US 20250275551A1

(19) United States

(12) Patent Application Publication

(10) Pub. No.: US 2025/0275551 A1

(43) Pub. Date: Sep. 4, 2025

YU et al.

(54) **COMPOSITIONS AND METHODS FOR PRODUCING PROTEIN**

(71) Applicant: **Terraferma Foods, Inc.**, San Francisco, CA (US)

(72) Inventors: **Ruby YU**, San Francisco, CA (US); **Victoria MO**, San Francisco, CA (US)

(21) Appl. No.: **19/036,854**

(22) Filed: **Jan. 24, 2025**

Related U.S. Application Data

(63) Continuation of application No. PCT/US2023/028715, filed on Jul. 26, 2023.

(60) Provisional application No. 63/392,613, filed on Jul. 27, 2022, provisional application No. 63/402,544, filed on Aug. 31, 2022.

Publication Classification

(51) **Int. Cl.**

A23J 3/04	(2006.01)
A23K 20/147	(2016.01)
A23K 50/40	(2016.01)
C12P 21/02	(2006.01)

(52) **U.S. Cl.**

CPC ***A23J 3/04*** (2013.01); ***A23K 20/147*** (2016.05); ***A23K 50/40*** (2016.05); ***C12P 21/02*** (2013.01)

(57) **ABSTRACT**

Disclosed herein is a food product comprising an animal-free composition comprising a meat protein.

Specification includes a Sequence Listing.

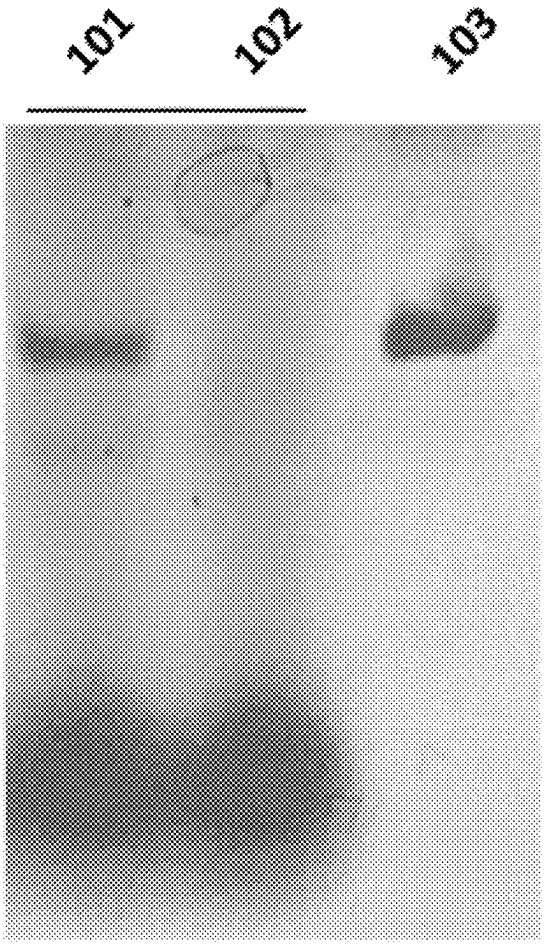

FIG. 1

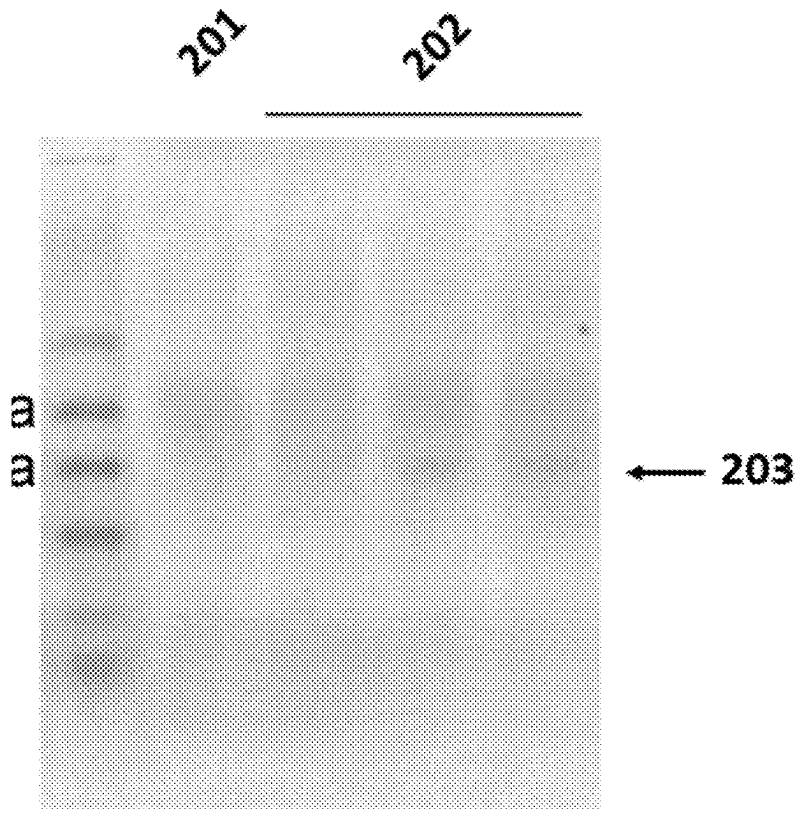

FIG. 2

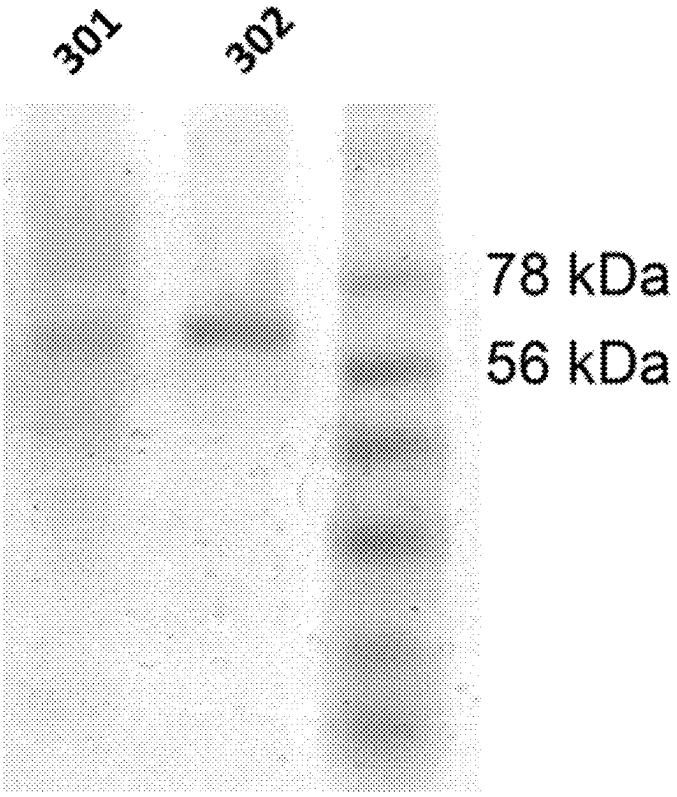

FIG. 3

COMPOSITIONS AND METHODS FOR PRODUCING PROTEIN

CROSS-REFERENCE

[0001] This application is a continuation of International Application No. PCT/US2023/028715, filed Jul. 26, 2023, which claims the benefit of U.S. Provisional Application No. 63/392,613, filed Jul. 27, 2022, and U.S. Provisional Application No. 63/402,544, filed Aug. 31, 2022, each of which is entirely incorporated herein by reference.

SEQUENCE LISTING

[0002] The instant application contains a Sequence Listing which has been submitted electronically in XML format and is hereby incorporated by reference in its entirety. Said XML copy, created on Jan. 21, 2025, is named 63116-701_601_SL.xml and is 458,343 bytes in size.

BACKGROUND

[0003] Livestock is the world's largest user of land resources on Earth. Land dedicated to the production of meat and dairy, which includes animal feed production, encompasses nearly 80% of global agricultural land. One-fifth of the word's meat production is consumed by companion animals (e.g., pets). Recombinant protein production can be used to supplement or offset meat production.

SUMMARY

[0004] An aspect of the present disclosure provides a method for producing a food product, the method comprising: a) contacting a plurality of *Pichia* cells with a recombinant nucleic acid; b) expressing a protein encoded by the recombinant nucleic acid; and c) isolating the protein from the plurality of *Pichia* cells, wherein the protein comprises an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 54-232.

[0005] In some embodiments, the food product is a cat food product. In some embodiments, the protein is an albumin protein. In some embodiments, the albumin protein is a rabbit albumin protein. In some embodiments, the albumin protein is a duck albumin protein. In some embodiments, the duck albumin protein is a mallard duck albumin protein. In some embodiments, the duck albumin protein is a Muscovy duck albumin protein. In some embodiments, the albumin protein is a cow albumin protein. In some embodiments, the albumin protein is a sheep albumin protein. In some embodiments, the albumin protein is a pig albumin protein. In some embodiments, the albumin protein is a bison albumin protein. In some embodiments, the albumin protein is a deer albumin protein. In some embodiments, the albumin protein is a chicken albumin protein. In some embodiments, the albumin protein is a turkey albumin protein. In some embodiments, the albumin protein is a salmon albumin protein. In some embodiments, the albumin protein is a cod albumin protein. In some embodiments, the albumin protein is a tuna albumin protein.

[0006] In some embodiments, the *Pichia* cells are *Pichia pastoris* cells. In some embodiments, the protein comprises an amino acid sequence that is at least 90% identical to any one of SEQ ID NOs: 54-232. In some embodiments, the protein comprises an amino acid sequence that is at least 95% identical to any one of SEQ ID NOs: 54-232. In some embodiments, the food product is substantially animal-free.

[0007] Another aspect of the present disclosure provides a food product comprising a recombinant protein, wherein the recombinant protein comprises an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 54-232. In some embodiments, the food product is substantially animal-free. In some embodiments, the recombinant protein comprises an amino acid sequence that is at least 90% identical to any one of SEQ ID NOs: 54-232. In some embodiments, the recombinant protein comprises an amino acid sequence that is at least 95% identical to any one of SEQ ID NOs: 54-232. In some embodiments, the food product is a cat food product.

[0008] In some embodiments, the recombinant protein is a recombinant albumin protein. In some embodiments, the recombinant albumin protein is a recombinant rabbit albumin protein. In some embodiments, the albumin protein is a recombinant duck albumin protein. In some embodiments, the recombinant duck albumin protein is a recombinant mallard duck albumin protein. In some embodiments, the recombinant duck albumin protein is a recombinant Muscovy duck albumin protein. In some embodiments, the recombinant albumin protein is a recombinant cow albumin protein. In some embodiments, the recombinant albumin protein is a recombinant sheep albumin protein. In some embodiments, the recombinant albumin protein is a recombinant pig albumin protein. In some embodiments, the recombinant albumin protein is a recombinant bison albumin protein. In some embodiments, the recombinant albumin protein is a recombinant deer albumin protein. In some embodiments, the recombinant albumin protein is a recombinant chicken albumin protein. In some embodiments, the recombinant albumin protein is a recombinant turkey albumin protein. In some embodiments, the recombinant albumin protein is a recombinant salmon albumin protein. In some embodiments, the recombinant albumin protein is a recombinant cod albumin protein. In some embodiments, the recombinant albumin protein is a recombinant tuna albumin protein.

[0009] Additional aspects and advantages of the present disclosure will become readily apparent to those skilled in this art from the following detailed description, wherein only illustrative embodiments of the present disclosure are shown and described. As will be realized, the present disclosure is capable of other and different embodiments, and its several details are capable of modifications in various obvious respects, all without departing from the disclosure. Accordingly, the drawings and description are to be regarded as illustrative in nature, and not as restrictive.

INCORPORATION BY REFERENCE

[0010] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference. To the extent publications and patents or patent applications incorporated by reference contradict the disclosure contained in the specification, the specification is intended to supersede and/or take precedence over any such contradictory material.

BRIEF DESCRIPTION OF THE DRAWINGS

[0011] The novel features of the invention are set forth with particularity in the appended claims. A better under-

standing of the features and advantages of the present invention will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention are utilized, and the accompanying drawings (also "figure" and "FIG." herein), of which:

[0012] FIG. **1** shows a Coomassie stained sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gel of an example media from *Pichia pastoris* cells including a construct for secreted recombinant rabbit albumin or wild-type cells.

[0013] FIG. **2** shows a Coomassie stained SDS-PAGE gel of an example media from *Pichia pastoris* cells including a construct for secreted recombinant rabbit Tropomyosin alpha-1 chain or wildtype cells.

[0014] FIG. **3** shows a Coomassie stained SDS-PAGE gel of an example recombinant rabbit albumin.

DETAILED DESCRIPTION

[0015] While various embodiments of the invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions may occur to those skilled in the art without departing from the invention. It should be understood that various alternatives to the embodiments of the invention described herein may be employed.

[0016] Generation of animal-free protein is a way to reduce global meat production and offset the agricultural and environmental consequences of animal farming. Animal free protein can be made using microorganisms designed to produce meat proteins using recombinant protein expression. Foods, such as pet food (e.g., cat food, dog food, or food for other companion animals) or foods for human consumption can be developed from animal-free meat proteins as an alternative or supplement to traditional meat-based foods.

[0017] Using recombinant protein expression methods, microorganisms (e.g., a yeast or bacteria) can be contacted with a DNA expression cassette. This DNA sequence can contain a gene or genes of interest, as well as regulatory sequences which will help in the production of the gene or genes of interest. Regulatory sequences (e.g., UTRs, promoters, signal sequences, etc.) can aid in, for example, transcription or translation efficiency, protein product yield, protein expression, or protein secretion.

[0018] Microorganisms can be induced to express and secrete meat proteins of specific animals (e.g., a rabbit protein or a duck protein). Microorganisms can alternatively or additionally be induced to express and secrete specific types of meat proteins (e.g., an albumin protein). Animal-free meat proteins can be purified and formulated into food products (e.g., a pet food or food for human consumption).

Definitions

[0019] Whenever the term "at least," "greater than," or "greater than or equal to" precedes the first numerical value in a series of two or more numerical values, the term "at least," "greater than" or "greater than or equal to" applies to each of the numerical values in that series of numerical values. For example, greater than or equal to 1, 2, or 3 is equivalent to greater than or equal to 1, greater than or equal to 2, or greater than or equal to 3.

[0020] Whenever the term "no more than," "less than," or "less than or equal to" precedes the first numerical value in a series of two or more numerical values, the term "no more than," "less than," or "less than or equal to" applies to each of the numerical values in that series of numerical values. For example, less than or equal to 3, 2, or 1 is equivalent to less than or equal to 3, less than or equal to 2, or less than or equal to 1.

[0021] The terms "companion animal" and "pet" are used interchangeably herein. A pet is a domestic or tamed animal that is kept for companionship. A pet can be, but is not limited to, a dog, a cat, a pig, a rat, a mouse, a snake, a turtle, a cow, a sheep, a llama, a donkey, a goat, a horse, a chicken, a parrot, a ferret, a guinea pig, a hamster, a rabbit, a fish, a frog, a toad, a salamander, an iguana, a lizard, a duck, a goose, or an insect.

[0022] The term "meat" can be defined as skeletal muscle and its associated tissues, including blood, derived from mammalian, avian, reptilian, amphibian, and aquatic species harvested for consumption. Edible offal comprising organs and non-skeletal muscle tissue can also be considered meat.

[0023] The term "animal-free" can be defined as something that is not harvested directly from an animal, living or dead (e.g., meat harvested from a dead animal).

Food Products

[0024] Described herein are food products comprising animal-free proteins. An animal-free protein can be a protein or a fragment or truncation of a protein that is not obtained from an animal (e.g., by harvesting meat from an animal). An animal-free protein can be obtained through in vitro or ex vivo methods such as cell-free protein synthesis (CFPS). Alternatively, an animal-free protein can be obtained through in vivo methods such as genetic engineering of cells for recombinant protein expression. Genetically engineered cells used for recombinant protein expression can be, but are not limited to, microorganism cells (e.g., bacteria, yeast), plant cells, insect cells, or animal cells.

[0025] In recombinant protein expression, a cell is contacted with a DNA sequence. The DNA sequence will be incorporated by the cell, and the cell's internal protein expression machinery will transcribe and translate the DNA sequence into a peptide.

[0026] A cell used in recombinant protein expression can be a microorganism cell. Microorganisms are microscopic organisms such as bacterium, viruses, or fungi (e.g., yeast). A microorganism cell used in recombinant protein expression can be a bacterial cell. A bacterial cell can be, but is not limited to, a cell from *Escherichia coli, Bacillus licheniformis, Bacillus subtilis*, or *Lactococcus lactis*. A microorganism cell used in recombinant protein expression can be a yeast cell. A yeast cell can be, but is not limited to, a cell from *Schizosaccharomyces pombe, Saccharomyces cerevisiae, Hansenula polymorpha* (also known as Ogataea *polymorpha*), *Pichia farinosa, Pichia anomala, Pichia heedii, Pichia guilliermondii, Pichia kluyveri, Pichia membranifaciens, Pichia norvegensis, Pichia ohmeri, Pichia pastoris* (also known as Komagataella phaffii), *Yarrowia lipolytica, Pichia methanolica*, or *Pichia subpelliculosa*. A microorganism cell used in recombinant protein expression can be a filamentous fungal cell. A filamentous fungal cell can be, but is not limited to, a cell from *Trichoderma reesei* (also known as Hypocrea jecorina), *Aspergillus niger, Aspergillus*

oryzae, Aspergillus tubingensis, or *Ashbya gossypii* (also known as Eremothecium *gossypii*).

[0027] *P. pastoris* cells can be used in recombinant protein expression methods. *P. pastoris* cells have robust protein secretion; the ability to grow to high biomass density, allowing higher yields; have a high protein yield with up to 35 g/L yield for secreted proteins; have a low endogenous secretome, allowing for less downstream purification; grow in relatively inexpensive feedstock/media (e.g., glycerol or methanol); have high genetic tractability; have a proven ability to properly express and fold distantly-related eukaryotic (e.g., plant or animal) proteins; have a 90 minute doubling time; have an ability to grow in large-scale fermenters; and have an approximately one-week fermentation time from start to harvest.

[0028] A cell used in recombinant protein expression can be chosen because it displays a desired characteristic such as having a rapid doubling time, the ability to properly fold a target recombinant protein; the ability to make appropriate post-translational modifications to a target recombinant protein; the ability to efficiently secrete the target recombinant protein; the ability to be grown cheaply; the ability to be grown on easily accessible media; or genetic tractability. Alternatively, a cell used in recombinant protein expression may not secrete or may not efficiently secrete the target recombinant protein. In an example, cells that do not secrete or do not efficiently secrete the target recombinant protein may be lysed to release the protein. Cell lysis may include mechanical disruption, liquid homogenization, sonication, freeze/thaw cycling, manual grinding, or any combination thereof. In some examples, cells that secrete the target recombinant protein may be further lysed to increase protein yield.

[0029] During recombinant protein expression, a cell is contacted with a DNA sequence that can comprise an expression cassette. An expression cassette can be a component of vector DNA which can comprise at least one gene and/or at least one regulatory sequence.

[0030] An expression cassette can comprise at least one gene of interest. An expression cassette can comprise at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10 or more genes of interest. Alternatively, or in addition to, an expression cassette can comprise at least one additional gene other than a gene of interest. An expression cassette can comprise at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10 or more additional genes other than a gene of interest. An expression cassette can be a linearized plasmid. Alternatively, an expression cassette can be a linear PCR product.

[0031] A DNA sequence (e.g., a vector or an expression cassette) can be transfected into a cell using various chemical and physical methods such as electroporation, calcium phosphate exposure, spheroplasting, polyethylene glycol (PEG) exposure, alkali cation exposure, or liposome-based transfection.

[0032] A gene of interest can be an animal protein-coding gene. An animal protein-coding gene can encode a full length or a fragment or truncated version of a naturally occurring protein. Alternatively, or in addition to, an animal protein-coding gene can encode a modified version (e.g., a codon-optimized version) of a naturally occurring protein. Alternatively, an animal protein-coding gene can encode a synthetic or novel protein.

[0033] An animal protein-coding gene can encode an antibody, a contractile protein, an enzyme, a hormonal protein, a structural protein, a storage protein, or a transport protein. An animal protein-coding gene can be specific to a particular type of tissue such as, but not limited to, connective tissue, epithelial tissue, muscle tissue, or nervous tissue.

[0034] An animal protein-coding gene can encode proteins represented in meat or other parts of animals that are eaten. Examples of proteins found in meat can include, but are not limited to: actin, myosin, myosin regulatory protein, albumin, carbonic anhydrase, myoglobin, collagen, elastin, titin, nebulin, sarcalumenin, creatin kinase, phosphoglucomutase, phosphoglycerate mutase, triosephosphate isomerase, alpha-actinin, glycogen phosphorylase, fructose-biphosphate adolase, glyceraldehyde-3-phosphate dehydrogenase, L-lactate dehydrogenase, tropomyosin, myosin-binding proteins, myomesin, beta-enolase, pyruvate kinase, phosphoglycerate kinase, plectin, ATP-dependent 6-phosphofructokinase, malate dehydrogenase, hemoglobin, filamin, aconitate hydratase, parvalbumin, heat shock cognate 71, troponin, calsequestrin, obscurin, glycogen debranching enzyme, elongation factor 1-alpha 2, aspartate aminotransferase, adenylate kinase isoenzyme, glucose-6-phosphate isomerase, desmin, AMP deaminase, and globulin.

[0035] An animal protein-coding gene can encode albumin. Albumin is a family of globular proteins, the most common of which are serum albumins. Albumin proteins are water-soluble and are commonly found in blood plasma. Albumin proteins differ from other blood proteins in that they are not glycosylated. Albumin proteins include, but are not limited to: serum albumin, alpha-fetoprotein, vitamin D-binding protein, afamin, ovalbumin and extracellular matrix protein 1.

[0036] An animal protein-coding gene can be a gene from a specific animal such as, but not limited to: a canine, a feline, a bird (e.g., a songbird, a pigeon, a chicken, a duck, a goose, a swan, a crow, a dove, a parrot, a turkey), a bovine (e.g., a cow, a bison), a camel, a goat, a pig, a deer, a sheep, a horse, a rodent (e.g., a mouse, a rat, a squirrel, a hamster, a guinea pig, a chipmunk, a gopher, a vole), a fish (e.g., a tuna, a salmon, a trout, a sardine, a tilefish, a halibut, a cod, a flounder), a rabbit, a lizard, a snake, or a frog.

[0037] An animal protein can comprise a mouse protein (Table 1). A mouse protein can comprise a mouse muscle protein. A mouse protein can comprise a mouse blood protein. A mouse protein can comprise a mouse albumin protein. A mouse protein can comprise a sequence that has at least or about 70%, at least or about 75%, at least or about 80%, at least or about 90%, at least or about 91%, at least or about 92%, at least or about 93%, at least or about 94%, at least or about 95%, at least or about 96%, at least or about 97% at least or about 98% at least or about 99%, or about 100% sequence identity to any one of SEQ ID NOs: 1-53.

[0038] A food product may comprise a fragment or a truncation of a mouse protein. The food product may include a fragment or truncation of any one of SEQ ID Nos: 1-53. A sequence of the fragment or truncation may be identical or substantially identical to a portion of any one of SEQ ID Nos: 1-53. A sequence of the fragment or truncation may be at least about 70%, 80%, 90%, 95%, 98%, or more identical to a portion of a sequence of any one of SEQ ID Nos: 1-53. A length of the fragment or truncation may be at least about 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of the length of any one of SEQ ID Nos: 1-53. A length of the fragment or truncation may be less than or equal to about 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, or less than a length of any one of SEQ ID Nos: 1-53.

TABLE 1

Mouse Proteins		
SEQ ID NO	Description	Sequence
1	Mouse TITIN	MTTQAPMFTQPLQSVVVLEGSTATFEAHVSGSPVPEVSWFRDGQVISTSTLPGVQISF SDGRARLMIPAVTKANSGRYSLRATNGSGQATSTAELLVTAETAPPNFSQRLQSMTV RQGSQVRLQVRVTGIPTPVVKFYRDGAEIQSSLDFQISQEGDLYSLLIAEAYPEDSGT YSVNATNSVGRATSTAELVVQGEEVVPAKKTKTIVSTAQISETRQTRIEKKIEAHFDA RSIATVEMVIDGATGQLPHKTPPRIPPKPKSRSPTPPSIAAKAQLARQQSPSPIRHSPSP VRHVRAPTPSPVRSVSPAGRISTSPIRSVKSPLLIRKTQTTTMATGPEVPPPWKQEGYV ASSTEAEMRETTMTSSTQIRREERWEGRYGVQEQVTISGAAAAAASASVSSSFTAGAI TTGTKEVKQETDKSAAVATVVAAVDMARVREPAISAVEQTAQRTTTTAVHIQPAQE QARKEAEKTAVTKVVVAADKAKEQELKSRTREVMVTTQEQTHISHEQIRKETEKAF VPKVVISATKAKEQETRITGEITTKQEQKRITQETIRQETEEIAASMVVVATAKSTKLE AAVGVQEETAIQQDQMHLTHEQMMKETRKTVVPKVIVATPKIKEQDLVSRSREAITT KRDQVQITQEKKRKEVETTALSTIAVATAKAKEQETVLRSREAMATRQEHIQVTHG QVGVGKKAEAVATVVAAVDQARVREPREPTHVEESHSQQTTLEYGYKEHISTTKVP EQPRRPASEPHVVPQAVKPAVIQAPSETHIKTTDQMGMHISSQVKKTTDISTERLVHV DKRPRTASPHFTVSKISVPKTEHGYEASIAGSAIATLQKELSATSSTQKITKSVKAPTV KPGETRVRAEPTPSPQFPPADMPPPDTYKSQAGIEVKKEVGVSISGSTVREEHFEVLR GREAKVTETARVPAPAEVPVTPPTLVSGLKNVTVIEGESVTLECHISGYPSPKVTWYR EDYQIESSIDFQITFQGGIARLMIREAFAEDSGRFTCSAVNEAGTVSTSCYLAVQVSEE FDKETTLTEKFATEEKRFVESRDVVMTDTSITEEQAGPGEPAAPFFISKPVVQKLVEG GSVVFECQIGGNPKPHVYWKKSGVPLTTGYRYKVSYNKQTGECRLVISMTFADDAG EYTIVIRNKHGETSASASLLEEADYEALVKTQQEMLYQTQMSTFIQEPKVGEIAPGFA YSEYEKEYEKEQALIRKKMAKDTVMVRTFVEDQEFHISSFEERLIKEIEYRIIKTTLEE LLEEDGEEKMAVDISESEAIESGFDIRIKNYRILEGMGVTFHCKMSGYPLPKIAWYKD GKRIRHGERYQMDFLQDGRASLRIPVVLPEDEGIYTAFASNIKGNAICSGKLYVEPAA PFSAPTYMPTPEAVSRIRSVSPRSLSRSPIRMSPAMSPARMSPARMSPARMSPARMSP GRRLEETDESQLERLYKPVFVLKPASFKCLEGQTARFDLKVVGRPMPETFWFHNGQ QIVNDYTHKVVIKEDGTQSLIIVPASPSDSGEWTVVAQNRAGKSTISVTLTVEAVEHQ IKPAFVEKLKNVNIKEGSRLEMKVRATGNPNPDIVWLKNSDIIVPHKYPRIRIEGTRGE AALKIDSIISQDSAWYTATAINKAGRDTTRCKVNVEVEFAEPEPERKLIIPRGTYRAKE IAAPELEPLHLRYGQEQWEEGDLYDKEKQQKPFFKKKLTSLRLKRFGPAHFECRLTPI GDPTMVVEWLHDGKPLEAANRLRMINEFGYCSLDYGAAYSRDSGVITCRATNKYG TDHTSATLIVKDEKSLVEESQLPDGKKGLQRIEELERMAHEGALTGVTTDQKEKQKP DIVLFPEPVRVLEGETARFRCRVTGYPQPKVNWYLNGQLIRKSKRFRVRYDGIHYLDI VDCKSYDTGEVKVTAENPEGVTEHKVKLEIQQREDFRSVLRRAPEPKPEFHVHEPGK LQFEVQKVDRPVDTSETKEVVKLKRAERITHEKVSEESEELRSKFKRRTEEGYYEAIT AVELKSRKKDESYEELLKKTKDELLHWTKELTEEEKKALAEEGKITIPTFKPERIELSP SMEAPKIFERIQSQTVGQGSDAHFRVRVVGKPDPECEWYKNGVKIERSDRIYWYWP EDNVCELVIRDVTAEDSASIMVKAINIAGETSSHAFLLVQAKQLITFTQELQDVVAKE KDTMATFECETSEPFIKVKWYKDGIEVHAGDKYRMHSDRKVHFLSVLTIDTSDAED YSCVLVEDENIKTTAKLIVEGAVVEFVKELQDIEVPESYSGELECIISPENIEGKWYHN DVELKSNGKYSITSRRGRQNLTVKDVTKEDQGEYSFVVDGKKTTCKLKMKPRPIAIL QGLSDQKVCEGDIVQLEVKVSLENVEGVWMKDGQEVQHSDRVHIVIDKQSHMLLIE DMTKEDAGNYSFTIPALGLSTSGNVSVYSVDVITPLKDVNVIEGTKAVLECKVSVPD VTSVKWYLNDEQIKPDDRVQSIVKGTKQRLVINRTHASDEGPYKLMVGRVETSCNL SVEKIKIIRGLRDLTCTETQNVVFEVELSHSGIDVVWNFKGKEIKPSSKYKIEAHGKIY KLTVLNMMKDDEGEYAFYAGENTTSGKLTVAGGAISTPLTDQTVAESQEAVFECEV ANPESEGEWLKDGKHLALSNNFRGESDGHKRRLVIAAAKLDDAGEYTYKVATSKTS AKLKVEAVKIKKTLRNLTVTETQDAVFSVELTHPDVKGVQWIKNGVVLDSNDKYEI SVKGTLYSLKIKNCAMADESVYGFKLGRLGASARLHVETVKIIKKPKDVTALENATV TFEVSVSHDTVPVKWFHKNVEIKPSDKHRLVSERKVHKLMLQSISPSDAGEYTAMV GQLECKAKLFVETLHITKTMKSIEVPETKAASFECEVSHFNVPSMWLKNGVEIEMSE KFKIVVQGKLHQLIIMNTSTEDSAEYTFVCGNDQVSATLTVTPIMITSMLKDINAEEK DTITFEVTVNYEGISYKWLKNGVEIKSTDRCQMRTKKLTHSLNIRNVHFGDAADYTF VAGKATSTATLYVEARHIEFRKHIKDIKVLEKKRAMFECEVSEPDITVQWMKDGQEL QIADRIKIQKEKYVHRLLIPSTRMSDAGKYTVVAGGNMSTANLFVEGRDVRIRSIKKE VQVIEKQRAVVEFEVNEDDVDAHWYKDGIEINFQVQERHQYVVERRIHRMFISETRH SDAGEYTFVAGRNRSSVTLYVNAPEPPQVLQELQPVTVQSGKPARFCAVISGRPQPKI SWYKEEQLLSTGFKCKFLHDGQEYTLLLIEAFPEDAAVYTCEAKNDYGVATTSASLS VEVPEVVSPDQEMPVYPPAIVTPLQDTVTSEGRPARFQCQVSGTDLKVSWYCKDKKI KPSRFFRMTQFEDTYQLEIAEAYPEDEGTYAFVANNAVGQVSSTATLRLEAPESILHE RIGQQIEMEMKEIASLLSAEEDFQTYSSDLRLPNANETLELLSEPPARSTQFDSRQEGA APVFIREISDVEISVEDVAKLSVTVTGCPKPKIQWFFNGMLLTPSADYKFVFDGDTHS LIILFTRFQDEGEYTCLASNEYGKAVCSAHLRISPRGERSTEMESGEKKALEKPKGPCP PYFFKELKPVHCGPGIPAVFEYSVHGEPAPTVLWFKEDMPLYTSVCYTIIHSPDGSGT FIVNDPQRGDSGLYLCKAQNLWGESTCAAELLVLPEDTDVPDASCKEESTLGVPGDF LETSARGPLVQGVDSRQEITAFAEGTISKAALIAEETLQLSYERSVDDSEVGTGVTIGA QKLPPVVLSTPQGTGELPSIDGAVHTQPGRGPPPTLNLQAVQAQTTLPKEATLQFEEP EGVFPGASSAAQVSPVTIKPLITLTAEPKGNYPQSSTAAPDHALLSSVAAETLQLGEK KIPEVDKAQRALLLSQSLAEGCVESLEVPDVAVSNMRSEPQVPFQHTCTEGKILMAS ADTLKSTGQDVALRTEEGKSLSFPLALEEKQVLLKEEQSEVVAVPTSQTSKSEKEPEA IKGVKEVREQELLSKETLFPSMPEEQRLHLKTQVRRALQAAVAREQANLFSEWLRNI

