Val	Glu	Asn	Lys	Ser	Leu	Val	Lys	Ser	Phe	Leu	Ala	Cys	Leu
				85					90					95	
Asp	Gly	Lys	Thr	Ile	Lys	Leu	Ser	Gly	Phe	Ser	Asp	Ile	Leu	Lys	Val
			100					105					110		
Arg	Ala	Ala	Glu	Phe	Lys	Ile	Asp	Phe	Pro	Thr	Arg	His	Asp	Trp	Asp
		115					120					125			
Ser	Phe	Phe	Arg	Asp	Ala	Lys	Asp	Met	Asn	Glu	Thr	Leu	Pro	Gly	Glu
	130					135					140				
Arg	Pro	Asp	Thr	Ile	His	Leu	Glu	Gly	Leu	Pro	Cys	Lys	Trp	Phe	Ala
145					150					155					160
Leu	Lys	Glu	Ser	Gly	Ser	Glu	Lys	Pro	Ser	Glu	Asp	Val	Leu	Val	Lys
				165					170					175	
Val	Phe	Glu	Lys	Phe	Gly	Glu	Ile	Arg	Asn	Val	Asp	Ile	Pro	Met	Leu
			180					185					190		
Asp	Pro	Tyr	Arg	Glu	Glu	Met	Thr	Gly	Arg	Asn	Phe	His	Thr	Phe	Ser
		195					200					205			
Phe	Gly	Gly	His	Leu	Asn	Phe	Glu	Ala	Tyr	Val	Gln	Tyr	Arg	Glu	Tyr
	210					215					220				
Met	Gly	Phe	Ile	Gln	Ala	Met	Ser	Ala	Leu	Arg	Gly	Met	Lys	Leu	Met
225					230					235					240
Tyr	Lys	Gly	Glu	Asp	Gly	Lys	Ala	Val	Ala	Cys	Asn	Ile	Lys	Val	Ser
				245					250					255	
Phe	Asp	Ser	Thr	Lys	His	Leu	Ser	Asp	Ala	Ser	Ile	Lys	Lys	Arg	Gln
			260					265					270		
Leu	Glu	Arg	Gln	Lys	Leu	Gln	Glu	Leu	Glu	Gln	Gln	Arg	Glu	Glu	Gln
		275					280					285			
Lys	Arg	Arg	Glu	Lys	Glu	Ala	Glu	Glu	Arg	Gln	Arg	Ala	Glu	Glu	Arg
	290					295					300				
Lys	Gln	Lys	Glu	Leu	Glu	Glu	Leu	Glu	Arg	Glu	Arg	Lys	Arg	Glu	Glu
305					310					315					320
Lys	Leu	Arg	Lys	Arg	Glu	Gln	Lys	Gln	Arg	Asp	Arg	Glu	Leu	Arg	Arg
				325					330					335	
Asn	Gln	Lys	Lys	Leu	Glu	Lys	Leu	Gln	Ala	Glu	Glu	Gln	Lys	Gln	Leu
			340					345					350		
Gln	Glu	Lys	Ile	Lys	Leu	Glu	Glu	Arg	Lys	Leu	Leu	Leu	Ala	Gln	Arg
		355					360					365			
Asn	Leu	Gln	Ser	Ile	Arg	Leu	Ile	Ala	Glu	Leu	Leu	Ser	Arg	Ala	Lys
	370					375					380				
Ala	Val	Lys	Leu	Arg	Glu	Gln	Glu	Gln	Lys	Glu	Glu	Lys	Leu	Arg	Leu
385					390					395					400
Gln	Gln	Gln	Glu	Glu	Arg	Arg	Arg	Leu	Gln	Glu	Ala	Glu	Leu	Arg	Arg
				405					410					415	
Val	Glu	Glu	Glu	Lys	Glu	Arg	Ala	Leu	Gly	Leu	Gln	Arg	Lys	Glu	Arg
			420					425					430		
Glu	Leu	Arg	Glu	Arg	Leu	Leu	Ser	Ile	Leu	Gln	Ser	Lys	Lys	Pro	Asp
		435				440					445				
Asp	Ser	His	Thr	His	Asp	Glu	Leu	Gly	Val	Ala	His	Gly	Pro	Ala	Ala
	450					455					460				
Ala	Arg	Pro	Gly	His	Pro	Ala	Asp	Arg	Val	Val	Arg	Leu	Cys	Glu	Arg

- 31/40 -

```

465          470          475          480
His His Ala Ala Pro Pro Arg Gly Pro Ala Pro Gly Arg Cys Pro Gln
          485          490          495
Gly Glu Pro Gly Pro Pro Arg Gly Arg Arg Arg Ser Gln Lys Arg Glu
          500          505          510
Arg Glu Arg Gly Arg Gly Gly Pro Met Gln Gly Gly Ser Glu Leu Leu
          515          520          525
Ser Cys Gly Pro Arg Gly Trp Leu Ser Arg Glu Glu Val Pro Gly Arg
          530          535          540
Arg Pro Leu Leu His Ser
545          550

```

<210> 36
 <211> 696
 <212> DNA
 <213> C. elegans coq-4 cDNA

<220>
 <223> clk

```

<400> 36
atgtctgctc aaaagctata cgcacgcgcg gttcccttgg caccgctgtc tcggatgctt      60
ctcgggattg gatcagcagt aacagcgatc tcggatccaa aaagaggaga tatggtggca      120
gcgatgggcg aaactactgc aattggacca gtttttagaaa atattcgaaa aagaatggaa      180
tctgatgttg ttggaaagcg acttcttctc gaaaaaccaa gaatttcaaa tggaacaatt      240
gatagaaagt ggctaagaca gttaccggat ggaacattag gaaaattgta ttcaaacttt      300
ctcgatcggt tgaacacatc tccagatgct cggcccactg tcaagtatat cgataatttg      360
gagcatcttt atgttatgca aaggtatcgc gaaacacacg acttcacca catcgcatg      420
gagcagaaaa cgaacatgct cggcgaggta acagtcaaat atttcgaagg aattcaatat      480
gggcttccaa tgtgtgtcac tgggtggaata tttggagggtg ccaggctttt aacaaaaaat      540
cgccaagaac ttgtcgaccg gaacctccct tgggttgtgg agcaggccac gaatgcacga      600
ttcttcatgg ctttcgactg ggaaaatcac tttgaaaagc agctcagcga ggtgcagaag      660
gagctaaata taacgccatt atctgtgaat atgtga      696

```

<210> 37
 <211> 263
 <212> PRT
 <213> unknown

<220>
 <223> C. elegans COQ-4 protein

```

<400> 37
Met Ser Ala Gln Lys Leu Tyr Ala Ser His Val Pro Leu Ala Pro Leu
  1          5          10          15
Ser Arg Met Leu Leu Gly Ile Gly Ser Ala Val Thr Ala Ile Ser Asp
          20          25          30
Pro Lys Arg Gly Asp Met Val Ala Ala Met Gly Glu Thr Thr Ala Ile
          35          40          45
Gly Pro Val Leu Glu Asn Ile Arg Lys Arg Met Glu Ser Asp Val Val
          50          55          60
Gly Lys Arg Leu Leu Leu Glu Lys Pro Arg Ile Ser Asn Gly Thr Ile
65          70          75          80
Asp Arg Lys Trp Leu Arg Gln Leu Pro Asp Gly Thr Leu Gly Lys Leu
          85          90          95

```

- 32/40 -

```

Tyr Ser Asn Phe Leu Asp Arg Leu Asn Thr Ser Pro Asp Ala Arg Pro
      100                      105                      110
Thr Val Lys Tyr Ile Asp Asn Leu Glu His Leu Tyr Val Met Gln Arg
      115                      120                      125
Tyr Arg Glu Thr His Asp Phe Thr His Ile Ala Leu Glu Gln Lys Thr
      130                      135                      140
Asn Met Leu Gly Glu Val Thr Val Lys Tyr Phe Glu Gly Ile Gln Tyr
145                      150                      155                      160
Gly Leu Pro Met Cys Val Thr Gly Gly Ile Phe Gly Gly Ala Arg Leu
      165                      170                      175
Leu Thr Lys Asn Arg Gln Glu Leu Val Asp Arg Asn Leu Pro Trp Val
      180                      185                      190
Val Glu Gln Ala Thr Asn Ala Arg Phe Phe Met Ala Phe Asp Trp Glu
      195                      200                      205
Asn His Phe Glu Lys Gln Leu Ser Glu Val Gln Lys Glu Leu Asn Ile
      210                      215                      220
Thr Pro Leu Ser Glu Leu Leu Asp Leu Pro Glu Met Glu Pro Asp Val
225                      230                      235                      240
Pro Asp Ile Leu Phe Ser Lys Gly His Pro Gly Phe Trp Arg Val Leu
      245                      250                      255
Gln Ala Val Asp Met Met Ile
      260

```

<210> 38

<211> 268

<212> PRT

<213> D. melanogaster COQ-4

<400> 38

```

Met Met Gln Arg Cys Leu Arg Leu Gln Lys Pro Leu Ala Leu Arg Arg
 1                      5                      10                      15
Gly Leu Arg Leu Ala Gln Ala Asn Ser Gln Ala Val Ala Thr Glu Ala
      20                      25                      30
Pro Glu Ala Glu Pro Leu Asp Ala Phe Glu Arg Gln Tyr Leu Lys Glu
      35                      40                      45
Arg Ile Glu Ile Ser Pro Phe Gln Arg Leu Phe Leu Gly Ala Gly Ser
      50                      55                      60
Ser Ile Ala Ala Leu Leu Asn Pro Arg Arg His Asp Met Ile Ala Cys
65                      70                      75                      80
Leu Gly Glu Thr Thr Gly Glu Asp Ala Leu Trp Thr Ile Leu Asp Thr
      85                      90                      95
Met Gln Ala Ser Glu Glu Gly Gln Arg Ile Met Ala Asp Lys Pro Arg
      100                      105                      110
Ile His Thr Ser Thr Ile Asp Phe Lys Tyr Leu Glu Thr Leu Pro Pro
      115                      120                      125
Asp Thr Phe Gly Ala Ala Tyr Val Lys Phe Leu Lys Asp Asn Gln Val
      130                      135                      140
Thr Pro Asp Ser Arg Met Ala Val Arg Phe Leu Glu Asp Pro Lys Leu
145                      150                      155                      160
Ala Tyr Leu Met Thr Arg Tyr Arg Glu Cys His Asp Leu Ile His Thr
      165                      170                      175
Val Leu Asp Met Pro Thr Asn Met Leu Gly Glu Val Ala Val Lys Trp
      180                      185                      190
Val Glu Ala Leu Asn Thr Gly Leu Pro Met Cys Tyr Gly Gly Ala Val
      195                      200                      205

```

- 33/40 -

Phe	Gly	Ala	Val	Arg	Leu	Arg	Pro	Lys	Gln	Arg	Arg	Ala	Tyr	Leu	Lys
210						215					220				
His	Tyr	Leu	Pro	Trp	Ala	Leu	Glu	Asn	Gly	Lys	Arg	Thr	Lys	Pro	Leu
225					230					235					240
Met	Pro	Val	Tyr	Trp	Glu	Lys	Arg	Trp	Glu	Gln	Asn	Ile	His	Glu	Leu
				245					250					255	
Arg	Ser	Glu	Leu	Gly	Ile	Thr	Val	Leu	Asn	Lys	Ala				
			260					265							

<210> 39

<211> 265

<212> PRT

<213> Homo Sapiens COQ-4

<400> 39

Met	Ala	Thr	Leu	Leu	Arg	Pro	Val	Leu	Arg	Arg	Leu	Cys	Gly	Leu	Pro
1				5					10					15	
Gly	Leu	Gln	Arg	Pro	Ala	Ala	Glu	Met	Pro	Leu	Arg	Ala	Arg	Ser	Asp
			20					25					30		
Gly	Ala	Gly	Pro	Leu	Tyr	Ser	His	His	Leu	Pro	Thr	Ser	Pro	Leu	Gln
		35					40					45			
Lys	Ala	Leu	Leu	Ala	Ala	Gly	Ser	Ala	Ala	Met	Ala	Leu	Tyr	Asn	Pro
	50					55				60					
Tyr	Arg	His	Asp	Met	Val	Ala	Val	Leu	Gly	Glu	Thr	Thr	Gly	His	Arg
65					70					75					80
Thr	Leu	Lys	Val	Leu	Arg	Asp	Gln	Met	Arg	Arg	Asp	Pro	Glu	Gly	Ala
				85					90					95	
Gln	Ile	Leu	Gln	Glu	Arg	Pro	Arg	Ile	Ser	Thr	Ser	Thr	Leu	Asp	Leu
			100					105					110		
Gly	Lys	Leu	Gln	Ser	Leu	Pro	Glu	Gly	Ser	Leu	Gly	Arg	Glu	Tyr	Leu
		115					120					125			
Arg	Phe	Leu	Asp	Val	Asn	Arg	Val	Ser	Pro	Asp	Thr	Arg	Ala	Pro	Thr
	130					135					140				
Arg	Phe	Val	Asp	Asp	Glu	Glu	Leu	Ala	Tyr	Val	Ile	Gln	Arg	Tyr	Arg
145					150					155					160
Glu	Val	His	Asp	Met	Leu	His	Thr	Leu	Leu	Gly	Met	Pro	Thr	Asn	Ile
				165					170					175	
Leu	Gly	Glu	Ile	Val	Val	Lys	Trp	Phe	Glu	Ala	Val	Gln	Thr	Gly	Leu
			180					185					190		
Pro	Met	Cys	Ile	Leu	Gly	Ala	Phe	Phe	Gly	Pro	Ile	Arg	Leu	Gly	Ala
		195					200					205			
Gln	Ser	Leu	Gln	Val	Leu	Val	Ser	Glu	Leu	Ile	Pro	Trp	Ala	Val	Gln
	210					215					220				
Asn	Gly	Arg	Arg	Ala	His	Cys	Val	Leu	Asn	Leu	Tyr	Tyr	Glu	Arg	Arg
225					230					235					240
Trp	Glu	Gln	Ser	Leu	Arg	Ala	Leu	Arg	Glu	Glu	Leu	Gly	Ile	Thr	Ala
				245					250					255	
Pro	Pro	Met	His	Val	Gln	Gly	Leu	Ala							
			260					265							

<210> 40

<211> 272

<212> PRT

<213> S. cerevisiae COQ-4

- 34/40 -

<400> 40

```

Met Phe Tyr Leu Asn Ala His Leu Glu Ile Asn Lys Val Val Asp Val
 1          5          10          15
Val Met Ser Leu Ser Lys Lys Phe Leu Lys Pro Ser Val Ala Ser Asn
          20          25          30
Gln Leu Arg Leu Leu Phe Thr Ala Ala Glu Arg Lys Val Asn Tyr Pro
          35          40          45
Gly His Val Pro Leu Ser Pro Leu Gln Arg Ile Phe Leu Val Ala Gly
          50          55          60
Ser Ala Ile Met Gly Leu Lys Ala Pro Trp Arg Gly Gly Asp Met Ile
65          70          75          80,
Ser Val Leu Gly Asp Ala Ser Gly Gln Pro Phe Phe Leu His Arg Leu
          85          90          95
Leu Asn Lys Met Leu Val Asp Lys Thr Gly Arg Glu Ile Leu Lys Asp
          100          105          110
Lys Pro Arg Met Thr Ser Lys Ser Leu Asn Leu Pro Phe Leu Arg Thr
          115          120          125
Leu Pro Pro Asn Thr Leu Gly Lys Ile Tyr Val Asp Trp Ile Asp Lys
          130          135          140
Glu His Val Gly Pro Asp Thr Arg Ser Pro Thr Arg Phe Val Asp Asp
145          150          155          160
Pro Glu Glu Ala Tyr Val Met Gln Arg Tyr Arg Glu Ser His Asp Phe
          165          170          175
Tyr His Ala Ile Cys Asn Met Pro Thr Asn Ile Glu Gly Glu Leu Ala
          180          185          190
Ile Lys Trp Leu Glu Phe Val Asn Met Gly Leu Pro Val Gly Ala Leu
          195          200          205
Ser Ala Leu Phe Gly Pro Leu Arg Leu Asn Cys Glu Gln Ala Ser Arg
          210          215          220
Phe Arg Arg Val Tyr Ile Pro Trp Ser Ile Arg Asn Gly Leu Asn Ala
225          230          235          240
Lys Thr Leu Ile Asn Val Tyr Trp Glu Lys Glu Leu Thr Asn Asp Ile
          245          250          255
Glu Asp Val Arg Arg Arg Ile Arg Ile Glu Ala Ala Pro Pro Leu Val
          260          265          270

```

<210> 41

<211> 226

<212> PRT

<213> Arabidopsis thaliana COQ-4

<400> 41

```

Met Ile Ile Glu Arg Ala Arg Val Pro Leu Ser Arg Trp Gln Gln Ala
 1          5          10          15
Ala Val Ala Met Gly Ser Ala Leu Gly Ala Leu Val Asp Pro Arg Arg
          20          25          30
Ala Asp Leu Ile Ala Ala Leu Gly Glu Thr Thr Gly Lys Pro Ala Phe
          35          40          45
Glu Met Val Leu Glu Arg Met Lys Lys Ser Glu Glu Gly Arg Ala Ile
          50          55          60
Leu Leu Glu Arg Pro Arg Val Val Ser Glu Gln Val Gly His Ala Trp
65          70          75          80
Asp Leu Pro Glu Asn Thr Phe Gly Ala Ala Tyr Ala Lys Phe Met Gly
          85          90          95
Ser Arg Asn Phe Ser Pro Asp Asp Arg Pro Pro Val Arg Phe Met Glu