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		DKVEVTAVNFTQEPKRILCTYLITSVSSLTEELTVTIEDIDPQMANLETGLKDALCSIV CEERNILMAEDPRIHEEDKIDVQGGRDHLSDAQKVETVIEAEADSKYLVSKEEVSWS KVESQLKDGDTNEVPQAETLKLAEESGTQKTSTEMSQEEAEGTLADLCPAVLKHLV DTISEEGDTVHLTSSISNAKEVHWYFKGNLVPSDGKFKCLKEQNAYTLVIEAVKTED EGEYVCEASNDSGKAKTSAKLTVGERAAPVIKRRIEPLEVALGHLAKFTCEIQGAPN VRFQWFKAGREIYESDKCSIRSSNYVSSLEILRTQVVDCGEYTCKASNEYGSVSCTAT LTVTEAYPPTFLSRPKALTTFVGKAAKFLCTVSGTPVIEIIWQKDGAALSPSPDCRVTD ADNKHSLELSNLTVQDRGIYSCKASNKFGADICQAELTIIDKPHFIKELEAVQSAINKK IHLECQVDEDRKVTITWSKDGQKLPAGKDYKIYFEDKIASLEIPLAKLKDSGTYTCTA SNEAGSSSSSAAVAVREPPSFVKKVDPSYLMLPGESARLHCKLKGSPVIQVTWFKNN KELSESNTVRMSFVNSEAILDITDVKVDDSGTYSCEATNDVGSDSCSTEVVIKEPPSFI KTLEPADIVRGANALLQCEIAGTGPFEVNWFKDKKQIRSSKKYRLFTQKTFVYLEISS FNSADVGDYECVVANEVGKCGCVATHLLKEPPTFVKKVDDFTALAGQTVTLQAAV RGSEPISVMWMKGQEVIKEDGKIKMSFSNGVAVLTIPDVQISLGGKYTCLAENEAGS QTSVGELIVKEPAKIIERAELIQVTAGDPATLEYTVSGTPELKPKWYKDGRPLVASKK YRISFKNNIAQLKFYSAELHDSGQYTFEISNEVGSSSCETTFTVLDRDIAPLFTKPLRN VDSVVGGACRLDCKIAGSLPMRVSWFKDGKELTASDRYQIAFVEGTASLEISRVDM NDAGNFTCRATNSVGSKDSSGALIVQEPPSFVTKPGSRDVLPGSAVCLKSAFQGSAPL TIKWFKGDKELVSGGSCYITKETSESSLELYAVKTSDSGTYTCKVSNVAGSVECSAD LFVKEPATFIEKLEPSQLLKKGDGTQLACKVTGTPPIKITWFANDRELRESSKHKMSF AESTAVLRLTDVAIEDSGEYMCEAQNEAGSDHCTGIVIVKESPYFTKEFKSIEVLKEY DVMLLAEVAGTPPPEITWFKDNTTLRSGRKYKTFLQDQLVSLQVLKFVAADAGEYQ CRVTNEVGSSTCSARVTLREPPSFIKKIEATSSLRGGTAAFQATLKGSLPITVTWLKDN DEITEDDNIRMTFENNVASLYLSGIEVKHDGKYVCQAKNDAGIQRCSALLSVKEPATI MEEAVSIDVTQGDPATLQVKFSGTKEISAKWFKDGQELTLGPKYKISVTDTVSILKIIS TEKKDSGEYTFEVQNDVGRSSCKASINVLDLIIPPSFTKKLRKMDSIKGSFIDLECIVA GSHPISIQWFKDDQEISASDKHKFSFHDNTAFLEISQLEGTDSGTYTCSATNKAGHSQ CSGHLTVKEPPYFVEKPQSQDVNPGTRVQLKALVGGTAPMTIKWFKDNKELHPGAA RSVWKDDTSTILELFSAKAADSGTYICQLSNDVGTTSSKATIFVKEPPQFIKKPSPVLV LRNGQSTTFECQVTGTPEIRVSWYLDGNEITDLRKYGISFVDGLATFQISNARVENSG TYVCEARNDAGTASCSIELKVKEPPIFIRELEPVEVVKDSDVELECEVMGTTPFEVTW LKNNKEIRSGKKYTMSEKMSVFYLHITKCDPSDVGEYQCIIANEGGSCACSARVALK EPPSFIKKIENVTTVLKSSATFQSTVAGSPPISITWLKDDQILEENDNVHISFEDSVATL QVRSVDNGHSGRYTCQAKNESGIERCYAFLLVQEPAQIIEKAKSVDVTEKDPVTLEC VVAGTPELKVKWLKDGKQIVPSRYFSMSFENNVASFRIQSVMKQDSGQYTFKVEND FGSSSCDAYLRVLDQDIPPSFTKKLTKMDKVLGSSIHMECKVSGSLPISAQWFKDGKE ISTSAKYRLVCHENTVSLEVSNLELEDTANYTCKVSNVAGDNACSGILTVKEPPSFLV KPERQQAIPDSTVEFKAVLKGTPPFKIKWLKDDVELVSGPKCFIGLEGSTSFLNLYSV DSSKTGQYTCQVTNDVGSDSCTTMLLVTEPPKFVKKLEASKIIKAGDSARLECKITGS PEIQVVWYRNEHELTASDKYQMTFIDSVAVIQMNSLGTEDSGDFICEAQNPAGSTSC STKVIVKEPPVFSSFPPIVETLKNTEVSLECELSGTPPFEVVWYKDKRQLRSSKKYKV ASKNFHASIHILNVESTDIGEYHCKAQNEVGSDACVCAVKLKEPPKFISKLNSLTVVA GEPAELQASIEGAQPISVQWLKEKEEVIRESENIRISFVNNVATLQFAKVEPANAGKYI CQVKNDGGVRENMATLTVLEPAVIIEKAGSMTVTVGETCALECKVAGTPELSVEWY KDGKLLTSSQKHKFSFYNKISSLKILSVEKEDAGTYTFQVQNNVGKSSCTAVVDVSD RMVPPSFTRRLKDTGGVLGTSCILECKVAGSSPISIAWFHEKTKIVSGAKYQTTFSDN VCTLQLNSLDSSDMGSYTCVAANVAGSDECRALLTVQEPPSFVKEPEPLEVLPGKNI TFTSVIRGTPPFKVGWFRGARELVKGNRCNIYFEDTVAELELFNIDISQSGEYTCVVS NNAGQASCTTRLFVKEPATFVKKLSDHSVEPGKSIILEGTYTGTLPISVTWKKDGVSI TPSERCNIVTTEKTCILEILSSTKGDAGHYSCEIENEAGRDACDALVSTLEPPYFVTELE PLEASVGDSVSLQCQVAGTPEITVSWFKGDTKLRSTPEYRTYFTNNVATLVFNKVSIN DSGEYTCMAENSIGTAASKTIFRIQERQLPPSFARQLKDIEQTVGLPVTLTCRLNGSAP IQVCWYRDGVLLRDDENLQMSFVDNVATLKILQTDLSHSGQYSCSASNPLGTASSTA RLTAREPKKSPFFDIKPVSIDVIAGESADFECHVTGAQPMRVTWSKDNKEIRPGGNYT ITCVGNTPHLRILKVGKGDSGQYTCQATNDVGKDMCSAQLSVKEPPKFIKKLDASKV AKQGESIQLECKISGSPEIKVVWFRNDSELHESWKYNMSFVNSVALLTINEASAEDTG DYICEAHNGVGDASCSTALKVKAPPVFTQKPPPVGALKGSDVILQCEISGTPPFEVVW VKDRKQVRSSKKFKITSKNFDTSLHIFNLEAPDIGEYHCKATNEVGSDTCACTVKFKE PPRFVKKLSDASTLIGDPVELQAVVEGFQPISVVWLKDKGEVIRESENVRISFVDNIAT LQLGSPEASQSGKYVCQIKNDAGMRECSAVLTVLEPATIVEKPEPMTVTTGNPFTLE CVVAGTPELSAKWFKDGRELSSGSRHHITFVRNLASLKIPSAEMNDKGLYTFEVENR VGKSSCTVSVHVSDRVVPPSFVRRLKDTSATLGASVVLECRVSGSAPISVGWFLDGN EIISSPKCQSSFADNVCTLTLSSLEPSDTGAYTCVAANVAGQDESSAVLTVQEPPSFEQ TPDSVEVLPGMSLTFTSVIRGTPPFKVKWFKGSRELVSGEACTISLEDFVTELELLEVE PGQSGDYSCLVTNDAGSASCTTHLFVKEPATFVKRLADTSVETGSPIVLEATYSGTPP ISVSWMKNEYPLSQSPNCGITTTEKSSILEILESTIEDYAQYACLIENEAGQDICEALVS VLEPPYFIEPLEHVEAAIGEPITLQCKVDGTPEIRISWYKEHTKLRSAPAYKMQFKNN VASLVINKVDHSDVGEYTCKAENSVGAVASSAVLVIKERKLPPSFARKLKDVHETLG FPVAFECRINGSEPLQVSWYKDGELLKDDANLQMSFVHHVATLQILQTDQSHVGQY NCSASNPLGTASSSAKLILSEHEVPPFFDLKPVSVDLALGESGSFKCHVTGTAPIKITW AKDNREIRPGGNYKMTLVENTATLTVLKVAKGDAGQYTCYASNVAGKDSCSAQLG

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		VQEPPRFIKKLDQSRIVKQDEYTRYECKIGGSPEIKVLWYKDEVEIQESSKFRMSFEDS VAILEMHSLSVEDSGDYTCEARNAAGSASSSTSLKVKEPPVFRKKPFPVETLKGADV HLECELQGTPPFQVSWHKDKRELRSGKKYKIMSENLLTSIHILNVDTADIGEYQCKA TNDVGSDTCVGSVTMKAPPQFVKKLTDISTIIGKEVQLQTTIEGAEPISVAWFKDKGE IVRESDNIWISYSENIATLQFSRAEPANAGKYTCQIKNDAGMQECYATLSVLEPAAIV EKPESIKVTTGDTCTLECTVSGTPELSTKWFKDGKELTSDNKYKISFFNKVSGLKIINV VPGDSGVYSFEVQNPVGKDSCKVSIQVSDRIIPPSFTRKLKETNGLSGSSVVMECKVY GSPPISVLWFHDGNEISSGRKYQTTLTDNTCALTVNMLEEADAGDYTCIATNVAGSD ECSAPLTVREPPSFVQKPDPMDVLTGSNVTFTSIVKGSPPFTVSWFKGSTELVPGARC NVSLQDSVGELELFDVDTSQSGEYTCIVSNEAGRASCTTRLFVKAPAIFVKRLNDYSI EKGKPLILEGTFSGTPPISVTWKKNGINVIASQRCNITTTEKSAILEILSSTVEDSGQYN CYIENASGKDSCSAQILILEPPYFVKQLEPVKVTVGDSASLQCQLAGTPEIGVSWYKG DTKLRPTATCKMHFKNNVATLVFTQVDSSDSGEYICRAENSVGEVSSSTFLTVQEQK LPPSFSRQLRDVQETVGLPVVFECAVSGSEPISVSWYKDGKPLKDSPNIQTSFLDNIAT LNIFKTDRSLSGQYSCTATNPIGSASSGAKLILTEGKNPPFFDIPLAPMDAVVGESADL ECHVTGTQPIKVTWAKDNREIRSGGNYQISYLENSAHLTIVKVDKGDSGQYTCYAV NEVGKDSCTAQLNIKERLIPPSFTKKLSETVEETEGNSFKLEGRVAGSQPITVAWYKN NVEIHPTSNCEIMPKNNALLLQVKRASMADAGLYTCKATNDAGSALCTSSIVIKEPK KPPVFDQHLAPVTASEGDSVQLSCHVQGSEPIRIQWLKAGREVKPSDRCSFSFASGTA MLELKETAKADSGDYVCKASNVAGSDTSKCKVTIKEKPAAAPAAKKAAVDGKLFF VSEPQSIRVVEKTTATFIAKVGGDPIPNVKWTKGKWRQLNQGGRILIHQKGDEAKLEI RDTTKTDSGLYRCVAFNKHGEIESNVNLQVDERKKQEKIEGDLRAMLKKTPALKKG SGEEEEIDIMELLKNVDPKEYEKYARMYGITDFRGLLQAFELLKQSQEEETHRLEIEE LEKSERDEKEFEELVAFIQQRLTQTEPVTLIKDIENQTVLKDNDAIFEIDIKINYPEIKLS WYKGTEKLEPSNKYEISIDGDRHTLRVKNCQPKDQGNYRLVCGPHIASAKLTVIEPA WERHLQDVTLKEGQTCTMTCQFSVPNVKSEWFRNGRVLKPQGRVKTEVEHKVHKL TIADVRAEDQGQYTCKHEDLETSAELRIEAEPIQFTKRIQNIVVSEHQSATFECEVSFD DAIVTWYKGPTELTESQKYNFRNDGRCHYMTIHNVTPDDEGVYSVIARLEPRGEARS TAELYLTTKEIKLEMKPPDIPDSRVPIPTMPIRAVPPEEIPPAVAPSIPLLLPLPEEKKPP AKRIEVTKKGVKKDTKKVVTKPKEEAPPPPVAKKPPPPTPMIPAKASEIIDVSSKAEE VKITTITRKKEVHKEKEAVYEREEAVYEKKVHIEPWEEPYEELETEPYTEPYEEPYYE EPDEDYEEIKVEAKKQVHEEWEEDFEEGQEYYEREEGYDEGEEEWEEIYHEREIIQV QKEVHEELHEKKIPAKVPEKKVPPPKVVKKPVVEKVEKTTRRMEEEKVQVIKVPEVS KKIVPQKPSRTPVQEEIIEVKVPAVHTKKMVISEEKMFFASHTEEEVSVSVPEVQKKT VTEEKIHVAVSKKIEPPPKVPEPPKKPVPEEVVPVPIPKKVEPPAAKVPEAPKKPVPEE KKPVPIPKKKEPAAPPQVPEAPKKPAPEEKIPVPVTKKKEAPPAKVPEVQKKVVTEEK IAIITQREESPPPAVPEIPKKKVPEEKRPVPRKEEVPPPKVPVPPKKAVPEAVVPAPIPK KAPPRAEVSKKTVVEEKRFAAEEKLSVAVPQRVELMRHEEEEWTYSEEEERVSVSV YREEERDEEEAEITEYEVLEEPEEYVVEEKMHFISKKVEVEPAKVPEKKIIPKPKVPAK IEEPPPTKVPEPPKKIVPEKKVPAPAPKKVPPAKAPEESKRPVPEKRAPAEEVGIEEPPP TKVAERHMKITQEEKVLVAVTKKEAPPRARVPEEPKKVAPEERFPKLKPRREEEPPA KVTEVRKRAVKEEKVSIEVPKREPRPTKEVTVTEEKKWSYTREEETVSEHREEEYED YEDYEEYKEFEEYEPTEEYDQYDEYAEREVEHYEEHEEYVTEPKKPVPVKPAQEPVP AKPKAPPPKVLKKAVPEEKAPLPIQKKLKPLPPKAPEEPKKVVEEKIQISITKREKQQV TEPVAKVPMKPKRVVPEAKIPAPTKEVAVPVRVPGVPKKRELEEVVVFKEEVEAHEE YIVEEEEEYVHEEEYVHKEEYVHEEEYVHKEEYIHEEEEHLHEEEETIAEEEVVPVAP VKVPVVPKKPVPEEKKPVPVPKKKEAPPAKVPEIPKKPEEKVPVPIPKKEKAPPAKVP EVPKKPVPEEKPPVPVPKKVEPPPAKVPEVPKKPVPEKKVPAPTPKKVEAPPAKVPEV PKKPIPEEKKPTALLKKMEAPPPKAPKKREVVPVPVALPREEEEEEVPFEEVPEEEILP EEEVPSEEEAPPEEVPPEEEEVLPEEEEVLPEEEEVLPEEEEVQPEEEALPEIKPKVPKP APVPKKTVPEKKVPVPVPKKVEPPPPPKVPEIKKKVPEKKVVVPKKEEAPPTKVPEVS KKVEERRIIPPKEEEVPPAEVYEEAEEPTPEEIPEEPPSIEEEEIVEEEEEEEEVLPPRAPE VVKKAVPEAPTPVPKKAEAPPAKVPKKIPEEKVPVPVQKKEAPPAKVPEKKKIPEKK VPVPKKEAVPPAKGKAVFEEKISVAYQQEELVQERIELELVEAKVEEAFEEEEFHEVQ EYFEEEEFHEVEEFIRVEERRFQEEHKVEEVHRVIEFLEAEEVEVYEKPKIPPKKGPEV SEKVIPPKKPPTKVIPRKEPPAKVPEVTKKTVVEEKIRAPEEPKVPAPKAPEVPKKITPE EKVREAVPKKPEVPPPKVPEVPKKIIQEEKLPVVLPEDTEIYMYEASEETVIEEEHVTL PQKARLKVAKVPAPPQTVVTEEKTYVTIRKTRETLALKESETTREAFPELKSYKAVPE IPEPPSPEDLEIIEDVLPEKRPPASKRRKTQLPTAPEAPREMPPEMNTFEEISVEPEMLPT QVLDTYQEATVEKKTLRISRKKPELPSDEEVPEAPREVVAKKKVLPPQVPEVVPVKV PGAPKEVVSERKSLEEPPKKPAVRPVTVPEEPKEVIPEKKVSLVPPKKPAAPPVTVPE APEEVFSEDEETLAPPQEPEAPPAKVPEAPKEVVPEKKVSVVPPKKPEAPPAKVPEAP KEAAPEKKVPVAPKKKPEAPPVKVPEAPKKVVPEKKLPVAAPKKPEAPAAEVPEVP KTAVPQKKIPEAIPPKPESPPLEVPEVPPKEVTPEKKVPAAPPTKPEIPPPKVPEAPQAA VVEEKTPEALPKKAEAAPVPVPQVQETVPEKTRPVGPPKKPEATTVPVPKVQKTIPEK TRPEAPPKRPEARTVPETVPEKTRPMAPPKKPEATTLPVPEVQETVPEKTRPVGPPKK PEATTVPVPEVQETIPEKTRPPKKPEATPVPVPEVQETVPEKTRPVGPPKKPEATTVSV PEVQETIPEKTRPAAPPKKPEATAVPETIPEKTRPEAPPKRPEATTVPVPEADQAVVPE KKVPRVPPKKVEAPPITVPEEPKEVIPEKKVSLVPPKKPAAPPVTVPEAPEEVFSEDEE TLAPPQEPEAPPAKVPEAPKEVVPEKKVSVVPPKKPEAPPAKVPEAPKEAAPEKKVP VAPKKKPEAPPVKVPEAPKKVVPEKKLPVAAPKKPEAPAAEVPEVPKAAVPQKKIPE

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		AIPPKPESPPPEVYEEPEEEIVPEEPPEEAVEEPVPAPPPKVTEPPKKPVPEKKAPPAVV KKPEPPPAKVPEVPKEAPPEKKVPPKKPEAPPAKVPEVPKEVVTEKKVAVPKKPEVPP AKVPEVPKKPVIEEKPAIPVVEKVASPPAEVYEEPEEVTAEEEEPAPAVEEEEYEAPPP PAPVPEEPKKVVPEKKFPVIKKPEAPPPKVPEVPKKAVPVKKVPVVKKPEPPEAEVPE VPKKLVPVKKEPVPVTKKTEVLPEKVPEAPKKITPEKKESVPVPEEPEAPPASVEETPE ETIYEEKATITIGRKETPPVEEREIEKFIQPEEPELEPEPEEIPVQEPEPEKKVIEKPKLKP RPPARPPSPPKEDVKEKMFQLKAVSKKKVPEKPEVVEKVEPAPLKVPTAEKKVRKLL PEPKPQPKEEVVLKSVLRKKPEEEEPKVEPKKVEKAKKPEEPQPPPKAVEVEAPPEPT PKERKVPEPAKVPEIKPAIPLPGPEPKPKPEPEVKTMKAPPIEPAPTPIAAPVTAPVVGK KAEAKPKDEAAKPKGPIKGVAKKTPSPIEAERKKLRPGSGGEKPPDEAPFTYQLKAV PLKFVKEIKDIVLTEAESVGSSAIFECLVSPSTAITTWMKDGSNIRESPKHRFIADGKD RKLHIIDVQLSDAGEYTCVLRLGNKEKTSTAKLIVEELPVRFVKTLEEEVTVVKGQPL YLSCELNKERDVVWRKDGKIVVEKPGRIVPGVIGLMRALTINDADDTDAGTYTVTV ENANNLECSSCVKVVEIIREWLVKPIRDQHVKPKGTAVFTCDIAKDTPNIKWFKGYD EIPLEPNDKTEILKEGNHLFLKVKNAMPEDIDEYAVEIEGKRYPAKLTLGEREVELLK PIEDVTIYEKESASFDAEISEEDIPGEWKLKGELLRPSPTCEIKAEGGKRFLTLHKVKL DQAGEVLYQACNAITTAILTVKEIELDFAVPLKDVTVPEKRQARFECVLTREANVIW SKGPDIIKASDKFDIIADGKKHILVINDSQFDDEGVYTAEVEGKKTSAQLFVTGIRLKF ISPLEDQTVKEGQTATFVCELSHEKMHVVWFKNDVKLHTTRTVLMSSEGKTYKLEIR ETTLDDISQIKAQVKNLSSTANLKVLEADPYFTVKLHDKTGVEKDEIILKCEVSKDVP VKWFKDGEEIVPSPKHSVKTDGLRRILKIKKAELKDKGEYVCDCGTDTTKANVTVE ARLIKVEKPLYGVEVFVGETARFEIELSEPDVHGQWKLKGEPLTASPDCEIIEDGKKH VLVLYNCQLDMTGEISFQAANAKSAANLKVKELPLIFITPLSDVKVFEKDEAKFECEV SREPKTFRWLKGTQEITGDDRFELIKDGTRHSLVIKSAAFEDEAKYMFEAEDKRTSG KLIIEGIRLKFLTPLKDVTAKERENAVFTVELSHDNIPVSWFKNDQRLHTSKRVSMHD EGKTHSITFKDLSIDDTSQIRVEAMGISSEAKLTVLEGDPYFTGKLQDYTGVEKDEVIL QCEISKADAPVKWFKDGKEIKPSKNVVIKADGKKRMLILKKALKSDIGQYTCDCGTD QTSGKLDIEDREIKLVRPLYSVEVMETETARFETEISEDDIHANWKLKGEALLQTPEC EIKEEGKIHVLILHNCRLDQTGGVDFQAANVKSSAHLRVKPRVIGLLRPLKDVTVTA GETATFDCELSYEDIPVEWYLKGKKLEPNDKVVTRSEGRVHTLTLRDVKLEDAGEV QLTAKDFKTQANLFVKEPPVEFTKPLEDQTVEEEATAVLECEVSRENAKVKWFKNG TEILKSKKYEIVADGRVRKLIIHGCTPEDIKTYTCDAKDFKTSCNLNVVPPHVEFLRPL TDLQVKEKETARFECEISKENEKVQWFKDGAEIKKGKKYDIISKGAVRILVINKCLLN DEAEYSCEVRTARTSGMLTVLEEEAVFTKNLANLEVSEGDTIKLVCEVSKPGAEVIW YKGDEEIIETGRFEILTDGRKRILIIQNAQLEDAGSYNCRLPSSRTDSKVKVHELAAEFI SKPQNLEILEGEKAEFVCTISKESFEVQWKRDDQTLESGDKYDIIADGKKRVLVVKD ATLQDMGTYVVMVGAARAAAHLTVIEKLRIIVPLKDTKVKEQQEVVENCEVNTEGA KAKWFRNEEAIFDSSKYIILQKDLVYTLRIRDARLDDQANFNVSLTNHRGENVKSAA NLIVEEEDLRIVEPLKDIETMEKKSVTFWCKVNRLNVTLKWTKNGEEVAFDNRISYRI DKYKHSLIIKDCGFPDEGEYVVTAGQDKSVAELLIIEAPTEFVEHLEDQTVTEFDDAV FSCQLSREKANVKWYRNGREIKEGKKYKFEKDGSIHRLIIKDCRLEDECEYACGVED RKSRARLFVEEIPVEIIRPPQDILEAPGADVIFLAELNKDKVEVQWLRNNMIVVQGDK HQMMSEGKIHRLQICDIKPRDQGEYRFIAKDKEARAKLELAAAPKIKTADQDLVVDA GQPLTMVVPYDAYPKAEAEWFKENEPLSTKTVDTTAEQTSFRISEAKKDDKGRYKIV LQNKHGKAEGFINLQVIDVPGPVRNLEVTETFDGEVSLAWEEPLTDGGSKIIGYVVER RDIKRKTWVLVTDRADSCEFTVTGLQKGGVEYLFRVSARNRVGTGEPVETDSPVEA RSKYDVPGPPLNVTITDVNRFGVSLTWEPPEYDGGAEITNYVIELRDKTSIRWDTAM TVRAEDLSATVTDVVEGQEYSFRVRAQNRIGVGKPSAATPFVKVADPIERPSPPVNL NASEQTQSSVQLTWEPPLKDGGSPILGYIIERREEGKDNWIRCNMKPVPELTYKVTGL QKGNKYLYRVSAENAAGVSDPSEILGPLTADDAFVEPTMDLSAFKDGLEVIVPNPIKI LVPSTGYPRPKATWTFGDQVLEEGDRVKMKTISAYAELVISPSERTDKGIYTLTLENP VKSISGEINVNVIAPPSAPKELKFSDITKDSVHLTWEPPDDDGGSPLTGYVVEKRDMS RKTWTKVMDFVTDLEFTVPDLVQGKEYLFKVCARNKCGPGEPAYTDEPVNMSAPA TVPDPPENVKWRDRTANSIFLTWDPPKNDGGSRIKGYIVEKCPRGSDKWVACGEPVP DTKMEVTGLEEGKWYAYRVKALNRQGASKPSKPTEEIQAVDTQEAPEIFLDVKLLA GITVKAGTKIELPATVTGKPEPKITWTKADTLLKPDQRITIENVPKKSTVTITDSKRSD TGTYIIEAVNVCGRATAVVEVNVLDKPGPPAAFDITDVTNESCLLTWNPPRDDGGSK ITNYVVERKATDSDVWHKLSSTVKDTNFKATKLTPNKEYIFRVAAENMYGVGEPVQ AAPIIAKYQFDPPGPPTRLEPSDITKDAVTLTWCEPDDDGGSPITGYWVERLDPDTDK WVRCNKMPVKDTTYRVKGLTNKKKYRFRVLAENLAGPGKPSRSTEPILIKDPIDPPW PPGKPTVKDIGKTSLVLNWTKPEHDGGAKIESYVIEMLKTGTDEWVRVAEGVPTTEH LLTGLMEGQEYSFRVRAVNKAGESEPSEPSDPVLCREKLYPPSPPRWLEVINITKNTA DLKWTVPEKDGGSPITNYIVEKRDVRRKGWQTVDTTVKDTKCTVTPLTEGSLYVFR VAAENAIGQSDYTEIGDSVLAKDTFTTPGPPYALTVVDVTKRHVDLKWEPPKNDGG RPIQRYIIEKKEKLGTRWVKAGKTSGPDCNFRVTDVIEGTEVQFQVRAENEAGVGHP SEPTEILSIEDPTSPPSPPLDLHVTDAGRKHIAIAWKPPEKNGGSPIIGYHVEMCPVGTE KWMRVNSRPIKDLKFKVEEGIVPDKEYVLRVRAVNAVGVSEPSEISENVVAKDPDC KPTIDLETHDIVVIEGEKLNIPVPFRAVPVPTVSWHKDGKEVKASDRLTMKNDHISAH LEVPKSVHADAGVYTITLENKLGSATASINVKVIGLPGPCKDIKASDITKSSCKLTWE PPEFDGGSPILHYVLERREAGRRTYIPVMSGENKLSWTVKDLIPNGEYFFRVKAVNKI GGGEYIELKNPVIAQDPKQPPDPPVDVEVHNPTAKAMTITWKPPLYDGGSKIMGYIIE