```

- 35/40 -

			100					105					110				
Thr	Asp	Glu	Leu	Ala	Tyr	Val	Ala	Thr	Arg	Ala	Arg	Glu	Val	His	Asp		
		115					120					125					
Leu	Trp	His	Thr	Leu	Phe	Gly	Leu	Pro	Thr	Asn	Leu	Ile	Gly	Glu	Ser		
	130					135					140						
Ser	Leu	Lys	Val	Ile	Glu	Phe	Glu	Gln	Met	Tyr	Leu	Pro	Met	Cys	Met		
145					150					155					160		
Leu	Ser	Val	Ile	Gly	Gly	Thr	Val	Arg	Phe	Asn	Glu	Lys	Gln	Arg	Ser		
			165					170						175			
Met	Phe	Leu	Lys	His	Tyr	Leu	Pro	Trp	Ala	Val	Arg	Ala	Gly	Arg	Gln		
		180						185					190				
Cys	Thr	Asp	Leu	Met	Cys	Val	Tyr	Tyr	Glu	Arg	His	Phe	Ser	Glu	Asp		
	195						200					205					
Leu	Glu	Gln	Val	Arg	Arg	Lys	Trp	Gly	Ile	Ile	Pro	Ala	Pro	Gln	His		
	210					215					220						
Pro	Lys																
225																	

<210> 42

<211> 167

<212> PRT

<213> unknown

<220>

<223> part of Mus musculus COQ-4 protein

<400> 42

Ala	Leu	Tyr	Asn	Pro	Tyr	Arg	His	Asp	Met	Val	Ala	Val	Leu	Gly	Glu		
1				5				10					15				
Thr	Thr	Gly	Cys	His	Thr	Leu	Lys	Phe	Leu	Arg	Asp	Gln	Met	Lys	Lys		
		20						25				30					
Asp	Pro	Glu	Gly	Ala	Gln	Ile	Leu	Gln	Glu	Arg	Pro	Arg	Ile	Ser	Leu		
	35					40					45						
Ser	Thr	Leu	Asp	Leu	Ser	Lys	Leu	Gln	Ser	Leu	Pro	Glu	Gly	Ser	Leu		
	50					55					60						
Gly	Arg	Glu	Tyr	Leu	Arg	Phe	Leu	Asp	Val	Asn	Lys	Val	Ser	Pro	Asp		
65				70				75						80			
Thr	Arg	Ala	Pro	Thr	Arg	Phe	Val	Asp	Asp	Glu	Glu	Leu	Ala	Tyr	Val		
			85					90						95			
Ile	Gln	Arg	Tyr	Arg	Glu	Val	His	Asp	Met	Leu	His	Thr	Leu	Leu	Gly		
		100						105					110				
Met	Pro	Thr	Asn	Met	Leu	Gly	Glu	Val	Val	Val	Lys	Trp	Phe	Glu	Ala		
	115					120						125					
Val	Gln	Thr	Gly	Leu	Pro	Met	Cys	Ile	Leu	Gly	Ala	Leu	Phe	Gly	Pro		
	130					135					140						
Ile	Arg	Leu	Arg	Thr	Gln	Ser	Leu	Gln	Val	Leu	Phe	Ser	Glu	Leu	Ile		
145					150					155					160		
Pro	Trp	Ala	Ile	Gln	Asn	Gly											
				165													

<210> 43

<211> 123

<212> PRT

<213> unknown

- 36/40 -

<220>

<223> part of Mus musculus COQ-4 protein

<221> VARIANT

<222> (1)...(123)

<223> Xaa = Any Amino Acid

<400> 43

Tyr	Pro	Asp	His	Ile	Pro	Thr	Thr	Pro	Leu	Gln	Lys	Met	Leu	Leu	Ala
1				5					10					15	
Ala	Gly	Ala	Ala	Gly	Met	Ala	Leu	Tyr	Asn	Pro	Tyr	Arg	His	Asp	Met
			20					25					30		
Val	Ala	Val	Leu	Gly	Glu	Thr	Thr	Gly	Cys	His	Thr	Leu	Lys	Phe	Leu
		35					40					45			
Arg	Asp	Gln	Met	Lys	Lys	Asp	Pro	Glu	Gly	Ala	Gln	Ile	Leu	Gln	Glu
	50					55					60				
Arg	Pro	Arg	Ile	Ser	Leu	Ser	Thr	Leu	Asp	Leu	Ser	Lys	Leu	Gln	Ser
65					70				75						80
Leu	Pro	Glu	Gly	Ser	Leu	Gly	Arg	Glu	Tyr	Leu	Arg	Phe	Leu	Asp	Val
				85				90					95		
Asn	Lys	Val	Ser	Pro	Asp	Thr	Arg	Ala	Pro	Thr	Arg	Phe	Val	Asp	Asp
				100				105					110		
Glu	Xaa	Leu	Ala	Tyr	Val	Asn	Gln	Lys	Tyr	Arg					
		115					120								

<210> 44

<211> 89

<212> PRT

<213> unknown

<220>

<223> part of Mus musculus COQ-4 protein

<400> 44

His	Leu	Pro	Thr	Ser	Tyr	Pro	Ser	Leu	Ser	Leu	Gly	Glu	Val	Val	Val
1				5					10					15	
Lys	Trp	Phe	Glu	Ala	Val	Gln	Thr	Gly	Leu	Pro	Met	Cys	Ile	Leu	Gly
			20					25					30		
Ala	Leu	Phe	Gly	Pro	Ile	Arg	Leu	Arg	Thr	Gln	Ser	Leu	Gln	Val	Leu
		35					40					45			
Phe	Ser	Glu	Leu	Ile	Pro	Trp	Ala	Ile	Gln	Asn	Gly	Arg	Arg	Ala	Thr
	50					55					60				
Cys	Val	Leu	Asn	Ile	Tyr	Tyr	Glu	Gln	Arg	Trp	Glu	Gln	Pro	Leu	Thr
65					70				75						80
Ala	Leu	Arg	Glu	Glu	Leu	Gly	Ile	Ser							
				85											

<210> 45

<211> 218

<212> PRT

<213> Unknown

<220>

<223> Mus musculus consensus COQ-4 proteein

- 37/40 -

<400> 45

```

Tyr Pro Asp His Ile Pro Thr Thr Pro Leu Gln Lys Met Leu Leu Ala
 1          5          10          15
Ala Gly Ala Ala Gly Met Ala Leu Tyr Asn Pro Tyr Arg His Asp Met
 20          25          30
Val Ala Val Leu Gly Glu Thr Thr Gly Cys His Thr Leu Lys Phe Leu
 35          40          45
Arg Asp Gln Met Lys Lys Asp Pro Glu Gly Ala Gln Ile Leu Gln Glu
 50          55          60
Arg Pro Arg Ile Ser Leu Ser Thr Leu Asp Leu Ser Lys Leu Gln Ser
 65          70          75          80
Leu Pro Glu Gly Ser Leu Gly Arg Glu Tyr Leu Arg Phe Leu Asp Val
 85          90          95
Asn Lys Val Ser Pro Asp Thr Arg Ala Pro Thr Arg Phe Val Asp Asp
 100         105         110
Glu Glu Leu Ala Tyr Val Ile Gln Arg Tyr Arg Glu Val His Asp Met
 115         120         125
Leu His Thr Leu Leu Gly Met Pro Thr Asn Met Leu Gly Glu Val Val
 130         135         140
Val Lys Trp Phe Glu Ala Val Gln Thr Gly Leu Pro Met Cys Ile Leu
 145         150         155         160
Gly Ala Leu Phe Gly Pro Ile Arg Leu Arg Thr Gln Ser Leu Gln Val
 165         170         175
Leu Phe Ser Glu Leu Ile Pro Trp Ala Ile Gln Asn Gly Arg Arg Ala
 180         185         190
Thr Cys Val Leu Asn Ile Tyr Tyr Glu Gln Arg Trp Glu Gln Pro Leu
 195         200         205
Thr Ala Leu Arg Glu Glu Leu Gly Ile Ser
 210         215

```

<210> 46

<211> 137

<212> PRT

<213> Glycine max COQ-4

<400> 46

```

Leu Pro Ala Asn Thr Phe Gly Ala Ala Tyr Ala Arg Phe Met Gly Ser
 1          5          10          15
Arg Asn Phe Ser Pro Asp Asp Arg Pro Pro Val Arg Phe Met Asp Thr
 20          25          30
Asp Glu Leu Ala Tyr Val Ala Met Arg Ala Arg Glu Val His Asp Phe
 35          40          45
Trp His Thr Leu Phe Asp Leu Pro Thr Asn Leu Ile Gly Glu Thr Ala
 50          55          60
Leu Lys Val Ile Glu Phe Glu Gln Met Gly Leu Pro Met Cys Leu Leu
 65          70          75          80
Ser Val Ile Gly Gly Thr Ala Arg Phe Ser Glu Lys Gln Arg Lys Leu
 85          90          95
Phe Tyr His His Tyr Phe Pro Trp Ala Ile His Ala Gly Met Pro Ser
 100         105         110
Thr Asp Leu Met Cys Val Tyr Tyr Glu Arg His Phe Asp Glu Asp Leu
 115         120         125
Glu Asp Val Arg Arg Lys Leu Gln Ile
 130         135

```

WO 01/98478

PCT/CA01/00913

- 38/40 -

<210> 47
 <211> 136
 <212> PRT
 <213> Bos taurus COQ-4

<400> 47
 Tyr Pro Glu His Ile Pro Thr Ser Val Leu Gln Lys Val Leu Leu Ala
 1 5 10 15
 Ala Gly Ser Ala Gly Met Ala Leu Tyr Asp Pro Tyr Arg His Asp Met
 20 25 30
 Val Ala Val Leu Gly Glu Thr Thr Gly Arg Arg Thr Leu Lys Val Leu
 35 40 45
 Arg Asp Gln Met Lys Arg Asp Pro Glu Gly Ala Gln Ile Leu Gln Glu
 50 55 60
 Arg Pro Arg Ile Ser Leu Ser Thr Leu Asp Met Gly Lys Leu Arg Ser
 65 70 75 80
 Leu Pro Glu Gly Ser Phe Gly Cys Ala Tyr Leu His Phe Leu Asp Val
 85 90 95
 Asn Arg Val Ser Pro Asp Thr Arg Ala Pro Thr Arg Phe Val Asp Asp
 100 105 110
 Glu Glu Leu Ala Tyr Val Ile Gln Arg Tyr Arg Glu Ile His Asp Met
 115 120 125
 Leu His Ala Leu Leu Gly Met Pro
 130 135

<210> 48
 <211> 120
 <212> PRT
 <213> Medicago trunculata COQ-4

<400> 48
 Asn Phe Ser Pro Asp Asp Arg Pro Pro Val Arg Phe Met Asp Thr Asp
 1 5 10 15
 Glu Leu Ala Tyr Val Ala Met Arg Ala Arg Glu Val His Asp Phe Trp
 20 25 30
 His Thr Leu Phe Asp Leu Pro Thr Asn Leu Ile Gly Glu Ser Ala Leu
 35 40 45
 Lys Val Ile Glu Phe Glu Gln Met His Leu Pro Met Cys Val Met Ser
 50 55 60
 Val Leu Gly Gly Thr Ala Arg Phe Ser Glu Lys Gln Arg Lys Leu Phe
 65 70 75 80
 Tyr Gln His Tyr Phe Pro Trp Ala Val Arg Ala Gly Thr Gln Cys Asn
 85 90 95
 Asp Leu Met Cys Val Tyr Tyr Glu Gln His Phe His Gln Asp Leu Glu
 100 105 110
 Asp Val Arg Arg Lys Leu Gly Ile
 115 120

<210> 49
 <211> 113
 <212> PRT
 <213> Ancylostoma caninum COQ-4

<400> 49
 Lys Phe Val Gln Asn Ser Glu His Leu Tyr Val Met Gln Arg Tyr Arg

- 39/40 -

1				5					10					15			
Glu	Thr	His	Asp	Phe	Asn	His	Val	Leu	Leu	Gln	Met	Pro	Thr	His	Met		
			20						25					30			
Leu	Gly	Glu	Val	Thr	Val	Lys	Tyr	Phe	Glu	Gly	Ile	Gln	Phe	Gly	Leu		
		35					40					45					
Pro	Met	Cys	Val	Thr	Ala	Gly	Leu	Phe	Gly	Ala	Ala	Arg	Leu	Arg	Lys		
	50					55					60						
Asn	His	Arg	His	Arg	Phe	Leu	Thr	Gln	His	Leu	Pro	Trp	Ile	Val	Glu		
65					70					75					80		
Gln	Ala	Thr	Lys	Gly	Arg	Phe	Phe	Met	Ala	Ile	Asp	Trp	Glu	Asn	His		
			85						90					95			
Trp	Glu	Glu	Thr	Ile	Pro	Ser	Leu	Gln	Glu	Gln	Phe	Gly	Ile	Thr	Pro		
			100					105					110				
Leu																	

<210> 50

<211> 113

<212> PRT

<213> Trypanosoma cruzi COQ-4

<220>

<221> VARIANT

<222> (1)...(113)

<223> Xaa = Any Amino Acid

<400> 50

Phe	Val	Ala	Ala	Thr	Thr	Arg	Ser	Ile	Trp	Asp	Pro	Val	Asn	Ala	Asp		
1				5					10					15			
Asp	Val	Ala	Ala	Val	Gly	Glu	Thr	Pro	Ala	Leu	Thr	Ala	Leu	Gly	His		
			20					25					30				
Met	Lys	Gln	Ser	Met	Met	Ser	Asp	Arg	Thr	Gly	Arg	Met	Ile	Leu	Arg		
		35				40					45						
Thr	Gln	Pro	Arg	Val	Thr	Asp	Glu	Thr	Leu	Glu	Phe	Ala	Ser	Arg	Gln		
	50				55					60							
Pro	Pro	Gly	Thr	Phe	Gly	His	Arg	Tyr	Ala	Gln	Phe	Met	Lys	Phe	Xaa		
65					70					75					80		
Xaa	Phe	Thr	Pro	Asn	Gly	Arg	Thr	Pro	Val	Ala	His	Ile	Ala	Asp	Pro		
			85					90					95				
Thr	Leu	Ala	Tyr	Val	Met	Gln	Arg	Gln	Arg	Glu	Thr	His	Asp	Phe	Leu		
			100					105					110				
His																	

<210> 51

<211> 99

<212> PRT

<213> Rattus rattus COQ-4

<400> 51

His	Pro	Tyr	Arg	His	Asp	Met	Leu	Pro	Val	Leu	Gly	Glu	Thr	Thr	Gly		
1				5					10					15			
Cys	His	Thr	Leu	Lys	Phe	Leu	Arg	Asp	Gln	Met	Lys	Lys	Asp	Pro	Glu		
			20					25					30				
Gly	Ala	Gln	Ile	Leu	Gln	Glu	Arg	Pro	Arg	Ile	Ser	Leu	Ser	Thr	Leu		

- 40/40 -

		35				40				45					
Asp	Leu	Ser	Lys	Leu	Gln	Ser	Leu	Pro	Glu	Gly	Ser	Leu	Gly	Arg	Glu
	50					55					60				
Tyr	Leu	Arg	Phe	Leu	Asn	Ala	Asn	Lys	Val	Ser	Pro	Asp	Thr	Arg	Ala
65					70					75					80
Pro	Thr	Arg	Phe	Val	Asp	Asp	Glu	Glu	Leu	Ala	Tyr	Val	Ile	Gln	Arg
				85					90					95	
Tyr	Arg	Glu													

<210> 52

<211> 122

<212> PRT

<213> Gossypium hirsutum COQ-4

<220>

<221> VARIANT

<222> (1)...(122)

<223> Xaa = Any Amino Acid

<400> 52

Ile	Lys	Leu	Ser	Pro	Trp	Gln	Gln	Ala	Ala	Val	Ala	Val	Gly	Ser	Ala
1				5					10					15	
Val	Gly	Ala	Leu	Leu	Asp	Pro	Arg	Arg	Ala	Asp	Leu	Ile	Ala	Ala	Leu
			20					25					30		
Gly	Glu	Thr	Thr	Gly	Lys	Pro	Ala	Phe	Glu	Arg	Val	Leu	Glu	Arg	Met
		35					40					45			
Arg	Arg	Ser	Pro	Glu	Gly	Lys	Thr	Xaa	Leu	Leu	Glu	Arg	Pro	Arg	Val
	50					55					60				
Ile	Ser	Ala	Asn	Val	Gly	His	Ala	Trp	Asp	Leu	Pro	Lys	Asn	Thr	Phe
65					70					75					80
Gly	Ala	Ala	Tyr	Ala	Arg	Phe	Leu	Gly	Ser	Xaa	Asn	Phe	Ser	Pro	Asp
				85					90					95	
Asp	Arg	Pro	Pro	Val	Arg	Phe	Met	Asp	Thr	Asp	Glu	Leu	Ala	Tyr	Val
			100					105					110		
Ala	Met	Arg	Ala	Arg	Glu	Val	His	Asp	Phe						
			115				120								

WHAT IS CLAIMED IS:

1. A *clk-2* gene which has a function at the level of cellular physiology involved in developmental rate, telomere length and longevity, wherein *clk-2* mutations cause a longer life, an altered cellular metabolism and/or an altered telomere length relative to the wild type, wherein *clk-2* overexpression leads to telomere shortening, and wherein *clk-2* gene has the identifying characteristics of nucleotide sequence set forth in SEQ ID NO:1.