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		KITKGEDRWKRCNEHLVPVLTYTAKGLEEGKEYQFRVRAENAAGIGEPSRATPPTKA VDPIDAPKVILRTSLEVKRGDEIALDATISGSPYPTITWIKDENVIVPEEIKKRAAPPVR RKKGEAEEEEPFSLPLTERLSINNSKQGESQLRIRDSLRPDHGQYMIKVENDHGVAKA PCSVSVLDTPGPPINFVFEDIRKDSVLCKWEPPLDDGGSEIINYTLEKKDKTKPDSDWI VITSTLRNCKYSVTKLIEGKEYLFRVRAENRFGPGPPCVSKPLLAKDPFEPPDAPDKPI VEDVTSNSMLVKWNEPKDNGSPILGYWLEKREVNSTHWSRVNKTLLSSLKTKVDGL LEGLTYVFRVCAENAAGPGKFSPPSDPKTARDPISPPGPPVPRVADTSSTTIELEWEPP AFNGGGEIMGYFVDKQLVGTNEWSRCTEKMIKVRQFTVKEIREGADYKLRVSAVNA AGEGPPGETEPVTVAEPQEPPTVELDVSVKGGIQIMAGKTLRIPAEVTGRPVPTKVWT IEEGELDKERVIIENVGTKSELIIKNALRKDHGRYVITATNSCGSKFAAVRVEVFDVPG PVLDLKPVVTNRKMCLLNWSDPADDGGSDITGFIIERKDAKMHTWRQPIETERSKCD ITGLIEGQEYKFRVIAKNKFGCGPPVEIGPILAVDPLGPPTSPERLTYTERTKSTITLDW KEPRSDGGSPIQGYIIEKRRHDKPDFERVNKRLCPTTSFLVENLDEHQMYEFRVKAVN DVGESEPSLPLNVVIQDDEVPPTIKLRLAVRGDTIKVKAGEPVNIPADVTGLPMPKIE WSKNEKVIDKPTDTLNITKEEVSRSEAKTELSIPKAAREDKGTYTITASNRLGSVFRN VHVEVYDRPSPPRNLAVTDIKAESCYLTWDAPLDNGGSEITHYIIDKRDASRKKSEW EEVTNTAVERRYGIWKLIPNGQYEFRVRAVNKYGISDECKSDKVVIQDPYRLPGPPG KPKVLERTKGSMLVSWTPPLDNGGSPITGYWLEKREEGGTYWSRVSRAPITKVGLK GVEFNVPRLIEGVKYQFRAMAINAAGIGPPSEPSDPAVAGDPIYPPGPPSCPEVKDKT KSSISLAWKPPAKDGGSPIKGYIVEMQEEGTTDWKPVNEPDKLLTACECVVPNLKEL RKYRFRVKAVNEAGESEPSDTTGEIPATDIQEVPEVFIDIGAQDCLVCKAGSQVKIPA VIKGRPTPKSSWEFDGKAKKAMKDGVHDIPEDAQLETAENSSVIIIPECTRAHSGKYS ITAKNKAGQKTANCRVKVMDAPGPPKDLKVSDITRGSCRLSWKMPDDDGGDRIKG YVIEKKTIDGKAWTKVNPNCGSTTFVVPDLISEQQYFFRVRAENRFGIGPPAETIQRTT ARDPIYPPDPPIKLKIGLITKNTVHLSWKPPKNDGGSPVTHYIVECLAWDPTGKKKEA WRQCNRRDVEELEFTVEDLIEGGEYEFRVKAVNEAGVSKPSATVGPVIVKDQTCPPA IELKEFMEVEEGTDVNIVAKIKGVPFPTLTWFKAPPKKPDSKEPVVYDTHVNKQVVD DTCTLVIPQSRRSDTGLYSITAVNNLGTASKEMRLNVLGRPGPPVGPIKFESISADQM TLSWLPPKDDGGSKITNYVIEKREANRKTWVRVSSEPKECMYTIPKLLEGHEYVFRI MAQNKYGIGEPLDSEPETARNLFSVPGAPDKPTVSSVTRNSMTVNWEEPEYDGGSPV TGYWLEMKDTTSKRWKRVNRDPIKAMTLGVSYKVTGLIEGSDYQFRVYAINAAGV GPASLPSDPVTARDPVAPPGPPFPKVTDWTKSSVDLEWSPPLKDGGSKITGYIVEYKE EGKEEWEKGKDKEVRGTKLVVTGLKEGAFYKFRVRAVNVAGVGEPGEVTDVIEMK DRIVSPDLQLDASVRDRIVVHAGGVIRIIAYVSGKPPPTVTWSMNERALPQEAAIETT AISSSMVIKNCQRSHQGVYSLLAKNEGGERKKTIIVDVLDVPGPVGIPFLSDNLTNDS CKLTWFSPEDDGGSPITNYVIQKREADRRAWTPVTYTVTRQNATVQGLIQGKSYFFR IAAENSIGMGPFVETPNALVIRDPITVPERPEDLEVKEVTKNTVSLTWNPPKYDGGSEI INYVLESRLIGTEKFHKVTNDNLLSRKYTVKGLKEGDTYEYRVSAVNIVGQGKPSFC TKPITCKDELAPPTLDLDFRDKLTVRVGESFALTGRYSGKPKPKIDWFKDEADVLED DRTHIKTTPTTLALEKTKAKRSDSGKYCVVVENSTGSRKGFCQVNVVDRPGPPVGPV IFDEVTKEYMVISWKPPLDDGGSEITNYIIEKKELGKDIWMPVTSASAKTTCKVPKLL EGKDYIFRIHAENLYGISDPLVSDSMKAKDRFRVPDAPEQPVVTEVTKDSALVTWNK PNDGGKPITNYILEKRETMSKRWVRVTKEPIHPYTKYRVPDLLEGCQYEFRVSAENEI GIGDPSPPSKPVFARDPIAKPSPPINPEAIDTTCNSVELTWQPPRHDGGSKILGYIVEYQ KVGDEEWRRANHTPESCPETKYKVTGLRDGQSYKFRVLAVNEAGESDPAHVPEPVL VKDRLEPPELILDANMAREQHIRVGDTLRLSAIIKGVPFPKVTWKKEDREAPTKAQID VTPVGSKLEIRNAAHEDGGIYSLTVENPAGTKTVSVKVLVLDKPGPPRDLEVSEIRKD SCYLTWKEPLDDGGSVVTNYVVERKDVATAQWSPLSTTSKKKSHMAKHLTEGNQY LFRVAAENQYGRGPFVETPKPIKALDPLHPPGPPKDLHHVDVDKTEVSLVWNKPDR DGGSPITGYLVEYQEEGDKDWIKFKTVKNLDCVVTGLQQGKTYRFRVKAENIIGLGL PDTTIPIECQEKLVPPSVELDVKLIEGLVVKAGTTVRFPAIIRGVPVPTAKWTTDGTEI KTDDHYTVETDSFSSVLTIKNCLRKDTGEYQLTVSNAAGTKTVAVHLTVLDVPGPPT GPINILDVTPEYMTISWQPPKDDGGSPVINYIVEKQDTRKGTWGVVSAGSSKLKLKV PHLQKGCEYVFRVKAENKMGVGPPLDSIPTVAKHKFSPPSPPGKPVVTDITENAATV SWTLPKSDGGSPITGYYVERREITGKWVRVNKTPIADLKFRVTGLYEGNTYEFRVFA ENLAGLSNPSPSSDPIKACRPIKPPGPPINPKLKDKSKESADLVWTKPLSDGGSPILGY VVEYQKPGTAQWDRINKDELIRQCAFRVPGLIEGNEYRFRIRAANIVGEGEPRELAES VIAKDILHPPEVELDVTCRDVITVRVGQTIRILARVKGRPEPDITWSKEGKVLVKDKR VDLIHDLPRVELQIKEAVRADHGKYIISAKNSSGHAQGSAIVNVLDRPGPCQNLKVSN VTKENCTISWENPLDNGGSEITNFIVEYRKPNQKGWSIVASDVTKRLVKANLLANNE YYFRVCAENKVGVGPTIETKTPILAINPIDRPGEPENLHIADKGKTFVYLKWRRPDYD GGSPNLSYHVERRLKGSADWERVHKGSIKETHYMVDKCVENQIYEFRVQTKNEGGE SDWVRTEEVVVKEDLQKPVLDLKLSGVLTVKAGDTIRLEAGVRGKPFPEVAWTKD KDATDLTRSPRVKIDTSAESSKFSLTKAKRSDGGKYVITATNPAGSFVAYATVNVLD KPGPVRNLKITDVSSDRCTIRWDPPEDDGGCEIQNYILEKCESKRMVWSTYSANVLT PSATVTRLIEGNEYIFRVRAENKIGTGPPTESKPVIAKTKYDRPGRPDPPEVTKVSKEE MTVVWNAPEYDGGKSITGYYLEKKEKHAVRWVPVNKSAIPERRLKVQNLLPGHEY QFRVKAENEVGIGEPSLPSRPVVAKDPIEPPGPPTNFKVVDTTKSSITLAWGKPVYDG GAPIIGYVVEMRPKIADASPDEGWKRCNAAAQLIRMEFTVTSLDENQEYEFRVCAQN QVGIGRPAELKEAIKPKEILEPPEIDLDASMRKLVVVRAGCPIRLFAIVRGRPAPKVTW RKVGIDNVVRKGQVDLVDTMAFLVIPNSTRDDSGKYSLTLVNPAGEKAVFVNVKVL

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		DTPGPVADLKVSDVTKTSCHVSWAPPENDGGSQVTHYIVEKREAERKTWSTVTPEV KKTSFNVTNLVPGNEYFFRVTAVNEYGPGVPTDIPKPVLASDPLSEPDPPRKLEVTEM TKNSATLAWLPPLRDGGAKIDGYIISYREEDQPADRWTEYSVVKDLSLIVTGLKEGK KYKFRVAARNAVGVSMPREAEGVYEAKEQLLPPKILMPEQITIKAGKKLRVEAHVY GKPNPICKWKKGDDEVVTSSHLAIHKADGSSVLIIKDVTRKDSGYYSLTAENSSGSDT QKIKVTVMDAPGPPQPPFDISEIDADACSLSWHIPLEDGGSNITNYIVEKCDVSRGDW VTALASVTKTSCRVGKLIPGQEYIFRVRAENRFGISEPLTSPKMLAKFPFDVPSEPKNA RVTKVNKDCIFVAWDRPDSDGGSPITGYLIERKERNSLLWVKANDTIVRSTEYPCAG LVEGLEYSFRIYALNKAGSSPPSKPTEYVTARMPVDPPGKPEVVDVTKNSASLIWARP KHDGGSRIIGYFVEACKLPGDKWVRCNTTPHQIPQEEYTATGLEENAQYQFRAIAKT AVNISQPSEPSDPVTILAENVPPRIELSVEMKSLLTVKAGTNVCLDATVFGKPMPTVS WKKDTTPIKQAEGIKMAMKRNLCTLELFSVNRKDSGDYTITAENSSGSKSATIKLKV LDKPGPPASVKINKMYADRAMLSWEPPLEDGGSEITNYIIDKRETSRPNWAQVSATV PITSCTVEKLIEGHEYQFRICAENKYGVGDPILTEPAIAKNPYDPPGRCDPPVISNITKD HMTVSWKAPADDGGSPITGYLVEKRETQAVNWTKVNRKPVIERTLKATGLQEGTEY EFRVTAINKAGPGKPSDASKAVYAQDPLYPPGPPAFPKVYDTTRSSVSLSWGKPAYD GGSPIIGYLVEVKRADSDHWVRCNLPEKLQKTRFEVTGLMENTEYQFRVYAVNKIG YSDPSDVPDKHCPKDILIPPEGELDAELRKTLILRAGVTMRLYVPVKGRPPPKITWSK PNVNLRERIGLDIKSTDFDTFLRCENVNKYDAGKYILTLENSCGKKEYTIVVKVLDTP GPPVNVTVKEVSKDSAYVTWDPPIIDGGSPIINYVVEKRDAERKSWSTVTTECSKTSF RVSNLEEGKSYFFRVFAENEYGIGDPGETRDAVKASETPGPVVDLKALAITKSSCTIG WKKPRSDGGSRITGYVVDFLTEENKWQRVMKSLSLQYSTKDLKEGKEYTFRVSAEN ENGEGTPSEIVVVAKDDVVAPDLDLKDLPDLCYLAKENSNFRLKIPIKGKPAPSVSW KKGEDPLATDTRVSVESTAVNTTLVVYDCQKSDAGKYTITLKNVAGTKEGTLSIKV VGKPGIPTGPIKFDEVTAEAMTLKWGPPKDDGGSEITNYVLEKRDSVNNKWVTCAS AVQKTTFRVTRLHEGIEYTFRVSAENKYGVGEGLKSEPIVAKHPFDVPDAPPPPNIVD VRHDSVSLTWTDPKKTGGSPITGYHIEFKERNSLLWKRANKTPIRMKDFKVTGLTEG LEYEFRVMAINLAGVGKPSLPSEPVVALDPIDPPGKPEVISVTRNSVTLIWTEPKYDG GHKLTGYIVEKRDLPSKSWMKANHVNVPDCAFTVTDLVEGGKYEFRIRAKNTAGAI SAPSESTGTIICKDEYEAPTIVLDPTIKDGLTVKAGDSIVLSAISILGKPLPKSSWSRAG KDIRPSDIAQITSTPTSSMLTVKYATRKDAGEYTITATNPFGTKEEHVKVSVLDVPGPP GPIEISNVSAEKATLTWTPPLEDGGSPIKAYVLEKRETSRLLWTVVSEDIQACRHVVT KLIQGNEYLFRVSAVNHYGKGEPVQSEPVKMVDRFGPPGPPGKPEISNVTKNTATVS WKRPTDDGGSEITGYHVERREKKGLRWVRATKTPVSDLRCKVTGLQEGNTYEFRVS AENRAGIGPPSDASNPVLMKDVAYPPGPPSNAHVTDTTKKSASLAWGKPHYDGGLE ITGYVVEHQKVGDDAWIKDTTGTALRITQFVVPDLQTKEKYNFRISAINDAGVGEPA VIPNVEIVEKEVAPDFELDAELRRTLVVRAGLSIRIFVPIKGRPAPEVTWTKDNINLKH RANIENTESFTLLIIPECNRYDTGKFVMTIENPAGKKSGFVNVRVLDTPGPVLNLRPTD ITKDSVTLHWDLPLIDGGSRITNYIVEKREATRKSYSTVTTKCHKCTYKVTGLTEGCE YFFRVMAENEYGVGEPTETTEPVRASEAPLPPDSLNIMDITKNTVSLAWPKPRHDGG SKITGYVIEAQRKGSDQWTHISTVKGLECVVRNLTEGEEYTFQVMAVNSAGRSAPRE SRPVIVKEQTMLPELDLRGIYQKLVIARAGDNIKVEIPVLGRPKPTVTWKKGDQILKQ TQRVNVENTATSTILNINECVRSDSGPYPLTAKNTVGEVGDVITIQVHDIPGPPTGPIK FDEVSSDFVTFSWEPPENDGGVPISNYVVEMRQTDSTTWVELATTVIRTTYKATRLT TGVEYQFRVRAQNRYGVGPGITSASVVANYPFKVPGPPGTPQVTAVTKDSMTISWH EPLSDGGSPILGYHIERKERNGILWQTVSKALVPGNIFKSTGLTDGIAYEFRVIAENMA GKSKPSKPSEPMFALDPIDPPGKPVPLNITRHTVALKWAKPEYTGGFKITSYVVEKRD LPNGRWLKANFSNILENEFTVSGLTEDAAYEFRVIAKNAAGAISPPSEPSDAITCRDDL EAPRIMVDVRFKDTITLKAGEAFKLEADVSGRPPPTMEWTKDGKELEGTGKLEIKIA DFSTHLINKDSSRTDSGAYILTATNPGGFAKHIFNVKVLDRPGPPEGPLAVSDVTSEK CVLSWLPPLDDGGAKIDHYIVQKRETSRLAWTNVATEVQVTKLKVTKLLKGNEYIF RVMAVNKYGVGEPLESEPVLAVDPYGPPDPPKNPEVTTITKDSMVVCWGHPDSDGG SEIINYIVERRDKAGQRWVKCNKKALTDLRFKVSGLTEGHEYEFRIMAENAAGISAPS ATSPFYKACDTVFKPGPPGNPRVLDTSRSSISIAWNKPIYDGGSEITGYMVEIALPEED EWQVVTPPAGLKATSYTITSLIENQEYKIRIYAMNSEGLGEPALVPGTPKAEERMLPP EIELDADLRKVVTIRACCTLRLFVPIKGRPAPEVKWAREHGESLDKASIESTSSYTLLV VGNVNRFDSGKYILTVENSSGSKSAFVNVRVLDTPGPPQNLKIKEVTKTSVTLTWEPP LLDGGSKIKNYIVEKRESTRKAYSTVATNCHKTSWKVDQLQEGCSYYFRVLAENEY GIGLPAETAESVKASERPLPPGKITLTDVTRNSVSLSWEKPEHDGGSRILGYIVEMQSK GSDRWATCATVKVTEATITGLIQGEEYSFRVSAQNEKGISDPRQLSVPVIAKDLVIPP AFKLLFNTFTVLAGEDLKIDVPFIGRPPPAVTWHKDDIPLKQTTRVNAESTENNSLLTI KEACREDVGHYTVKLTNSAGEATETLNVIVLDKPGPPTGPVKMDEVTADSVTLSWE PPKYDGGSSINNYIVEKRDTSTTAWQIVSATVARTTIKACRLKTGCEYQFRIAAENRY GKSTYLNSEPVVAQYPFKVPGPPGTPFVTLASKDSMEVQWHEPVSDGGSKVIGYHLE RKERNSILWVKLNKTPIPQTKFKTTGLEEGIEYEFRVSAENIVGIGKPSKPSECYAAHD PCDPPGRPEAIIVTRNSVTLQWKKPTYDGGSKITGYIVEKKELPDGRWMKASFTNIID TQFEVTGLLEDHRYEFRVIARNAAGVFSEPSESTGAITARDEVEPPRISMDPKYRDTV VVQAGESFKIDADIYGKPIPTTQWVKGDQELSSTARLEIKSTDFATSLSVKDAVRVDS GNYILKAKNVAGEKSVTINVKVLDRPGPPEGPVAISGVTAEKCTLAWKPPLQDGGSD ITNYIVERRETSRLVWTLVDANVQTLSCKVLKLLEGNEYIFRIMAVNKYGVGEPLESE SLIAKNPFVVPDAPKAPEVTAVTKDSMIVVWERPASDGGSEILGYVLEKRDKEGIRW

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		TRCHKRLIGELRLRVTGLLENHNYEFRVSAENAAGLSEPSPPSAYQKACDPIYKPGPP NNPKVMDVTRSSVFLSWTKPIYDGGCEIQGYIVEKCDVSVGEWTMCTPPTGINKTNL EVEKLLEKHEYNFRICAINKAGVGEHADVPGPVMVEEKLEAPDIDLDLELRKVINIRA GGSLRLFVPIKGRPTPEVKWGKVDGDIRDAAIIDVTSSFTSLVLDNVNRYDSGKYTLT LENSSGTKSAFVTVRVLDTPSPPVNLKVTEITKDSVSITWEPPLLDGGSKIKNYIVEKR EATRKSYAAVVTNCHKNSWKIDQLQEGCSYYFRVTAENEYGIGLPARTADPIKVAE VPQPPGKITVDDVTRNSVSLSWTKPEHDGGSKIIQYIVEMQAKNTDKWSECARVKSL DAVITNLTQGEEYLFRVIAVNEKGRSDPRSLAVPIIAKDLVIEPDVRPAFSSYSVQVGQ DLKIEVPISGRPKPSISWTKDGMPLKQTTRINVTDSLDLTTLSIKETHKDDGGQYGITV SNVVGQKTAAIEIITLDKPDPPKGPVKFDEISAESITLSWNPPLYTGGCQITNYIVQKR DTTTTVWDVVSATVARTTLKVTKLKTGTEYQFRIFAENRYGQSFALESEPVVAQYPY KEPGPPGTPFVTAISKESMVVQWHEPINNGGSPVIGYHLERKERNSILWTKVNKTIIH DTQFKALNLEEGIEYEFRVYAENIVGVGKASKNSECYVARDPCDPPGTPEAIIVKRNE ITLQWTKPVYDGGSMITGYIVEKRDLPEGRWMKASFTNVIETQFTVSGLTEDQRYEF RVIAKNAAGAISKPSDSTGPITAKDEVELPRISMDPKFRDTIVVNAGETFRLEADVHG KPLPTIEWLRGDKEIEESARCEIKNTDFKALLIVKDAIRIDGGQYILRASNVAGSKSFP VNVKVLDRPGPPEGPVQVTGVTAEKCTLAWSPPLQDGGSDISHYVVEKRETSRLAW TVVASEVVTNSLKVTKLLEGNKYIFRIMAVNKYGVGEPLESAPVLMKNPFVLPGPPK SLEVTNIAKDSMTVCWNRPDSDGGSEIIGYIVEKRDRSGIRWIKCNKRRITDLRLRVT GLTEDHEYEFRVSAENAAGVGEPSPATVYYKACDPVFKPGPPTNAHVVDTTKNSITL AWSKPIYDGGSEILGYVVEICKADEEEWQIVTPQTGLRVTRFEIAKLIEHQEYKIRVC ALNKVGLGEAASVPGTVKPEDKLEAPELDLDSELRKGIVVRAGGSARIHIPFKGRPTP EITWSKEEGEFTDKVQIEKGINFTQLSIDNCDRNDAGKYILKLENSSGSKSAFVTVKV LDTPGPPQNLAVKEVRKDSVLLVWEPPIIDGGAKVKNYVIDKRESTRKAYANVSSKC NKTSFRVENLTEGAIYYFRVMAENEFGVGVPTETSDAVKASEPPSPPGKVTLTDVSQ TSASLMWEKPEHDGGSRILGYVVEMQPKGTEKWSVVAESKVCNAVVSGLSSGQEY QFRVKAYNEKGKSDPRVLGIPVIAKDLTIQPSFKLPFNTYSVQAGEDLKIEIPVIGRPRP KISWVKDGEPLKQTTRVNVEETATSTILHIKESSKDDFGKYSVTATNSAGTATENLSV IVLEKPGPPVGPVKFDEVSADFVVISWEPPAYTGGCQISNYIVEKRDTTTTNWQMVS ATVARTTIKISKLKTGTEYQFRIFAENRYGKSTPLDSKPVVVQYPFKEPGPPGTPFVTS ISKDQMLVQWHEPVNDGGSKVTGYHLEQKEKNSILWVKLNKIPIQDTKFKTTGLDE GLEYEFRVSAENIVGIGKPSKVSECYVARDPCDPPGRPEAIVITRNSVTLKWKKPVYD GGSKITGYIVEKKDLPDGRWMKASFTNVVETEFTVTGLVEDQRYEFRVIARNAADN FSEPSESSGAITARDEIDAPNASLDPKYRDVIIVHAGETFVLEADIRGKPIPDIIWSKDG NELEETAARMEIKSTLQKTTLIVKDCIRTDGGQYTLKLSNVGGTKTIPITVKVLDRPG PPEGPLKVTGVTAEKCYLAWNPPLQDGGASISHYIIEKRETSRLSWTQVSNEVQALN YKVTKLLPGNEYIFRVMAVNKYGIGEALESEPVIACNPYKRPGPPSTPEASAITKDSM VLTWTRPVDDGGAEIEGYILEKRDKEGIRWTKCNKKTLTDLRFRVTGLTEGHSYEFR VAAENAAGVGEPSEPSVFYRACDALYPPGPPSNPKVTDTSRSSVSLAWNKPIYDGGA PVRGYVIELKKAAADEWTTCTPPSGLQGKQFTVTKLKENTEYNFRICAFNTEGVGEP ATIPGSVVAQERMEAPEIELDADLRKVVTLRASATLRLFVTIKGRPEPEVKWEKAEGI LTERAQIEVTSSYTMLVIDNVTRFDSGRYNLTLENNSGSKTAFVNVRVLDSPSAPVNL TIREVKKDSVTLSWEPPLIDGGAKITNYIVEKRETTRKAYATITNNCTKNTFKIENLQE GCSYYFRVLASNEYGIGLPAETAEPVKVSEPPLPPGRVTLVDVTRNTATIKWEKPESD GGSKITGYVVEMQTKGSEKWSACTQVKTLETTISGLTAGEEYVFRVAAVNEKGRSD PRQLGVPVIAKDIEIKPSVELPFNTFNVKANDQLKIDIPFKGRPQATVAWKKDGQVLR ETTRVNVASSKTVTTLSIKEASREDVGTYELCVSNTAGSITVPITVIVLDRPGPPGPIRI DEVSCDNVSISWNPPEYDGGCQISNYIVEKRETTSTTWQVVSQAVARTSIKIVRLTTG SEYQFRVCAENRYGKSSYSESSAVVAEYPFSPPGPPGTPKVVHATKSTMVVSWQVPV NDGGSQVIGYHLEYKERSSILWSKANKVLIADTQMKVSGLDEGLMYEYRVYAENIA GIGKCSKACEPVPARDPCDPPGQPEVTNITRKSVSLKWSKPRYDGGAKITGYIVERRE LPDGRWLKCNFTNVQETYFEVTELTEDQRYEFRVFARNAADSVSEPSESTGPITVKD DVEAPRIMMDVKFRDVIIVKAGEVLKINADIAGRPLPVISWAKDGVEIEERAKTEIVS TDYTTTLTVKDCVRRDTGQYVLTLKNVAGTRTMAVNCKVLDKPGPPAGPLEINGLT AEKCSLSWGRPQEDGGADIDYYIVEKRETSRLAWTICEAELRTTSCKVTKLLKGNEY IFRVTGVNKYGVGEPLESMAVKALDPFTTPSPPTSLEITSVTKDSMTLCWSRPETDGG SDISGYIIERREKNSLRWMRVNKKPVYDLRVKSTGLREGCEYEYRVFAENAAGLSLP SETSPLVRAEDPVFLPSPPSKPKIVDSGKTTITIGWVKPLFDGGAPITGYTVEYKKSEET DWKVAIQSFRGTEYTMSGLTTGDEYVFRVRSLNKMGASDPSDSSDPQVAKEREEEP VFDVDSEMRKTLIVKAGSSFTMTVPFRGRPIPNVSWSKPDTDLRTRAYIDSTDSRTLL TIENANRNDSGKYTLTIQNVLSAASMTFVVKVLDSPGPPANITVREVTKETAMLSWD VPENDGGAPVKNYHIEKREASKKAWVSVTNNCNRLSYKVTNLQEGAIYYFRVSGEN EFGVGVPAETKEGVKITEKPSPPEKLGVTSVSKDSVSLSWLKPEHDGGSRIIHYVVEA LEKGQKTWVKCAVVKTTHHVVSGLRESHEYFFRVFAENQAGLSDPRELLLPVLIKD QLEPPEIDMKNFPSHTVYVRAGSNLKVDIPISGKPLPKVTLSRDGVPLKATMRENTEIT AENLTINLKESVTTDAGRYEITAANSSGTTKTFINIIVLDRPGPPTGPVAISDITEESVTL KWEPPKYDGGSHVTNYIVLKRETSTAVWSEVSATVARTMIKVMKLTTGEEYQFRIK AENREGISDHIDSVCVVVKLPYTTPGPPSTPWVSNVTRESITVGWHEPVSNGGSAVTG YHLEMKDRNSILWQKANKMIIRTTHFKVTTISAGLIYEFRVYAENAAGIGKPSHPSEP VLAIDACEPPRNVRITDISKNSVNLSWQQPAFDGGSKITGYIVERRDLPDGRWTKASF TNVIETQFTVSGLTQNSQYEFRVFARNAVGSVSNPSEVVGPITCIDSYGGPVIDLPLEY

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		TEVVKYRAGTSVKLRAGISGKPEPTIEWYKDDKELQTNALVCVENSTDLASILIKDA NRLNSGSYELKLRNAMGSASATIRVQILDKPGPPGGPIEFKTVTAEKITLLWRPPADD GGAKITHYIVEKRETSRVVWSMVAENLEECIVTTTKIIKGNEYVFRVRAVNKYGIGEP LESEPVVAKNAFVTPGPPSIPEVTKITKNSMTVVWDRPTVDGGSEINGYFLERRDKKS LAWLKVLKETIRDTRQKVTGLTENSDYQYRVCAVNAAGVGPFSEPSDFYKAADPID PPGPPAKIRIADSTKSSITLGWSKPVYDGGSDVTGYVVEMKQGDEEEWTIVSTRGEV RTTEYVVSNLKPGVNYYFQVSAVNCAGQGEPITMTEPAQAKDVLEEPEIDLDVALRT SVIAKAGEDVQLLIPFKGRPPPTVTWRKDEKNLGSDTRYSIQNTDSSSLLVIPQVTRN DTGKYILTIENGVGQPKSSTVSVKVLDTPAACQKLQVKHVSLGTVTLLWDPPLIDGG SPIINYVIEKRDATKRTWSVVSHKCSGTSFKVTDLSEKTPFFFRVLAENEIGIGEPCETT EPVKAAEVPAPIRDLSMKDSTKTSVVLSWTKPDFDGGSIITDYLVERKGKGEQAWSH AGISKTCEIEIGQLKEQSVLEFRVSARNEKGQSDPVTIGPLTVKELVITPEVDLSEIPGA QISVRIGHNVHLELPYKGKPKPSISWLKDGLPLKESEYVRFSKTENKITLSIKNSKKEH GGKYTVILDNAVCRNSFPITIITLGPPSKPKGPIRFDEIKADSAIMSWDIPEDDGGGEIT CYSIEKREASQTNWKMVCSSVARTTFKVSNLVKDSEYQFRVRAENRYGVSEPLASNI IVAKHQFRIPGPPGKPVIYNVTSDGMSLTWDAPVYDGGSEVTGFHVEKKERNSILWQ RVNTSPISGREYRATGLIEGLDYQFRVYAENSAGLSSPSDPSKFTLAVSPVDPPGTPDY IDVTRETITLKWNPPLRDGGSKIVAYSIEKRQGSDRWVRCNFTDVSECQYTVTGLSPG DRYEFRIIARNAVGTISPPSQSSGLIMTRDENVPPTVEFGPEYFDGLVIKSGDSLRIKAL VQGRPVPRVTWFKDGVEIERRMNMEITDVLGSTSLFVRDATRDHRGVYTVEAKNVS GSTKAEVTVKVQDTPGKVVGPIRFTNITGEKMTLWWEAPLNDGCAPVTHYIIEKRET SRLAWALIEDNCEALSYTAIKLITGNEYQFRISAVNKFGVGRPLESDPVVAQIQYTIPD APGVPEPSNVTGNSITLTWTRPESDGGSEIQHYILERREKKSTRWVKVISKRPISETRF KVTGLVEGNEYEFHVMAENAAGIGPASGISRLIKCREPVNPPSAPSVVKVTDTSKTTV SLEWARPVFDGGMEIIGYIIEMCKADLGDWHKVNTEPCVKTRYTVTDLQAGEEYKF RVSAVNGAGKGDSCEVTGTIKAVDRLSAPELDIDANFKQTHIVRAGVSIRLFIAYQGR PTPTAVWSKPDSNLSIRADIHTTDSFSTLTVENCNRNDAGKYTLTVENNSGKKSITFT VKVLDSPGPPGPITFKDVTRGSATLMWDAPLLDGGARIHHYVIEKREASRRSWQVVS EKCTRQILKVSELTEGVPYYFRVSAENEYGVGEPYEMPEPIVATEQPAPPRRLDVVDT SKSSAVLAWLKPDHDGGSRITSYLLEMRQKGSDFWVEAGHTKQLTFTVERLVENTE YEFRVKAKNDAGYSEPREAFSSVIIKEPQIEPTADLTGITNQLITCKAGSTFTIDVPISG RPAPKVTWKLEEMRLKETDRMSIATTKDRTTLTVKDSMRGDSGRYFLTLENTAGVK TFTITVVVIGRPGPVTGPIEVSSVSAESCVLSWTEPKDDGGTEITNYIVEKRESGTTAW QLINSSVKRTQIKVTHLTKYKEYCFRVSSENRFGVSKPLESVPIVAEHPFVPPSAPTRP EVYYVSANAMSIRWEEPYHDGGSKIVGYWVEKKERNTILWVKENKVPCLECNYKV TGLVEGLEYQFRTYALNAAGVSKASEASRPIMAQNPVDPPGRPEVTDVTRSTVSLIW SAPVYDGGSKVVGYIIERKPVSEVGDGRWLKCNYTIVSDNFFTVTALSEGDTYEFRV LAKNAAGVISKGSESTGPVTCRDEYAPPKAELDARLQGDLVTIRAGSDLVLDAAVG GKPEPKIIWTKGDKELDLCEKISLQYTGKRATAVIKYCDRSDSGKYTLTVKNASGTK SVSVMVKVLDSPGPCGKLTVSRVTEEKCTLAWSLPQEDGGAEITHYIVERRETSRLN WVIVEGECLTASYVVTRLIKNNEYTFRVRAVNKYGLGVPVESEPIVARNSFTIPSQPGI PEEVGAGKEHIIIQWTKPESDGGNEISNYLVDKREKKSLRWTRVNKDYVVYDTRLK VTSLMEGCDYQFRVTAVNAAGNSEPSEASNFISCREPSYTPGPPSAPRVVDTTKRSISL AWTKPMYDGGTDIIGYVLEMQEKDTDQWCRVHTNATIRNNEFTVPDLKMGQKYSF RVAAVNAKGMSDYSETTAEIEPVERLEIPDLELADDLKKTVIVRAGASLRLMVSVSG RPSPVITWSKKGIDLANRAIIDNTESYSLLIVDKVNRYDAGKYTIEAENQSGKKSATV LVKVYDTPGPCPSVSVKEVSRDSVTITWEIPTIDGGAPVNNYIIEKREAAMRAFKTVT TKCSKTLYRISGLVEGTMYYFRVLPENIYGIGEPCETSDAVLVSEVPLVPTKLEVVDV TKSTVTLAWEKPLYDGGSRLTGYVLEACKAGTERWMKVVTLKPTVLEHTVISLNEG EQYLFRVRAQNEKGVSEPREIVTPVTVQDLRVLPTIDLSTMPQKTIHVPAGRPIELVIPI TGRPPPTASWFFAGSKLRESERVTVETHTKVTKLTIRETTIRDTGEYTLELKNVTGTTS ETIKVIILDKPGPPTGPIKIDEIDATSVTISWEPPELDGGAPLSGYVVEQRDAHRPGWLP VSESVTRPTFKFTRLTEGNEYVFRVAATNRFGIGSYLQSEVIECRSSISIPGPPETLQIFD VSRDGMTLTWYPPEDDGGSQVTGYIVERKEVRADRWVRVNKVPVTMTRYRSTGLI EGLEYEHRVTAINARGTGKPSRPSKPTVAMDPIAPPGKPQNPRVTDTTRTSVSLAWS VPEDEGGSKVTGYLIEMQKVDQREWTKCNTTPTKIREYTLTHLPQGAEYRFRVLAC NAGGPGEPAEVPGTVKVTEMLEYPDYELDERYQEGVFVRQGGVIRLTIPIKGKPFPV CKWTKEGQDISKRAMIATSETHTELVIKEADRNDSGTYDLVLENKCGKKTVYIKVK VIGSPNTPEGPLEYDDIQARSVRVSWRPPADDGGADILGYILERREVPKAAWYTIDSR VRGTSLVVKGLKENVEYHFRVSAENQFGISKPLKSEEPVIPKTPLNPPEPPSNPPEVLD VTKSSVSLSWSRPKDDGGSRVTGYYIERKETSTDKWVRHNKTQITTTMYTVTGLVP DAEYQFRIIAQNDVGLSETSPASEPVVCKDPFDKPSQPGELEILSISKDSVTLQWEKPE CDGGKEILGYWVEYRQSGDSAWKKSNKERIKDRQFTIGGLLEATEYEFRVFAENETG LSRPRRTAMSVKTKLTSGEAPGVRKEMADVTTKLGEAAQLSCQIVGRPLPDIKWYRF GKELIQSRKYKMSSDGRTHTLTVMTDEQEDEGVYTCVATNEVGEVESSSKLLLQAA PQFHPGYPLKEKYYGAVGSTLRLHVMYIGRPVPAMTWFHGQKLLQNSEKITIENTEH YTHLVMKNVQRKTHAGKYKVQLSNAFGTVDATLDVEIQDKPDKPTGPIVIEALLKN SVVISWKPPADDGGSWITNYVVEKCEAKEGAEWQLVSSAISVTTCRIVNLTENAGYY FRVSAQNTFGISEPLEVASIVIIKSPFEKPGVPGKPTITAVTKDSCVVAWKPPASDGGA KIRNYYLERREKKQNKWIAVTTEEIRETVFSVQNLIEGLEYEFRVKCENLGGESEWSE ISEPVTPKSDVPIQAPHFKEELRNLNVRYQSNATLVCKVTGHPKPIVKWYRQGKEIIA

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		DGLKYRIQEFKGGYHQLIIASVTDDDATVYQVRATNQGGSVSGTASLEVEVPAKIHL PKTLEGMGAVHALRGEVVSIKIPFSGKPDPVITWQKGQDLIDNNGHYQVIVTRSFTSL VFSNGVERKDAGFYVVCAKNRFGIDQKTVELDVADVPDPPRGVKVSDVSRDSVNLT WTEPASDGGSKVTNYIVEKCATTAERWLRVGQARETRYTVINLFGKTSYQFRVIAEN KFGLSKPSEPSEPTVTKEDKTRAMNYDDEVDETREVTTTKASHSKTKELYEKYMIAE DLGRGEFGIVHRCVETSSKRTFMAKFVKVKGTDQVLVKKEISILNIARHRNILYLHES FESMEELVMIFEFISGLDIFERINTSAFELNEREIVSYVRQVCEALEFLHSQNIGHFDIRP ENIIYQTRKNSTIKIIEFGQARQLKPGDNFRLLFTAPEYYAPEVHQHDVVSTATDMWS LGTLVYVLLSGINPFLAETNQQMIENIMNAEYTFDEEAFKEISLEAMDFVDRLLVKER KSRMTASEALQHPWLKQRIDRVSTKVIRTLKHRRYYHTLIKKDLNMVVSAARISCGG AIRSQRGVSVAKVKVASIEIGPVSGQIMHAIGEEGGYVKYVCKIENYDQSTQVTWYF GVRQLENSEKYEITYEDGVATMYVKDITKFDDGTYRCKVVNDYGEDSSYAELFVKG VREVYDYYCRRTKKVKRRTDAMRLLERPPEFTLPLYNKTAYVGENVRFGVTITVHP EPRVTWYKSGQKIKPGDDEKKYTFESDKGLYQLTINSVTTDDDAEYTVVARNKHGE DSCKAKLTVTLHPPPTETTLRPMFKRLLANAECHEGQSVCFEIRVSGIPAPTLKWEKD GQPLSLGPHIEIVHEGLDYYALHIRDTLPEDTGYYRVTATNTAGSTSCQAHLQVERLR YVKQEFQSKEERERHVQKQIDKTLRMAEILSGTETVPLTPVAQEALREAAILYKPAVS TKTVKGEYRLQTEEKKEERKLRMPYEVPEPRRFKQATVEEDQRIKQFVPMSDMKWY KKIRDQYEMPGKLDRVVQKRPKRIRLSRWEQFYVMPLPRITDQYRPKWRIPKLTQD DLEMVRPARRRTPSPDYDLYYYRRRRRSLGDMSDEELLLPIDDYLAMKRTEEERLRL EEELELGFSASPPSRSPPRFELSSLRYSSPPAHVKVEDRRRDFRYSTYHVPTKEETSTS YAELRERHAQASYRQPKLRQRIMAEKEEEELLRPVTTTQRLSEYKSELDYMSKEEKS KKKSKRQRQVTEITEIEEEYEISRRAQRESSSSVSRLLRRRRSLSPTYIELMRPVSELIRS HPRPAEEYEDDAERRSPTPERTRPRSPSPVSSERSLSRFERSARFDIFSRYESMKAALK TQKTSERKYEVLSQQPFTLDHAPRITLRMRSHRVPCGQNTRFILNVQSKPTAEVKWY HNGVELQESSKIHYTNTSGVLTLEILDCQTEDGGTYRAVCTNYKGEASDYATLDVTG GAYTTYASQRRDEEVPKSVFPELTKTEAYAVSSFKRTSELEAASSVREVKSQMTETR ESLSTYEHYASAEMKSATSEEKSLEEKATVRKIKTTLAARILTKPRSITVHEGESARFS CDTDGEPVPTVTWLREGQVVSTSARHQVTTTKYKSTFEISSVQASDEGNYSVVVENS DGKQEAQFTLTVQKARVIEKAVTSPPRVKSPEPRVKSPETVKSPKRVKSPEPVTSHPK AVSPTETKPTEKGQHLPVSAPPKITQSLKAEASKDIAKLTCAVESSALCAKEVAWYK DGKKLKENGHFQFHYSADGTYELKIHNLSESDCGEYVCEVSGEGGTSKTSFQFTGQS FKSIHEQVSSISETTKSVQKTAESPEAKKQEPIAPESISSKPVIVTGLRDTTVSSDSVAKF TIKVTGEPQPTITWTKDGKAIAQGSKYKLSSNKEEFILEILKTETSDGGLYACTVTNSA GSVSSSCKLTIKAVKDTEAQKVSTQKTSEVTSQKKASAQEEISQKALTSEEIKMSEVK SHETLAIKEEASKVLIAEEVKRSAAASLEKSIVHEEVTKTSQASEEVKTHAEIKTLSTQ MNITKGQRATLKANIAGATDVKWVLNGTELPNSEEYRYGVSGSDQTLTIKQASHRE EGILSCIGKTSQGVVKCQFDLTLSEELSDAPSFITQPRSQNINEGQNVLFSCEVSGEPSP EIEWFKNNLPISISSNISVSRSRNVYTLEIRNAAVSDSGKYTIKAKNFHGQCSATASLT VLPLVEEPPREVVLKTSSDVSLHGSVSSQSVQMSASKQEASFSSFSSSSASSMTEMKF ASMSAQSMSSMQESFVEMSSSSFMGKSSMTQLESSTSRMLKAGGRGIPPKIEALPSDI SIDEGKVLTVACAFTGEPTPEITWSCGGRKIQNQEQQGRFHIENTDDLTTLIIMDVQK QDGGLYTLSLGNEFGSDSATVNINIRSM
2	Mouse MYH4 (Myosin-4)	MSSDAEMAVFGEAAPYLRKSEKERIEAQNKPFDAKSSVFVVDAKESYVKATVQSRE GGKVTAKTEGGATVTVKDDQVFSMNPPKYDKIEDMAMMTHLHEPAVLYNLKERY AAWMIYTYSGLFCVTVNPYKWLPVYNPEVVAAYRGKKRQEAPPHIFSISDNAYQFM LTDRENQSILITGESGAGKTVNTKRVIQYFATIAVTGDKKKEEATSGKMQGTLEDQII SANPLLEAFGNAKTVRNDNSSRFGKFIRIHFGATGKLASADIETYLLEKSRVTFQLKA ERSYHIFYQIMSNKKPELIEMLLITTNPYDFAYVSQGEITVPSIDDQEELMATDTAVDI LGFSADEKVAIYKLTGAVMHYGNMKFKQKQREEQAEPDGTEVADKAAYLTSLNSA DLLKALCYPRVKVGNEYVTKGQTVQQVYNSVGALAKSMYEKMFLWMVTRINQQL DTKQPRQYFIGVLDIAGFEIFDFNTLEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGI DWEFIDFGMDLAACIELIEKPMGIFSILEEECMFPKATDTSFKNKLYEQHLGKSNNFQ KPKPAKGKAEAHFSLVHYAGTVDYNIIGWLDKNKDPLNETVVGLYQKSGLKTLAFL FSGGQAAEAEGGGGKKGGKKKGSSFQTVSALFRENLNKLMTNLKSTHPHFVRCLIP NETKTPGAMEHELVLHQLRCNGVLEGIRICRKGFPSRILYADFKQRYKVLNASAIPEG QFIDSKKASEKLLGSIDIDHTQYKFGHTKVFFKAGLLGTLEEMRDEKLAQLITRTQAV CRGYLMRVEFKKMMERRESIFCIQYNVRAFMNVKHWPWMKLYFKIKPLLKSAETE KEMANMKEDFEKAKEDLAKSEAKRKELEEKMVALMQEKNDLQLQVQAEADGLAD AEERCDQLIKTKIQLEAKIKELTERAEDEEEINAELTAKKRKLEDECSELKKDIDDLEL TLAKVEKEKHATENKVKNLTEEMAGLDENIAKLTKEKKALQEAHQQTLDDLQAEE DKVNTLTKAKTKLEQQVDDLEGSLEQEKKLRMDLERAKRKLEGDLKLAQESTMDIE NDKQQLDEKLKKKEFEMSNLQSKIEDEQALGMQLQKKIKELQARIEELEEEIEAERA SRAKAEKQRSDLSRELEEISERLEEAGGATSAQIEMNKKREAEFQKMRRDLEEATLQ HEATAAALRKKHADSVAELGEQIDNLQRVKQKLEKEKSELKMEIDDLASNMETVSK AKGNLEKMCRTLEDQLSEVKTKEEEQQRLINELSTQKARLHTESGEFSRQLDEKDA MVSQLSRGKQAFTQQIEELKRQLEEESKAKNALAHALQSARHDCDLLREQYEEEQE AKAELQRAMSKANSEVAQWRTKYETDAIQRTEELEEAKKKLAQRLQDAEEHVEAV NSKCASLEKTKQRLQNEVEDLMIDVERSNAACAALDKKQRNFDKVLAEWKQKYEE TQAELEASQKESRSLSTELFKVKNAYEESLDQLETLKRENKNLQQEISDLTEQIAEGG

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		KHIHELEKIKKQIDQEKSELQASLEEAEASLEHEEGKILRIQLELNQVKSEIDRKIAEKD EEIDQLKRNHLRVVESMQSTLDAEIRSRNDALRIKKKMEGDLNEMEIQLNHANRQA AEAIRNLRNTQGMLKDTQLHLDDALRGQDDLKEQLAMVERRANLMQAEIEELRAS LEQTERSRRVAEQELLDASERVQLLHTQNTSLINTKKKLETDLRISQIQGEMEDIVQEAR NAEEKAKKAITDAAMMAEELKKEQDTSAHLERMKKNMEQTVKDLQHRLDEAEQL ALKGGKKQIQKLEARVRELENEVENEQKRNIEAVKGLRKHERRVKELTYQTEEDRK NVLRLQDLVDKLQTKVKAYKRQAEEAEEQSNVNLAKFRKIQHELEEAEERADIAES QVNKLRVKSREVHTKVISEE
3	Mouse E9Q1W3 (Nebulin)	MADEEEYEEVIEYYTEETVYEEVPGETITEVYETTTTRTSDYGQSETATPALAQPQPA KPVERKKVIRKKVDSSKFMTPYIQHSQKMQDLFSSNKYKENYEKAKGQPYAITSDTP ELRRIKKVQDQLSEVKYRMDGDVAKTICHVDEKAKDIEHAKKVSQQVSKVLYKQN WEDTKDKYLLPPDAPELVQAVKNTAMFSKKLYTEDWEADKALFYPYNDSPELRRV AQAQKALSDIAYKKGLTEQQTQFTSLPDPPEVEFAKKVTNQVSKQKYKEDYEKKVK GKWSETPCFEIATARMNADNLSSRKYQEDFENMKDQIYFMQTETPEYKMNKQAGV AASKVKYKQDYEKNKGKADYNVLPASENPLLRQLKTAGNVLSDKLYKENYEKTKA RSINYCETPKFQLDTVLQNFSSDTKYKDSYLKNILGHYVGSFEDPYHTHCMKVTAQN SDKNYKAEYEEDRGKGFFPQTITQEYEAIKKLDQCKDHAYKVHPDKTKFTQVTDSP VLVQAQVNSKQLSDLNYKAKHESEKFKCHIPPDAPAFIQHRVNAYNLSDNVYKQDW EKSKAKKFDIKVDAIPLLAAKANSKNASDVMYKKDYEKSKGKMIGALSINDDPKML HSLKTAKNQSDRLYRENYEKTKAKSMNYCETPKYQLDTLLKNFSEAKYKDSYVKN MLGHYVGSFEDPYHTHCMKVSAQNSDKSYKAEYEEDKGKCYFPQTITQEYESIKKL DQCKDHTYKVHPDKTRFTAVTDTPVLLQAQLNTKQLSDLNYKAKHEGEKFKCSVP ADAPEFIQHRVNAYNLSDNVYKHDWEKSKAKKFDIKVDAIPLLAAKANSKNASDV MYKKDYEKSKGKMIGALSINDDPKMLHSLKTAKNQSDREYRKDYEKSKTIYTAPLD MIQVTQAKKSQAIASDVDYKHVLHNYSYPPDSINVDLAKKAYALQSDVEYKADYNS WMKGCGWVPFGSLEMEKAKRASEILNEKKYRQHPDTLKFTSIEDAPITVQSKINQAQ RSDIAYKAKGEEILHKYNLPADLPQFIQAKVNAYNISENMYKADLKDLSKKGYDLRI DAIPIKAAKAARQAASDIHYKKDYEKAKGKMVGFQSLQDDPKLVHYMNVAKIQSD REYKKDYEKTKTKYNTPHDMFDVVAAKKAQDVISNTNYKHSLHHYTYLPDAMDLE LSKNMMHIQSDNAYKEDYNTWMKGIGWIPIGSLDQEKVKKAGDALSEKKYRQHPD TLKFTSIVDSPVMVQAKQNTQQASDILYKAKGEDVKHKYTLNPDLPQFLQAKCNAY NISDVCYKRDWHDLIAKGNNVLADAIPITAAKSSRNIASDYKYKEAYEKAKGKHVG FRSLQDDPKLVHYMNVAKLQSDREYKKNYENTKTSYHTPGDMVSITAARMAQDVA TNVNYKQPLHHYTYLPDALSLEHTRNANQIQSDNVYKDEYNSFFKGIGWIPIGSLEV EKAKKAGEALNERKYRQHPDTIKFTSVPDSMGMVLAQQNTKQLSDLNYKVEGEKM KHKYTMDPELPQFIQAKVNAINMSDAHYKADWKKTLAKGYDLRPDAIPIVAAKSSR NIASDCKYKEAYEKARGKQIGFLSLQDDPKLVHYMNVAKIQSDREYKKGYEASKTR YHTPLDMFSVTAAKKSQEVATNSNYRQPYHNYTLLPDALNVEHSRNAMQIQSDNLY KSDFTHWMKGIGWVPIDSLEVEKAKKAGEILSERKYRQHPEKLKFTYSMDTMEQAL NKSNKLTMDKRLYTEKWNKDKTSIHVMPDTPDILLSRVNQITMSDKLYKAGWEEEK TKGYDLKPDAIAIKAAKASRDIASDYKYKQAYEQSKGKHIGFRSLEDDPKLVHFMQ VAKIQSDREYKKGYEKSKTSFHTPVDMFSVVAAKKSQEVATNANYRNIIHTYNMLP DAMSFELAKNMMQIQSDNQYKADYADFMKGIGWLPLGSLEAEKNKKAMEIISEKK YRQHPDTLKYSTLMDSMNMVLAQNNAKIMNKNLYKQAWEADKTKIHIMPDIPEIVL AKANAINMSDKLYKLSLEESRKKGYDLRTDAIPIKAAKASRDIASDYKYKHSYEKEK GKMVGFRSLEDDPKLVHSMQVAKMQSDREYKKNYEKTKTSYHTPADMLSVTAAK DAQANITNTNYKRLIHKYILLPDAMNIELTRNMNHIQSDNEYKQDYREWYRGLGWS PAGSLEVEKAKKATEYASDQKYRQHPSTFQFKKLTDSMDMVLAKQNAHTMNKHLY TIDWNKDKTKIHVMPDTPDILQAKQNQTLYSQKSYRLGWEEALKKGYDLPVDAISV QLAKTSRDIASDFKYKQGYRKQLGHHIGFLSVQDDPKLVLSMNVAKMQSDREYKK DFEKWKTKYTSPVDMLGVVLAKKCQALVSDADYRNYLHQWTCLPDQNDVIQAKK VYELQSENMYKSDLEWLRGIGWSPLGSLEAEKNKRASEIISEKKYRQPPDRNKFTSIP DAMDIVLAKTNAKNRSDRLYREAWDKDKTQIHIMPDTPDIILAKANLINTSDKLYKM GYEELKKKGYDLPVDAIPIKAAKASREIASEYKYKEGFRKQLGHHIGARNIKDDPKM MWSMHVAKIQSDREYKKDFEKWKTKYSSPVDMLGVVLAKKCQTLVSDADYRNYL HQWTCLPDQNDVIHARQAYDLQSDNVYKADLQWLKGIGWIPSGSLEDEKNKRATQI LSDHVYRQHPNKFKFSSLMDSMPMVLAKNNAITMNHHLYTEAWNKDKTTVHIMPD TPEVILAKQNKVNYSEKLYKLGLEEAKRKGYDMRLDAIPIKTAKASRDIASDFKYKE GYRKQLGHHIGARGIHDDPKMMWSMHVAKIQSDREYKKDFEKWKTKYSSPVDML GVVLAKKCQTLVSDADYRNYLHQWTCLPDQNDVIHARQAYDLQSDNMYKSDLQW MRGIGWVPIGSLDVEKCKRATEILSDKIYRQPPDKFKFTSVTDSLEQVLAKNNAITMN KRLYTEAWDKDKTQIHIMPDTPEIMLARMNKVNYSESLYKLANEEAKKKGYDLRSD AIPIVAAKASRDIASDYKYKDGYRKQLGHHIGARNIKDDPKMMWSMHVAKIQSDRE YKKDFEKWKTKYSSPVDMLGVVLAKKCQILVSDIDYKHPLHEWTCLPDQNDVIHAR QAYDLQSDNVYKSDLQWMRGIGWVPIGSLDVVKCKRAAEILSDNIYRQPPNKFKFT SVTDSLEQVLAKSNALNMNKRLYTEAWDKDKIQIHVMPDTPEIMLARQNRINYSESL YKLANEEAKKKGYDLRVDAIPIVAAKASRDIASDYKYKVGYRKQLGHHIGARNIED DPKMMWSMHVAKIQSDREYKKDFEKWKTKYSSPVDMLGVVLAKKCQTLVSDVDY RNYLHQWTCLPDQNDVIHARQAYDLQSDNVYKSDLQWMRGIGWVPIGSLDVVKCK RAAEILSDNIYRQPPNKFKFTSVTDSLEQVLAKSNALNMNKRLYTEAWDKDKTQIHI