2. Use of a *clk-2* gene to alter a function at the level of cellular physiology involved in the regulation of developmental rates, telomere length and longevity, wherein *clk-2* mutations cause a longer life, altered cellular metabolism and physiological rates and/or an altered telomere length relative to the wild type, wherein *clk-2* overexpression leads to telomere shortening, and wherein *clk-2* gene has the identifying characteristics of nucleotide sequences set forth in SEQ ID NO:1, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, or SEQ ID NO:24, or wherein said gene codes for a protein sequence as set forth in SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31 or SEQ ID NO:32.

- 40 -

3. The *clk-2* gene of claim 1, which codes for a CLK-2 protein having the amino acid sequence set forth in SEQ ID NO:2.

4. The use of a *clk-2* gene to alter function at the level of cellular physiology involved in the regulation of developmental rates, telomere length and/or longevity, wherein *clk-2* mutations cause a longer life, altered cellular metabolism and physiological rates and an altered telomere length relative to the wild type, wherein *clk-2* overexpression leads to telomere shortening, and wherein said gene codes for a protein having a sequence as set forth in SEQ ID NO:32.

5. A CLK-2 protein which has a function at the level of cellular physiology involved in the regulation of developmental rate, telomere length and longevity, wherein said CLK-2 protein is encoded by the gene of claim 1.

6. Use of a CLK-2 protein to alter a function at the level of cellular physiology involved in the regulation of developmental rate, telomere length and longevity, wherein *clk-2* overexpression leads to telomere shortening, and wherein said CLK-2 protein is encoded by a gene as defined in claim 2.

7. A mutant CLK-2 protein which has the amino acid sequence set forth in SEQ ID NO:31.

- 41 -

8. A CLK-2 protein which has the amino acid sequence set forth in SEQ ID NO:2.

9. Use of CLK-2 protein to alter a function at the level of cellular physiology involved in the regulation of developmental rates, telomere length and longevity, wherein said CLK-2 protein has the amino acid sequence as set forth in SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31 or SEQ ID NO:32.

10. A *clk-2* gene which has the nucleotide sequence set forth in SEQ ID NO:1.

11. A mouse which comprises a gene knockout of the murine *clk-2* gene homologue to a *clk-2* gene as defined in claim 2.

12. A method to increase the life span of multicellular organism which comprises altering the function of telomeres and/or regulating telomere length.

13. The method of claim 12, wherein said multicellular organism is a metazoan.

14. A method to increase the life span of multicellular organism which comprises altering the

- 42 -

mechanisms of sub-telomeric silencing and/or regulating telomere length.

15. The method of claim 14, wherein said multicellular organism is a metazoan.

16. The use of *clk-2* gene as defined in claim 1, 3 or 10 and homologues thereof, to manipulate the physiological rates and/or telomere biology, whereby life span of an organism is altered.

17. Use of *clk-2* gene as defined in claim 1, 3 or 10, or CLK-2 protein as defined in claim 5, 7 or 8 and homologues thereof, for screening drugs which decrease or increase the life span of a multicellular organism.

18. The use of claim 17, wherein said drug enhances or suppresses the expression of the *clk-2* gene or activity of the protein CLK-2, and homologues thereof.

19. Use of a compound for the manufacture of a medicament for increasing and/or decreasing physiological rates of tissues, organ, and/or whole organism of a host; wherein said compound is interfering with activity of CLK-2 protein of claim 5, 7 or 8, and homologues thereof.

20. Use of a compound which promotes tissue and/or organ specific reduction or increase of *clk-2* activity for the manufacture of a medicament for the treatment of pathological conditions causing increase of physiological rate of tissue and/or organ in an

individual, wherein said compound is interfering with activity of CLK-2 protein of claim 5, 7 or 8, and homologues thereof.

21. Use of a compound which promotes tissue and/or organ specific reduction or increase of *clk-2* activity for the manufacture of a medicament for the treatment of pathological conditions causing decrease of physiological rate of tissue and/or organ in an individual, wherein said compound is interfering with activity of CLK-2 protein as defined in claim 5, 7 or 8, and homologues thereof.

22. A *clk-2* co-expressed gene which comprises a *cex-7* gene having the nucleotide sequence set forth in SEQ ID NO:33 which codes for a CEX-7 protein having the amino acid sequence set forth in SEQ ID NO:34 wherein said gene is located in the *clk-2* operon and said *cex-7* gene is transcriptionally co-expressed with *clk-2* gene present in said operon.

23. A human homologue of *cex-7* gene of claim 22, wherein said gene codes for a protein having a sequence as set forth in SEQ ID NO:35.

24. Use of a human homologue of *cex-7* gene of claim 22 and homologues thereof, to alter a function at the level of cellular level physiology involved in the regulation of developmental rates and longevity wherein said gene codes for a protein having a sequence as set forth in SEQ ID NO:35.

- 44 -

25. A mouse which comprises a gene knock out of the murine *cex-7* gene homologue of the human gene as set forth in SEQ ID NO:35.

26. Use of a compound for the manufacture of a medicament for increasing and/or decreasing physiological rates of tissues, organ, and/or whole organism of a host; wherein said compound is interfering with activity of CEX-7 as defined in claim 22 or 23, and homologues thereof.

27. Use of a compound which promotes tissue and/or organ specific reduction or increase of *cex-7* activity for the manufacture of a medicament for the treatment of pathological conditions causing increase of physiological rate of tissue and/or organ in an individual, wherein said compound is interfering with activity of CEX-7 as defined in claim 22 or 23, and homologues thereof.

28. Use of a compound which promotes tissue and/or organ specific reduction or increase of *cex-7* activity for the manufacture of a medicament for the treatment of pathological conditions causing decrease of physiological rate of tissue and/or organ in an individual, wherein said compound is interfering with activity of CEX-7 as defined in claim 22 or 23, and homologues thereof.

29. A *coq-4* gene which has a function at the level of cellular physiology involved in the regulation of developmental rate and longevity, wherein *coq-4*

- 45 -

mutations cause altered cellular metabolism and physiological relative to the wild type, wherein *coq-4* gene has the identifying characteristics of nucleotide sequence set forth in SEQ ID NO:36.

30. A *coq-4* gene which has a function at the level of cellular physiology involved in the regulation of developmental rate and longevity, wherein *coq-4* mutations cause altered cellular metabolism and physiological relative to the wild type, wherein *coq-4* gene has the identifying characteristics of nucleotide sequence set forth in SEQ ID NO:36, wherein said gene codes for a protein having a sequence as set forth in SEQ ID NO:37.

31. Use of *coq-4* gene to alter a function at the level of cellular physiology involved in the regulation of developmental rates, wherein *coq-4* mutations cause an altered cellular metabolism and physiological rates relative to the wild type, wherein said gene codes for a protein having a sequence as set forth in SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51 or SEQ ID NO:52 and homologues thereof.

32. A COQ-4 protein which has a function at the level of cellular physiology involved in the regulation of developmental rate and longevity, wherein said COQ-4 protein is encoded by the gene of claim 29.

- 46 -

33. A mouse which comprises a gene knock out of the murine *coq-4* gene as set forth in SEQ ID NO:45.

34. Use of a compound for the manufacture of a medicament for increasing and/or decreasing physiological rates of tissues, organs and/or whole organism of a host; wherein said compound is interfering with activity of COQ-4 protein as defined in claim 32, and homologues thereof.

35. Use of a compound which promotes tissue and/or organ specific reduction or increase of *coq-4* activity for the manufacture of a medicament for the treatment of pathological conditions causing increase of physiological rate of tissue and/or organ in an individual, wherein said compound is interfering with activity of COQ-4 protein as defined in claim 32, and homologues thereof.

36. Use of a compound which promotes tissue and/or organ specific reduction or increase of *coq-4* activity for the manufacture of a medicament for the treatment of pathological conditions causing decrease of physiological rate of tissue and/or organ in an individual, wherein said compound is interfering with activity of COQ-4 protein as defined in claim 32, and homologues thereof.

37. Use of a compound which promotes tissue and/or organ specific reduction or increase of *clk-2* activity for the manufacture of a medicament for the treatment of pathological conditions due to altered telomere

length in tissue and/or organ in an individual, wherein said compound is interfering with activity of CLK-2 protein as defined in claim 5, 7 or 8, and homologues thereof.

1	CTCAAGATGA	ATTTACGAAG	TCGCCTGGTA	AATGCCACGG	AACGTGCTGT	GCTTTTTCAA
61	ATTTTCAAAG	ATGTGCAGAA	TGATCCGGAA	AAGTACGACA	ATGCAGTTGA	GGCGATCTGT
121	GAATCAATCG	ACTATTTTGG	CAAATTTTGG	ACCGATAGCG	AGTATCTTAC	ACAAATCAAG
181	CCGATTCTTG	ATACACAGTG	CCCAACAAAG	TCGATAAATT	GCTTCTCGAA	ATGTTTGACA
241	AAAGTGAGCA	CAGATATAAA	TACTACCACA	TTTCGAGATG	TGATCACCAT	GCTCGACTGG
301	TTGAAGTATG	TCGTTGAAAA	ATCGCTGACA	AGTGCTATTT	GTAGCAGTCT	GAAAGTTAAA
361	GAAACTGATG	TCAGTGCAGT	TCAGTTGTAT	CGAGAAATTCG	CATCAGCATG	TTCAAATATT
421	CCGGAGAAAG	TTTCGGAATTG	TTGTGCAAAAG	GCAATTGTCTG	GCGAGCATGT	CAAATATATC
481	AACACGGTTA	AGTGGATATT	CAAAATGAAT	CTGGTGCAAG	GAATTCAAAA	GGCTATGCTT
541	CTTGCTCACG	ACGACATTGT	AAC TGCTGCC	CCGTTCACTT	CATTCTACGG	ATCCGGTGGT
601	CCTTATATGA	AGACTGTTCG	AGAAATTATT	TCATCTGGAA	GAAACATAGA	TATCACCAAC
661	AAGGATGGGC	TTCTAGTTCA	AATGATTGAA	TGGATTGGTT	CACTAAACAA	TTTTGATTCT
721	CAATGGCGTC	GGATGATGTT	TCTCATCTTC	CAAGAGCCCA	CATATCAGGG	AATTCAAAGT
781	CATGAATCAC	TACTGACAA	ATTGTTCCCTA	ATTTCGAAAA	GTGACCAAAT	CTTGAAACGA
841	TGTATCGAAG	CCACTGATCT	GACTGGAACA	CTGAAGCGTG	TAGTGATGGT	TAAGCTCCCG
901	TTTCAGCGAG	TTCTCAAACG	AAAGACCATC	GAGATTCTTA	TCAAATTTGT	TTATCGAACT
961	AAGGAACAA	TTGCCATCCA	GCTATTAGAG	ACTTCTGTGA	AAATCTGGAG	TGATCTCAAT
1021	TACGCAAAA	GTGCTCCAGA	ATCACAAAGAA	CGACACATAG	TCAGAAATGAT	ATTATACTTG
1081	GTTCATCTTT	TCAGAACATG	TTCTTCAATC	GATTGGGAGT	CAC TCTTCCT	GAACTCTATG
1141	GATGGAGTTC	ATTGTCGAAT	GAGCATGTTG	CCTATGTACG	TCCAAAGTGG	TATTTTGTGT
1201	AATCAAGCAC	TGTGCAAGCA	AGCGACAAAG	CATCGATCGA	AAACGCACGG	ATCAGATGAG
1261	CAACCTCCAG	AGACTCTAGA	AGAAAAACAAA	TTCGTTTCAA	GTGAAGTGGG	AAAAATATGG
1321	TTTGAAGAGA	TGACGTCAAT	TTTGGAAACAT	GGATTTAATT	CTTCTACAGT	GAAAGATTCT

1 / 3 5

Fig. 1A

2 / 3 5

1381	GAGCGAGTTC	GAGAAACCGC	CAACGAAATA	ACCAAAGACG	ATTCGGGTGA	AGAAATTGAA
1441	GAAACCAATG	CACAGCGTCT	TCAAAACAAC	AAAGATTCCG	CAGCAATCAC	ATCGAAAAAC
1501	AATCTACGTT	TAGATTCTGA	TGATGACGAA	GACTTTCCTG	ACTATCAAGT	TAATGAATCA
1561	GAAAAGATCT	TCAAGAAATTT	AGAAATTGGA	GAAGAACCGA	AAAAATAAAGT	GACACCTCCA
1621	GCATACATTG	CAGATGCTTT	CGAAATGCTA	TTAGAGAAAG	AAAAATACGA	GGTTTTTGAA
1681	GCAGCTTTCT	TTAATAATTAC	GAATTTGATC	AATCGCCGGC	CAATTGGATT	TCCACAAATC
1741	GCTGAGAAGT	TGTTCATCCG	AATCCTCCAT	CTTCAAAACA	ATTTTGGAAC	GCCTAAATTC
1801	AAGGAAACTG	TTGATGAAAT	TGCAGTTGCA	TGTATCACTC	AGCGTCCAGA	AATTGTCCCA
1861	TCTGTAGTGC	GTCTGATCAT	TGCACCCAGGT	CAAGGTTTCA	GTATCAAAACA	ACGTCTTCTT
1921	CATTACATTC	ACAAATGCTGC	TGATGGAATG	GGTGCATTGG	ATAAGAAACT	TGAAGAGTGT
1981	GTAATGGCGC	AACAATTGAG	AATTGGTGGT	CCAACGTTAA	GTATTATTCT	TCATCGAACT
2041	ATAAACACTG	ATTATGACGA	TGAGGATGAA	GATCCCCACA	GACTTTTAGT	TCCTGAAATG
2101	CGACGAAATG	TGGATGCTCG	CATTGCTGCA	AATACCAGAA	GAATTGGAAC	GACGCGAGAG
2161	CCACCAAGAG	CCGGAGTTGT	CAATCGTCTC	GCACAAGCTG	CCAAATATAT	GTTTTATCCT
2221	TTGTTGGTTT	TGCCACGTGG	TGAGAACGCA	AGTCTTTTGG	GCAAGGACTC	CGATCTACTC
2281	GCCTCACTCA	TCATGGTTGC	ATCGATGGTT	TATGTGAGAT	GTGGCGTATG	TCCTCAAAAT
2341	CATCGAATGT	CAAGTGAGCT	TATATCATAT	GCAACGCCCTC	ATCGATTCTC	TGAAAAACGCA
2401	AAACTACGGA	CTGCCCTGCAT	CATTGCCCAT	CTGAATGTGA	CGACTTTGTT	GCCTGGAGAT
2461	CTTATGGATG	AAC TGTTTGA	TGTACCGGCA	CTTATTGGAT	GGTTTGATTG	GGCCAATTCT
2521	GTA CTGGTTA	ATGCATCTTC	ATCACAATTG	GAAAAGGATA	TGACTCGCCA	GTTTGGT CAT
2581	AGTGTTACAA	AACACCTTCA	ACGTTATCAT	CCAGCTGTAC	TGCAACACCA	AGACGTTTAA
2641	ATAGTTACTA	TTCACTTGTT	TTCCCTTCTTT	TCAAGTACTG	TATCATTTCT	TACTATCTTG
2701	CCAACACTTT	GATCTCTACC	TCGTTCACTT	CTTGCTTTGC	CACCCGTTGA	TATCACCTGT
2761	CTCATTCATT	TATCAGCATG	TTCATAATAT	CAAAAATAAA	ATCTTATCAA	ATGT

(SEQ ID NO:1)

Fig. 1B

3/35

MNLRSRLVNATERAVLFQIFKDVQNDPEKYDNAVEAICESIDYFGKFLTDSE
YLTQIKPILDTQCPTKSIICFSKCLTKVSTDINTTTFRDVITMLDWLKYVVE
KSLTSAICSSSLKVKETDVSAVQLYREFASACSNIPEKVSNCCAKALSGEHVK
YINTVKWIFKMNLVQGIQKAMLLAHDDIVTAAPFTSFYGS GGPYMKTVAEII
SSGRIDITNKDGLLVQMIEWIGSLNNFDSQWRRMMFLIFQEPTYQGIQVHES
LLTTLFLISKSDQILKRCIEATDLTGTLKRVMVKLPFQRVLKRKTIEILIN
FVYRTKEQFAIQLLET SVKIWSDLNYAKSAPESQERHIVRMILYLVHLFRTC
SSIDWESLFLNSMDGVHCRMSMLPMYVQSGIFVNQALCKQATKHR SKTHGSD
EQPPETLEENKFVSSEVGKIWFEEMTSILEHGFNSSTVKDSERVRETANEIT
KDDSGEEFEETNAQRLQNNKDSAAITSKNNLR LDSDDDEDFPDYQVNESEKI
FKNLEIGEEPKNKVTPPAYIADAFEM LLEKEKEYEVFEAAFFNITNLINRRPI
GFPQIAEKLFIRILHLQNNFGTPKFKETVDEIAVACITQRPEIVPSVVRLII
APGQGFSIKQRLLHYIHNAADGMGALDKKLEECVMAQQLRIGGPTLSIILHR
TINTDYDDEDEDPHRLLVPEWRRMVDARIAANTRRIGTTREPPRAGVVNRLA
QAAKYM FYPLLVLPRGENASLLGKDSDLLASLIMVASMVYVRCGVCPQIHRM
SSELISYATPHRFSENAKLRTACIIAHLNVTTLLPGDLMDELFDVPALIGWF
DWANSVLVNASSSQLEKDMTRQFGHSVTKHLQRHHPAVLQHQDV