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		MPDTPEIMLARQNKINYSESLYRQAMEEAKKEGYDLRSDAIPIVAAKASRDIASDYK YKEAYRKQLGHHIGARAIHDDPKMMWSLHIAKVQSDREYKKEFEKYKTRYSSPVD MLGIVLAKKCQTLVSDVDYKHPLHEWTCLPDQNDVIQARKAYDLQSDNLYKADLE WMKGIGWVPIGSLEVVKAKRATEILSDNIYRQRPDTLKFTSITDSLEQVLAKNNAITM NKRLYTEAWDNDKKTIHVMPDTPEIMLAKLNRINYSDKLYKVSLEEARRQGYDLRV DAIPIRSAKASREIASDYKYKEGYRKQLGHHIGARNIKDDPKMMWSMHVAKIQSDR EYKKDFEKWKTKYSSPVDMLGVVLAKKCQILVSDIDYKHPLHEWTCLPDQNDVIQA RKAYDLQSDAVYKADLEWLRGIGWVPIGSVEVEKVKRAGEILSDRKYRQPADQLKF TCITDTPEIVLAKSNAITMSKHLYTEAWDADKASIHVMPDTPEILLAKSNSANISHKL YTKGWDESKMKDYDLREDAISIKSAKASRDIASDYKYKEAYEKQKGHHIGAQSVED DPRILCAMHAGKIYSDREYKKEFQKWKTKFSSPVDMLGVVLAKKCQTLVSDVDYR NYLHNWTCLPDQNDIIQAKKAYELQSDSVYKADLEWLRGIGWMPEGSIEMTRVKG AQDLVNERLYRTRPEALKFTSIVDTPEVVLAKANSLQMSERLYQEAWNKDKTNISLP LDTPAMLQAQINALQISNKLYQKDWDEAKQKGYDLRADAIEIKHAKASREIASEYK YKEGYRKQLGHHVGFRSLQDDPKLVWSIHAAKIQSDREYKKAYEKSKGIYNTPLDM MSIVQAKKCQVLVSDVDYRNYLHQWTCLPDQNDVIQAKKAYELQSDAIYKADLEW LRGIGWMPEGSPEVLRVKNAQEILCDSVYRTPVVKLKYTSIVDTPEVVLAKSNAENIS IPKYREVWDKDKTSIHIMPDTPEINLARANALNVSNKLYREGWDEVKMSCDVRLDAI PIQAAKASREIASDYKYKLDHEKQKGHYVGTLTARDDNKIRWALIAGKIQNEREYRL QWAKWKTKFQSPVDMLSILHSKKCQTLVSDIDYRNYLHEWTCMPDQNDVIQAKKA YELQSDAVYKADLEWLRGIGWMPNDSVSVNHAKHAADIFSEKKYRTKIETLNFTPV DDRVDYVTAKQSSEIRDDIKYRKDWNAIKSKYTLTETPLLHTAQEAARILDQYLYKE GWEKQKATGYILPPDAVPFVHAHHSSDVQSELKYKAEHVKQRGHYVGVPTMRDDP KLVWFEHAGQIQNDRLYKESYHKSKAKINIPADMVSVVAAKEGQNLVSDIDYRQYL HQWTCHPDQNDVIQAKIAYDLQSDNVYKADLEWLRGIGWIPLDSVDHVRVTKNQE MVNQIKYKKDALANYPNFTSVEDPPEIVLAKINSVNQSDVKYKETFNKAKGKYIFSP DTPYISHSKDMGKLYSTILYKGAWEGTKAYGYTLDERYIPIVGAKHADLVNSELKYK ETYEKQKGHYLAGKEISEFPGVVHCLDFQQMRSALNYRRDYEDTKANVHIPNDMM NHVLAKRCQYILSDLEYRHYFHQWTSLPEEPNVIRARNAQEILSDNVYKDDLNWLK GIGCYVWDTPQILHAKKSYDLQSQLQYTAASKENLQNYNLVTDTPVYVTALQSGIN ASEVKYKENYHQIKDKYTTVLETADYDRTKHLKSLYSSNIYKEAWDKVKATSYNLP TSTLSLTHAKNQKHLASHIKYREEYEKFKALYTLPKSVDDDPNTARCLRVGKLNIDR LYRSVYEKNKMKINIVPDMVEMVTAKDSQKKVSEIDYRLHLHEWICHPDLQVNSHV RKVTDQISDIVYKDDLTWLKGIGCYVWDTPEILHAKHAYDLRNDIKYKAHVQKTRN DYKLVTDTPVYVQAVKSGKQLSDAVYHYDYVHSVRGKVAPTTKTVDLDRALHAY KLQSENLYKKGLRFLPTGYRLPVDTPHFNHTKDTRYMSSYFKYKEAYEHIKANGYT LGPKDVPFVHVRRVNNVTSERLYRELYHKLKDKIHTTPDTPEIRQVKKTQEAVSELIY KSDFFKMQGHMISLPYTPQVLHCRYVGDITSDIKYKEDLQILKGLGCFLYDTPDMVR SRHLRKLWSNYLYTDNARKMRDKYKVVLDTPEYRKVQELKTHLSELVYRASGKKQ KSIFTSVPDTPDLLRAKRGQKLQSQYLYVELATKERPHHHAGNQTTALQHAKDVKD MVSEKKYKIQYEKMKDKYTPVADTPILIRAKRAYWNASDLRYKETFQKTKGRYHT VKDALDIVYHRKVTDDISKVKYKENYMSQLGIWSSIPDRPEHFHHRAVTDAVSDVK YKEDLTWLKGIGCYAYDTPDLTLAEQNKTLYSKYKYKEVFERTKSNFKYVADCPIN RHFKYATQLMNEKKYRADYEQRKDKYHLVVDEPRHLLAKTAGDQISQIKYRKKYE KSKDKFTSVVDTPEHLRTTKVNKQISDILYKLEYNKARPRGYTTIHDTPMLLHVRKV KDEVSDLKYKEVYQRNKSNCTIKPDAVHIKAAKDAYKVNTNLDYKKQYEANKAY WKWTPDRPDFIQAAKSSLQQSDFEYKLDREFLKGCKLSVTDDKDMVLALRNSLIESD LKYKEKHVKERGSCHAVPDTPQILLAKTVSNLVSENKYKDHVKKHLAQGSYTTLPE TRDTVHVKAVTKHVSDTNYKKKFVKEKGKSNYSIMQEPPEVKHAMEVAKKQSDVA YKKDAKENLHYTTVADRPDIKKATQAAKQASEVEYRAKHRKEGSHGLSMLGRPDIE MAKKAAKLSSQVKYRENFDKEKGKTPKYNPKDSQLYKTMKDANNLASEVKYKAD LKKLHKPVTDMKESLIMHHVLNTSQLASSYQYKKNYEKSKGHYHTIPDNLEQLHHR EATELQSIVKYKEKYEKERGKPMLDFETPTYITAKESQQMQSGKEYRKDYEESIKGR NLTGLEVTPALLHVRHATKIASEKEYRKDLEESIRGRGLTEMEDTPDMLRAKNATQI LNEKEYKRDLELEVKGRGLNAMANETPDFMRARNATDIASQIKYKQSAEMEKANFT SVVDTPEIIHAQQVKNLSSQKKYKEDAEKSMSYYETVLDTPEMQRVRENQKNFSLL QYQCDLKNSKGKITVVQDTPEILRVKENQKNFSSVLYKEDVSPGTAIGKTPEMLRVK QTQDHISSVKYKEAIGQGTPIPDLPEVKRVKETQKHISSVMYKENPGTGIPTTVTPEIE RVKRNQENFSSVLYKESLGKGTPTAITPEMERVKRNQENFSSVLYKENMGKGTPLPV TPEMERVKHNQENISSVLYKENVGKATPTPVTPEMQRVKRNQENISSVLYKENMGK ATPTPFTPEMERVKRNQENFSSVLYKENMRKATPTPVTPDMERAKRNQENISSVLYS DSFRKQIQGKAAYVLDTPEMRRVRETQRHISTVKYHEDFEKHKGCFTPVVTDPITER VKKNTQDFSDINYRGIQRKVVEMEQKRNDQDQDTITGLRVWRTNPGSVFDYDPAED NIQSRSLHMINVQAQRRSRDQSRSASALSISGGEEKSEHSEAADQRLSTYSDGGAFFS ATSTAYKHVKTTQLPQQRSSSVATQQTTLSSIPSHPSTAGKIFRAIYDYIAADADEVSF KDGDAIVNAQAIDEGWMYGTVQRTGRTGMLPANYVEAI
4	Mouse ACTS (Actin)	MCDEDETTALVCDNGSGLVKAGFAGDDAPRAVFPSIVGRPRHQGVMVGMGQKDSY VGDEAQSKRGILTLKYPIEHGIITNWDDMEKIWHHTFYNELRVAPEEHPTLLTEAPLN PKANREKMTQIMFETFNVPAMYVAIQAVLSLYASGRTTGIVLDSGDGVTHNVPIYEG YALPHAIMRLDLAGRDLTDYLMKILTERGYSFVTTAEREIVRDIKEKLCYVALDFEN

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		EMATAASSSSLEKSYELPDGQVITIGNERFRCPETLFQPSFIGMESAGIHETTYNSIMK CDIDIRKDLYANNVMSGGTTMYPGIADRMQKEITALAPSTMKIKIIAPPERKYSVWIG GSILASLSTFQQMWITKQEYDEAGPSIVHRKCF
5	Mouse KCRM (Creatine Kinase)	MPFGNTHNKFKLNYKPQEEYPDLSKHNNHMAKVLTPDLYNKLRDKETPSGFTLDDV IQTGVDNPGHPFIMTVGCVAGDEESYTVFKDLFDPIIQDRHGGYKPTDKHKTDLNHE NLKGGDDLDPNYVLSSRVRTGRSIKGYTLPPHCSRGERRAVEKLSVEALNSLTGEFK GKYYPLKSMTEQEQQQLIDDHFLFDKPVSPLLLASGMARDWPDARGIWHNDNKSFL VWVNEEDHLRVISMEKGGNMKEVFRRFCVGLQKIEEIFKKAGHPFMWNEHLGYVL TCPSNLGTGLRGGVHVKLANLSKHPKFEEILTRLRLQKRGTGGVDTAAVGAVFDISN ADRLGSSEVEQVQLVVDGVKLMVEMEKKLEKGQSIDDMIPAQK
6	Mouse ACTN3 (Alpha-Actinin-3)	MMMVMQPEGLGAGEGPFSGGGGGEYMEQEEDWDRDLLLDPAWEKQQRKTFTAW CNSHLRKAGTQIENIEEDFRNGLKLMLLLEVISGERLPRPDKGKMRFHKIANVNKAL DFIASKGVKLVSIGAEEIVDGNLKMTLGMIWTIILRFAIQDISVEETSAKEGLLLWCQR KTAPYRNVNVQNFHTSWKDGLALCALIHRHRPDLIDYAKLRKDDPIGNLNTAFEVA EKYLDIPKMLDAEDIVNTPKPDEKAIMTYVSCFYHAFAGAEQAETAANRICKVLAVN QENEKLMEEYEKLASELLEWIRRTVPWLENRVGEPSMSAMQRKLEDFRDYRRLHKP PRVQEKCQLEINFNTLQTKLRLSHRPAFMPSEGKLVSDIANAWRGLEQVEKGYEDW LLSEIRRLQRLQHLAEKFQQKASLHEAWTRGKEEMLNQHDYESASLQEVRALLRRH EAFESDLAAHQDRVEHIAALAQELNELDYHEAASVNSRCQAICDQWDNLGTLTQKR RDALERMEKLLETIDQLQLEFARRAAPFNNWLDGAIEDLQDVWLVHSVEETQSLLT AHEQFKATLPEADRERGAILGIQGEIQKICQTYGLRPKSGNPYITLSSQDINNKWDTV RKLVPSRDQTLQEELARQQVNERLRRQFAAQANAIGPWIQGKVEEVGRLAAGLAGS LEEQMAGLRQQEQNIINYKSNIDRLEGDHQLLQESLVFDNKHTVYSMEHIRVGWEQ LLTSIARTINEVENQVLTRDAKGLSQEQLNEFRASFNHFDRKRNGMMEPDDFRACLIS MGYDLGEVEFARIMTMVDPNAAGVVTFQAFIDFMTRETAETDTAEQVVASFKILAG DKNYITPEELRRELPAEQAEYCIRRMAPYKGSGAPSGALDYVAFSSALYGESDL
7	Mouse MYH1 (Myosin-1)	MSSDAEMAVFGEAAPYLRKSEKERIEAQNKPFDAKSSVFVVDAKESFVKATVQSRE GGKVTAKTEGGTTVTVKDDQVYPMNPPKYDKIEDMAMMTHLHEPAVLYNLKERY AAWMIYTYSGLFCVTVNPYKWLPVYNAEVVAAYRGKKRQEAPPHIFSISDNAYQFM LTDRENQSILITGESGAGKTVNTKRVIQYFATIAVTGEKKKEEATSGKMQGTLEDQIIS ANPLLEAFGNAKTVRNDNSSRFGKFIRIHFGTTGKLASADIETYLLEKSRVTFQLKAE RSYHIFYQIMSNKKPDLIEMLLITTNPYDYAFVSQGEITVPSIDDQEELMATDSAIDILG FTSDERVSIYKLTGAVMHYGNMKFKQKQREEQAEPDGTEVADKAAYLQNLNSADL LKALCYPRVKVGNEYVTKGQTVQQVYNSVGALAKAVYEKMFLWMVTRINQQLDT KQPRQYFIGVLDIAGFEIFDFNSLEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIE WEFIDFGMDLAACIELIEKPMGIFSILEEECMFPKATDTSFKNKLYEQHLGKSNNFQK PKPAKGKVEAHFSLVHYAGTVDYNIAGWLDKNKDPLNETVVGLYQKSSMKTLAYL FSGAAAAAEAESGGGGGKKGAKKKGSSFQTVSALFRENLNKLMTNLRSTHPHFVRC IIPNETKTPGAMEHELVLHQLRCNGVLEGIRICRKGFPSRILYADFKQRYKVLNASAIP EGQFIDSKKASEKLLGSIDIDHTQYKFGHTKVFFKAGLLGLLEEMRDDKLAQLITRTQ AMCRGYLARVEYQKMVERRESIFCIQYNVRAFMNVKHWPWMKLYFKIKPLLKSAE TEKEMANMKEEFEKAKENLAKAEAKRKELEEKMVALMQEKNDLQLQVQSEADSL ADAEERCDQLIKTKIQLEAKIKEVTERAEDEEEINAELTAKKRKLEDECSELKKDIDD LELTLAKVEKEKHATENKVKNLTEEMAGLDETIAKLTKEKKALQEAHQQTLDDLQA EEDKVNTLTKAKIKLEQQVDDLEGSLEQEKKIRMDLERAKRKLEGDLKLAQESTMD VENDKQQLDEKLKKKEFEMSNLQSKIEDEQALGMQLQKKIKELQARIEELEEEIEAE RASRAKAEKQRSDLSRELEEISERLEEAGGATSAQIEMNKKREAEFQKMRRDLEEAT LQHEATAATLRKKHADSVAELGEQIDNLQRVKQKLEKEKSEMKMEIDDLASNMEVI SKSKGNLEKMCRTLEDQVSELKTKEEEQQRLINELTAQRGRLQTESGEYSRQLDEKD SLVSQLSRGKQAFTQQIEELKRQLEEEIKAKSALAHALQSSRHDCDLLREQYEEEQEA KAELQRAMSKANSEVAQWRTKYETDAIQRTEELEEAKKKLAQRLQDAEEHVEAVN AKCASLEKTKQRLQNEVEDLMIDVERTNAACAALDKKQRNFDKILAEWKQKYEET HAELEASQKESRSLSTELFKIKNAYEESLDHLETLKRENKNLQQEISDLTEQIAEGGK RIHELEKIKKQIEQEKSELQAALEEAEASLEHEEGKILRIQLELNQVKSEIDRKIAEKDE EIDQLKRNHIRVVESMQSTLDAEIRSRNDAIRLKKKMEGDLNEMEIQLNHSNRMAAE ALRNYRNTQGILKDTQLHLDDALRGQEDLKEQLAMVERRANLLQAEIEELRATLEQ TERSRKIAEQELLDASERVQLLHTQNTSLINTKKKLETDISQIQGEMEDIVQEARNAEE KAKKAITDAAMMAEELKKEQDTSAHLERMKKNLEQTVKDLQHRLDEAEQLALKG GKKQIQKLEARVRELEGEVENEQKRNVEAIKGLRKHERRVKELTYQTEEDRKNVLR LQDLVDKLQSKVKAYKRQAEEAEEQSNVNLAKFRKIQHELEEAEERADIAESQVNK LRVKSREVHTKIISEE
8	Mouse PYGM (glycogen Phosphorylase)	MSRPLSDQDKRKQISVRGLAGVENVSELKKNFNRHLHFTLVKDRNVATPRDYYFAL AHTVRDHLVGRWIRTQQHYYEKDPKRIYYLSLEFYMGRTLQNTMVNLALENACDE ATYQLGLDMEELEEIEEDAGLGNGGLGRLAACFLDSMATLGLAAYGYGIRYEFGIFN QKICGGWQMEEADDWLRYGNPWEKARPEFTLPVHFYGRVEHTSQGAKWVDTQVV LAMPYDTPVPGYRNNVVNTMRLWSAKAPNDFNLKDFNVGGYIQAVLDRNLAENIS RVLYPNDNFFEGKELRLKQEYFVVAATLQDIIRRFKSSKFGSRDPVRTNFDAFPDKVA

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		IQLNDTHPSLAIPELMRILVDLERLDWDKAWDVTVKTCAYTNHTVLPEALERWPVH LMETLLPRHLQIIYEINQRFLNRVAAAFPGDVDRLRRMSLVEEGAVKRINMAHLCIA GSHAVNGVARIHSEILKKTIFKDFYELEPHKFQNKTNGITPRRWLVLCNPGLAEVIAE RIGEDYISDLDQLRKLLSYVDDEAFIRDVAKVKQENKLKFSAYLEREYKVHINPNSLF DVQVKRIHEYKRQLLNCLHIITLYNRIKREPNRFMVPRTIMIGGKAAPGYHMAKMIIK LITAIGDVVNHDPAVGDRLRVIFLENYRVSLAEKVIPAADLSEQISTAGTEASGTGNM KFMLNGALTIGTMDGANVEMAEEAGEENFFIFGMRVEDVERLDQRGYNAQEYYDRI PELRQIIEQLSSGFFSPKQPDLFKDIVNMLMHHDRFKVFADYEEYIKCQDKVSELYKN PREWTRMVIRNIATSGKFSSDRTIAQYAREIWGVEPSRQRLPAPDEKI
9	Mouse ALDOA (fructose-bisphosphate aldolase A)	MPHPYPALTPEQKKELSDIAHRIVAPGKGILAADESTGSIAKRLQSIGTENTEENRRFY RQLLLTADDRVNPCIGGVILFHETLYQKADDGRPFPQVIKSKGGVVGIKVDKGVVPL AGTNGETTTQGLDGLSERCAQYKKDGADFAKWRCVLKIGEHTPSALAIMENANVLA RYASICQQNGIVPIVEPEILPDGDHDLKRCQYVTEKVLAAVYKALSDHHVYLEGTLL KPNMVTPGHACTQKFSNEEIAMATVTALRRTVPPAVTGVTFLSGGQSEEEASINLNAI NKCPLLKPWALTFSYGRALQASALKAWGGKKENLKAAQEEYIKRALANSLACQGK YTPSGQSGAAASESLFISNHAY
10	Mouse MYL 1 (myosin light chain)	MAPKKDVKKPAAAPAPAPAPAPAPAKPKEEKIDLSAIKIEFSKEQQEDFKEAFLLFDR TGECKITLSQVGDVLRALGTNPTNAEVKKVLGNPSNEEMNAKKIEFEQFLPMMQAIS NNKDQGGYEDFVEGLRVFDKEGNGTVMGAELRHVLATLGEKMKEEEVEALLAGQE DSNGCINYEAFVKHIMSV
11	Mouse G3P (glyceraldehyde-3-phosphate dehydrogenase)	MVKVGVNGFGRIGRLVTRAAICSGKVEIVAINDPFIDLNYMVYMFQYDSTHGKFNG TVKAENGKLVINGKPITIFQERDPTNIKWGEAGAEYVVESTGVFTTMEKAGAHLKGG AKRVIISAPSADAPMFVMGVNHEKYDNSLKIVSNASCTTNCLAPLAKVIHDNFGIVE GLMTTVHAITATQKTVDGPSGKLWRDGRGAAQNIIPASTGAAKAVGKVIPELNGKL TGMAFRVPTPNVSVVDLTCRLEKPAKYDDIKKVVKQASEGPLKGILGYTEDQVVSC DFNSNSHSSTFDAGAGIALNDNFVKLISWYDNEYGYSNRVVDLMAYMASKE
12	Mouse TPM1 (tropomyosin alpha-1)	MDAIKKKMQMLKLDKENALDRAEQAEADKKAAEDRSKQLEDELVSLQKKLKGTE DELDKYSEALKDAQEKLELAEKKATDAEADVASLNRRIQLVEEELDRAQERLATAL QKLEEAEKAADESERGMKVIESRAQKDEEKMEIQEIQLKEAKHIAEDADRKYEEVAR KLVIIESDLERAEERAELSEGKCAELEEELKTVTNNLKSLEAQAEKYSQKEDKYEEEI KVLSDKLKEAETRAEFAERSVTKLEKSIDDLEDELYAQKLKYKAISEELDHALNDMT SI
13	Mouse MYPC2 (myosin-binding protein C)	MPEAKPAAKKASKGKDAPKEAPAKQTPEEPPKEAPPEDQSPTAEEPTGIFLKKPDSVS VETGKDAVILAKVNGKELPGKPTIKWFKGKWQELGSKSGARFIFKESHDSTSNVYTV ELHIGKVVLGDRGDYRLEIKAKDVCDSCSFNVDVEAPRQDSSGQSLESFKRSGDGKS EDAGELDFSGLLKKREVVEEEKKKKKDDDDLGIPPEIWELLKGAKKSEYEKIAFQYG ITDLRGMLKRLKKAKVEVKKSAAFTKKLDPAYQVDRGNKIKLVVEISDPDLPLKWF KNGQEIKPSSKYVFENVGKKRILTINKCTLADDAAYEVAVQDEKCFTELFVKEPPVLI VTPLEDQQVFVGDRVEMSVEVSEEGAQVMWMKDGVEMTREDSYKARYRFKKDGK RHILIYSDVAQEDGGRYQVITNGGQCEAELIVEEKQLEVLQDIADLTVKAAEQAVFK CEVSDEKVTGKWYKNGVEVRPSKRITISHVGRFHKLVIDDVRPEDEGDYTFVPDGYA LSLSAKLNFLEIKVEYVPKQEPPKIHLDCSGKTSDNSIVVVAGNKLRLDVAITGEPPPT ATWLRGDEVFTATEGRTHIEQRPDCSSFVIESAERSDEGRYTIKVTNPVGEDVASIFLR VVDVPDPPEAVRVTSVGEDWAILVWEPPKYDGGQPVTGYLMERKKKGSQRWMKIN FEVFTDTTYESTKMIEGVLYEMRVFAVNAIGVSQPSMNTKPFMPIAPTSAPQHLTVED VTDTTTTLKWRPPDRIGAGGIDGYLVEYCLEGSEEWVPANKEPVERCGFTVKDLPTG ARILFRVVGVNIAGRSEPATLLQPVTIREIVEQPKIRLPRHLRQTYIRKVGEALNLVIPF QGKPRPQVVWTKGGAPLDTSRVNVRTSDFDTVFFVRQAARSDSGEYELSVQIENMK DTATIRIRVVEKAGPAENVMVKEVWGTNALVEWQPPKDDGNSEITGYFVQKADKK TMEWFNVYEHNRHTSCTVSDLIVGNEYYFRIFSENICGLSDSPGVSKNTARILKTGITL KPLEYKEHDFRTAPKFLTPLMDRVVVAGYTAALNCAVRGHPKPKVVWMKNKMEIH EDPKFLITNYQGILTLNIRRPSPFDAGTYSCRAFNELGEALAECKLDVRVPQ
14	Mouse MYPC3 (myosin-binding protein C)	MPEPGKKPVSAFNKKPRSAEVTAGSAAVFEAETERSGVMVRWQRDGSDITANDKY GLAAEGKRHTLTVRDASPDDQGSYAVIAGSSKVKFDLKVTEPAPPEKAESEVAPGAP EEVPAPATELEESVSSPEGSVSVTQDGSAAEHQGAPDDPIGLFLMRPQDGEVTVGGSI VFSARVAGASLLKPPVVKWFKGKWVDLSSKVGQHLQLHDSYDRASKVYLFELHITD AQTTSAGGYRCEVSTKDKFDSCNFNLTVHEAIGSGDLDLRSAFRRTSLAGAGRRTSD SHEDAGTPDFSSLLKKRDSFRRDSKLEAPAEEDVWEILRQAPPSEYERIAFQHGVEAC HRPLKRLKGMKQDEKKSTAFQKKLEPAYQVNKGHKIRLTVELADPDAEVKWLKNG QEIQMSGSKYIFESVGAKRTLTISQCSLADDAAYQCVVGGEKCSTELFVKEPPVLITR SLEDQLVMVGQRVEFECEVSEEGAQVKWLKDGVELTREETFKYRFKKDGRKHHLII NEATLEDAGHYAVRTSGGQSLAELIVQEKKLEVYQSIADLAVGAKDQAVFKCEVSD ENVRGVWLKNGKELVPDNRIKVSHIGRVHKLTIDDVTPADEADYSFVPEGFACNLSA KLHFMEVKIDFVPRQEPPKIHLDCPGSTPDTIVVVTGNKLRLDVPISGDPAPTVVWQK TVTQGKKASAGPHPDAPEDAGADEEWVFDKKLLCETEGRVRVETTKDRSVFTVEGA

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		EKEDEGVYTVTVKNPVGEDQVNLTVKVIDVPDAPAAPKISNVGEDSCTVQWEPPAY DGGQPVLGYILERKKKKSYRWMRLNFDLLRELSHEARRMIEGVAYEMRVYAVNAV GMSRPSPASQPFMPIGPPGEPTHLAVEDVSDTTVSLKWRPPERVGAGGLDGYSVEYC QEGCSEWTPALQGLTERRSMLVKDLPTGARLLFRVRAHNVAGPGGPIVTKEPVTVQ EILQRPRLQLPRHLRQTIQKKVGEPVNLLIPFQGKPRPQVTWTKEGQPLAGEEVSIRNS PTDTILFIRAARRTHSGTYQVTVRIENMEDKATLILQIVDKPSPPQDIRIVETWGFNVA LEWKPPQDDGNTEIWGYTVQKADKKTMEWFTVLEHYRRTHCVVSELIIGNGYYFRV FSHNMVGSSDKAAATKEPVFIPRPGITYEPPKYKALDFSEAPSFTQPLANRSIIAGYNA ILCCAVRGSPKPKISWFKNGLDLGEDARFRMFCKQGVLTLEIRKPCPYDGGVYVCRA TNLQGEAQCECRLEVRVPQ
15	Mouse Q14BI5 (myomesin 2)	MSLVAVPFYQKRHKHFDQSYRNIQTRYLLDQYALKKQATTQSSSQRSLTERSSSKRA SSQSSAGAMTCRLCAKRVSASEEEEVENENRYRSQAASYGEAKRQRFLSELAQLEEN VHLARSQARDKLDKYFMEQTVEDNLAWERHSFEDRMSRAPEILVRLRSHTIWERMS VRLCFTVQGFPTPVVQWYKDGSLICQAGEPGKYRIESRYGVHTLEINRANFEDTATY SAVATNSHGQVSTNAAVVVRRYRGEEEPFHSVGLPIGLPLSSVIPYTHFDVQFLEKFG VTFRREGETVTLKCTLLVTPDLKRVQPRAEWYRDDVLLKESKWTKMFFGEGQASLS FSHLNKDDEGLYTLRIVSRGGVSDHSAFMFVRDADPLVTGAPGAPMDLQCHDANRD YVIVTWKPPNTTTESPVIGYFIDKCEVGTNNWVQCNDAPVKICKYPVTGLFEGRSYV FRVRAVNNAGISRPSRISDAVAALDPVDLRRLQAIHLEGEKEIVIYQDDLEGDVQIPGP PTNVQASEVSRNYVVLSWDPPSPRGKDPLMYFIEKSAVGSGSWQRVNAQTAVRSPR YAVFDLAEGKSYVFRVLSANKHGLSDPSEITPPIQAQDMIVVPSAPGRVLASRNTKTS VVVQWDRPKHEEDLLGYYVDCCVAGTNMWEPCNHKPIGYNRFVVHGLTTGEQYIF RVKAVNAVGTSENSQESEVIKVQAALTVPSHPYGITLLNCDGHSMTLGWKVPKFSG GSAIIGYYLDKREVHHKNWHEVNSSPVKERILTVEGLTEGSLYEFKIAATNLAGIGQP SDPSEHFKCEAWTDPEPGPAYDLTFCEVRDTSLVILWKAPVYSGSSPVSGYFVDFKEE DSGEWKTTSEAATPNRYLKVCDLQQGKTYVFRIRAVNASGPGKPSDTSEPVLVEARP GTKEISAGVDEEGNIYLGFDCQEMTDASQFTWCKAYEEIADEERFEVHTEGDHSKLY FKNPDKEDLGTYSVSVSDTDGVSSSFVLDEEELERLMALSNEIKNPTIPLKSELAYEIF DKGQVRFWLQAEHLSPDASFRFIINDREVSDSDTHRIKCDRSTGMIEMVMDRFTIENE GTYTVQIQDGKAKNQSSLVLIGDAFKAVLEEAEFQRKEFLRKQGPHFAEYLHWDVT EECEVRLVCKVANTKRETVFKWLKDDVLYETETPPDLEKGVCELLIPKLSKKDHGE YKATLKDDRGQDVSVLEVGGKVYEDMILAMSRVCGASASPLKVLCTPEGIRLQCFM KYFTEEMKVSWYHKEAKISSSEHMRIGGSEEMAWLQICEPTEKDKGKYTFEIFDGKD SHQRSLDLSGQAFDEAYAEFQQLKAAAFAEKNRGKVIGGLPDVVTIMEGKTLNLTC TVFGNPDPEVVWFKNDKDIELSEHFLVKMEQSKYVSLTIQGVTAEDSGKYSINVKNK YGGEKIDVTVSVYKHGEKIPDISPPQQAKPKLIPASTSSD
16	Mouse ALBU (albumin)	MKWVTFLLLLFVSGSAFSRGVFRREAHKSEIAHRYNDLGEQHFKGLVLIAFSQYLQK CSYDEHAKLVQEVTDFAKTCVADESAANCDKSLHTLFGDKLCAIPNLRENYGELAD CCTKQEPERNECFLQHKDDNPSLPPFERPEAEAMCTSFKENPTTFMGHYLHEVARRH PYFYAPELLYYAEQYNEILTQCCAEADKESCLTPKLDGVKEKALVSSVRQRMKCSS MQKFGERAFKAWAVARLSQTFPNADFAEITKLATDLTKVNKECCHGDLLECADDR AELAKYMCENQATISSKLQTCCDKPLLKKAHCLSEVEHDTMPADLPAIAADFVEDQ EVCKNYAEAKDVFLGTFLYEYSRRHPDYSVSLLLRLAKKYEATLEKCCAEANPPAC YGTVLAEFQPLVEEPKNLVKTNCDLYEKLGEYGFQNAILVRYTQKAPQVSTPTLVEA ARNLGRVGTKCCTLPEDQRLPCVEDYLSAILNRVCLLHEKTPVSEHVTKCCSGSLVE RRPCFSALTVDETYVPKEFKAETFTFHSDICTLPEKEKQIKKQTALAELVKHKPKATA EQLKTVMDDFAQFLDTCCKAADKDTCFSTEGPNLVTRCKDALA
17	Mouse ENOB (beta-enolase)	MAMQKIFAREILDSRGNPTVEVDLHTAKGRFRAAVPSGASTGIYEALELRDGDKARY LGKGVLKAVEHINKTLGPALLEKKLSVVDQEKVDKFMIELDGTENKSKFGANAILG VSLAVCKAGAAEKGVPLYRHIADLAGNPDLVLPVPAFNVINGGSHAGNKLAMQEF MILPVGASSFKEAMRIGAEVYHHLKGVIKAKYGKDATNVGDEGGFAPNILENNEAL ELLKTAIQAAGYPDKVVIGMDVAASEFYRNGKYDLDFKSPDDPARHISGEKLGELYK NFIQNYPVVSIEDPFDQDDWATWTSFLSGVDIQIVGDDLTVTNPKRIAQAVEKKACN CLLLKVNQIGSVTESIQACKLAQSNGWGVMVSHRSGETEDTFIADLVVGLCTGQIKT GAPCRSERLAKYNQLMRIEEALGDKAVFAGRKFRNPKAK
18	Mouse MLRS (myosin regulatory light chain 2)	MAPKKAKRRAGAEGSSNVFSMFDQTQIQEFKEAFTVIDQNRDGIIDKEDLRDTFAAM GRLNVKNEELDAMMKEASGPINFTVFLTMFGEKLKGADPEDVITGAFKVLDPEGKG TIKKQFLEELLTTQCDRFSQEEIKNMWAAFPPDVGGNVDYKNICYVITHGDAKDQE
19	Mouse KPYR (pyruvate kinase PKLR)	MSVQENELPQQLWPWIFKSQKDLAKSALSGAPGGPAGYLRRASVAQLTQELGTAFF QQQQLPAAMADTFLEHLCLLDIDSEPVAARSTSIIATIGPASRSVDRLKEMIKAGMNI ARLNFSHGSHEYHAESIANIREAAESFATSPLSYRPVAIALDTKGPEIRTGVLQGGPES EVEIVKGSQVLVTVDPKFRTRGDAKTVWVDYHNITQVVAVGGRIYIDDGLISLVVRK IGPEGLVTEVEHGGFLGNRKGVNLPNAEVDLPGLSEQDLLDLRFGVEHYVDIIFASFV RKASDVVAVRDALGPEGRGIKIISKIENHEGVKKFDEILEVSDGIMMARGDLGIEIPAE KVFLAQKMMIGRCNLAGKPVVCATQMLESMITKARPTRAETSDVANAVLDGADCI

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		MLSGETAKGSFPVEAVKMQHAIAREAEAAVYHRQLFEELRRAAPLSRDPTEVTAIGA VEASFKCCAAAIIVLTKTGRSAQLLSRYRPRAAVIAVTRSAQAARQVHLSRGVFPLLY REPPEAVWADDVDRRVQFGIESGKLRGFLRVGDLVIVVTGWRPGSGYTNIMRVLTIS
20	Mouse CAH3 (carbonic anhydrase 3)	MAKEWGYASHNGPDHWHELYPIAKGDNQSPIELHTKDIKHDPSLQPWSASYDPGSA KTILNNGKTCRVVFDDTYDRSMLRGGPLSGPYRLRQFHLHWGSSDDHGSEHTVDGV KYAAELHLVHWNPKYNTFGEALKQPDGIAVVGIFLKIGREKGEFQILLDALDKIKTK GKEAPFTHFDPSCLFPACRDYWTYHGSFTTPPCEECIVWLLLKEPMTVSSDQMAKLR SLFSSAENEPPVPLVGNWRPPQPVKGRVVRASFK
21	Mouse LDHA (L-lactate dehydrogenase A)	MATLKDQLIVNLLKEEQAPQNKITVVGVGAVGMACAISILMKDLADELALVDVMED KLKGEMMDLQHGSLFLKTPKIVSSKDYCVTANSKLVIITAGARQQEGESRLNLVQRN VNIFKFIIPNIVKYSPHCKLLIVSNPVDILTYVAWKISGFPKNRVIGSGCNLDSARFRYL MGERLGVHALSCHGWVLGEHGDSSVPVWSGVNVAGVSLKSLNPELGTDADKEQW KEVHKQVVDSAYEVIKLKGYTSWAIGLSVADLAESIMKNLRRVHPISTMIKGLYGIN EDVFLSVPCILGQNGISDVVKVTLTPEEEARLKKSADTLWGIQKELQF
22	Mouse G3UW82 (myosin, heavy polypeptide 2)	MSSDAEMAVFGEAAPYLRKSEKERIEAQNRPFDAKTSVFVAEPKESFVKGTIQSKDA GKVTVKTEAGATLTVKEDQIFPMNPPKYDKIEDMAMMTHLHEPAVLYNLKERYAA WMIYTYSGLFCVTVNPYKWLPVYNPEVVAAYRGKKRQEAPPHIFSISDNAYQFMLT DRENQSILITGESGAGKTVNTKRVIQYFATIAVTGDKKKEEATSGKMQGTLEDQIISA NPLLEAFGNAKTVRNDNSSRFGKFIRIHFGTTGKLASADIETYLLEKSRVTFQLKAER SYHIFYQITSNKKPELIEMLLITTNPYDYPFVSQGEISVASIDDQEELMATDSAIDILGFT NDEKVSIYKLTGAVMHYGNMKFKQKQREEQAEPDGTEVADKAAYLQGLNSADLLK ALCYPRVKVGNEYVTKGQTVEQVTNAVGALAKAMYEKMFLWMVTRINQQLDTKQ PRQYFIGVLDIAGFEIFDFNSLEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEWTF IDFGMDLAACIELIEKPMGIFSILEEECMFPKATDTSFKNKLYEQHLGKSANFQKPKV VKGKAEAHFSLIHYAGTVDYNITGWLDKNKDPLNETVVGLYQKSSVKTLAYLFSGA QTAEAEASSGGAAKKGAKKKGSSFQTVSALFRENLNKLMTNLRSTHPHFVRCIIPNE TKTPGAMEHELVLHQLRCNGVLEGIRICRKGFPSRILYADFKQRYKVLNASAIPEGQY IDSKKASEKLLGSIDIDHTQYKFGHTKVFFKAGLLGLLEEMRDDKLAQLITRTQAMC RGFLARVEYQKMVERRESIFCIQYNIRAFMNVKHWPWMKLFFKIKPLLKSAETEKE MATMKEEFQKTKDDLAKSEAKRKELEEKMVSLLKEKNDLQLQVQAEAEGLADAEE RCDQLIKTKIQLEAKIKEVTERAEDEEEINAELTAKKRKLEDECSELKKDIDDLELTLA KVEKEKHATENKVKNLTEEMAGLDETIAKLTKEKKALQEAHQQTLDDLQAEEDKV NTLTKAKIKLEQQVDDLEGSLEQEKKLRMDLERAKRKLEGDLKLAQESIMDIENEKQ QLDERLKKKEFEMSNLQSKIEDEQAIGIQLQKKIKELQARIEELEEEIEAERASRAKAE KQRSDLSRELEEISERLEEAGGATSAQIEMNKKREAEFQKMRRDLEEATLQHEATAA TLRKKHADSVAELGEQIDNLQRVKQKLEKEKSEMKMEIDDLASNVETVSKAKGNLE KMCRTLEDQVSELKSKEEEQQRLINDLTTQRGRLQTESGEFSRQLDEKEALVSQLSR GKQAFTQQIEELKRQLEEEVKAKNALAHALQSSRHDCDLLREQYEEEQESKAELQR ALSKANSEVAQWRTKYETDAIQRTEELEEAKKKLAQRLQAAEEHVEAVNAKCASLE KTKQRLQNEVEDLMLDVERTNAACAALDKKQRNFDKILAEWKQKYEETHAELEAS QKEARSLGTELFKMKNAYEESLDQLETLKRENKNLQQEISDLTEQIAEGGKRIHELEK IKKQVEQEKCELQAALEEAEASLEHEEGKILRIQLELNQVKSEIDRKIAEKDEEIDQLK RNHIRVVESMQSTLDAEIRSRNDAIRIKKKMEGDLNEMEIQLNHSNRMAAEALRNYR NTQGILKDTQLHLDDALRGQEDLKEQLAMVERRANLLQAEIEELRATLEQTERSRKI AEQELLDASERVQLLHTQNTSLINTKKKLETDISQIQGEMEDIVQEARNAEEKAKKAI TDAAMMAEELKKEQDTSAHLERMKKNMEQTVKDLQLRLDEAEQLALKGGKKQIQ KLEARVRELEGEVESEQKRNAEAVKGLRKHERRVKELTYQTEEDRKNILRLQDLVD KLQAKVKSYKRQAEEAEEQSNTNLSKFRKIQHELEEAEERADIAESQVNKLRVKSRE VHTKIISEE
23	Mouse Q6P6L5 (myosin-binding protein C)	MPEPTKKEENEVSAPAPPPEEPNKEKEAGTAPAKEWSLGESPAGGEEQDKQNANSQL STLFVEKPQTGSVKVGANITFVAKVRAEDLLRKPTVKWFKGKWMDLASKAGKHLQ LKETFERQTRVYTFEMQIIKAKENYAGNYRCEVTYKDKFDSCSFDLEVHESTGTTPNI DIRSAFKRSGEGQEDAGELDFSGLLKRREVKQQEEEPEIDVWELLKNANPNEYEKIAF QYGITDLRGMLKRLKRMRRVVEKKSAAFAKILDPAYQVDKGGKVRFVVELADPKLE VKWFKNGQEIRPSTKYIFEHKGNERIMFINNCSLTDDSEYYVTAGDEKCSTELFVREP PIMVTKQLEDMNAYCGERVEMEVEVSEDDANVKWFKNGEEIFPGPKSRYKIKVEGK KHTLIIEGATKADSAEYSAMTTGGQSSAKLSVDLRPLKITTPLTDQTVKLGKEVCLKC EISENVPGKWTKNGLPVQEGERLKVVHKGRIHKLVIANALIEDEGEYVFTPDAYNVP LSAKVHVIDPPKIILDGLDADNTVTVIAGSKLRLEIPVTGEPPPKAIWSRADKAIMEGS GRIRAESYPDSSTLVIDVAERDDSGVYNINLKNEAGEAHASIKIKVVDIPDPPVAPNVT EVGDDWCIMNWEPPAYDGGSPILGYFIERKKKQSSRWMRLNFDLCKETTFEPKKMIE GVAYEVRIPAVNAIGISKPSMPSKPFVPLAVTSPPTLLAVDSVTDSSVTMKWRPPDQI GAAGLDGYVLEYCFEGTEDWIVANTDLIDKTKFTINGLPTDAKIFVRVKAINAAGAS EPKYYSQPILVKEIIEPPKIRIPRHLKQTYIRRVGEAVNLVIPFQGKPRPELTWKKDGEE IDKNQINIRNSETDTIIFIRKAERSHSGKYDLQVKVDKYVENASIDIQIVDRPGPPQTVT IEDVWGENVALTWTPPKDDGNAAITGYTIQKADKKSMEWFTVIEHYHRTNATITEL VIGNEYYFRVFAENMCGLSEDATMTKESAVIAKDGKIYKNPVYEDFNFTEAPMFTQP

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		LVNTYAIAGYNATLNCSVRGNPKPKITWMKNKVAIVDDPRYRMFSNQGVCTLEIRK PSPYDGGTYCCKAVNDLGTVEIECKLEVKVVAQ
24	Mouse PGM1 (phosphogluco-mutase-1)	MVKIVTVKTQAYPDQKPGTSGLRKRVKVFQSNANYAENFIQSIVSTVEPALRQEATL VVGGDGRFYMTEAIQLIVRIAAANGIGRLVIGQNGILSTPAVSCIIRKIKAIGGIILTASH NPGGPNGDFGIKFNISNGGPAPEAITDKIFQISKTIEEYAICPDLKVDLGVLGKQQFDLE NKFKPFTVEIVDSVEAYATMLRNIFDFNALKELLSGPNRLKIRIDAMHGVVGPYVKKI LCEELGAPANSAVNCVPLEDFGGHHPDPNLTYAADLVETMKSGEHDFGAAFDGDG DRNMILGKHGFFVNPSDSVAVIAANIFSIPYFQQTGVRGFARSMPTSGALDRVANAT KIALYETPTGWKFFGNLMDASKLSLCGEESFGTGSDHIREKDGLWAVLAWLSILATR KQSVEDILKDHWQKFGRNFFTRYDYEEVEAEGANKMMKDLEALMLDRSFVGKQFS ANDKVYTVEKADNFEYSDPVDGSISKNQGLRLIFADGSRIIFRLSGTGSAGATIRLYID SYEKDVAKINQDPQVMLAPLISIALKVSQLQERTGRTAPTVIT
25	Mouse TPM2 (tropomyosin beta chain)	MDAIKKKMQMLKLDKENAIDRAEQAEADKKQAEDRCKQLEEEQQALQKKLKGTE DEVEKYSESVKDAQEKLEQAEKKATDAEADVASLNRRIQLVEEELDRAQERLATAL QKLEEAEKAADESERGMKVIENRAMKDEEKMELQEMQLKEAKHIAEDSDRKYEEV ARKLVILEGELERSEERAEVAESKCGDLEEELKIVTNNLKSLEAQADKYSTKEDKYE EEIKLLEEKLKEAETRAEFAERSVAKLEKTIDDLEDEVYAQKMKYKAISEELDNALN DITSL
26	Mouse TPIS (triosephosphate isomerase)	MAPTRKFFVGGNWKMNGRKKCLGELICTLNAANVPAGTEVVCAPPTAYIDFARQKL DPKIAVAAQNCYKVTNGAFTGEISPGMIKDLGATWVVLGHSERRHVFGESDELIGQK VSHALAEGLGVIACIGEKLDEREAGITEKVVFEQTKVIADNVKDWSKVVLAYEPVW AIGTGKTATPQQAQEVHEKLRGWLKSNVNDGVAQSTRIIYGGSVTGATCKELASQP DVDGFLVGGASLKPEFVDIINAKQ
27	Mouse PGAM2 (phosphogly-cerate mutase 2)	MTTHRLVMVRHGESLWNQENRFCGWFDAELSEKGAEEAKRGATAIKDAKIEFDICY TSVLKRAIRTLWTILDVTDQMWVPVVRTWRLNERHYGGLTGLNKAETAAKHGEEQ VKIWRRSFDTPPPPMDEKHNYYTSISKDRRYAGLKPEELPTCESLKDTIARALPFWNE EIAPKIKAGQRVLIAAHGNSLRGIVKHLEGMSDQAIMELNLPTGIPIVYELDQNLKPT KPMRFLGDEETVRKAMEAVAAQGKAK
28	Mouse SRCA (sarcalumenin)	MKALLLLCCFLASLLLSGQAEVEDASEEAPLRDRSHIDKTLMLNEDKPADDYSAVLQ RLRKIYHTSIKPLEQSYKYNELRQHEITDGEITSKPMVLFLGPWSVGKSTMINYLLGL EDTRYQLYTGAEPTTSEFTVLMHGPKLKTIEGIVMAADSARSFSPLEKFGQNFLEKLI GIEVPHKLLERVTFVDTPGIIENRKQQERGYPFNDVCQWFIDRADLIFVVFDPTKLDV GLELEMLFRQLKGRESQIRIILNKADNLATQMLMRVYGALFWSLAPLINVTEPPRVY VSSFWPQDYKPDTHRELFLKEEISLLEDLNQVIENRLENKIAFIRQHAIRVRIHALLVD RYLQTYKDKMTFFSDGELVFKDIVEDPDKFYIFKTILAKTNVSKFDLPNREAYKDFFG INPISNFKLLSQQCSYMGGCFLEKIERAITQELPSLLGSIGLGKNPGAPNCDKTGCGET PKNRYKKH
29	Mouse PGK 1 (phosphoglycerate kinase 1)	MSLSNKLTLDKLDVKGKRVVMRVDFNVPMKNNQITNNQRIKAAVPSIKFCLDNGAK SVVLMSHLGRPDGVPMPDKYSLEPVAAELKSLLGKDVLFLKDCVGPEVENACANPA AGTVILLENLRFHVEEEGKGKDASGNKVKAEPAKIDAFRASLSKLGDVYVNDAFGT AHRAHSSMVGVNLPQKAGGFLMKKELNYFAKALESPERPFLAILGGAKVADKIQLIN NMLDKVNEMIIGGGMAFTFLKVLNNMEIGTSLYDEEGAKIVKDLMSKAEKNGVKIT LPVDFVTADKFDENAKTGQATVASGIPAGWMGLDCGTESSKKYAEAVGRAKQIVW NGPVGVFEWEAFARGTKSLMDEVVKATSRGCITIIGGGDTATCCAKWNTEDKVSHV STGGGASLELLEGKVLPGVDALSNV
30	Mouse PLEC (plectin)	MVAGMLMPLDRLRAIYEVLFREGVMVAKKDRRPRSLHPHVPGVTNLQVMRAMAS LKARGLVRETFAWCHFYWYLTNEGIDHLRQYLHLPPEIVPASLQRVRRPVAMVIPAR RRSPHVQTMQGPLGCPPKRGPLPAEDPAREERQVYRRKEREEGAPETPVVSATTVGT LARPGPEPAPATDERDRVQKKTFTKWVNKHLIKHWRAEAQRHISDLYEDLRDGHNL ISLLEVLSGDSLPREKGRMRFHKLQNVQIALDYLRHRQVKLVNIRNDDIADGNPKLT LGLIWTIILHFQISDIQVSGQSEDMTAKEKLLLWSQRMVEGYQGLRCDNFTTSWRDG RLFNAIIHRHKPMLIDMNKVYRQTNLENLDQAFSVAERDLGVTRLLDPEDVDVPQPD EKSIITYVSSLYDAMPRVPGAQDGVRANELQLRWQEYRELVLLLLQWIRHHTAAFEE RKFPSSFEEIEILWCQFLKFKETELPAKEADKNRSKVIYQSLEGAVQAGQLKIPPGYHP LDVEKEWGKLHVAILEREKQLRSEFERLECLQRIVSKLQMEAGLCEEQLNQADALLQ SDIRLLASGKVAQRAGEVERDLDKADGMIRLLFNDVQTLKDGRHPQGEQMYRRVY RLHERLVAIRTEYNLRLKAGVGAPVTQVTLQSTQRRPELEDSTLRYLQDLLAWVEEN QRRIDSAEWGVDLPSVEAQLGSHRGMHQSIEEFRAKIERARNDESQLSPATRGAYRD CLGRLDLQYAKLLNSSKARLRSLESLHGFVAAATKELMWLNEKEEEEVGFDWSDRN TNMAAKKESYSALMRELEMKEKKIKEIQNTGDRLLREDHPARPTVESFQAALQTQW SWMLQLCCCIEAHLKENTAYFQFFSDVREAEEQLQKLQETLRRKYSCDRTITVTRLE DLLQDAQDEKEQLNEYKGHLSGLAKRAKAIVQLKPRNPAHPVRGHVPLIAVCDYKQ VEVTVHKGDQCQLVGPAQPSHWKVLSGSSSEAAVPSVCFLVPPPNQEAQEAVARLE AQHQALVTLWHQLHVDMKSLLAWQSLSRDIQLIRSWSLVTFRTLKPEEQRQALRNL ELHYQAFLRDSQDAGGFGPEDRLVAEREYGSCSRHYQQLLQSLEQGEQEESRCQRCI SELKDIRLQLEACETRTVHRLRLPLDKDPARECAQRIAEQQKAQAEVEGLGKGVARL

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		SAEAEKVLALPEPSPAAPTLRSELELTLGKLEQVRSLSAIYLEKLKTISLVIRSTQGAEE VLKTHEEQLKEAQAVPATLQELEATKASLKKLRAQAEAQQPVFNTLRDELRGAQEV GERLQQRHGERDVEVERWRERVTQLLERWQAVLAQTDVRQRELEQLGRQLRYYRE SADPLSAWLQDAKRRQEQIQAVPIANCQAAREQLRQEKALLEEIERHGEKVEECQKF AKQYINAIKDYELQLITYKAQLEPVASPAKKPKVQSGSESVIQEYVDLRTRYSELTTL TSQYIKFISETLRRMEEEERLAEQQRAEERERLAEVEAALEKQRQLAEAHAQAKAQA ELEAQELQRRMQEEVARREEAAVDAQQQKRSIQEELQHLRQSSEAEIQAKAQQVEA AERSRMRIEEEIRVVRLQLETTERQRGGAEGELQALRARAEEAEAQKRQAQEEAERL RRQVQDESQRKRQAEAELALRVKAEAEAAREKQRALQALDELRLQAEEAERRLRQ AEAERARQVQVALETAQRSAEVELQSKRASFAEKTAQLERTLQEEHVTVAQLREEA ERRAQQQAEAERAREEAERELERWQLKANEALRLRLQAEEVAQQKSLAQADAEKQ KEEAEREARRRGKAEEQAVRQRELAEQELEKQRQLAEGTAQQRLAAEQELIRLRAE TEQGEQQRQLLEEELARLQHEATAATQKRQELEAELAKVRAEMEVLLASKARAEEE SRSTSEKSKQRLEAEAGRFRELAEEAARLRALAEEAKRQRQLAEEDAARQRAEAER VLTEKLAAISEATRLKTEAEIALKEKEAENERLRRLAEDEAFQRRRLEEQAALHKADI EERLAQLRKASESELERQKGLVEDTLRQRRQVEEEIMALKVSFEKAAAGKAELELEL GRIRSNAEDTMRSKEQAELEAARQRQLAAEEEQRRREAEERVQRSLAAEEEAARQR KVALEEVERLKAKVEEARRLRERAEQESARQLQLAQEAAQKRLQAEEKAHAFVVQ QREEELQQTLQQEQNMLDRLRSEAEAARRAAEEAEEAREQAEREAAQSRKQVEEAE RLKQSAEEQAQAQAQAQAAAEKLRKEAEQEAARRAQAEQAALKQKQAADAEMEK HKKFAEQTLRQKAQVEQELTTLRLQLEETDHQKSILDEELQRLKAEVTEAARQRSQV EEELFSVRVQMEELGKLKARIEAENRALILRDKDNTQRFLEEEAEKMKQVAEEAARL SVAAQEAARLRQLAEEDLAQQRALAEKMLKEKMQAVQEATRLKAEAELLQQQKEL AQEQARRLQEDKEQMAQQLVEETQGFQRTLEAERQRQLEMSAEAERLKLRMAEMS RAQARAEEDAQRFRKQAEEIGEKLHRTELATQEKVTLVQTLEIQRQQSDHDAERLRE AIAELEREKEKLKQEAKLLQLKSEEMQTVQQEQILQETQALQKSFLSEKDSLLQRERF IEQEKAKLEQLFQDEVAKAKQLREEQQRQQQQMEQEKQELMASMEEARRRQREAE EGVRRKQEELQHLEQQRQQQEKLLAEENQRLRERLQRLEEEHRAALAHSEIATTQA ASTKALPNGRDAPDGPSVEAEPEYTFEGLRQKVPAQQLQEAGILSQEELQRLAQGHT TVAELTQREDVYRYLKGRSSIAGLLLKPTNEKLSVYTALQRQLLSPGTALILLEAQAA SGFLLDPVRNRRLTVNEAVKEGVVGPELHHKLLSAERAVTGYKDPYTGEQISLFQA MKKDLIVRDHGVRLLEAQIATGGIIDPVHSHRVPVDVAYKRGYFDEEMNRILSDPSD DTKGFFDPNTHENLTYLQLLERCVEDPETGLRLLPLTDKAAKGGELVYTDTEARDVF EKATVSAPFGKFQGRTVTIWEIINSEYFTAEQRRDLLQQFRTGHITVEKIIKIVITVVEE HERKGQLCFEGLRALVPAAELLDSGVISHELYQQLQRGERSVREVAEADSVRQALR GTNVIAGVWLEEAGQKLSIYEALKKDLLQPEVAVALLEAQAGTGHIIDPATSARLTV DEAVRAGLVGPELHEKLLSAEKAVTGYRDPYSGQSVSLFQALKKGLIPREQGLRLLD AQLSTGGIVDPSKSHRVPLDVAYARGYLDKETNRALTSPRDDARVYHDPSTQEPVTY SQLQQRCRSDQLTGLSLLPLSEKAVRARQEEVYSELQARETLEQAKVEVPVGSFKGR AMTVWELISSEYFTEEQRQELLRQFRTGKVTVEKVIKIVITIVEEVETRRQERLSFSGL RAPVPASELLDAKILSRAQFDQLKDGKTSVKELSEVGSVRTLLQGSGCLAGIYLEDSK EKVTIYEAMRRGLLRPSTATLLLEAQAATGFLVDPVRNQRLYVHEAVKAGVVGPEL HEKLLSAEKAVTGYKDPYSGNTISLFQAMKKGLVLRDHAIRLLEAQVATGGIIDPVH SHRLPVDVAYQRGYFDEEMNRVLADPSDDTKGFFDPNTHENLTYLQLLERCVEDPE TGLRLLPLKGAEKTEVVETTQVYTEEETRRAFEETQIDIPGGGSHGGSSMSLWEVMQ SNMIPEDQRARLMADFQAGRVTKERMIIIIIEIIEKTEIIRQQNLASYDYVRRRLTAEDL YEARIISLETYNLFREGTKNLREVLEMESAWRYLYGTGAVAGVYLPGSRQTLTIYQA LKKGLLSAEVARLLLEAQAATGFLLDPVKGERLTVDEAVRKGLVGPELHDRLLSAE RAVTGYRDPYTEQTISLFQAMKKELIPAEEALRLLDAQLATGGIVDPRLGFHLPLEVA YQRGYLNKDTHDQLSEPSEVRSYVDPSTDERLSYTQLLKRCRRDDPSGQMLLLLSDA RKLTFRGLRKQITVEELVRSQVMDEATALQLQEGLTSIEEVTKNLQKFLEGTSCIAGV FVDATKERLSVYQAMKKGIIRPGTAFELLEAQAATGYVIDPIKGLKLTVEEAVRMGI VGPEFKDKLLSAERAVTGYKDPYSGKLISLFQAMKKGLILKDHGIRLLEAQIATGGII DPEESHRLPVEVAYKRGLFDEEMNEILTDPSDDTKGFFDPNTEENLTYLQLMERCITD PQTGLCLLPLKEKKRERKTSSKSSVRKRRVVIVDPETGKEMSVYEAYRKGLIDHQTY LELSEQECEWEEITISSSDGVVKSMIIDRRSGRQYDIDDAITKNLIDRSALDQYRAGTL SITEFADMLSGNAGGFRSRSSSVGSSSSYPISSAGPRTQLASWSDPTEETGPVAGILDT ETLEKVSITEAMHRNLVDNITGQRLLEAQACTGGIIDPSTGERFPVTEAVNKGLVDKI MVDRINLAQKAFCGFEDPRTKTKMSAAQALKKGWLYYEAGQRFLEVQYLTGGLIEP DTPGRVSLDEALQRGTVDARTAQKLRDVSAYSKYLTCPKTKLKISYKDALDRSMVE EGTGLRLLEAAAQSSKGYYSPYSVSGSGSTAGSRTGSRTGSRAGSRRGSFDATGSGFS MTFSSSSYSSSGYGRRYASGPSASLGGPESAVA
31	Mouse ACTN2 (alpha-actinin-2)	MNQIEPGVQYNYVYDEDEYMIQEEEWDRDLLLDPAWEKQQRKTFTAWCNSHLRKA GTQIENIEEDFRNGLKLMLLLEVISGERLPKPDRGKMRFHKIANVNKALDYIASKGVK LVSIGAEEIVDGNVKMTLGMIWTIILRFAIQDISVEETSAKEGLLLWCQRKTAPYRNV NIQNFHTSWKDGLGLCALIHRHRPDLIDYSKLNKDDPIGNINLAMEIAEKHLDIPKML DAEDIVNTPKPDERAIMTYVSCFYHAFAGAEQAETAANRICKVLAVNQENERLMEE YERLASELLEWIRRTIPWLENRTPEKTMQAMQKKLEDFRDYRRKHKPPKVQEKCQL EINFNTLQTKLRISNRPAFMPSEGKMVSDIAGAWQRLEQAEKGYEEWLLNEIRRLER LEHLAEKFRQKASTHETWAYGKEQILLQKDYESASLTEVRALLRKHEAFESDLAAH