(SEQ ID NO:2)

Fig. 2

4 / 35

MEPAPSEVRLAVREAIHALSSSEDGGHIFCTLESLKRYLGEMEPPALPREKE
EFASAHFSPVLRCLASRLSPAWLELLPHGRLEELWASFFLEGPADQAFLVLM
ETIEGAAGPSFRLMKMARLLARFLREGRLAVLMEAQCRQQTQPGFILLRETL
LGKVVALPDHLGNRLQQENLAEFFPQNYFRLLGEEVVRVLQAVVDSLQGGLD
SSVSFVSQVLGKACVHGRQQEILGVLVPRLAALTQGSYLHQRVCWRLVEQVP
DRAMEAVLTGLVEAALGPEVLSRLLGNLVVKNKKAQFVMTQKLLFLQSRLTT
PMLQSLLGHLAMDSQRRPLLLQVLKELLETTWGSSSAIRHTPLPQQRHVSKAV
LICLAQLGEPELRDSRDELLASMMAGVKCRLDSSLPPVRRLGMIVAEEVVSAR
IHPEGPPLKFQYEEDELSLELLALASPQPAGDGASEAGTSLVPATAEPPAET
PAEIVDGGVPQAQLAGSDSDLDSDDEFVPYDMSGDRELKSSKAPAYVRDCVE
ALTTSEDIERWEAALRALEGLVYRSPTATREVSVELAKVLLHLEEKTCVVGF
AGLRQRALVAVTVTDPAPVADYLT SQFYALNYS LRQRM DILDVLT LAAQELS
RPGCLGRTPQPGSPSPNTPCLPEAAVSQPGSAVASDWRVVVEERIRSKTQRL
SKGGPRQGPAGSPSRFNSVAGHFFFPLLQRFDRPLVTFDLLGEDQLVLGRLA
HTLGALMCLAVNTTVAVAMGKALLEFVWALRFHIDAYVRQGLLSAVSSVLLS
LPAALLEDLMDELLEARSWLADVAEKDPDEDCRTLALRALLLLQRLKNRLLP
PASP

(SEQ. ID NO:3)

Fig. 3

1
61
121
181
241
301
361
421
481
541
601
661
721
781
841
901
961
1021
1081
1141
1201
1261
1321
1381
1441
1501
1561
1621

GGGCCCCAGA
GGGGCCAACC
CGGCTGCAGC
GTCCAGGTCG
GAGGTTCCGAC
CACATCTTCT
CTCCCCGAGG
GCCAGCAGGC
TGGGCCAGCT
ATCGAGGGTG
TTCCTGCGCG
CCCGGCTTCA
CTGGGCAACC
CTGCTCGGCG
CTGGATTCCCT
CAGCAGGAGA
CTGCACCAGC
GTGCTGACCG
AACCTGGTGG
CAGTCCCCGC
CAGCGGCGCC
AGTGCCCATCC
TGCCCTGGCG
ATGATGGCGG
ATGATCGTGG
CAGTACGGAAG
GGTGACGGCG
GAGACCCCCG

GGCTCAAGAA
GGCTGGGCCG
GAATCGGTGG
TCTTCCCCGTG
TCGCCGTCCG
GCACCCCTGGA
AGAAGGAGGA
TGAGCCCAGC
TCTTCCCTGGA
CTGCGGGCCC
AGGCCCGGCT
TCCCTGCTCCG
GCCCTGCAGCA
AGGAGGTCGT
CCGTGTCCTT
TCCTGGGCGT
GCGTCTGCTG
GGCTGGTGGA
TGAAGAACAA
TCACGACGCC
CGCTCCTGCT
GCCACACTCC
AACTCGGGGA
GCGTGAAGTG
CAGAGGTCTGT
AGGATGAACCT
CCTCGGAGGC
CAGAGATCGT

AACCCGCGGG
CGATCGCGTT
CGCGCGGCGC
ACGCCCAGAT
GGAAGCCATT
GTCCCTGAAG
GTTTGCCCTCG
CTGGCTGGAG
GGCCCGGCG
CAGCTTCCGG
GGCAGTGCTG
GGAGACGCTG
GGAGAACTTG
CCGGGTGCTG
CGTGTCTCAG
GCTGGTACCC
GCGCCTGGTG
GGCCGCACCTG
GAAGGCCCAG
CATGCTGCAG
GCAGGTGCTG
CCTGCCGCAG
GCCGGAACCTG
CCGCCCTGGAC
TAGTGCCCCG
GAGCCTCGAG
GGGCACGTCC
GGATGGCGGC

AGCTTCGCCC
TCGTCCGGGG
CTGAGCGCGC
CTGTCCCTGCA
CATGCCCTCT
CGGTATCTCG
GCCCACTTCT
CTGCTGCCCC
GACCAAGCCT
CTGATGAAGA
ATGGAGGCGC
CTGGGCAAGG
GCCGAGTTCT
CAGGCGGTTG
GTCCTTGGGA
CGCTGGCAG
GAGCAAGTGC
GGCCTGAGG
TTTGTGATGA
AGCCTGCTGG
AAGGAGCTGT
CAGCGCCACG
CGGGACAGCC
AGTAGCCTGC
ATCCACCCCG
CTGCTGGCCT
CTCGTTCCAG
GTCCCCCAAG

GGACCCAGGA
CCGCGCGGGC
TGCAGTCACC
GGATGGAGCC
CGTCTTCGGA
GTGAGATGGA
CGCCTGTCCCT
ATGGCCGCCCT
TCCTGGTGTT
TGGCGCGGCT
AGTGTCTGGCA
TGGTGCCCT
TCCCCCAGAA
TGGACTCTCT
AAGCCTGTGT
CGCTCACCCA
CGGACCGGGC
TCCTTTTCGAG
CCCAGAAGCT
GCCATCTGGC
TGGAGACGTG
TCAGCAAGGC
GGGATGAACT
CCCCCGTGCG
AGGGGCCCTCC
TGGCCTCCCC
CCACGGCAGA
CACAGCTGGC

ACTCGTGCTC
CGTGGGAAT
CGGAGCCCG
AGCACCCCTCA
GGATGGCGGC
GCCCTCAGCG
CAGATGTCTT
GGAGGAGCTG
GATGGAGACC
GCTGGCCAGA
GCAGACGCGG
GCCGATCAC
CTACTTCCGC
CCAAGGTGGC
CCACGGGAGG
GGCAGCTAC
CATGGACAGC
GGCAGCAGC
TGTCCTCATC
GCTGGCCAGC
ACGCCCTGGC
CCTGAAATTC
CCAGCCTGCG
GCCCCCTGCA
GGGCTCTGAC

5 / 3 5

Fig. 4A

6 / 3 5

1681	TCGGACCTGG	ACAGCGATGA	TGAGTTTGTC	CCCTACGACA	TGTCGGGGGA	CAGAGAGCTG
1741	AAGAGCAGCA	AGGCTCCTGC	CTACGTCCGG	GACTGCGTGG	AAGCCCTGAC	CACGTCTGAG
1801	GACATAGAGC	GCTGGGAGGC	AGCCCTGCGG	GCCCTTGAGG	GCCTGGTCTA	CAGGAGCCCC
1861	ACAGCCACTC	GGAGGTGAG	CGTGAGCTG	GCCAAGGTGC	TTCTGCACTC	GGAGGAGAAAG
1921	ACCTGTGTGG	TGGATTGTC	AGGCTGCGC	CAGAGAGCCC	TGGTGGCCGT	CACGGTCACA
1981	GACCCGGCCC	CGGTGGCCGA	CTATCTGACC	TCACAGTTCT	ATGCCCTCAA	CTACAGCCCTC
2041	CGGCAGCGCA	TGGACATCCT	GGATGTGCTG	ACTCTGGCTG	CCCAGGAGCT	GTCTAGGCCT
2101	GGGTGCCCTCG	GGAGGACTCC	CCAACCTGGC	TCCCCAAGTC	CCAACACCCC	GTGCCTGCCA
2161	GAGGCAGCCG	TCTCTCAGCC	TGGCAGTGCC	GTGGCGTCTG	ACTGGCGGGT	GGTGGTGGAG
2221	GAGCGGATCA	GAAGCAAGAC	CCAGCGGCTC	TCCAAGGGTG	GCCCAGAGGA	GGGCCCGGCA
2281	GGCAGCCCCA	GCAGATTCAA	CTCCGTGGCC	GGCCACTTCT	TCTTCCCCCT	CCTTCAGCGC
2341	TTTGACAGGC	CTCTGGTGAC	CTTCGACCTC	TTGGGAGAAAG	ACCAGCTGGT	TCTCGGAAGG
2401	CTGGCGCAC	CCTTAGGGGC	CCTGATGTGC	CTGGCTGTTA	ACACCACGGT	GGCTGTGGCC
2461	ATGGGCAAGG	CCCTGCTGGA	ATTCTGTGTT	GCCCTTCGCT	TCCACATCGA	TGCCCTACGTG
2521	CGCCAGGGGC	TGTTGTCTGC	CGTCTCCTCC	GTCTGTCTCA	GCCTGCCCTGC	TGCGCGCCTG
2581	CTGGAGGACC	TGATGGACGA	GCTGCTGGAA	GCCCGGTCTT	GGCTGGCGGA	CGTGGCTGAG
2641	AAAGACCCCG	ACGAGGACTG	CAGGACGCTG	GCACTGAGGG	CCCTGCTGCT	TCTGCAGAGA
2701	CTCAAGAACA	GGCTCCTCCC	ACCCGCGTCT	CCCTAGTCCC	TGGAGGCCTC	CCCAGGACCA
2761	CCCTCGCCGA	CAGCAAGGCA	GGCGGCTGAG	CAGCGGCTTG	GAGCAGCAGA	GCCAGGCTTT
2821	GTAGCGAGGC	CAGGTCTTCG	GCCGCATCCG	GTACGGGAGAG	TGCAGATGCA	GGAAGGCCCCG
2881	GCCTGCCGCT	ATTATAGTG	CAGCCAGTCC	GCTAAAATA	CACCTGGGCT	GGCCTGCTCC
2941	CGCCGGGACA	TGGCAGCCTG	GACGTGGGGC	TGGGGCTGTG	GGCGCTGCTG	GCGGGTTGA
3001	CTCTTCCAGT	GAGGGCAGAA	CCAGGCTGGC	AGGAGGGGAG	GACGGTGTAC	CTGCTGCTCA
3061	GAGCCCCCAA	GGCTCTCCTC	TGAGAGCCAC	CAAGCAGGAC	AGAGCAGCTC	TTGTCCCAGG
3121	TCCCTCGGGC	TGAGCGCCGT	GTCAACCAGGA	GAATAGTGCT	CACAGCCCAG	GCAGGGTGTG
3181	TGGCTCCTGG	ATGGGCTCGT	GGGCGGGAT	GGGACAGGGC	ACGGGCTCTC	AGAAAAATAAA
3241	CTGCTTTATT	GGAAAAAAA	AAAAAAA	AAAAAAA	AAAAAAA	AAAAAAA
3301	AAAAAAA	AAAAAAA				

(SEQ ID NO: 4)

Fig. 4B

7/35

GTGACCCGCAAGCTCCTGCTCCTGCAGTACCAGCACACGACACCCATGGTGC
AGAGCCTGTTGGGGTACTTGGCTCTAGACAGTCAGCGGCGGCCACTCCTCAT
ACAGGTGCTTAAGGAGCTGCTGGAGACCTGGGGCTGCAGCAGTGCTGTCCGA
CACACACCCCTGGAGCAGCAGTGTTACATCAGCAAGGCCATCCTTGTCTGCC
TGGCACACCTCGGGGAGCCGGAGCTGCAGGACATCCGGGATGAATTGCTGGC
CAGCATGATGGCAGGCGTGAAGTGCCGCCTGGATAGCAGCCTGCCCCCTGTG
CGCCGCTTGGGCATGATTGTGGCCGAGGTCATCAGCTCCAGGATCCACCCTG
AGGGGCCTCTCCTGAAATTCCAATATGAAGATGACGAGATGAGCCGTGAGTT
GCTGGCCTTGGCTACCCCAGAGCCTGCGGGTGACTGCTCCTCGGTGTCAC

(SEQ ID NO:5)

Fig. 5

TGTCCCATGTGCTGACTTCTGAGGTGGCTGTGCTAGTGGGTAAGGCCCTGCT
GGAGTTTGTATGGGCCCTTCGCTTCCACGTTGACATTTACGTGCGCCGGGGC
TTGCTGTCTGCTGTGTCCTCTGTCCTCCTCAGTGTACCCACAGAGCGGCTGC
TGGGGGACCTGCCAGATGAGCTACTGGAAGCCAGATCCTGGTTGGCAGATGT
GGCTGAGAAGGATGTGGACGAGGACTGTAGGGAGCTGGCAGTAAGGGCTCTG
CTGCTTC

(SEQ ID NO:6)

Fig. 6

8/35

TCACAGTTCT ATGGCCTAAA CTATAGCCTC CGCCAGCGCA TGGACATCCT
GGACGTCCTT GTTCTGGCTG CCCAGGCACT GTCTCGGCCA AAGAGCCTGC
AGAGACGTTC CCAGCACGGT CCCCCCGTTC CTGGCACCAT GTGTTCACCA
GCACTAGCCG TTTCTCAGAC TGGCAATGTC GCTGCTCCTG ACTGGCAGGT
GGTTGTGGAG GAGCGGATCA GAAGCAAGAC CCGGAGGTTC TCGAAGGGCT
GTCCTCAGCG GGACGTGTCA GGCGTCCCCA ATGAATTCAG CTCTGTGGCT
GGCTACTTCT TCTTCCCCCT CCTTCAGCAC TTTGACA (SEQ ID NO:7)

Fig. 7

VTRKLLLLQYQHHTPMVQSLLGYLALDSQRRPLLIQVLKELLETWGCSSAVR
HTPLEQQCYISKAILVCLAHLGEPELQDIRDELLASMMAGVKCRDSSLPPV
RRLGMIVAEVISSRIH PEGPLLKFQYEDDEMSRELLALATPEPAGDCSS
(SEQ ID NO:8)

Fig. 8

SQFYGLNYSRLRQMDILDVLVLAAQALSRLQRRSQHGPPVPGTMCSPAL
AVSQTGNVAAPDWQVVVEERIRSKTRRFSGKCPQRDVSGVPNEFSSVAGYFF
FPLLQHFD (SEQ ID NO:9)

Fig. 9

LTSEVAVLVGKALLEFVWALRFHVDIYVRRGLLSAVSSVLLSVPTERLLGDL
PDELLEARSWLADVAEKDVDEDCRELAVRALLL (SEQ ID NO:10)

Fig. 10

9/35

VTRKLLLLQYQHHTTPMVQSLLGYLALDSQRRPLLIQVLKELLETWGCSSAVR
HTPLEQQCYISKAILVCLAHLGEPELQDIRDELLASMMAGVKCRLDSSLPPV
RRLGMIVAEVISSRIH PEGPLLKFQYEDDEMSRELLALATPEPAGDCSS//
SQFYGLNYSRLRQMDILDVLVLAQAALSRPKSLQRRSQHGPPVPGTMCSPAL
AVSQTGNVAAPDWQVVVEERIRSKTRRF SKGCPQRDVSGVPNEFSSVAGYFF
FPLLQHFD//
LTSEVAVLVGKALLEFVWALRFHVDIYVRRGLLSAVSSVLLSVPTERLLGDL
PDELLEARSWLADVAEKDVDEDCRELAVRALLL (SEQ ID NO:11)

Fig. 11

KAPVYVRDCVEALTASEDWERWEAALRALEGLVFRSPAAAREVSVELAKVLL
HLEEKTA VAGFEGLRQRALVAVTVTD PARVAEYLTAQFYALNYSRLRQMDIL
DVLTLAAQELSRPGRLGR (SEQ ID NO:12)

Fig. 12

10/35

LISLPAQVANRLGRRLPETFAPVSYQKLLLRQWLKSLHFVLQCDDNREYFDL
EPYSWLLSQAINLIYDVSTLESLLRVLKDYAVAPRGRKVVHTILKELDPAAC
LKTAQSALSAGLNLYVLIGAATLETPHWKHCLLQKLPLQRTFVDNKQLITLA
SYLNAVAPAQLQVLLNQLLGIWSKRISLQKLGSQEHLAISKLLVLAQLDSDD
DEPLDEDDDELKPYDMSNDTTTTIDQRPKFVIDLLHLLREKVENYQVFEGAL
GTAEQLIRGQLAKHDTQLALDLLQLFLVMEMQFYEEQFERTQFKCCVAICVA
HPGPCAELYLCRQFHTDNSFYSSVRILILQVLAATAKELSGDENMQNEMEIV
DVIPPAAKHPRKFEFQQEEESPAARLAAAQRIIRDRLRAKTKRYFSKPKAGD
QMEKANPFHPVAGTFFFSLVRGQRTQMLYVKYEIDTQLLVNLLNTMSVLVM
CSQNCPLLPAMTREIFDLCAFVRFNAEARVRAATLQLIGIALVTTPAHVLAQ
HFAESLNELQRWLNDFIRSPLVGGETSEECRELA (SEQ ID NO:13)

Fig. 13

11/35

MAEGTKQERTLENNLLHKVGEAVSAISDAKHVDQVISAIHSVAVLLFPVDPS
LFSGNFEMLHIVRGSGTFGLLMILYLGSAQVCSSVPSADERNEWLETFYRG
VAFPTFARVLLLLDVASDWLSCFPISVQKHLYDKFFLDGSVIEVVQVLVPFLH
HVGDGGVNANSVQTNVERLLILCLENLGVKITKEIGNIYQGHNSSNGSLK
PLLSRLSQILTSIPDKARASCTEANTVIVLSFVGEVFSRICRRGLSDLCLS
EVTPHVLAQVRLLNSKIGAEVDTFQLDPTTRIWSKTMEAVTDPYAVEKMA
EQLLHQLYAEHPSDVEAFWTIWTLFHRNVIHQASVRQAKCFLWQLDSFFRYP
FFFFHHPNAVKQCVLECPPVTNTLAKGDVTQGLLETTQRLASVWSKREFLQ
SVQLEQQAYLQFLFPVTDISDITAALGLCLENMSREELDRTKDVMHSILQGV
SCRLENPGDLVRKMASSIAFMFSKVIDPKNPLYLDDSI TDNAIDWEFGLQTA
SITNTMENG DGENKRSASLTEVNESSRRNKQKENRKSKNISAFVLADPNEIV
DLATLNCDTESDKDDGDDDDASVSSDNSSVTSLEPYDLMDDDKDLGKQF'THLV
DVVGALRKTD DADGVEKAIYVAEKLVRASPDELTHIAGDLARTLVQVRCSDI
AIEGEEDSAEEKRQRALIALLVTRPFESLET LNNILYSPNVDVSQRIMILDV
MAEAARELAN SKTLKPKHEARGPLISNISDPQPWYLPSNASTPWKKVSETGS
FHLNWANRFERELQSKPGQTKKGKSRRWSLKSADRDQNSTDWSQNRFP LYAA
AFMLPAMKEFDKKRHGVDLLGRDFVVLGKLVHMLGVCMQCASMHP EASALAI
SLLDMLQRREVCNHPEAYVRRAVLFAASSVLVSLHPSYIVSTLVEGNLDLSR
ALEWIRTWALQIADSDIDRDCYTMALSCLQLHAEMALQTSRALESTGGSSSS
SSIRPMNISLPSGISKLTSIKLPSSNVHL (SEQ ID NO:14)

Fig. 14

12/35

GTTTCTACTGGGTTTGGGAATGGATACGCACATGGGCTCTTCCATGTTGCAGA
AACAGATCCTGATACAGAATGCACATCAATGGCTATGACCCTCCCTGCGGCT
TTCATTCTGAGATGGCCCTTCAGACATCACGAGCACTGGAATCGGCGGATCA
CAGCAAGGCCAGCAGCAGCAGCAGCAGGTCGCTACCTTCTAAGCTTGATAAC
ATCATCATACCATTTGCCAACATGATGTGATCGATGTACTAGATTAAGTTGT
AAATAAGCATATATTATCTTGCCATGTAATATTAGAAATTCCAGGTCAGCTA
ATTCTGATTAGGCTTGATATTGTCATATTACAGGGTTTTTTTAATCTGGGATT
TGAAAATGCTAGAATTTTTGTGATTCGTGTGGAGATTGTATGGAAGTTGTC
CTGATCCAGATTAAAGAGAATTGATGAAAATTGTACCCAAAAAAAAAAAAAAAAA
A (SEQ ID NO:15)

Fig. 15

AGTTCCTAGTGATCAGGGACCTGCTGGTGCAGGTCTNTGGAGGGAAGTCTCA
GAATCAGGAACACTTCTGAATTGGTCACACCGGTATGAAAGAGAAGTTCCAT
CTAGATCTGGTCAAGTTAAATCAGGAAAATCTCGTAAATGGGGTCTTGGAAA
AGCTAAAGATTTGCAGACAGAGTGGTCAAAAAACAGATTTCTTTTATATGCT
GCTGCTTTTATGCTCCCTGTTATGCAAGGATATGATAAAAGATCACATGGTG
TTGACTTGCTCAATCGGGACTTTGTTGTCCTAGGTAAATTGATATACATGCT
TGGTGTCTGTATGAAGTGCATGGCAATGCATCCAGAAGCATCAGCTGTTGCC
CCAGCTCTTCTT (SEQ ID NO:16)

Fig. 16

13/35

PWREVSESGTL-LNWSHRYEREVPSRSGQVKSGKSRKWGLGKAKDLQ--
TEWSK

NRFLYAAAFMLPVMQGYDKRSHGVDLLNRDFVVLGKLIYMLGVCMKCMAMH
PEASAVAPALLDMIRSRAV (SEQ ID NO:17)

Fig. 17

LEWIRTWALPCCRNYRMHINGYDPPCGFHSEMALQTSRALESADHSKASSSS
SRSLPSKLDNIIIPFANM (SEQ ID NO:18)

Fig. 18

PWREVSESGTL-LNWSHRYEREVPSRSGQVKSGKSRKWGLGKAKDLQ--
TEWSKN
RFPLYAAAFMLPVMQGYDKRSHGVDLLNRDFVVLGKLIYMLGVCMKCMAMHP
EASAVAPALLDMIRSRAV//
LEWIRTWALPCCRNYRMHINGYDPPCGFHSEMALQTSRALESADHSKASSSS
SRSLPSKLDNIIIPFANM (SEQ ID NO:19)

Fig. 19

14 / 35

TTT TAGGAATTGGTCAAATAGCTATGAGAGGGAACTTCCCCCAAACCTAAT
CAGGTCAAGAAAGGGAAAACACGCCGGTGGAGCCTACAATCTCCCACACAAC
AAAACCAGATGGAGTATTCTCATAATAAGTTACCCATGTATGCTGCTGCATT
CATGCTTCCTGCCATGGAGGGATATGATAAAAAAAGGCAAGGTGTTGACTTG
CTTGGAAGAGATTTTATTGTCTTGGGGAAACTCATTTATATGCTTGGGGTCT
GTATGAAATCTGTAGCCATGCATCCAGAAGCTTCTATGCTGGCTCCTTCCCT
CCTAAATATGTTAAGATCCAGGGAGGTATGCCATCACCAGGAAGCATATGTG
AGAAGAGCCGTGCTTTTTGCAGCTGCATGTGTATTGGTTGCCCTTCATCCTA
CTTACATTTTCATCCACCTTACTCGAAGGAAATGCTGAAATTTGACTG

(SEQ ID NO:20)

Fig. 20

TCCGGAGAAGGACTCTGATTCCCCTTCCAATAAAGAGAGAAAAGTATTTGTTTA
AAGGGTAAAAAGAAGTTATTGGACTTTAATGCGCTTGATCCAGATGAGATTA
TTGATCCAGCATCACTGAATCTTGAATCAGACGATAGCGATGAGGATGCTGA
CGATGGTGCTAGTGAGAATTCATATTCTTCAAGTGATTCATCTTTACGGCCA
TATGATTTGTCAGATGATGACTCAGATTTGAAAAGAAAAATTTACAGTTGG
CTGATGTAGTTGCAGCTCTTAGAAAATCCAATGATGCCGATGGGGTGGAAAG
GGCTATTGATGTAGCTGAAAAGCTCATAAGAGCATCCCCCGATGAACTAAAA
CATGCAGCAAGGGATATGACCAGAACTCTTGTTTCAGGTTCGGTGCTCTGATA
TAGCTTTAGAAGGTGCAGAAGAATCAACTGAAGACAAAAGACAAAGATCATT
AGTTGCCTTAGTAGTTACCTGCCCATTGGAATCACTTGAGTCACTAAACAAC

C

(SEQ ID NO:21)

Fig. 21

15/35

ATNNATNNNNNTAGCAGAAATAATTAGCAGTTTAGCACGTTTAGAATTTAT
ATTCTGGGCTAGAAGCTATTCACAATAGCTAGAAGCATCACTAATCCCCATT
TTAGTAGGGAATTTTAATGGTCACTTTGGATGCATCAGAAGGAAGAACAGGA
CCTGCCTTGAGTGAACCTTCTTACTGACTCCAATGCTCGGGAAGTTTGAAGAG
CCATCTCGGCATGGAGTTGTATACATGTCATAGCCATCGTATAGCATTCTTT
ATCTGTGTCTGACTCGGCTATGTCAAGTGCCCATGTGCGAATCCATTCAAGG
CCAGTCGAAATTTTCAGCATTTCCTTCGAGTAAGGTGGATGAAATGTAAGTAG
GATGAAGGGCAACCAATACACATGCAGCTGCAAAAAGCACGGCTCTTCTCAC
ATATGCTTCCTGGTGATGGCATACTCCCTGGATCTTAACATATTTAGGAGG
GAAGGAGCCAGCATAGAAGCTTCTGGATGCATGGCTACAGATTTTCATACAGA
CCCCAAGCATATAAATGAGTTTCCCCAAGACAATAAAATCTCTTCCAAGCAA
GTCAACACCTTGCCTTTTTTTTATCATATCCCTCCATGGCAGGAAGCATGAAT
GCAGCAGCATACATGGGTAACTTATTATGAGAATACTCCATCTGGTTTTGT
GTGTGGGNNNNNGTNNNCNCCACNNNCGTGTTTTCCCTTTCTTGACCTGATT
NGNTTTTGNNNGA (SEQ ID NO:22)

Fig. 22

16/35

TTGAAATTTCAACTGGCCTTGAATGGATTCGCACATGGGCACTTGATGTAGC
CGAGTCGGACACAGATAAAGAATGCTATACGATGGCTATGACATGTATACAG
CTCCATGTTGAGATGGCTCTTCAAACCTCCCGAGCATTGGAGTCAGTAAGAA
ATCACTCAAGGCAGGTCCTGTTCTTCCTTCTGATGCATCCAAAGTGACCAT
TAAAATTCCCCACTTAAATGGGGATTAGTGATGCTTCTAGCTATTGTGGATG
GATTCTAGCCCAGAAATACAAATTCTATATCATGCTAAACTGCTAATTATTTTC
TGCTACTCTATTAATTTAATGTTTTATGGGGTCCTTGTTTTTCGTAAGAAATT
GTATATTTTAGGATTGTATATATTCTCACAAGGGGACGAGGTTGATGTCAAT
GCTCCCCTACACACCCACCCTTGATGTAGCACGTGTTACT

(SEQ ID NO:23)

Fig. 23

AAACTCGCGGCGGAAGGCAGCAGAGTGGCGCCGATGGAGGAGGGTTTAGAGA
AGAGAGAATTGGAAGGCGAAGTAGTTACCAAGGTTGTCGAAGTGGTTTTCGGC
TATAAAGAATGCGAAGCACGTCGATCAAGTCATTCGCGCGCTTCATTCCTTA
GTCACCCTTCTTTTTCCCCTTTGACTCTTCACTCCTATCAGATAGCATTGACC
AGAGTTACCGAGACCAGGTCGAAGTTCCTTCTGCAGAAAAACGACATGCTTG
GTGGCGTGCGTTTTATCGAGGAGCTGCTTTTCTTACACTGGCTAGGTTTTTA
TTACTTGATGTTGCCTCGAACTGGTTGGGTTGTTTTCCCTTTATGGCGCAGA
AGTACATTTATGATGTTTTCTTTGTTTCGTGGATTGGTCACTGACGTTCTGCA
GATTCTGGTTCCTTTTTCTTCAGCTGAGTGCGAGTGATGG

(SEQ ID NO:24)

Fig. 24

17/35

KXNQVKKGKTRXWXXXPTQ-
QNQMEYSHNKLPMYAAAFMLPAMEGYDKKRQGVDLLG
RDFIVLGKLIYMLGVCMKSVAMHPEASMLAPSLNMLRSREVCHHQEAYVRRAVLF
AAACVLVALHPTYISSTLLEGNAEISTGLEWIRTWALDIAESDTDKECYTMAMTCI
QLHAEMALQTSRALESVRSSLKAGPV-----LPSDASKVT-IKIP
(SEQ ID NO:25)

Fig. 25

NWSNSYERELPPKPNQVKKGKTRRWSLQSPTQ-
QNQMEYSHNKLPMYAAAFMLP
AMEGYDKKRQGVDLLGRDFIVLGKLIYMLGVCMKSVAMHPEASMLAPSLNML
LRSREVCHHQEAYVRRAVLFAAACVLVALHPTYISSTLLEGNAEIS
(SEQ ID NO:26)

Fig. 26

EKDSDSPSNKEKSICLKGGKKLLDFNALDPDEIIDPASLNLESDDSDDEDADD
GASENSYSSSDSSLRPYDLSDDDSDLKRKISQLADVVAALRKSNADGVERA
IDVAEKLIRASPDELKHAARDMTRTLVQVRCSDIALEGAEESTEDKRQRSLV
ALVVTCPFESLESLNN
(SEQ ID NO:27)

Fig. 27

18/35

MEEGLEKRELEGEVVTKVVEVVSAIKNAKHVDQVIRALHSLVTLLFPFDSSL
 LSDSIDQSYRDQVEVPSAEKRHAWWRAFYRGAAFPPTLARFLLLDVASNWLGC
 FPFMAQKYIYDVFFVRGLVTDVLQILVPFL (SEQ ID NO:28)

Fig. 28

EISTGLEWIRTWALDVAESDTDKECYTMAMTCIQLHVEMALQTSRALESVRN
 SLKAGPV-----LPSDASKVT-IKIPHLN (SEQ ID NO:29)

Fig. 29

MEEGLEKRELEGEVVTKVVEVVSAIKNAKHVDQVIRALHSLVTLLFPFDSSL
 LSDSIDQSYRDQVEVPSAEKRHAWWRAFYRGAAFPPTLARFLLLDVASNWLGC
 FPFMAQKYIYDVFFVRGLVTDVLQILVPFL//
 EKDSDSPSNKEKSICLKGGKKLLDFNALDPDEIIDPASLNLESDDSDDEDADD
 GASENSYSSSDSSLRPYDLSDDDSDLKRKISQLADVVAALRKSNDADGVERA
 IDVAEKLIRASPDELKHAARDMTRTLVQVRCSDIALEGAEESTEDKRQRSLV
 ALVVTCPFESLESLNN//
 EISTGLEWIRTWALDVAESDTDKECYTMAMTCIQLHVEMALQTSRALESVRN
 SLKAGPVLPSDASKVTIKIPHLNNWSNSYERELPPKPNQVKKGKTRRWSLQS
 PTQQNQMEYSHNKLPMYAAAFMLPAMEGYDKKRQGVDLLGRDFIVLGKLIYM
 LGVCMKSVAMHPEASMLAPSLNMLRSREVCHHQEAYVRRRAVLFAAACVLVA
 LHPTYISSTLLEGN (SEQ ID NO:30)

Fig. 30

19/35

MNLR SRLVNA TERA VLFQIF KDVQNDPEKY D NAVEAICES
IDYFGKFLTD SEYLTQIKPI LDTQCPTKSI ICFSKCLTKV
STDINTTTFR DVITMLDWLK YVVEKSLTSA ICSSLKVKET
DVSAVQLYRE FASACSNIFE KVSNCCKAKAL SGEHVKYINT
VKWIFKMNLV QGIQKAMLLA HDDIVTAAPF TSFYGSGGPY
MKTVAEIISS GRIDITNKDG LLVQMIEWIG SLNNFDSQWR
RMMFLIFQEP TYQGIQVHES LLTTLFLISK SDQILKRCIE
ATDLTGTLKR VVMVKLPFQR
VLKRKTIEIL INFVYRTKEQ FAIQLLETSV KIWSDLNYAK
SAPESQERHI VRMILYLVHL FRTCSSIDWE SLFLNSMDGV
HCRMSMLPMY VQSGIFVNQA LCKQATKHSR KTHGSDEQPP
ETLEENKFVS SEVGKIWFEE MTSILEHGFN SSTVKDSERV
RETANEITKD DSGEEFEETN AQRLQNNKDS AAITSKNNLR
LDSDDDEDFP DYQVNESEKI FKNLEIGEEP KNKVTPPAYI
ADAFEMLLEK EKYEVFEAAF FNITNLINRR PIGFPQIAEK
LFIRILHLQN NFGTPKFKET
VDEIAVACIT QRPEIVPSVV RLIIAPGQGF SIKQRLLHYI
HNAADGMGAL DKKLEECVMA QQLRIGGPTL SIILHRTINT
DYDDEDEDPH RLLVPEWRRM VDARIAANTR RIGTTREPPR
AGVVNRLAQA AKYMFYPLL V LPRGENASLL GKDSDLLASL
IMVASMVYVR YGVCPQIHRM SSELISYATP HRFSENAKLR
TACIIAHLNV TTLLPGDLMD ELFDVPALIG WFDWANSVLV
NASSSQLEKD MTRQFGHSVT KHLQRHHPAV LQHQDV

(SEQ ID NO:31)

Fig. 31

SUBSTITUTE SHEET (RULE 26)