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		QDRVEQIAAIAQELNELDYHDAVNVNDRCQKICDQWDRLGTLTQKRREALERTEKL LETIDQLHLEFAKRAAPFNNWMEGAMEDLQDMFIVHSIEEIQSLITAHEQFKATLPEA DGERQSILAIQNEVEKVIQSYSIRISSSNPYSTVTMDELRNKWDKVKQLVPVRDQSLQ EELARQHANERLRRQFAAQANAIGPWIQNKMEEIARSSIQITGALEDQMNQLKQYEH NIINYKNNIDKLEGDHQLIQEALVFDNKHTNYTMEHIRVGWELLLTTIARTINEVETQI LTRDAKGITQEQMNEFRASFNHFDRRKNGLMDHEDFRACLISMGYDLGEAEFARIM TLVDPNGQGTVTFQSFIDFMTRETADTDTAEQVIASFRILASDKPYILAEELRRELPPD QAQYCIKRMPPYSGPGSVPGALDYTAFSSALYGESDL
32	Mouse PFKAM (ATP-dependent 6-phosphofru-nctokiase)	MTHEEHHAAKTLGIGKAIAVLTSGGDAQGMNAAVRAVVRVGIFTGARVFFVHEGY QGLVDGGEHIREATWESVSMMLQLGGTVIGSARCKDFRERGRLRAAHNLVKRGIT NLCVIGGDGSLTGADTFRSEWSDLLNDLQKDGKITAEEATKSSYLNIVGLVGSIDNDF CGTDMTIGTDSALHRIVEIVDAITTTAQSHQRTFVLEVMGRHCGYLALVTSLSCGAD WVFIPECPPDDDWEEHLCRRLSETRTRGSRLNIIIVAEGAIDKNGKPITSEDIKNLVVK RLGYDTRVTVLGHVQRGGTPSAFDRILGSRMGVEAVMALLEGTPDTPACVVSLSGN QAVRLPLMECVQVTKDVTKAMDEKRFDEAIKLRGRSFMNNWEVYKLLAHVRPPVS KGGLHTVAVMNVGAPAAGMNAAVRSTVRIGLIQGNRVLVVHDGFEGLAKGQIEEA GWSYVGGWTGQGGSKLGTKRTLPKKNLEQISANITKFNIQGLVIIGGFEAYTGGLEL MEGRKQFDELCIPFVVIPATVSNNVPGSDFSIGADTALNTICTTCDRIKQSAAGTKRR VFIIETMGGYCGYLATMAGLAAGADAAYIFEEPFTIRDLQVNVEHLVQKMKTTVKR GLVLRNEKCNENYTTDFIFNLYSEEGKGIFDSRKNVLGHMQQGGSPTPFDRNFATKM GAKAMNWMSGKIKESYRNGRIFANTPDSGCVLGMRKRALVFQPVTELKDQTDFEH RIPKEQWWLKLRPILKILAKYEIDLDTSDHAHLEHISRKRSGEAAV
33	Mouse MDHM (malate dehydrogenase)	MLSALARPAGAALRRSFSTSAQNNAKVAVLGASGGIGQPLSLLLKNSPLVSRLTLYDI AHTPGVAADLSHIETRANVKGYLGPEQLPDCLKGCDVVVIPAGVPRKPGMTRDDLF NTNATIVATLTAACAQHCPEAMVCIIANPVNSTIPITAEVFKKHGVYNPNKIFGVTTL DIVRANTFVAELKGLDPARVNVPVIGGHAGKTIIPLISQCTPKVDFPQDQLATLTGRIQ EAGTEVVKAKAGAGSATLSMAYAGARFVFSLVDAMNGKEGVVECSFVQSKETECT YFSTPLLLGKKGLEKNLGIGKITPFEEKMIAEAIPELKASIKKGEDFVKNMK
34	Mouse MYG (myoglobin)	MGLSDGEWQLVLNVWGKVEADLAGHGQEVLIGLFKTHPETLDKFDKFKNLKSEED MKGSEDLKKHGCTVLTALGTILKKKGQHAAEIQPLAQSHATKHKIPVKYLEFISEIIIE VLKKRHSGDFGADAQGAMSKALELFRNDIAAKYKELGFQG
35	Mouse HBBB1 (hemoglobin subunit beta-1)	MVHLTDAEKAAVSCLWGKVNSDEVGGEALGRLLVVYPWTQRYFDSFGDLSSASAI MGNAKVKAHGKKVITAFNDGLNHLDSLKGTFASLSELHCDKLHVDPENFRLLGNMI VIVLGHHLGKDFTPAAQAAFQKVVAGVATALAHKYH
36	Mouse FLNC (filamin-C)	MMNNSNYSDASGLGLVDEADEMPSTEKDLAEDAPWKKIQQNTFTRWCNEHLKCVG KRLTDLQRDLSDGLRLIALLEVLSQKRMYRKFHPRPNFRQMKLENVSVALEFLEREH IKLVSIDSKAIVDGNLKLILGLIWTLILHYSISMPMWEDEDDEDARKQTPKQRLLGWI QNKVPQLPITNFNRDWQDGKALGALVDNCAPGLCPDWEAWDPNQPVQNAREAMQ QADDWLGVPQVIAPEEIVDPNVDEHSVMTYLSQFPKAKLKPGAPVRSKQLNPKKAIA YGPGIEPQGNTVLQPAHFTVQTVDAGVGEVLVYIEDPEGHTEEAKVVPNNDKDRTY AVSYVPKVAGLHKVTVLFAGQNIERSPFEVNVGMALGDANKVSARGPGLEPVGNV ANKPTYFDIYTAGAGTGDVAVVIVDPQGRRDTVEVALEDKGDNTFRCTYRPVMEGP HTVHVAFAGAPITRSPFPVHVAEACNPNACRASGRGLQPKGVRVKEVADFKVFTKG AGSGELKVTVKGPKGTEEPVKVREAGDGVFECEYYPVVPGKYVVTITWGGYAIPRS PFEVQVSPEAGAQKVRAWGPGLETGQVGKSADFVVEAIGTEVGTLGFSIEGPSQAKI ECDDKGDGSCDVRYWPTEPGEYAVHVICDDEDIRDSPFIAHIQPAPPDCFPDKVKAF GPGLEPTGCIVDRPAEFTIDARAAGKGDLKLYAQDADGCPIDIKVIPNGDGTFRCSYV PTKPIKHTIIVSWGGVNVPKSPFRVNVGEGSHPERVKVYGPGVEKTGLKANEPTYFT VDCSEAGQGDVSIGIKCAPGVVGPVEADIDFDIIKNDNDTFTVKYTPPGAGHYTIMVL FANQEIPASPFHIKVDPSHDASKVKAEGPGLSRTGVEVGKPTHFTVLTKGAGKAKLD VHFAGAAKGEAVRDFEIIDNHDYSYTVKYTAVQQGNMAVTVTYGGDPVPKSPFVV NVAPPLDLSKVKVQGLNSKVAVGQEQAFSVNTRGAGGQGQLDVRMTSPSRRPIPCK LEPGGGAEAQAVRYMPPEEGPYKVDITYDGHPVPGSPFAVEGVLPPDPSKVCAYGPG LKGGLVGTPAPFSIDTKGAGTGGLGLTVEGPCEAKIECQDNGDGSCAVSYLPTEPGE YTINILFAEAHIPGSPFKATIQPVFDPSKVRASGPGLERGKAGEAATFTVDCSEAGEAE LTIEILSDAGVKAEVLIQNNADGTYHITYSPAFPGTYTITIKYGGHPIPKFPTRVHVQPA VDTSGIKVSGPGVEPHGVLREVTTEFTVDARSLTATGGNHVTARVLNPSGAKTDTYV TDNGDGTYRVQYTAYEEGVHLVEVLYDEVAVPKSPFRVGVTEGCDPTRVRAFGPGL EGGLVNKANRFTVETRGAGTGGLGLAIEGPSEAKMSCKDNKDGSCTVEYIPFTPGDY DVNITFGGQPIPGSPFRVPVKDVVDPGKVKCSGPGLGTGVRARVPQTFTVDCSQAGR APLQVAVLGPTGVAEPVEVRDNGDGTHTVHYTPATDGPYTVAVKYADQEVPRSPFK IKVLPSHDASKVRASGPGLNASGIPASLPVEFTIDARDAGQGLLTVQILDPEGKPKKA NIRDNGDGTYTVSYLPDMSGRYTITIKYGGDEIPYSPFRIHALPTGDASKCLVTVSIGG HGLGACLGPRIQIGEETVITVDAKAAGKGKVTCTVSTPDGAELDVDVVENHDGTFDI YYTAPEPGKYVITIRFGGEHIPNSPFHVLACDPLPHVEEPAEMLQMRQPYAPLRPGTC PTHWATEEPVVPVEPLESMLRPFNLVIPFTVQKGELTGEVRMPSGKTARPNITDNKD

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		GTITVRYAPTEKGLHQMGIKYDGNHIPGSPLQFYVDAINSRHVSAYGPGLSHGMVNK PATFTIVTKDAGEGGLSLAVEGPSKAEITCKDNKDGTCTVSYLPTAPGDYSIIVRFDD KHIPGSPFTAKITGDDSMRTSQLNVGTSTDVSLKITEGDLSQLTASIRAPSGNEEPCLL KRLPNRHIGISFTPKEVGEHVVSVRKSGKHVTNSPFKILVGPSEIGDASKVRVWGKGL SEGQTFQVAEFIVDTRNAGYGGLGLSIEGPSKVDINCEDMEDGTCKVTYCPTEPGTYI INIKFADKHVPGSPFTVKVTGEGRMKESITRRRQAPSIATIGSTCDLNLKIPGNWFQM VSAQERLTRTFTRSSHTYTRTERTEISKTRGGETKREVRVEESTQVGGDPFPAVFGDF LGRERLGSFGSITRQQEGEASSQDMTAQVTSPSGKTEAAEIVEGEDSAYSVRFVPQE MGPHTVTVKYRGQHVPGSPFQFTVGPLGEGGAHKVRAGGTGLERGVAGVPAEFSIW TREAGAGGLSIAVEGPSKAEIAFEDRKDGSCGVSYVVQEPGDYEVSIKFNDEHIPDSP FVVPVASLSDDARRLTVTSLQETGLKVNQPASFAVQLNGARGVIDARVHTPSGAVEE CYVSELDSDKHTIRFIPHENGVHSIDVKFNGAHIPGSPFKIRVGEQSQAGDPGLVSAY GPGLEGGTTGVSSEFIVNTQNAGSGALSVTIDGPSKVQLDCRECPEGHVVTYTPMAP GNYLIAIKYGGPQHIVGSPFKAKVTGPRLSGGHSLHETSTVLVETVTKSSSSRGASYSS IPKFSSDASKVVTRGPGLSQAFVGQKNSFTVDCSKAGTNMMMVGVHGPKTPCEEVY VKHMGNRVYNVTYTVKEKGDYILIVKWGDESVPGSPFKVNVP
37	Mouse ACON (aconitate hydratase)	MAPYSLLVTRLQKALGVRQYHVASVLCQRAKVAMSHFEPSEYIRYDLLEKNINIVRK RLNRPLTLSEKIVYGHLDDPANQEIERGKTYLRLRPDRVAMQDATAQMAMLQFISSG LPKVAVPSTIHCDHLIEAQVGGEKDLRRAKDINQEVYNFLATAGAKYGVGFWRPGS GIIHQIILENYAYPGVLLIGTDSHTPNGGGLGGICIGVGGADAVDVMAGIPWELKCPK VIGVKLTGSLSGWTSPKDVILKVAGILTVKGGTGAIVEYHGPGVDSISCTGMATICNM GAEIGATTSVFPYNHRMKKYLSKTGRTDIANLAEEFKDHLVPDPGCQYDQVIEINLN ELKPHINGPFTPDLAHPVADVGTVAEKEGWPLDIRVGLIGSCTNSSYEDMGRSAAVA KQALAHGLKCKSQFTITPGSEQIRATIERDGYAQILRDVGGIVLANACGPCIGQWDRK DIKKGEKNTIVTSYNRNFTGRNDANPETHAFVTSPEIVTALAIAGTLKFNPETDFLTG KDGKKFKLEAPDADELPRSDFDPGQDTYQHPPKDSSGQRVDVSPTSQRLQLLEPFDK WDGKDLEDLQILIKVKGKCTTDHISAAGPWLKFRGHLDNISNNLLIGAINIENGKANS VRNAVTQEFGPVPDTARYYKKHGIRWVVIGDENYGEGSSREHAALEPRHLGGRAIIT KSFARIHETNLKKQGLLPLTFADPSDYNKIHPVDKLTIQGLKDFAPGKPLKCVIKHPN GTQETILLNHTFNETQIEWFRAGSALNRMKELQQ
38	Mouse PRVA (parvalbumin alpha)	MSMTDVLSAEDIKKAIGAFAAADSFDHKKFFQMVGLKKKNPDEVKKVFHILDKDKS GFIEEDELGSILKGFSSDARDLSAKETKTLLAAGDKDGDGKIGVEEFSTLVAES
39	Mouse HBA (hemoglobin subunit alpha)	MVLSGEDKSNIKAAWGKIGGHGAEYGAEALERMFASFPTTKTYFPHFDVSHGSAQV KGHGKKVADALASAAGHLDDLPGALSALSDLHAHKLRVDPVNFKLLSHCLLVTLAS HHPADFTPAVHASLDKFLASVSTVLTSKYR
40	Mouse HBAZ (hemoglobin subunit zeta)	MSLMKNERAIIMSMWEKMAAQAEPIGTETLERLFCSYPQTKTYFPHFDLHHGSQQLR AHGFKIMTAVGDAVKSIDNLSSALTKLSELHAYILRVDPVNFKLLSHCLLVTMAARF PADFTPEVHEAWDKFMSILSSILTEKYR
41	Mouse HSP7C (heat shock cognate 71)	MSKGPAVGIDLGTTYSCVGVFQHGKVEIIANDQGNRTTPSYVAFTDTERLIGDAAKN QVAMNPTNTVFDAKRLIGRRFDDAVVQSDMKHWPFMVVNDAGRPKVQVEYKGET KSFYPEEVSSMVLTKMKEIAEAYLGKTVTNAVVTVPAYFNDSQRQATKDAGTIAGL NVLRIINEPTAAAIAYGLDKKVGAERNVLIFDLGGGTFDVSILTIEDGIFEVKSTAGDT HLGGEDFDNRMVNHFIAEFKRKHKKDISENKRAVRRLRTACERAKRTLSSSTQASIEI DSLYEGIDFYTSITRARFEELNADLFRGTLDPVEKALRDAKLDKSQIHDIVLVGGSTRI PKIQKLLQDFFNGKELNKSINPDEAVAYGAAVQAAILSGDKSENVQDLLLLDVTPLS LGIETAGGVMTVLIKRNTTIPTKQTQTFTTYSDNQPGVLIQVYEGERAMTKDNNLLG KFELTGIPPAPRGVPQIEVTFDIDANGILNVSAVDKSTGKENKITITNDKGRLSKEDIER MVQEAEKYKAEDEKQRDKVSSKNSLESYAFNMKATVEDEKLQGKINDEDKQKILD KCNEIISWLDKNQTAEKEEFEHQQKELEKVCNPIITKLYQSAGGMPGGMPGGFPGGG APPSGGASSGPTIEEVD
42	Mouse TNNI2 (troponin I)	MGDEEKRNRAITARRQHLKSVMLQIAATELEKEESRRESEKENYLSEHCPPLHIPGSM SEVQELCKQLHAKIDVAEEEKYDMEVKVQKSSKELEDMNQKLFDLRGKFKRPPLRR VRMSADAMLKALLGSKHKVCMDLRANLKQVKKEDTEKERDLRDVGDWRKNIEEK SGMEGRKKMFESES
43	Mouse CASQ1 (calsequestrin-1)	MRATDRMGARAVSELRLALLFVLVLGTPRLGVQGEDGLDFPEYDGVDRVINVNAK NYKNVFKKYEVLALLYHEPPEDDKASQRQFEMEELILELAAQVLEDKGVGFGLVDS EKDAAVAKKLGLTEEDSVYVFKGDEVIEYDGEFSADTLVEFLLDVLEDPVELIEGER ELQAFENIEDEIKLIGYFKSKDSEHYKAYEDAAEEFHPYIPFFATFDSKVAKKLTLKLN EIDFYEAFMEEPMTIPDKPNSEEEIVSFVEEHRRSTLRKLKPESMYETWEDDLDGIHIV AFAEEADPDGYEFLETLKAVAQDNTENPDLSIIWIDPDDFPLLVPYWEKTFDIDLSAP QIGVVNVTDADSIWMEMDNEEDLPSADELEDWLEDVLEGEINTEDDDDDDDDDDD DDDDDD

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
44	Mouse OBSCN (obscurin)	MDHSFSGAPRFLTRPKAFVVSVGKDATLSCQIVGNPTPHVSWEKDRQPVEAGARFRL AQDGDVYRLTILDLALGDSGQYVCRARNAIGEAFAAVGLRVDSEGTCAEQAPHFLL RPTSIRVREGADATFRCRVGGSPQPAVSWSKDGRRLGPPDAPHVRVEEHGESSALRI RSARPRDGGTYEVRAENPLGSASAAAALVVDSDAEVAGPPGTSTATLLAHLQQRRE AMRAEGIPPSPPGAGTRTCTVTEGKHARLSCFVTGEPKPETVWKKDGQLVTEGRRH VVYEDEQENFVLKILFCKQSDRGLYTCTASNLVGQTYSSVLVVVREPTVPFKKRLQD LEVREKESATFQCEVAQPATEAAWFKEETRLWASAKYDIEEEGTERRLTVRNVSAD DDAVYICETTEGSRTVAELSVQGNLTRKLPRKTAVRTGDTAIFWVELAVPEGPVRWL RNQEEMVAGGRVAITAEGTCHTLTIFQCTLEDMGEVAFVSGGCRTTTQFCVSAPRRP PLYPPADPVVKAKTESSVTLSWSAPPHGDRPVTIDGYVLEKRKLGAYAWSRCHEAG WLATTEFTITGVAEEGDFQFRVSAINHFGQGPYLEFPGTMHLVPMLAVKTPLKAVEA VEGGEVTFSVDLTVASSGEWELDGEALKASSIYVIRCDRTRHMLTIREVPARLHGAQ LKFVANGIETSIQMVVRAALGLPSSKLPAAAAREVLAQLHEEAQLLAELSDQAAAVT WLKDGRELSLGPKYEMQVSAGRRALLVRDVAQDDAGLYECVSRGSRTAYQLLVQD ITDGYRDWGPAGPQKHMCKCAGAKIARYLGSSCYRFLQYDKGVWHWLEAALDTR QGKGTSSCSLHEKPKLVFAKGQQAHSEVKAEAGNSATLSCEVTQAQTEVTWFKDGK KLSSSSKVRMEASGCSRRLVVQQAGKADAGEYSCEAGGQKLSFRLDVAEPKLVFAK GQQAHSEVKAEAGASATLSCEVAQAQTEVTWFKDGKKLSSSSKVRMEASGCSRRL VVQQAGKADAGEYSCEAGGQKLSFRLDVAEPKMVFAKEQQARSEVKAEAGASATL SCEVAQAQTEVTWFKDGKKLSSSSKVRMEASGCSRRLVVQQAGKADAGEYSCEAG GQKLSFRLDVTEPKLVFAKEQQARSEVKAEVGNSATLSCEVAQAQTEVTWFKDGKK LSSSSKVRMEASGCSRRLVVQQAGKADAGEYSCEAGGQKLSFRLDVAEPKLVFAKE QQARSEVKAEAGNSATLSCEVAQAQTEVTWFKDGKKLSSSSKVRMEASGCSRRLVV QQAGKADAGEYSCEAGGQKLSFHLDVTEPKLVFAKEQQAHSEVKAEAGASATLSCE VAQAQTEVTWFKDGKKLSSSSKVRMEASGCSRRLVVQQAGKADAGEYSCEAEGQK LSFRLDVAEPKLVFAKEQQARSEVKAEAGASATLSCEVAQAQTEVTWFKDGKKLSS SSKVRVEASGCSRRLVVQQAGKADAGEYSCEAGGQKLSFRLDVAEPKLVFAKEQQA NSEVKAEAGASATLSCEVAQAQTEVTWFKDGKKLSSSSKVRVEASGCSRRLVVQQA GKADAGEYSCEAGGQKLSFRLDVAEPKLAFAKEQQAHSEVKAEAGASATLSCEVAQ AQTEVTWFKDGKKLSSSSKVRVEASGCSRRLVVQQAGKADAGEYSCEAGGQKLSF RLDVAEPKLAFAKEQQAHSEVKAEAGASATLSCEVAQAQTEVTWFKDGKKLSSSSK VRVEASGCSRRLVVQQAGKADAGEYSCEAGGQKLSFRLDVAEPKLAFAKEQQAHS EVKAEAGASATLSCEVAQAQTEVTWFKDGKKLSSSSKVRVEASGCSRRLVVQQAG KADAGEYSCEAGGQKLFFRLDVAEPKLMFAKEQQAHSEVKAEAGASATLSCEVAQ AQTEVTWFKDGKKLSSSSKVRVEASGCSRRLVVQQAGKADAGEYSCEAGGQKLSF RLDVAEPKLVFAKEQQAHSEVKAEAGASATLSCEVAQAQTEVTWFKDGKKLSSSSK VRVEASGCSRRLVVQQVGKADAGEYSCEARGQKLSFRLDVADTRLMFAKEQQART EVKAEAGNSATLSCEVAQAQTEVTWFKDGKKLSSSSKVRVEASGCSRRLVVQQAG KADAGEYSCEAGGQKLSFRLDVAEAESQIPERPSRREPLVVKEHETIILTATIAAPSVA AVTWLKDGVEIRRSKRHEATSLGDTHTLTVRGAQVLDSAIYSCRVGKEGQDFPVQV EEVAAKFSKPLEPVEGELGGTVMLVCELSPEQAEVVWRCGNTQLRPGKRFQMTSEG PRRTLTVSGLREDDAEEYVCESRDDRTSARLTVKVPRVVKFTSGLSAMVAEEGQEA TFQCVVSPSDAGVTWYKDGMQLQPSEKFVMVESGASRSLTILGLTLEDAGQVTVEA EGASSSAALRVREAPVLFKKKLEPQTVEERTSVTLEVELTRPWPEVKWTRNAAVLTP SENVEIRAEGARHCLVLRSVGFADRGFFGCETPDDKTQAKLNVEMRQVRLVRGLQE VEAKEQGTASMDVELSHAEVEGSWTRDGLRLQPGPKCHLAVQGPVHILTLSALQPQ DSGLVAFRAEGVHTSARLIVTELPVSFTRVLQDVVATQKEKVTLECELSRPVDVRWL KDGVELRAGKAIGIVAQGTCRSLVIYRCETGDQGVYVCDALDAQTSASLRVQGRTY TLIFRRVLAEDAGEVKFVAENAESRAHLRVKELPVTLLRPLRDKIAMEKHRGVLECQ VSRASAQVRWFKGGVELQSGPKYEVVSDGLYRKLVINDVQPEDEDTYTCDAGNVK TSAQFFVEEQSITIVRGLKDMTVMEPAPAWFECETSIPSVRPPKWLLGKTVLQAGGN VGLEQDGTVHRLTLHKTCSTMTGPVHFTIGKSRSSAQLVVSDIPVVLTRPLEPKAGRE LQSVVLSCDFRPAPKAVQWYKDDTPLSPSEKFKMALEGQMAELRILRLTPADAGVY RCQAGSAQSSAEVTVEAREVTVIQPLQDAEAMEEGRVCFSCELSHKDEDIEWSLNGT PLYNDSFHEISHEGCLHTLVLKSVRQADTGTVCATSPKVSVSARLVVKGKPVVFLKA LDDVSAEERGTLTLQCEVSDPEARVVWRKDGVELGPSDKYDFLHKAGARGLTVHD LSHEDAGLYTCQVGSKETQSKVSVHDLHVGITKRLKTVEVLEGESCSFECVLSHESPS DPAVWTVGGKTVGGSGHFHAVRQGRKYTLTVKDAALSDAGEVVFSVLGLTSKASLI IRERPVDITKPLEDQRTTLGEDVMLSCELSRAGTSVRWLKDGKAIRKSQKYDLLSEG TRAVLVVRKASLKDSGEYTCETEASKSTAKLCVEEKANRFTEELADLQVEEKGRAV FTCKTEHPASVVTWRKGLLELRASGKHVPSQEGLTLKLTINALERTDSDTYTCDIGQ ARTQARLLVHGQKVRVIEDLEDTAVQEGSSAKFCCRIAPADYGPVHWFLDKTPLHS NELNEITVQPGGYHVLTLRQLALKDSGTVYFEAGDQRTSAALRVTEKPSIFSRPLTDV TVTEGEDLTLVCETTTVDSSVRWTKDGKTLRPSARCQLSHEGCQAQLLITATTPQDG GRYKCEIGGASSSSIVRVHALPVRFRESLKDVEVPEGKAATLRCVLSSVAAPVEWRH GDDVLKSSNKYSLRQEGAVLELVIRDLQPQDSGQYSCSFGDQTTSATLTVKTSSAQF VGKLRNKEATEGTTVTLRCELTKEAPVEWKKGTETLRNGDKYSLKQDGAVCELQIC SLLVADAGEYSCVCGQEKTSATLTVKALLVHFVRRLRSKEATEGDTTTLQCELSKAA PVEWRKGTETLRDGDRYSLKQDGAVCELQIRSLTIADAGEYLCTCGQEKTSATLTVR ALPAKFKDSLKNEEATEGTTATLSCELSKAAPVTWKKGPKTLQSGDKYVLRQDGAV CGLQIHGLTMADAGEYSCVCGQEKTSATLTVRGLPAKFIEDLRSQEATEGATAILRC

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		ELSKAAPVEWRKGSETLKDGDRYTLRQDGAVCELQIRGLAVVDTGTYSSLPTKFTE GLRNEEATEGTMATLRCQMSKAAPVEWRKGSETLRDGDRYSLRQDGAMCELQIRG LTIEDSGEYTCVCGQEKTSATLSVKALPSRFIEDLRSQEATEGTMATLRCQMSKTAPV EWKKGSETLRDGGRYSLRQDGPVCELQICDLAVEDAGEYSCVCGQEKTSATLSIKAL PPRFIEDLRSQEATEGTMATLRCQMSKAAPVEWRKGSETLGDGGRYSLRQNGAVCE LQIHDLAVEDTGEYSCVCGQEKTSATLNVKALPPRFIEDLRSQEATEGTMATLRCQM SKAAPVEWRKGSETLRDGGRYSLRQDGAVCELQIHDLDVEDAGQYSCVCGQEKTS AVLTVDALPPKFTEGLKKEEATEGTMVTLRCQMSKEATVEWRKGAKTLSDGGRYS LRQDGAMCELQICGLAVEDAGEYSCVCGQEKTSATLSVKALPPRFIEDLRSQEATEG TMATLRCQMSKAAPVEWRKGSETLRDGDRYSLRQDGAVCELQIRDLAVEDAGEYL CVCGQEKTSATLSVKALPPRFIEDLRSQEATEGTMATLRCQMSKAAPVEWRKGSKT LRDGDRYSLRQDGAMCELQICDLAVEDTGDYSCVCGQEKTSATLSVKALPPRFIEDL RSQEAREGTVATLRCRMSKAAPVEWRKGSETLKDGDRYSLRQEGNLCELQIRDLAV EDTEEYSCVCGQEKTSATLSVKALPAKFIEDLRSQEAPESSTVTLRCKLSKKASVVW KKGSETLRNGARYSLRQDGAVCELEIRDLTVEDTGEYSCTCGQERTSATLSIMAPQV VFQQPLQNLQAEEGSMASLRCELSVPNAAMVWSKGGLELQGDTRREARQQGCVAE LLLRDLRREDAGEYSCTCGSQTTSATLMVTAAPVRFLRELQAQDVDEGATARLRCE LSREAVSVEWRKGSLQLFPCAKYQMVQEGTTAELLVHGVEQEDAGEYTCDAGHTQ SIARLSVRAPKPKFKTDLQSTEQEAGGTARLCCQLSEAEPGTPVQWLKEGVELHVGS KYEMRRQGAVCELLIHGLEAKDTGEYACLVGGQKTLASLRVKEPEVTIVRGLVDME VQADEDVEFTCKVSQAGATDVQWHLQGLPLQSNEVTEVAVLADGCTHVLQLKGVT LEDAGTVSFHVGGLSSSAQLTVRVPEVTVLEPLKDVQLSEGQDAHFRCRLSRASGQE ARWALGGVPLQCNEMNDITVEQGTLYLLTLHKVTLEDAGTITLQVGSCSSEAQLKV TAKNTVLRGLENVDALEGGEALFECQLSQPEVAAHTWLLDDEPVHTSEKVEVVYFE NGLRHLLLLKNLKPQDSCRVTFLAGDVVTSAFLTVRGWRLEVLEPPHDASVKAGMQ VRFTCILSEAVPVGEATWYINGAAIQPDDTDWTVTTDGSHHALTLSNAQPQHAGEVT FAARDAVASARLSVLALPDPPEDAEVVGRSDHSVTLSWVAPMSDGGGGLCGYRVE MKEASTGQWQLCHDLVPGPECVVDDLVPGKTYRFRVAAVGPAGAGEPVHLPQMV KIAPAPAPAPAPAPAPAPETRQAVVGEDICLELEVAADGGEVVWHKGTERIQPGGHF EVLSRGQRQMLVIKGFRTEDQGEYRCGPIQGLPSSGAATFNVVMTSGSGDEVPAQPS LPPEAAQEGDLHLLWEALARKRRMSREPTLDSISELPEEDSRVQHLRQEAEETAPDLS EGYSTADELARTGEADLSHTSSDDESRAGTPSLVTYLKKAGGPGISPLASKHEAQVT TSVKPQKQQEPVVPTCPPPGDLSAADLMDPSLDKAAVKIQAAFKGYKVRKEMKQQE GPVFSRTFGDTEAQVGDVLRLECVLATKTDMRACWLKDGIELTDGRHYHIDQLKDG TCSLLVTGLAPTDSGRYTCQVSTKSGRVSHSACVVVSGTESEAESSSGGELDDAFRR AARRLHRLFRTKSPAELSEEELFLSADEGPGEPEEPADWQTYREDENFVCIRFESLAE ARQAVTCFRNMFATMGIGVEISLGEQGPRGVEMRIGKVAPTVTPAVPLAKTPGLQTS DAAPVFLTELQNQDVQDGYPMSFDCVVTGQPVPSVRWFKDGKLLEEDDHYMINED QQGGHQLIITAVVPADMGVYRCLAENSMGVSSTKAELRVELTSTDYDTAADATETS SYFSAQGYLSSREQEGTESDEGQLPQVLEELKDLQVAPGTRLAKFQLKVKGYPAPKL YWFKDGQPLTTSDHIRMTDKKTLHTLEIVSVTREDSGQYAAYISNAVGAAYSSARLL VRGPSEPEEKPASDVHERLVPPRILEKFTPKKVKRGSSITFSVKVEGHPAPSVHWLKE EAEKGVLWIGPDTPGYTMASSSKQHSLVLLDVGRQHQGTYTCIATNPAGQALCSAS LHISGLAKEEEQERVKEALISSFLQGTSQAVSAQMSESAGFADLVGQSKGESLVAEEA HSHLSLAEVGTEEFLQKLTSQITEMVSAKISQAKLQVPGGDSDEETKTPSASPRHGRS RPSSSVQESSSESEDGDSRGEIFDIYVVTADYLPLGAEQDAIILREGQYVEVLDSAHPL RWLVRTKPTKSSPSRQGWVSPAYLDKRLKLSPEWGPTEAPEFPGEAVSEDEYRTRLS SVIQELLSSEQAFVGELQFLESHHMKHLERSPRVPAAVASQKTVIFRNVQDISHFHSSF LKELQGCGTDDDVAMCFIKNQEAFEKYLEFLVGRVQAESVVVSTPVQEFYKKYAEE TLSAKDPTQPPPPPPLQHYLEQPVERVQKYQALLKELIRNKARNRQNCALLEQAYAVV SALPQRAENKLHVSLMENYPGTLEALGEPIRQGHFIVWEGAPGARMPWKGHNRHVF LFRNYLVICKPRRDSRTDTFSYVFRNMMKLSSIDLNDQVEGDDRAFEVWHEREDSV RKYLLQARTVIIKNSWVKEICGIQQRLAQPVWRPPEFEEELADCTAELGETVKLACR VTGTPKPIVSWYKDGKPVEVDPHHILIEDPDGSCTLILDNLTGIDSGQYMCFAASAAG NASTLGKILVQVPPRFVNKVRATPFVEGEDAQITCTVEGAPYPQIRWYKDGTLLAPG NRYRMLNEPRSGVLVLVIQAASKEDLGHYECELVNRLGSTRGGGELYMQSPALRAR DQHHREQIVAAVEDTSVEGSAHSAQDGADQQAASVLWRLLGSEALGPSPGDLPNTR QSEPPAFEEAASQIPGAASGTPEVSQPGTHKGLEQETTSSGSQGWTVPIRVEGTAWPG AGTGQLLLDVHSQVIMETTQRTYVCQAPDTGVTRAPSMQVTIEDVQVQVGDMAQF DAVIEGHPPPIVTWYKGSTQLTSSARLSQRQDGTTYSLVLTDVAPHDAGVYTCVANN AGGQVLCKAELLVHGGDKLDAENQVYRRKLHSFYDVQEEIGRGVFGFVKRVQHKG NKMFCAAKFIPLRSKTRAQAYQERDILATLGHPLVTGLLDQFETRKTLILILELCSSEE LLDRLFKKGVVTEAEVKVYIQQLVEGLHYLHSHGILHLDIKPPNILMVHPAREDIKIC DFGFAQKITPSEPQYSKYGSPEFVSPEIIEQNPVSEGSDIWAMGVISYLSLTCSSPFAGE SDRATLLNVLEGRVSWSSPTAAHLSEDAKDFIKATLQRTPRARPSTSQCLAHPWFLK SMPAEEAHFINTKQLKFLLARSRWQRSLMSYKSILVMRSIPELLQGPPDSPSLGVARH LRGEASGASSSSSSSDNELAPFARAKSLPPSPVTHSPLLHPRGFLRPSASLPEETEASMP TADAAVPASPQSAGPPASPGCVPRHSVISSLFYQQAGEGAERGNKTSGAKRHPARRR HLLKGGYIARALPGLREPLMEYSLLEEEAAREEQASLMTKTPSFETALRLPSSSVREV PGRSHSLDNPPVTTGPSPEACKEQLLFPPSTGLTHETTAKDRGHKEGFLQESVPFPPM SGDSRPGKQEGSSQDSCRGKPASSCHSELGSGSQEGCGPPSSQSLGSLPPQSLKKELST