20/35

MVLETLKQGL DSSQIHEALI QLDSPREPV DLDASMVLK FVIPVYPSLP ERSKVILRRL
ASKSFTFLCQ IVTFSRTISG RDGLQEIRIY QEILEDIISF EPGCLTFYK ASTTSKADRD
SIKALFFGSK LFNVLANRID MAKYLGYLRL QWKFLLESNE TDPPGFLGEW LVSSFLLNPV
LAADMMLGEL FLLKESYFFS FQKIISASSL IDQKRLLIAKF LLPYIQVIVT LENLNDVRKI
LRRFDLDKII SLSVLFEIQS LPLKEVIVRL MSNHSSTKFV SALVSKFADF TDEEVDTKTC
ELLVLFVAVHN LNHSQREEIA HDERFLNGVT KHLGSNEREA RERAMFIAKL LSGGHLKYES
DFKINIPNVK FESNSDDKII DFQSLKNPSI CNTQTDVGKD KITEVSGHVQ SLTLDCSDSD
DEDENDEREI VKRIVFLKDL MKEYEKTGES RKAPLIPLLK QTVKLIRQKA DFQLEVGYYA
QGILSSIVCL NNEFDEPLFE QWRINALTSI LVVLPEKVNG AINILFNSEL SLQQRMSLLS
ALGLSARELR GLDDPTIVKP KFDFPTNRLP WDDQSHHNSR LVEVQESTSM IKKTKTVWKS
RKLGKDREKG TQNRFRKYAG LFFYPLAHGW LNGIDVGTYN QLFKSHYLT LRIIYSCANP
VHDFESMTEL MNHIISSAIE EGISLNKG (SEQ ID NO:32)

Fig. 32

21/35

TCGAACTTCAATTTATATCGATGTTTCTTCTATTTTAGTGCAAATTTTAATCAAAAAAATT
AAACCTTTTGTCAACATGGCGCGAAATTTTCCCCTGCTCTCGATTACCTGGGATCAGAAG
CTGAAGATTTTAACAAGGCGCAACACTTGTATCTGAAACCGATGGCTGTTATAAAAATTAC
CGTAGTGCTGCCTCGGATGACGATCCCGGGACAATCAATATCTAATTGGGATCTCATGGAA
AGACTAAAACGTGCAATTGATCCAATTCAAATGGATTCTTGCAAAGTTCGTGAGAGCAATA
TCGACAGTGTTATTTTTGAAGCGGAACCTCTTTCGCTAGGAATCATGCAGAAAACGATGAA
GATTCTCGATGGATTCTCTATGAAAGTGTCTGGATTTGCTGAGCCCTTGAAAGTTAAAACA
AAAGAGGCAAAGCTTGATTTTCCAAGTCGTCATGATTGGGATGATTGGTTTATGAAACATA
AGATGGACGAGATGAAACCCGGAGAACGTCCTGATACAGTTTATTTGGCAAGAATTCCAGT
GAAATGGTTCTGCGATGGTTACAATGATCTTCCTTCTGAACGACGTCTTCGCGTTGCAATG
GAAGCGTTCGGATCTGTTAGAGTTGTCGACATTCCAATTTGTGACCCACTCCGATCTCGAA
TGAATTCAAAAATCTCTGGTATTCAGCAAAAAGGTTTCGGATTAGGGCAGGATGTATTTTT
TGAAGCATATGTGCAATTTATGGAGTATAAAGGTTTCGCCACTGCCATGGATTCCCTGAGA
AATCGTAAATGGGCAAAACGAATTGATGGACGGTTCTTTCAGGCTAATGTCAAGGTGGATT
TCGATCGATCACGTCATCTCAGCGAAGTTCAAATTGCAAAGCGAGCGGAAGAACGACGTCA
AATTGAAACGGAACGACTGCGGCAAGAAGAAGAGGAGTTGAATATCAGACGTCAAGAAGAA
CTGAAAGTTAAACAAGAACCTCGATGATAAAGATAGAAGACGAGAAGATAGAGAACGGAAAC
GCCGCGAAAAGAGGGAACTCGAACGAATGGCTGAAGAAGAGAAAAAACGCCTTGAAAAGGA
ACGACTTGAAGCAGAGCAACGATCACGAGCCACTCGCCGTTTGCAAGGTGTTTCGTCTTTTG
AAGTTTCTGTTTGAAAAAATTGAAGCACGAGAAGAGAGACGAAAGAGAAAAGAAGAAGAAA
AGTTGAAAGATGAGCTGAGCAAAATCAAAGAACTTGAAGAACAACCTGTAGAACAAGAAGA
CGCATTGAGGCAAGCTCTACTTCAGCAGAGGGAAATTGGAATGCGAGAACGCTTGAAAGAG
AAGATGAAAGCTTCAGGAGCAGAGAAGGATAAGAAGAGAGATAAAAATAAGACTTCTAGAA
GAAGACGAACATAATAGACATGACACTTCATCCTCTTCATCCGAAGATTCTGATTCTCCATC
TGATTCTCCCCACTCTCGTCGACAAAGAAAACGTCAAAGTGAAGGAGATGATTTC

(SEQ ID NO:33)

Fig. 33

22/35

MARNFPLYLGSEAEDFNKAQHLYLKPMAVIKITVVLPRMTIPGQSI SNWDLMER
 LKRAIDPIQMDSCKVRESNIDSVIFEAEALLSLGIMQKTMKILDGFSMKVSGFAE
 PLKVKTKEAKLDFPSRHDWDDWFMKHKMDEMKPGERPDTVYLARI PVKWFCDGY
 NDLPSEERRLRVAMEAFGSVRVVDIPICDPLRSRMNSKISGIQQKGFGLGQDVFF
 EAYVQFMEYKGFATAMDSLNRNRKWAKRIDGRFFQANVKVDFDRSRHLSEVQIAK
 RAEERRQIETERLRQEEEEELNIRRQEELKVKQELDDKDRRREDRERKRREKREL
 ERMAEEKKRLEKERLEAEQRSRATRRLQGVRLKFLFEKIEAREERRKRKEEE
 KLKDELSKIKELEEQPVEQEDALRQALLQOREIRMRERLKEKMKASGAEKDKKR
 DKNKTSRRRRRTNRHDTSSSSSESDSPSDSPHSRRQRKRQSEGDDF (SEQ ID NO:34)

Fig. 34

MAAATIVHDT	SEAVELCPAY	GLYLKPITKM	TISVALPQLK	QPGKSISNWE
VMERLKGMVQ	NHQFSTLRIS	KSTMDFIRFE	GEVENKSLVK	SFLACLDGKT
IKLSGFSDIL	KVRAAEFKID	FPTRHDWDSF	FRDAKDMNET	LPGERPDTIH
LEGLPCKWFA	LKESGSEKPS	EDVLVKVF EK	FGEIRNVDIP	MLDPYREEMT
GRNFHTFSFG	GHLNFEAYVQ	YREYMGFIQA	MSALRGMKLM	YKGEDGKAVA
CNIKVSFDST	KHLSDASIKK	RQLERQKLQE	LEQQREEQKR	REKEAEERQR
AEERKQKELE	ELERERKREE	KLRKREQKQR	DRELRRNQKK	LEKLQAEQK
QLQEKIKLEE	RKLLLAQRNL	QSIRLIAELL	SRAKAVKLRE	QEQKEEKLRL
QQQEERRRLQ	EAELRRVEEE	KERALGLQRK	ERELRERLLS	ILQSKKPDDS
HTHDELGVAH	GPAAARPGHP	ADRVVRLCER	HHAAPPRGPA	PGRCPQGEPC
PPRGRRRSQK	RERERGRGGP	MQGGSELLSC	GPRGWLSREE	VPGRRPLLHS

(SEQ ID NO :35)

Fig. 35

23/35

ATGTCTGCTC AAAAGCTATA CGCATCGCAT GTTCCCTTGG CACCGCTGTC
TCGGATGCTT CTCGGGATTG GATCAGCAGT AACAGCGATC TCGGATCCAA
AAAGAGGAGA TATGGTGGCA GCGATGGGCG AAACACTGTC AATTGGACCA
GTTTTAGAAA ATATTCGAAA AAGAATGGAA TCTGATGTTG TTGGAAAGCG
ACTTCTTCTC GAAAAACCAA GAATTTCAAA TGGAACAATT GATAGAAAGT
GGCTAAGACA GTTACCGGAT GGAACATTAG GAAAATTGTA TTCAAACTTT
CTCGATCGTT TGAACACATC TCCAGATGCT CGGCCCACTG TCAAGTATAT
CGATAATTTG GAGCATCTTT ATGTTATGCA AAGGTATCGC GAAACACACG
ACTTCACCCA CATCGCATTG GAGCAGAAAA CGAACATGCT CGGCGAGGTA
ACAGTCAAAT ATTTCGAAGG AATTCAATAT GGGCTTCCAA TGTGTGTCAC
TGGTGGGAATA TTTGGAGGTG CCAGGCTTTT AACAAAAAAT CGCCAAGAAC
TTGTCGACCG GAACCTCCCT TGGGTTGTGG AGCAGGCCAC GAATGCACGA
TTCTTCATGG CTTTCGACTG GGAAAATCAC TTTGAAAAGC AGCTCAGCGA
GGTGCAGAAG GAGCTAAATA TAACGCCATT ATCTGTGAAT ATGTGA

(SEQ ID NO:36)

Fig. 36

MSAQKLYASHVPLAPLSRMLLGIGSAVTAISDPKRGMVAAMGETTAIGPVL
ENIRKRMESDVVGKRLLEKPRI SNGTIDRKWLRQLPDGTLGKLYSNFLDRL
NTSPDARPTVKYIDNLEHLYVMQRYRETHDFTHIALEQKTNMLGEVTVKYFE
GIQYGLPMC VTGGIFGGARLLTKNRQELVDRNLPWVVEQATNARFFMAFDWE
NHFEKQLSEVQKELNITPLSELLDLPEMEPDVPDILFSKGHPGFWRVLQAVD
MMI

(SEQ ID NO:37)

Fig. 37

hCLK-2	MEPAPSEVRLAVREAIHALSSS.....EDGGHIFCTLESLKRY.....
CLK-2	MNLRSLVNATERAVLFQIFKDVQNDPEKYDNAVEAICESIDYFG.....KFLTDSEYLTQIKPILDTQCPTKSIICFSKCLTKKVSTDINTTFRDVS
Tel2p	MVLETLKQGLDS.....SQIHEALIQLDSPREPVDLDASMV
AtCLK-2	MAEGTKQERTLENNLLHKVGEAVSAISDAKHVDQVISAIHSVAVLLFFPVDPSLFSGNFEMLHIVRSGSGTFGLLMILYLG.....
hCLK-2	LGEMEPALPREKEE.....FASAHFSPVLRCLASRLSP.....AWLELLPHGRLEELWASFFLEG.....PADQAFVLVLMETI
CLK-2	ITMLDWLKYVVEKSLTSAICSSSLKVKETDVSQVLYREFASACSNIPKVS.....CCAKALSGEHVKYINTVKWIFKMNLVQGIQKAMLLAHDDIVTAAPFT
Tel2p	LIKFVIPVPSLPERSKVI LRRLASKSFTFLCQIVTFSTRTISGRDGLQEIRIYQEILEDIISFEPCCLTFYLKASTTSKADRDSEIKALFFGSKLFENVLANRI
AtCLK-2	SAQVCSSVVPADERNLEWLETFYRGVAFPTFARVLLLDVAS.....DWLSCFPISVQKHLYDKFFLDG.....SVIEVVQVLVPFL
hCLK-2	EGAAG....PSFRMLKMARLLARFLRE..GRLAVLMEAQCRQQTRPGFI....LLRETLGKVVALPDHLGNRLQQENLAE..FFPQNYFRLLGEEVVRVL
CLK-2	SFYGSGGYPYMKTVAEIISGRIDITNKD.GLLVQMIIEWIGSLNNFDSQWRRMMFLIFQEPTYQGIQVHESLLTTLFLISKDQILKRCIEATDLTGTLKRVV
Tel2p	DMAKYLG...YLRQWKFLLESNETDPP.GFLGEWLVSSEFLNPVLAAD....MLLGELFLLKESYFFSFQKIISASSLID..QKRLLIAKFLLPYIQVIVT
AtCLK-2	HHVGDGGVNANSVQTNVERLLILCLLENDGVLKITKEIGNIYQGHN.....SSNGSLKPLLSRLSQILTSI..PDKARASCTEANCCTVIVL
hCLK-2	QAVVDSLQGGLDSSVSFVSQVLGKACVHGRQQEILGVLVPRLAALTQGSYLHQVVCWRLVEQVPPDRAMEAVLTGLVEAALGPEVLS
CLK-2	MVKLPFQORVLKRKT.....I EILINFVYRTKEQFAIQLLETSVKIWSDLNYAKSAPESQERHIVRMILYLVHLFR.....
Tel2p	LENLNDVRKIILRR.....FDLDKIISLSVLFEIQSLPLKEVIVRLMSNHSSSTKFVSALVSKFADFTDEEVDI.....
AtCLK-2	SFVGEVFSRICRRGLSDLLLSEVTPHVLAQVRRLLNSKIGAIIEVDTFQLDPTTRIWSKTMEAVTDPYAVEKMAEQLLHQLYAEHPSDVEAFWTIWTLFHRNV
hCLK-2	RLLGNLVVKNKKAQFVMTQKLLFLQSRLLTTPMLQSLLGHLAMDSQRRPRLLLQVLKELLETWGSSSA
CLK-2	TCSSIDWESLFLNSMDGVHCRMSMLPVMYVQSGIFVNVQALCKQATKHRSKTHGSDEQ
Tel2p	KTCELLVLFAVHNLNHSQ.....REE.....
AtCLK-2	IHQASVRQAKCFLWQLDSEFFRYPPFFFFHHPNAVKQCVCLECPPTNTLAKGDVTQGLLETTQRLASVWSK.....REFLQSVQLEQQAYLQFLFPVTDISD

Fig. 38A

hCLK-2	IRHTPLPQQRHVSKAVLICLAQLGEPRLDRSRDELLASMMAGVKCRLDSSLPPVRRLGMIVAEVVSARIHPEG.....PPLKFQYEED.....ELSLELL
CLK-2	PPETLE.....ENKFVSSEVGKIWFREMTSILEHGFNSSTVKDSER.....VRETANEITTKDDSGEEF
Tel2p	IAHDER.....FLNGVTKHLGSNEREARERAMFIAKLLSGHLKYESDFKINIPNVKFESNSDDKIIDFQSLKNP
AtCLK-2	ITAALG.....LCLENMSR.EELDRTKDVMHISILQGVSCRLENPGDLVRKMASSIAFMFSKVIDPKN.....PLYLDDDSITDNAIDWEFGLQTA
hCLK-2	ALASPPQAGDGASEAGTSLVPATAEPP.....AETPAEIVDGGVPQAQLAGSDSDLDSDDEFVPYDMS...GDRELKSS.....KAPA
CLK-2	EETNAQORLQNN.....KDSAAITSKNNLRLDSDDDDEFDPDYQVNESEKIFKNLEIGEEP.KNKVTPPA
Tel2p	SICNTQTDVGK.....DKITEVSGHVQSLTLDSCSDSDDDEDN...DERE..IVKRIVFLKDDLKMEYEKTGE
AtCLK-2	SITNTMENGDGENKRSASLTEVNESSRRNKQKENRKSKNISAFVLADPNIEIVDLATLNCDE.SDKDDGDDDASVSSDNSSVTSLEPYDLMDDDKDLGKQFT
hCLK-2	YVRDCVEALTTSEDIERWEAALRALEGLVYRSP.....TATREVSVELAKVLLHLEKTCVVGFAGLRQRALVAVTVTDPAPVADYLT SQFYALNYSL
CLK-2	YIADAFEMLLLEKEKYEVEFAAFNITNLINRRPIGFPQIAEKLFIIRILHLQNNFGTPKFKETVDEIAVACITQRPEIVPSVVRLIIAPGQGFSIKQRLLLHYI
Tel2p	SRKAPLIPLLKQTVKLIRQKADFQLEVGYYAQG.....ILSS..IVCLNNEFDEPLFEQWRINALTSILVLPKEKVNAIN.ILFNSELSL
AtCLK-2	HLVDVVGALRKTDDADGVEKAIYVAEKLVRASPDDELTHIAGDLARTLVQVRCSDIAIEGEEDSAEE...KRQRALIALLVTRPFESLETINNIIYSPNVDV
hCLK-2	RQRM DILDVLTAAQELS....RPGCLGRTPQGPS.....PNTPCLPEAAVSQPGSAVASDWRVVVEERIRSKTQRLSKGGPRQGPAGSPR.....
CLK-2	HNAADGMGALDKKLEECVMAQQLRIG...GPTLSIILHR.....TINTDYDDEDEDP.....HRLLPPEWRRMVDARIAANTRRIGTTREPPR.....
Tel2p	QQRMSLLSALGLSARELRG.....LDDPTIVKPKFDFP.....TNRLPWDDQSHHN.....SRLVEVQESTSMIKKTKTVWKSRLGKDR...E
AtCLK-2	SQRIMILDVMAEAAARELANSKTLKPKHEARGPLISNISDPQPWYLPASNASTPWKKVSETG.SFHLNWANRFERELQSKPGQTKGKSRRWLSLKSADRDQNST
hCLK-2	FNSVAGHFFFP LLQRFDRPLVTFD LLGEDQLVGLR LAHTLGALMCLAVNTTVAV.....AMGKALLEFVWA..LRFHIDAYVRQGLLSAVSSVLLSL
CLK-2	AGVVNRLAQAAKYMFP LLVLP RG.....ENASLLGKDSDLLASLIMVASMVVVRGVCPCQIHRMSSELISYATPHRFSENAKLRTACTIIAHLNVTTLL
Tel2p	KGTQNRFRKYAGLFFYP LAHGWLN...GIDVGTYNQLEFKSHYLTTLRIIYSCAN.....PVHDFESMTELMNHHIISSAIEEGISLNKG.....
AtCLK-2	DWSQNRFFPLYAAAFMLPAMKEFDKRRHGVDLLGRDFVVLGKLVHMLGVCMQCASMHP EAS...ALAISLIDMLQRREVCNHP EAYVVRRAVLFAASSVLVSL
hCLK-2	P.....AARLLEDLMDELL EARSWLADVAEKDPDED CRTLALRALLL.....LQRLKNRLLPPASP.....
CLK-2	PG.....DLMDELF DVPALIGWFDWAN.....SVLVNASSSQLEKDMTRQFGHSVTKHLQRHHPAVLQH QDV.....
Tel2p	
AtCLK-2	HPSYIVSTLVEGNLDLSRALEWIRTWALQIADSDIDRDCYTMALSCQLHAEMALQTSRALESTGGSSSSSSIRPMNISLP SGISKLTSIKLPSSNVHL

Fig. 38B

D.m.		0
H.s.	MEPAPSEVRLAVREAIHALSSSEHGGHIFCTILESJKRYLGEMEPPAI[REKEEEFASAHFSVLRCLASRLSPAWLELLPHGRLEELWAS[EFFLEGPADQAFVLMEITI	107
C.e.	MNLRSLVNATERAVLFQIFKDVQNDPEKVDNAVEATCESIDYFGKELIDSEYLTQIKPILDTCQ[FKSLICFSKCLITKVSIDIN.....TTT[ERDVITIMLDWLKYVVEKS	106
D.m.		73
H.s.	EGNAGHSFRIMKMARLIARFI[REGRLAVIMEAOCROQTRPGFIILRETIILGKVVALLPDHLGNRLQOENLAHFF[EPQNYFRLLIGE[EVVRVLQAVVLSLQGGILDSSVSFVSQVLGKACVHGRQQEITIG	232
C.e.	LTSALCSSLKVKEIDVSAVQIYREFASACSNIPERKVSNOCAKALSGEHVKYINIVKWLFKMNI[VG.....IQKAMLI[AHDDIVTAPFTISFYSGCGPYMKTVAEI[SSGRIDITTNKDGIL	222
D.m.	STLRV[KOYAVAPRGRKV[HTTIKELDP.....AACTIKTAQSALSAGINIVYLIGAAITETPHWKHCLIQKLP[OR.....ITFV[NKQITITILASYINAVAFACQVLINQILGWS.KR[ISLQ	185
H.s.	VILVPRLAALTQGSYLHGRVQWRIV[EQV[PDRA.....MEAVITGLIVEALGPEVILSRLLGNLVKXKKAQFVMTQKIL[ELQSRITITPMLOSRLITPMLOSILGHLAWDS.QRR[PLILQVIR[ELILETWGSSA[IRHT	352
C.e.	VQMI[EWIGSLNNFDSQMRMFIILFQ[EFITYQGIQVHESILITILFLISKSDQILKR[CI[EAIDITG.TIKR[VVMVKLPFQR.....VLKHKI[TEI[ILINFVYRKEQFAIQILLETISVKIWSDI[NYAKS	341
D.m.	K[IGSQEHLAISKLIVLAQID.....SID	208
H.s.	FUPQ[QRHVSKAVILICLAQIGEP[ELRDSRDEILASMA[GVKCR[LSLSPVRRIGMIVAEVVSARIHPEGPPLKFOVEEDELSEILALASPOPAGDGAS[AGTISVPATAEPPA.....ETPA	470
C.e.	APESQ[ERHIVRMILIVLH[FRITCSSIDWESILFTNSM[GVHCR.NSMLEMYVQSGIFVNOALCKOATKRSKTHGSD[EQPPEILEENKFV[SEVGKIWFTEMTSIL[EHGNSSTVKDSERVREITAN	465
D.m.		287
H.s.	ETVDSGGVPQ.....DEFLDEDDDELKPYIMSNIDYT.....TTIDQ[REKFVIDLILHILR..EKVENYQVFECAIGTAFBQLIRGQAKHDITQIALDLILQIFL	565
C.e.	ELTKDDSGEEFE[INAQRLQNKDSALITTSKNILRLDSDDDEDEP[DYQVNESEKIFKNLEIGE[EPKNKVTFEAYADAFEMILEK[KEYEVFEAAFFNTJNLI[NRFPICGFPQIAEKILFTRILHIQN	590
D.m.	VMEMQFVMEQFERITQFKCQVAICVAHPG[CAEYLCRQ[EHINSEFYSASVRITILQVILAAIAKELSGDENM[QNEMEIVDVIT[PPAAKH[ERKFE[QOEE[.SPA[ARLAAQRI[TRURLAKTIRYFSK	411
H.s.	KTCVVGHAGLRQR...ALVAVTVIDPAFVACVILTSQFVAIN.YSLRQ[MDILIVITLAAJELSRPGCLGRTPQ[GSPTSNTPCILPEAAVSQHG...S.AVASI[WRV[VEERIRSKTQRLS..	677
C.e.	NFGITPKFKEITVDEIAVACITIQRPEITVESVVRILITAPGQGFSTKQRI[HYITHNAD[GMALDKI[EECVMAQOLRIGGPTLSIILHRTINIDVDEDEDPHRLLVPEWRRVWDARTIANIRRIG..	713
D.m.	PRAGQOMERAN...P[HEHVAGHFFF[SLVRGQRIRQMLYVKY[EIFDTQLIVNILNTMSVIMCSQNC[HLFAMITREIFDLCAFVRENAA[FRANITOLIGIATITTEFAVILAQHFAESTINELQRWI	533
H.s.	.KSGPRQGFAGSPSRFNSVAGHFFFELIQRFDRLP[ITFDLIGEDQIVLGRIAHTILGALMCLAVNTTIVAVAMKALILEFWAURFHIDAVRQGHILSAVSSVILSLFAARHIEDIMDELL[EARSWI	801
C.e.	..TTIREPPRAGVVRNRLAQA[KVMFYFLLVLP..RGENASLIGKDSILLASIMVASMVYVRGVC[QIHRMSSELTSVATPHR[SENAKLRDACTI[AHINVTITILPGDILMDEIFDVPAIIG...W	831
D.m.	NDFIR[SPLVGGETIS.....EECREIA	554
H.s.	ADV[AEKDPD.....ELCRILALRALLILQRLKNRILPPAS.P	837
C.e.	ELWANSVLVNASSSQLEKQMTIRQFCHSVYKHLQRHHPAVILQHODV	876

26/35

Fig. 39

H.s.	MEPAPSEVRLAVREAIHALSSSEGGHIFCTLESIKRYLGEMEPALPREKEEFASAHFSPVLRCLASRLSPAWLELPHGRLEELWASFFLEGPADQAFTVIMETIEGAAGPSFRIMKMARLLA	125
M.m.		0
S.s.		0
H.s.	RFLREGRLAVLMEAQCRQOTRPGFILLIRETLLGKVVALPDHLGNRLQOENLAEFFPQNYFRLLGEEVVRVLQAVVDSLQGGLDSSVSFVSQVLGKACVHGRQOEILGVLVPRLAALTQGSYLHQR	250
M.m.		0
S.s.		0
H.s.	VCWRLVEQVDPORAMEAVLTGLVEAALGPEVLSRLLGVLVKNKKAQFVMTQKLLLELQSRLTTEMLOSILGHILAMDSQRRPLILCVLKELLETWGSSSAILRHTPIEQQRHVSKAVILICLALGEPE	375
M.m.	VTRKLLILQYQHITPMVQSLLGYLALDSQRRPLILCVLKELLETWGSSSIVRHTPIEQQCYTISKATILVCLAHLGEPE	77
S.s.		0
H.s.	IFDSRDELLASMMAGVKCRIDSSLPFVRRIGMIVAENV SARITHEGPEHLKFQYHEDELSIELLAIASQPAGDCGASEAGTSLVPATAEPPTAEIVDGGVPOAQLAGSDSDLDSDDEFVPYDM	500
M.m.	LQDHRDELLASMMAGVKCRIDSSLPFVRRIGMIVAENVISSRITHEGPHILKFQYEDDENSRRELLAIATPEPAGDCS.....	152
S.s.		0
H.s.	SGDRRLKSSKRAHAYVVRDCVEALITISEDIERWEAALRALEGI VRSSEFIRREVSVELAKVLLHLEEKTCVGHAGLRQRAIVAVTVTDEPAFVADYLT SQFVALNYSLRQRMIDILDVLTIAAQELSR	625
M.m.		182
S.s.	KAPVYVRDCVEALITASEDWEHWEAALRALEGI VRSSPAAREVSVELAKVLLHLEEKTA VAGHEGLRQRAIVAVTVTDEPAFVADYLT SQFVALNYSLRQRMIDILDVLTIAAQELSR	116
H.s.	PGCLGRTPQFGSPSPNTPCILPEAAVSQFGSAVASLWRVVVEERIRSKTQRLSKGGRQGPAGSPSRFNSVAGHFFFPILLQRFDRPLVTFDLIGEDQLVLGRLAHTIGALMCLAVNITTVAVAMGKA	750
M.m.	PKSTQRRSQHGHPVEGTMCSFAI AVSQTGNVVAFLWVVVVEERIRSKTIRRESKGCQFQRDVSGVNEHSSVAGYFFFPILLQHED.....LTSEVAVLVGKA	277
S.s.	PGRLGR	122
H.s.	LLEFVWALRFHIDA YVRQGLLSAVSSVLLSTPAARLHEDIMDELLEARSWLADVAEKDEDEDCEJLALRALLLILQRLKNRLLPPASP	837
M.m.	LLEFVWALRFHVDIYVVRGLLSASVSSVLLSMPTERLIGDILDELLEARSWLALVAEKDVDEDCRELA VRALLI	350
S.s.		

Fig. 40

28 / 35

Fig. 41

29/35

G. max		0
M.truncatula		0
A. thaliana	MIIERARV[ELSRWQQAAMGSA]LCAIVDERR.ADLIAALGETTG	44
G. hirsutum	IKLSPWQQAAMAVGSAVCAILDERR.ADLIAALGETTG	37
S.pombe	MFYLNALHEINKVVDVMSLSKKFLKPSVASNQLRLLFATAERKVNYPGHV[ELSP]QRIFLVAGSALMGKAPMRGGDMISVLGDSG	88
M. musculus	YPDHIE[ETTP]QKMLIAAGAAAGVALYNBYR.HDMVAVLGETTG	41
R. rattus	HBYR.HDMIEVLGETTG	16
H. sapiens	MATLLRPVLRRLCGLPGLQRPAAEMPLRARS DGACPLYSHHI[ETSP]QKAI[IAAGSAAMALYNBYR.HDMVAVLGETTG	78
B. taurus	YPEHI[ETSP]QKVI[IAAGSAGMALYD]PYR.HDMVAVLGETTG	41
C. elegans	MSAQKLYASHV[ELAPLSRMILGIGSAVTAISDEKR.GDMVAVMGETTA	47
A. caninum		0
D. melanogas	MMQRCIRLQKPLALRRGLRLAQANSQAVATEAPEAEPLDAFERQYLKERIEISPE[QRLFI]GAGSSIAALINERR.HDMIA[CLGETTG	86
S. cerevisia	MLRLSLRSTATLPVKCQRRGLILPAAAMYTLGSLIFGKEARLADAMERGELHNKNVDYAKEAEERTELRIRALANTRPMEPRYNCHV[ELHR]YEKLII[FAISGNSEFFHED.GYNIVQIGETTA	124
T. cruzi	FVNA[TT]RSIWDPVNA[DDVAAVGETTPA	26
G. max	LPANTFGAAAYARFV[SGSRNFSPDDREPPVRFMDTDELA]YVAMRAREVHDEFWH[ILFDLP]TNLIGETA[KVIEFEQ]MGLP[MC]LLSVICG	85
M.truncatula	NFSPPDDREPPVRFMDTDELA]YVAMRAREVHDEFWH[ILFDLP]TNLIGESA[KVIEFEQ]MHL[PMQ]VMSVLGG	68
A. thaliana	RVVSEQVGHWDL[RENTFTGAAAYAKF]MGSRNFSPDDREPPVRFMETDELA]YVATRAREVHDLW[ILFGLP]TNLIGESSIKVIEFEQ]MYL[PMQ]MLSVICG	166
G. hirsutum	RVISANVGHWDL[RENTFTGAAAYARFTGSXNFSPDDREPPVRFMDTDELA]YVAMRAREVHDEFWH[ILFDLP]TNLIGESSIKVIEFEQ]MYL[PMQ]MLSVICG	122
S.pombe	QPF[ILHRL]NKMLV[KTGRE]ILKDKPRMTSKSINLP[ILR]ILPNTILGKIYVDWIDKEHVGP[TRSE]TRFVDDPE[IA]YVMQRYRESHD[EF]HA[ICNMPTN]IEGELA[KWIEFVNMGLPV]GALSALFG	213
M. musculus	CHT[KFLRDQ]KKDPEGAQILQERPRISLSTIDLSKIQSLPEGSLGREYLRFDVNVKVSPTTRAP[TRFVDD]DELA]YV[IQRYREVHDM]IEH[ILFGLP]TNLIGESA[KVIEFEQ]MHL[PMQ]VMSVLGG	165
R. rattus	CHT[KFLRDQ]KKDPEGAQILQERPRISLSTIDLSKIQSLPEGSLGREYLRFDVNVKVSPTTRAP[TRFVDD]DELA]YV[IQRYREVHDM]IEH[ILFGLP]TNLIGESA[KVIEFEQ]MHL[PMQ]VMSVLGG	99
H. sapiens	HR[TKVLRDQ]RRDPEGAQILQERPRISLSTIDLSKIQSLPEGSLGREYLRFDVNVKVSPTTRAP[TRFVDD]DELA]YV[IQRYREVHDM]IEH[ILFGLP]TNLIGESA[KVIEFEQ]MHL[PMQ]VMSVLGG	202
B. taurus	RR[TKVLRDQ]KKDPEGAQILQERPRISLSTIDLSKIQSLPEGSLGREYLRFDVNVKVSPTTRAP[TRFVDD]DELA]YV[IQRYREVHDM]IEH[ILFGLP]TNLIGESA[KVIEFEQ]MHL[PMQ]VMSVLGG	136
C. elegans	IGPV[ENIRKRMESDV]GKRLILIEKPRI[SNGTIDRKWLRQLPDGTLGKLYSNFLDR]LNT[SPDAREPTV]KYIDNLEH[ILYVMQRYRETHD]ETHALEQ[KTNNM]GEVTVKYFEQ]YQGL[PMQ]VTGGIFG	172
A. caninum	KEVQNSEH[ILYVMQRYRETHD]ENH[ILQMP]THMGEVTVKYFEQ]YQGL[PMQ]VTAGLFG	58
D. melanogas	EDALWTILDITMOASEGQRI[MA]DKPRI[HTSTIDFKY]IE[ILP]PDFTFGAAVVKFT[KDNQV]IPDSRM[AVR]IE[EDPK]LAYIMTR[RECHD]LIH[ILV]MLMPTNNMGEVAVKWW[EA]LNTGL[PMQ]CYGGAFFG	210
S. cerevisia	LPVFIENLKQTM[LS]SSGRRI[IK]EQ[ENITTEILHMDK]LAKLPHNTFGVYVQWIKRENVSPTTRAPVKEIDDPMHAYIFKRYRQCHDFYHA[ITNM]EIIIEGELTTKALEGANI[GV]PMAILGGIILA	249
T. cruzi	LTALGHMKQSNMSDRI[GRM]ILRTQPRVTD[ETLEFASRQ..PPGT]FGHRYAQFMKFXXT[PN]GRTPVAHIADPTLA]YVMQ[QRETHD]ELH	113
G. max	TARFSEKQ[RLKLFYHYFP]WALHA[CMPS]TDL[MCVY]YERHFEDE[LE]DVRRK[ILQI	137
M.truncatula	TARFSEKQ[RLKLFYHYFP]WAVRAGTQCNDLMCVY]YEQHFHODLE[DVRRK]ILGI	120
A. thaliana	TVRFNEKQ[RS]MFLKHYL[PWAVRAGRCQCTDLMCVY]YERHFESE[LEQV]EPKMGIIIPAPQHPK	226
G. hirsutum		
S.pombe	PIRINCEQASRFRRVYIPW[SIRNGLNAKTLIN]VYWEKELTNDIEDVRRRIRIEAAPLV	272
M. musculus	PIRIRTQSLQVLFSELIPWALQNGRRATCVLN]IYV]EQRWEQPLTAIR[EEH]GIS	218
R. rattus		
H. sapiens	PIRIGAQSLQVLVSELIPWAVQNGERRAHCVLN]IYV]ERRRWEQSLRAIR[EEH]GITAPPMHVQGLA	265
B. taurus		
C. elegans	GARILTKNRQELVDRNI[PWVVEQ]ATNARFFMAFDWENHFEKQ[SEVQKH]NITPLSVNM	231
A. caninum	AARIKKNH[HRFLTOH]IPMLVEQATKGRFFMAIDWENHWEETIPSLQEQEGITPL	113
D. melanogas	AVRIRPKQRRAYLKHYL[PWALENGKRTKPLMPVYWEKRWEQNIHELRS]ELGITVLNKA	268
S. cerevisia	PIR[IKK]VQRKRLYN]IYIPWAVRTGLSCKPLIN]VYWEEMLEKDVTAIRK[ELKH]ITLLRI	306
T. cruzi		

Fig. 42

30/35

MMQRCLRLQK PLALRRGLRL AQANSQAVAT EAPEAEPLDA FERQYLKERI EISPFQRLFL
GAGSSIAALL NPRRHDMIAC LGETTGEDAL WTILDTMQAS EEGQRIMADK PRIHTSTIDF
KYLETLPPDT FGAAYVKFLK DNQVTPDSRM AVRFLDPKL AYLMTRYREC HDLIHTVLDM
PTNMLGEVAV KWVEALNTGL PMCYGGAVFG AVRLRPKQRR AYLKHYLPWA LENGKRTKPL
MPVYWEKRWE QNIHELRLSEL GITVLNKA (SEQ ID NO:38)

Fig. 43

MATLLRPVLR RLCGLPGLQR PAAEMPLRAR SDGAGPLYSH HLPTSPLQKA LLAAGSAAMA
LYNPYRHDMV AVLGETTGHR TLKVLRDQMR RDPEGAQILQ ERPRISTSTL DLGKLQSLPE
GSLGREYLR F LDVNRVSPDT RAPTRFVDDE ELAYVIQRYR EVHDMMLHTLL GMPTNILGEI
VVKWFEAVQT GLPMCILGAF FGPIRLGAQS LQVLVSELIP WAVQNGRRRAH CVLNLYYERR
WEQSLRALRE ELGITAPPMH VQGLA (SEQ ID NO :39)

Fig. 44

MFYLNHAHLEI NKVVDVVM SL SKKFLKPSVA SNQLRLLFTA AERKVNYPGH VPLSPLQRI F
LVAGSAIMGL KAPWRGGDMI SVLG DASGP FFLHRLLNKM LVDKTGREIL KDKPRMTSKS
LNL PFLRTL P PNTLGKIYVD WIDKEHVGPD TRSPTRFVDD PEEAYVMQRY RESHDFYHAI
CNMPTNIEGE LAIKWLEFVN MGLPVGALSA LFGPLRLNCE QASRFRRVYI PWSIRNGLNA
KTLIN VYWEK ELTNDIEDVR RRIRIEAAPP LV (SEQ ID NO :40)

Fig. 45

31/35

MIIERARVPL SRWQQAAMVAM GSALGALVDP RRADLIAALG ETTGKPAFEM VLERMKKSEE
 GRAILLERPR VVSEQVGHAW DLPENTFGAA YAKFMGSRNF SPDDRPPVRF METDELAYVA
 TRAREVHDLW HTLFGLPTNL IGESSLKVIE FEQMYLPMCM LSVIGGTVRF NEKQSRMFLK
 HYL PWAVRAG RQCTDLMCVY YERHFSEDL E QVRRKWGIIP APQHPK (SEQ ID NO:41)

Fig. 46

ALYNPYRHDM VAVLGETTGC HTLKFLRDQM KKDPEGAQIL QERPRISLST LDLSKLQSLP
 EGSLGREYLR FLDVNVKSPD TRAPTRFVDD EELAYVIQRY REVHDMMLHTL LGMPTNMLGE
 VVVKWFEAVQ TGLPMCILGA LFGPIRLRTQ SLQVLFSELI PWAIQNG (SEQ ID NO:42)

Fig. 47A

YPDHIPTTPL QKMLLAAGAA GMALYNPYRH DMVAVLGETT GCHTLKFLRD QMKKDPEGAQ
 ILQERPRISL STLDLSKLQS LPEGSLGREY LRFLDVNVKVS PDTRAPTRFV DDEXLAYVNQ
 KYR (SEQ ID NO:43)

Fig. 47B

HLPTSYP SLS LGEVVVKWFE AVQTGLPMCI LGALFGPIRL RTQSLQVLFS ELIPWAIQNG
 RRATCVLNIY YEQRWEQPLT ALREELGIS (SEQ ID NO:44)

Fig. 47C

YPDHIPTTPL QKMLLAAGAA GMALYNPYRH DMVAVLGETT GCHTLKFLRD QMKKDPEGAQ
 ILQERPRISL STLDLSKLQS LPEGSLGREY LRFLDVNVKVS PDTRAPTRFV DDEELAYVIQ
 RYREVHDMMLH TLLGMPTNML GEVVVKWFEA VQTGLPMCIL GALFGPIRLR TQSLQVLFS
 LIPWAIQNGR RATCVLNIYY EQRWEQPLTA LREELGIS (SEQ ID NO:45)

Fig. 47D

32/35

LPANTFGAAY ARFMGSRNFS PDDRPPVRFM DTDELAYVAM RAREVHDFWH TLFDLPTNLI
GETALKVIEF EQMGLPMCLL SVIGGTARFS EKQRKLFYHH YFPWAIHAGM PSTDLMCVYY
ERHFDEDLED VRRKLQI (SEQ ID NO:46)

Fig. 48

YPEHIPTSVL QKVLLAAGSA GMALYDPYRH DMVAVLGETT GRRTLKVLRD QMKRDPEGAQ
ILQERPRISL STLDMGKLRS LPEGSFGCAY LHFLDVNRVS PDTRAPTRFV DDEELAYVIQ
RYREIHDMLH ALLGMP (SEQ ID NO:47)

Fig. 49

NFSPDDRPPV RFMDTDELAY VAMRAREVHD FWHTLFDLPT NLIGESALKV IEFEQMHLPM
CVMSVLGGTA RFSEKQRKLF YQHYFPWAVR AGTQCNDLMC VYEQHFHQD LEDVRRKLGI
(SEQ ID NO:48)

Fig. 50

KFVQNSEHLY VMQRYRETHD FNHVLLQMPT HMLGEVTVKY FEGIQFGLPM CVTAGLFGAA
RLRKNHRHRF LTQHLPWIVE QATKGRFFMA IDWENHWEET IPSLQEQFGI TPL
(SEQ ID NO:49)

Fig. 51

33/35

FVAATTRSIW DPNADDVAA VGETPALTAL GHMKQSMMSD RTGRMILRTQ PRVTDETFEF
ASRQPPGTFG HRYAQFMKFX XFTPNGRTPV AHIADPTLAY VMQRQRETHD FLH
(SEQ ID NO :50)

Fig. 52

HPYRHDMLPV LGETTGCHTL KFLRDQMKKD PEGAQILQER PRISLSTLDL SKLQSLPEGS
LGREYLRFLN ANKVSPDTRA PTRFVDDEEL AYVIQRYRE (SEQ ID NO :51)

Fig. 53

IKLSPWQQAA VAVGSAVGAL LDPRRADLIA ALGETTGKPA FERVLERMRR SPEGKTXLLE
RPRVISANVG HAWDLPKNTF GAAYARFLGS XNFSPDDRPP VRFMDTDELA YVAMRAREVH
DF (SEQ ID NO:52)

Fig. 54

34/35

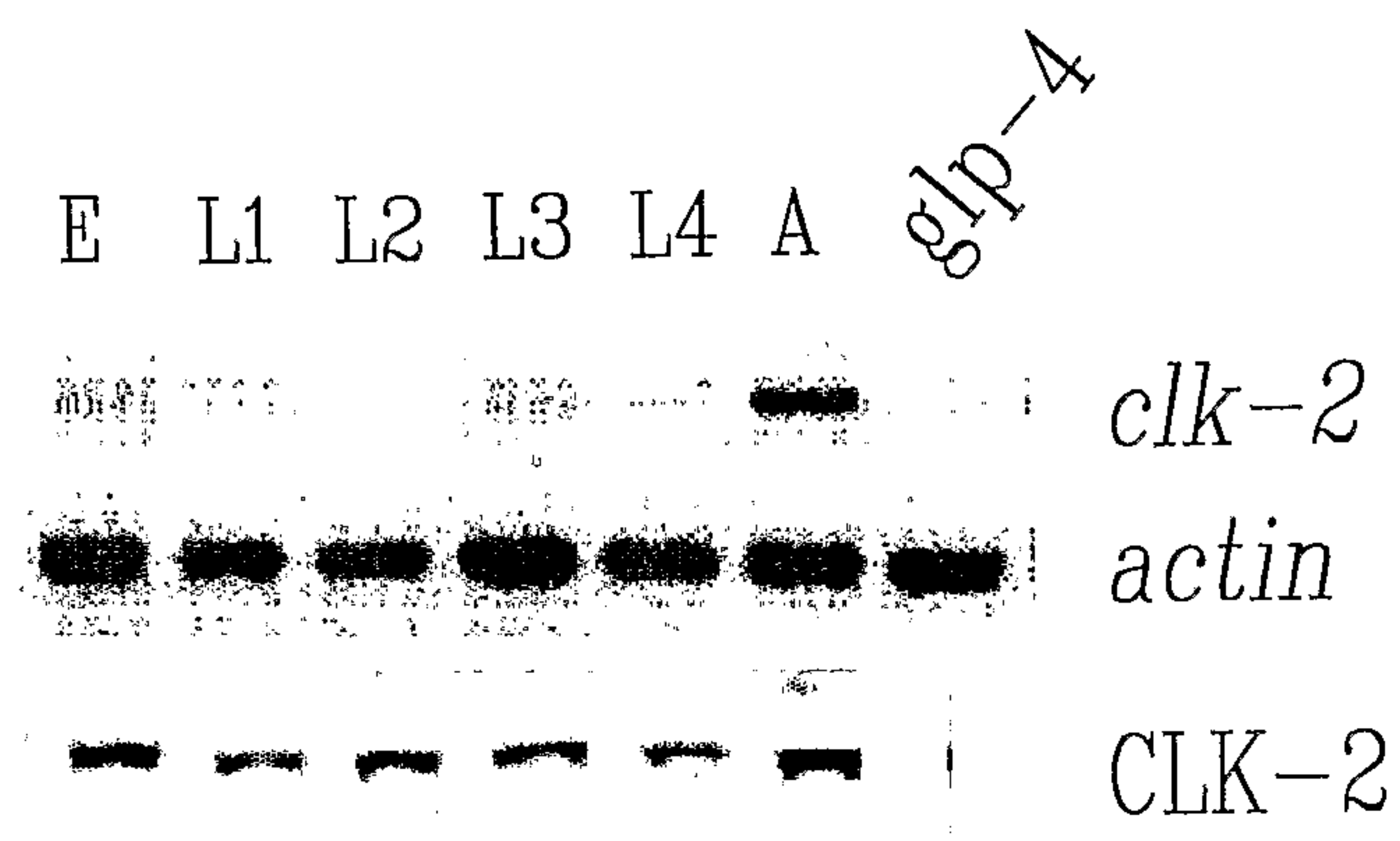


Fig. 55 A

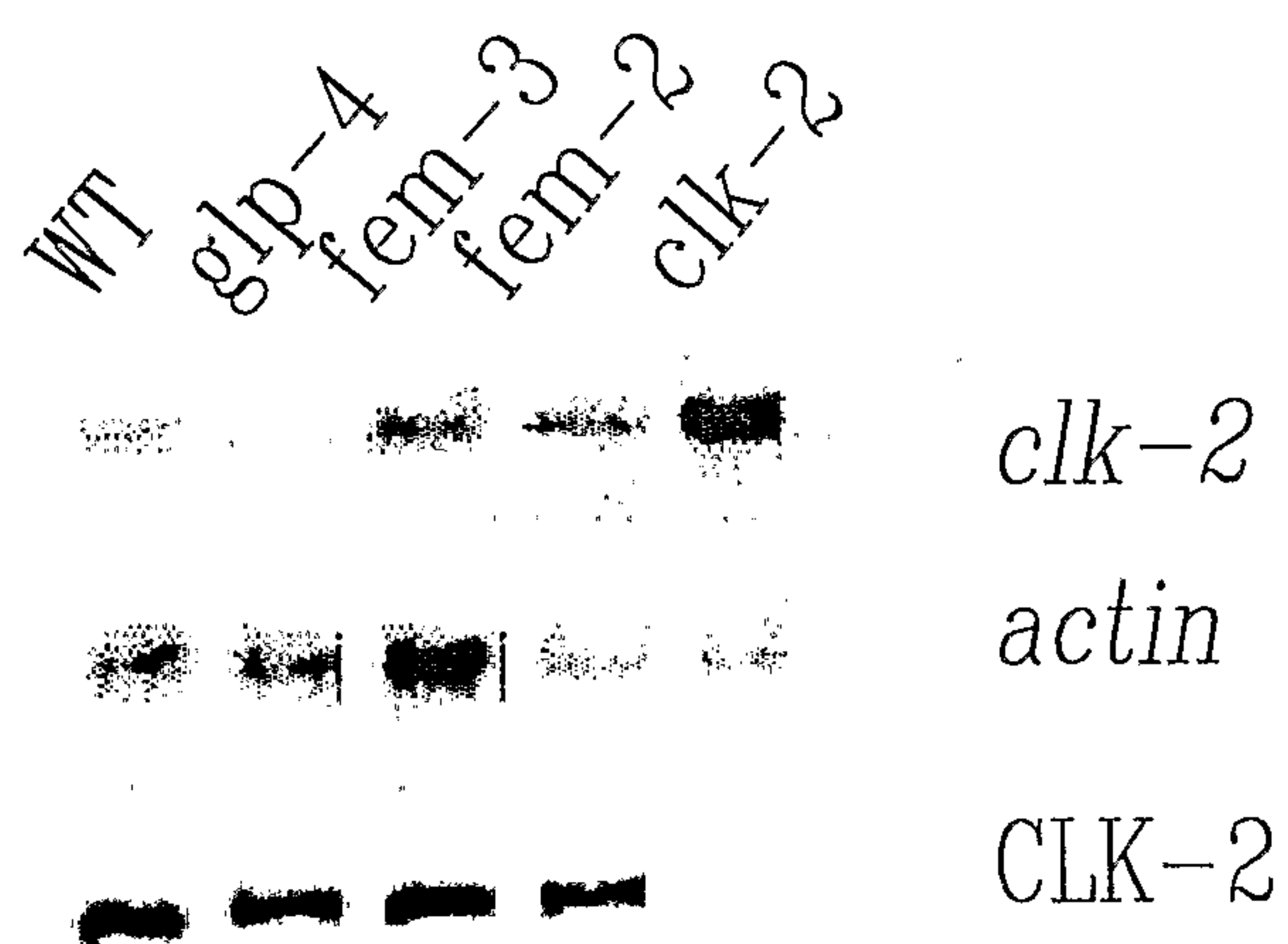


Fig. 55 B

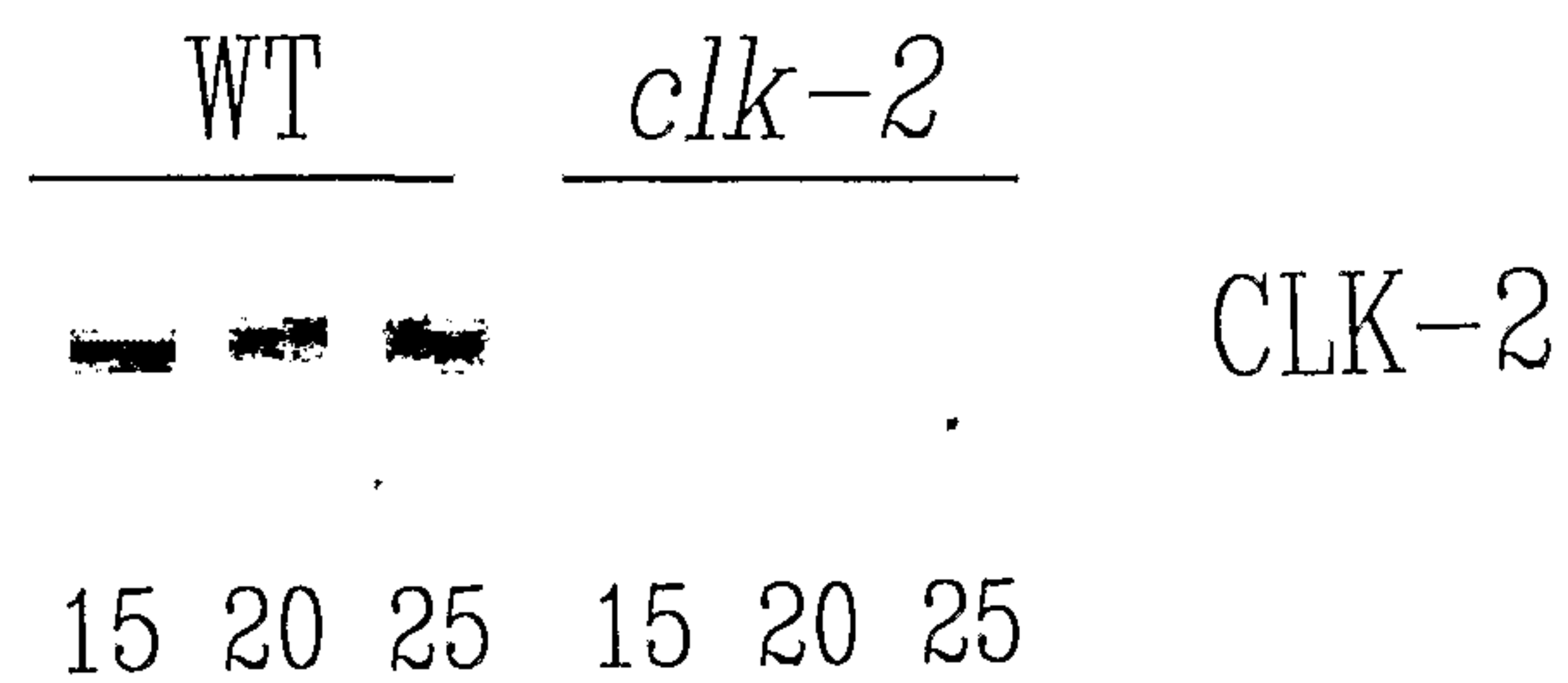


Fig. 55 C

35/35