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		SCGPLFSEQPQAAPFPTQVSPLLGSEKEPQDGSLSEGPVPVPSSSPGSASQVDASLDTE GLSEAGDTCDFTPPPQRPQEQATTRKFSLESRGGYAGVAGYGTFAFGGDAGGMLGQ GPLWARMAWAVSQSSEEQDEAATESPQPLESLGPIAEASGVPLRTSPSLTPWEEVEQ VSLVQIRDLSGDAEAADTISLDISEVDPAYLNLSDLYDIKYLPFEFMIFRRVPKPIEQPE SPGSETEAGQGLADFLEEAAWPWPGELGLRAGLEITEEPEEPGDLEALLGEAAVGRK RKWSPSRGLFQFPGRCLSGEEPVELGLRQRVKASMAHISRILKGRPEGPEREGPPRKK AGLASFRLSGLKGRDQELSDEAVVLGQSVTLACQVLAQPTAQATWSKDGVLLESSG HLLISSTLKNFQLLTILVVKEEDLGTYTCCVSNPLGTAVTTGVLRKAERPSSSPRPEVG ELYKDAVLLVWKPVESCGPVTYIVQCCIEGGSWTTLASDISDCCYLTGKLSRGGMYI FRTACVSKAGMGPYSSPSEQVLLGGPNHLASEEESSRGRPAQLLPSTKTFAFQMQIRR GRFSVVRQCREKASGRALAAKIVPYQPEDKTAVLREYEALKRLHHPHLAQLHAAYL SPRHLVLILELCSGPELLPSLAERESYSESDVKDYLWQMLSATQYLHAQHILHLDLRS ENMMVTEYNLLKVIDLGNAQSLDQEKVPAPENFKDYLETMAPELLEGQGAVPQTDI WAIGVTAFIMLSGEYPESSEGTRDLQKGLRKGLIRLSRCYAGLSGGAVAFLQSSLCA QPWGRPCASTCLQCGWLTEEGPTGSRPTPVTFPTVRLRAFVREREKRRALLYKKHNL AQVR
45	Mouse F8VPN4 (glycogen debranching enzyme)	MEHSKQIRILLLNEMEKLEKTLFRLEQGFELQFRLGPTLQGKAVTVYTNYPLPGEAFN REKFRSLDWENPTEREDDSDKYCKLHLQQSGSFQYYFLQGNEKSGGGYIVVDPILRI GVDNHVLPLDCVTLQTFLAKCLGPFDEWESRLRVAKESGYNMIHFTPLQTLGLSRSC YSLADQLELNPDFSRPSKRYTWSDVGQLVEKLKREWNILCITDVVYNHTATNSKWIL EHPESAYNLVNSPHLKPAWVLDRALWHFSCDVADGKYREKGVPALIENDQHMNCI RKIIWEDIFPRIQLWEFFQVDVHKAVEQFRRLLSQENRRVTKSEPKEHLKIIQDPEYRR RGCAVDMDTALATFIPHDNGPAAIEECCNWFRKRLEELNSEKHHLTSCHQEQAVNC LLGNVFYERLAGHGPKLGPVTRKYPLVTRYFTPPFGEMALSAEEALIHLPDKACFLM AHNGWVMGDDPLRNFAEPGSDVYLRRELICWGDSVKLRYGNKPEDCPYLWAHMK KYTEITATHFQGVRLDNCHSTPLHVAEYMLDAARKLQPNLYVVAELFTGSEELDNIF VTRLGISSLIREAMSAYNSHEEGRLVYRYGGEPVGSFVQPCLRPLMPAIAHALFMDIT HDNECPIVHRSAYDALPSTTVVSMACCASGSTRGYDELVPHQISVVAEERFYTKWNP GASPADTGDVNVHSGIIAARCAINRLHQELGAKGFIQVYVDQVDEDIVAVTRHSPSIH QSVVAVSRTAFRNPKTSFYSKEVPQMCIPGKIEEVVLEARTIERNTKPYKKDENSING MPNMTVELREHIQLHESKIVRQAGVATKGPNEYIQEIEFENLSPGSVIIFRVSLDPHAQ VAVGILRNHLTQFSSHFKSGSLAVDNADPILKIPFASIASKLTLAELNQVLYRCESEEQ EDGGGCYDIPNWSSLKYAGLQGLMSVLAEIRPKNDLGHPFCENLRSGDWMIDYVSG RLISRSGSIAEVGKWLQAMFFYLKQIPRYLIPCYFDAILIGAYTTLLDVAWKQMSSFV QNGSTFVKHLSLGSVQMCGVGKCPCLPLLSPSLLDVPCRLNEITKEKEQCCASLAAG LPHFSSGLFRCWGRDTFIALRGMLLVTGRYLEARNIILAFASTLRHGLIPNLLGEGTY ARYNCRDAVWWWLQCIQDYCRTVPNGLDILKCPVSRMYPTDDSAPLPAGTLDQPLF EVIQEAMQRHMQGIQFRERNAGPQIDRNMKDEGFNITAGIDEETGFVYGGNRFNCGT WMDKMGESDRARNRGIPATPRDGSAVEIVGLCKSAVRWLLELSKKNIFPYHEVRVK RHGKVVAVSYDEWNRKIQNNFEKLFHVSEDPSDPNEKHPNLVHKRGIYKDSYGASS PWCDYQLRPNFTIAMVVAPELFTAEKAWKALEIAEKKLLGPLGMKTLDPDDMVYC GVYDNALDNDNYNLARGFNYHQGPEWLWPIGYFLRAKLYFSKMMGPETAAKTVFL VKNVLSRHYVHLERSPWKGLPELTNENGQYCPFSCETQAWSMAVVLETLYDL
46	Mouse EF1A2 (elongation factor 1-alpha 2)	MGKEKTHINIVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKEAAEMGKGSFKYA WVLDKLKAERERGITIDISLWKFETTKYYITIIDAPGHRDFIKNMITGTSQADCAVLIV AAGVGEFEAGISKNGQTREHALLAYTLGVKQLIVGVNKMDSTEPAYSEKRYDEIVK EVSAYIKKIGYNPATVPFVPISGWHGDNMLEPSPNMPWFKGWKVERKEGNASGVSL LEALDTILPPTRPTDKPLRLPLQDVYKIGGIGTVPVGRVETGILRPGMVVTFAPVNITT EVKSVEMHHEALSEALPGDNVGFNVKNVSVKDIRRGNVCGDSKADPPQEAAQFTSQ VIILNHPGQISAGYSPVIDCHTAHIACKFAELKEKIDRRSGKKLEDNPKSLKSGDAAIV EMVPGKPMCVESFSQYPPLGRFAVRDMRQTVAVGVIKNVEKKSGGAGKVTKSAQK AQKAGK
47	Mouse AATC (aspartate aminotransferase	MAPPSVFAQVPQAPPVLVFKLTADFRDDPDPRKVNLGVGAYRTDESQPWVLPVVRK VEQKIANDNSLNHEYLPILGLAEFRSCASRLVLGDNSLAIRENRVGGVQSLGGTGAL RIGADFLGRWYNGTDNKNTPIYVSSPTWENHNAVFSAAGFKDIRPYCYWDAEKRGL DLQGFLNDLENAPEFSIFVLHACAHNPTGTDPTPEQWKQIAAVMQRRFLFPFFDSAY QGFASGDLEKDAWAIRYFVSEGFELFCAQSFSKNFGLYNERVGNLTVVGKESDSVLR VLSQMEKIVRITWSNPPAQGARIVAATLSDPELFKEWKGNVKTMADRILTMRSELRA RLEALKTPGTWSHITEQIGMFSFTGLNPKQVEYLVNEKHIYLLPSGRINMCGLTTKNL DYVATSIHEAVTKIQ
48	Mouse KAD1 (adeny late kinase isoenzyme 1)	MEEKLKKAKIIFVVGGPGSGKGTQCEKIVQKYGYTHLSTGDLLRAEVSSGSERGKKL SAIMEKGELVPLDTVLDMLRDAMLAKVDSSNGFLIDGYPREVKQGEEFEQKIGQPTL LLYVDAGAETMTQRLLKRGETSGRVDDNEETIKKRLETYYNATEPVISFYDKRGIVR KVNAEGTVDTVFSEVCTYLDSLK

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
49	Mouse G6PI (glucose-6-phosphate isomerase)	MAALTRNPQFQKLLEWHRANSANLKLRELFEADPERFNNFSLNLNTNHGHILVDYS KNLVNKEVMQMLVELAKSRGVEAARDNMFSGSKINYTENRAVLHVALRNRSNTPI KVDGKDVMPEVNRVLDKMKSFCQRVRSGDWKGYTGKSITDIINIGIGGSDLGPLMV TEALKPYSKGGPRVWFVSNIDGTHIAKTLASLSPETSLFIIASKTFTTQETITNAETAKE WFLEAAKDPSAVAKHFVALSTNTAKVKEFGIDPQNMFEFWDWVGGRYSLWSAIGL SIALHVGFDHFEQLLSGAHWMDQHFLKTPLEKNAPVLLALLGIWYINCYGCETHALL PYDQYMHRFAAYFQQGDMESNGKYITKSGARVDHQTGPIVWGEPGTNGQHAFYQL IHQGTKMIPCDFLIPVQTQHPIRKGLHHKILLANFLAQTEALMKGKLPEEARKELQAA GKSPEDLEKLLPHKVFEGNRPTNSIVFTKLTPFILGALIAMYEHKIFVQGIMWDINSFD QWGVELGKQLAKKIEPELEGSSAVTSHDSSTNGLISFIKQQRDTKLE
50	Mouse DESM (desmin)	MSQAYSSSQRVSSYRRTFGGAPGFSLGSPLSSPVFPRAGFGTKGSSSSMTSRVYQVSR TSGGAGGLGSLRSSRLGTTRAPSYGAGELLDFSLADAVNQEFLATRTNEKVELQELN DRFANYIEKVRFLEQQNAALAAEVNRLKGREPTRVAELYEEEMRELRRQVEVLTNQ RARVDVERDNLIDDLQRLKAKLQEEIQLREEAENNLAAFRADVDAATLARIDLERRI ESLNEEIAFLKKVHEEEIRELQAQLQEQQVQVEMDMSKPDLTAALRDIRAQYETIAA KNISEAEEWYKSKVSDLTQAANKNNDALRQAKQEMMEYRHQIQSYTCEIDALKGT NDSLMRQMRELEDRFASEANGYQDNIARLEEEIRHLKDEMARHLREYQDLLNVKM ALDVEIATYRKLLEGEESRINLPIQTFSALNFRETSPEQRGSEVHTKKTVMIKTIETRD GEVVSEATQQQHEVL
51	Mouse KCRS (creatine kinase S-type)	MASAFSKLLTGRNASLLFTTLGTSALTTGYLLNRQKVSADAREQHKLFPPSADYPDL RKHNNCMAECLTPTIYAKLRNKMTPSGYTLDQCIQTGVDNPGHPFIKTVGMVAGDE ESYEVFADLFDPVIKLRHNGYDPRVMKHPTDLDASKITHGQFDERYVLSSRVRTGRSI RGLSLPPACSRAERREVENVAITALEGLKGDLAGRYYKLSEMTEQDQQRLIDDHFLF DKPVSPLLTCAGMARDWPDARGIWHNYDKTFLIWINEEDHTRVISMEKGGNMKRVF ERFCRGLKEVERLIQERGWEFMWNERLGYILTCPSNLGTGLRAGVHVRIPKLSKDPR FSKILENLRLQKRGTGGVDTAAVADVYDISNIDRIGRSEVELVQIVIDGVNYLVDCEK KLERGQDIKVPPPLPQFGRK
52	Mouse E9PWQ3 (collagen, type VI, alpha 3)	MRKHRHLPLVAVFSLLLSGIATTHAQQHGDVKNGAAADIVFLVDSSWSAGKDRFLL VQEFLSDVVESLAVGDNDFHFALVRLNGNPHTEFLLNTYHSKQEVLSHIVNMSYIGG SNQTGKGLEYVIHSHLTEASGSRAADGVPQVIIVLTDGQSEDGFALPSAELKSADVN VFAVGVEGADERALGEVASEPLSMHVFNLENVTSLHGLVGNLVSCIHSSVNPERAG DKESLKDITAQDSADIIFLIDGSQNTGNANFDVIRDFLVNVLERLSVGNQQVQVGVV QYSEEPITMFSLNSYPSKAAVLDAVKGLSLVGGESANIGQALDFVVENHFTRAGGSR VEEGVPQVLVLISAGPSSDEIRDSVVALKQASVFSFGLGAQAASRAELQHIATDDSLV FTVPEFRSFGDLQEQILPYLVGVAQRHIVLQPPAIVTQVMEVNKRDIVFLVDGSSSLG PSNFNAIRDFVTRVIQRLEIGQDLVQVSVAQYADTVKPEFYLNSYTNKRDAITAVRK MRALNGSALYTGSSLDFVRNNLFTSSAGHRAAEGVPKLLVLITGGKSLDEVSQPAQE LKRGSIMALAVGSKAADEDELKEIAFDSSLVFIPAEFRPAPLQNMLPSLMAPLRTLTG TTEVHVNKRDIIFLLDGSDNVGKNNFPYVRDFVTNLVNSLDVGSDNIRVGLVQFSDT PVTEFSLDTYQTKSELLAHLRRLQLKGGSGLNAGSALSYIHANHFTEAGGSRTREHV PQLLLLLMAGPSEDAYLQAANALVRSGVLTFCVGTNRADKAELEHIAFNPSLVYLM DDFRSLPSLPQQLIQPLTTYVSGGVQEVPLSQPESKRDILFLFDGSVNVLGQFPAVRDF LYRIIEELDVKPDGTRVAIAQFSDDVRLESRFSEHQTKAEILNLVKKMKLKTGKALNL GYALDYALRNIFVRSAGSRIEDNVQQFLVLLVAGRSSDAVAGPASSLKQRGVVPFIF QAKNANPSELEQIVLSPAFILAAESLPKIGDLQSQIVSLLKAEQGSGPVSGEKDVVFLI DGSEGVRSGFPLLKDFVQRVVESLDVGPDRVRVALVQYSDRTRPEFYLNSHMDQQG VISAIRRLTLLGGPTPNTGAALEFVLRNILTSSTGSRIAEGVPQLLIVLTAEPSGDDVRG PSVVLKQGGAVPIGIGIGNADISEMQTISFIPDFAVAIPTFRELGTIQQVISERVIQLNRE ELSSLKPILTPSTGAGVGSKKDVVFLIDGSRNAGPEFQYIRTLIERIVEYLDIGFDTTRV AVIQFSEDSKMEFPLNAHFSKDEVQNAVRRLRPKGGSQVYIGNALEYVLKNIFQRPL GSRIEEGVPQFLVLISSGKSDDEVDDSAVELKQFGVAPLTIARHTDQEELVKISLSPEY VYSVSTFRELPRLEQKLLTPITTLTSQQIHQILASTRYPPSVVESDAADIVFLIDSSDAV KPDGIAHIRDFVSRIVRRLNIGPSKVRIGVVQFSNDVFPEFYLKTHKSQSSVLEAIRRLR FKGGSPLNTGRALEFVARNLFVKSAGSRIEDGVPQHLVLFLGGKSQDDVARHAQVIS SSGIVSLGIGDRNIDRTDLQTITNDPRLVFTVREFRELPNIEERVMLSFGPSGATPQPPG VDLPSPSRPEKKKADIVFLLDGSINFRRDSFQEVLRFASEIVDTVYEDGDSIRVGLVQY NSDPTDEFFLRDFSTKRQIIDAINKVVYKGGRHANTRVGIEHLLRNHFVPEAGSRLDE RVPQIAFVITGGKSVEDAQDVSLALTQKGVKVFAVGVRNIDSEEVGKIASNSATAFR VGSVQELSELSETVLETLHDAMHETLCPGVTDVSKACNLEVILGFDGSRDQNVFVSQ KGLESKVDIILNRISQIQRISCSGNQLPTVRVSVMANTPSGPVEAFDFAEYQPELFEKF RNMRSQRPYVLTADTLKLYQNKFRQSSPDTVKVVIHFTDGADGDMADLYRASEELR QAGAQALILVGLERVANLERLMHLEFGRGFMYDRPLRLNLLDLDYELAEQLDNIAE KACCGVPCKCSGERGDRGPIGSIGPKGISGEDGYRGYPGDEGGPGERGPPGVNGTQG FQGCPGQRGVKGSRGFPGEKGELGEIGLDGLDGEEGDKGLPGSSGEKGSPGRRGDK GPKGDKGERGDVGIRGDPGDSGRDSQQRGPKGETGDIGPMGLPGRDGIPGSPGDPGK DGGSGRRGPAGAKGNRGGPGQPGFEGEQGTRGSQGPPGPIGPPGLIGEQGIPGPRGG GGTAGAPGERGRTGPLGRKGEPGEPGPKGSIGNRGPRGETGDDGRDGVGSEGRRGK

TABLE 1-continued

Mouse Proteins		
SEQ ID NO	Description	Sequence
		KGERGFPGYPGPKGTPGEPGADGPPGPKGIRGRRGNSGPPGATGQKGDPGYPGPSGH KGNRGDSVDQCALIQSIRDKCPCCYGPLECPVFPTELAFALDTSEGVTQDTFSRMREV LLGIVGDLTIAESNCPRGARVAVVTYNNEVTTEIRFADSKKKSALLDSIQNLQVALTS KQQSLETAMSFVARNTFKRVRSGFLMRKVAVFFSNKPTRASPQLREAVLKLSDAGIT PLFLTSQEDRQLINALQINNTAVGHALVLPARRDLTDFLKNVLTCHVCLDICNIDPSC GFGSWRPSFRDRRAAGSDVDIDLAFILDSSEATTLFQFNEMKKYIGYVIRQLDLSPDP KASQHFARVAVVQQSTYESVDNASVPPVKVEFSLTDYGAKEKLLDFLSRRMTQLQG TMGLGNAIEYTIENIFESAPNPRDLKIMVLMLTGDMQRQQLEEAQRAILQAKCKGYF FVVLGIGRKVNIKEVYSFASEPNDVFFKFVDKSTELNEEPLMRFGRLLPSFVSSENAF YLPPDLRKQCDWFQGDQPVKNGVKFGHKQINPPHTANSSLTSKVVTTMKPVTTTKP TAIVNLPPAKPAPARPAPAQPVLAKPDPAKPAQARPAPAKPASAKLVPPQPVHVQPA PAQTASVRPAPAKPAPPQPAAAKPVPAKPAVPAQPAPPQPAAAKPVPAKPAVPAQPA PPQPAAAKPVPAKPAVPAQPAAAQPMPAQPVLTKSAAVKPASANKPVAAKPVATNT ATATARPALAAKPAAAKPAATRPLAAAIRPVATKPEAPRQQAKPAATKPATTKPLAR VSREVQVSEVTENSARLHWERPEPSSSFFYDLTVTSAHDQSLVLRQNLTVTDRVIGG LLAGQLYHVVVVSYLQSQVRAIYQGSFNTKKTQPPPLQAAHRASSSTINLMVNTEPL FLTKTDICKLSRDAGTCVDFKLLWHYDLESKSCKRFWYGGCGGNENRFHSQEECEK MCSPDLLV
53	Mouse AMD1 (AMP deaminase 1)	MPLFKLTGQGKQIDDAMRSFAEKVFASEVKDEGGRHEISPFDVDEICPISLHEMQAHI FHMENLSMDGRRKRRFQGRKTVNLSIPQSETSSTKLSHIEEFISSSPTYESVPDFQRVQI TGDYASGVTVEDFEVVCKGLYRALCIREKYMQKSFQRFPKTPSKYLRNIDGEALVG NESFYPVFTPPPKKGEDPFRTEDLPANLGYHLKMKAGVIYIYPDEAAANRDEPKPYP YPNLDDFLDDMNFLLALIAQGPVKTYAHRRLKFLSSKFQVHQMLNEMDELKELKNN PHRDFYNCRKVDTHIHAAACMNQKHLLRFIKKSYHIDADRVVYSTKEKSLTLKELFA KLNMHPYDLTVDSLDVHAGRQTFQRFDKFNDKYNPVGASELRDLYLKTDNYINGE YFATIIKEVGADLVEAKYQHAEPRLSIYGRSPDEWNKLSSWFVCNRIYCPNMTWMIQ VPRIYDVFRSKNFLPHFGKMLENIFLPVFEATINPQAHPDLSVFLKHITGFDSVDDESK HSGHMFSSKSPKPEEWTMENNPSYTYYAYYMYANITVLNSLRKERGMNTFLFRPHC GEAGALTHLMTAFMIADNISHGLNLKKSPVLQYLFFLAQIPIAMSPLSNNSLFLEYAK NPFLDFLQKGLMISLSTDDPMQFHFTKEPLMEEYAIAAQVFKLSTCDMCEVARNSVL QCGISHEEKAKFLGNNYLEEGPVGNDIRRTNVAQIRMAYRYETWCYELNLIAEGLKA TE

[0039] An animal protein can comprise a rabbit protein (Table 2). A rabbit protein can comprise a rabbit muscle protein. A rabbit protein can comprise a rabbit blood protein. A rabbit protein can comprise a rabbit albumin protein. A rabbit protein can comprise a sequence that has at least or about 70%, at least or about 75%, at least or about 80%, at least or about 90%, at least or about 91%, at least or about 92%, at least or about 93%, at least or about 94%, at least or about 95%, at least or about 96%, at least or about 97% at least or about 98% at least or about 99%, or about 100% sequence identity to any one of SEQ ID NOs: 54-78. A rabbit albumin protein can comprise a sequence that has at least or about 70%, at least or about 75%, at least or about 80%, at least or about 90%, at least or about 91%, at least or about 92%, at least or about 93%, at least or about 94%, at least or about 95%, at least or about 96%, at least or about 97% at least or about 98% at least or about 99%, or about 100% sequence identity to SEQ ID NO: 54-78.

[0040] A food product may comprise a fragment or a truncation of a rabbit protein. The food product may include a fragment or truncation of any one of SEQ ID Nos: 54-78. A sequence of the fragment or truncation may be identical or substantially identical to a portion of any one of SEQ ID Nos: 54-78. A sequence of the fragment or truncation may be at least about 70%, 80%, 90%, 95%, 98%, or more identical to a portion of a sequence of any one of SEQ ID Nos: 54-78. A length of the fragment or truncation may be at least about 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of a length of any one of SEQ ID Nos: 54-78. A length of the fragment or truncation may be less than or equal to about 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, or less than a length of any one of SEQ ID Nos: 54-78.

TABLE 2

Rabbit Proteins		
SEQ ID NO	Description	Amino Acid Sequence
54	Rabbit KCRM (Creatine Kinase M-type)	FGNTHNKYKLNYKSEEEYPDLSKHNNHMAKVLTPDLYKKLRDKETPSGFTL DDVIQTGVDNPGHPFIMTVGCVAGDEESYTVFKDLFDPIIQDRHGGFKPTDKH KTDLNHENLKGGDDLDPHYVLSSRVRTGRSIKGYTLPPHCSRGERRAVEKLS VEALNSLTGEFKGKYYPLKSMTEQEQQQLIDDHFLFDKPVSPLLLASGMARD WPDARGIWHNDNKSFLVWVNEEDHLRVISMEKGGNMKEVFRRFCVGLQKIE EIFKKAGHPFMWNEHLGYVLTCPSNLGTGLRGGVHVKLAHLSKHPKFEEILT

TABLE 2-continued

Rabbit Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		RLRLQKRGTGGVDTAAVGSVFDISNADRLGSSEVEQVQLVVDGVKLMVEME KKLEKGQSIDDMIPAQK
55	Rabbit G1U4H8 (Actinin Alpha 3)	MMMVMQPEGLGAREGPFAGGGGGGEYMEQEEDWDRDLLLDPAWEKQQRK TFTAWCNSHLRKAGTQIENIEEDFRNGLKLMLLLEVISGERLPRPDKGKMRFH KIANVNKALDFIASKGVKLVSIGAEEIVDGNLKMTLGMIWTIILRFAIQDISVE ETSAKEGLLLWCQRKTAPYRNVNVQNFHTSWKDGLALCALIHRHRPDLIDY AKLRKDDPIGNLNTAFEVAEKYLDIPKMLDAEDIVNTPKPDEKAIMTYVSCFY HAFAGAEQAETAANRICKVLAVNQENEKLMEEYEKLASELLEWIRRTVPWL ENRVGEPSMSAMQRKLEDFRDYRRLHKPPRVQEKCQLEINFNTLQTKLRLSH RPAFMPSEGKLVSDIANAWRGLEQVEKGYEDWLLSEIRRLQRLQHLAEKFQQ KASLHETWTRGKEDMLTQRDYESASLQEVRALLRRHEAFESDLAAHQDRVE HIAALAQELNELDYHEAASVNSRCQAICDQWDNLGTLTQKRRDALERMEKL LETIDQLQLEFARRAAPFNNWLDGAIEDLQDVWLVHSVEETQSLLTAHDQFK ATLPEADRERGAILGIQGEIQKICQTYGLRPSSTNPYINLSPQDINTKWDTVRK LVPSRDQTLQEELARQQMNERLRRQFAAQANAVGPWIQAKVEEVGRLAAGL AGSLEEQMAGLRQQEQNIINYKSNIDRLEGDHQLLQESLVFDNKHTVYSMEH IRVGWEQLLTSIARTINEVENQVLTRDAKGLSQEQLNEFRASFNHFDRKRNG MMEPDDFRACLISMYDLGEVEFARIMTMVDPNAAGVVTFQAFIDFMTRETAE TDTAEQVVASFKILAGDKNYITPEELRRELPAEQAEYCIHRMAPYKGAGAPA GALDYVAFSSALYGESDL
56	Rabbit PYGM (Glycogen Phosphorylase, muscle form)	SRPLSDQEKRKQISVRGLAGVENVTELKKNFNRHLHFTLVKDRNVATPRDYY FALAHTVRDHLVGRWIRTQQHYYEKDPKRIYYLSLEFYMGRTLQNTMVNLA LENACDEATYQLGLDMEELEEIEEDAGLGNGGLGRLAACFLDSMATLGLAA YGYGIRYEFGIFNQKICGGWQMEEADDWLRYGNPWEKARPEFTLPVHFYGR VEHTSQGAKWVDTQVVLAMPYDTPVPGYRNNVVNTMRLWSAKAPNDFNL KDFNVGGYIQAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYFVVAATLQD IIRRFKSSKFGCRDPVRTNFDAFPDKVAIQLNDTHPSLAIPELMRVLVDLERLD WDKAWEVTVKTCAYTNHTVLPEALERWPVHLLETLLPRHLQIIYEINQRFLN RVAAAFPGDVDRLRRMSLVEEGAVKRINMAHLCIAGSHAVNGVARIHSEILK KTIFKDFYELEPHKFQNKTNGITPRRWLVLCNPGLAEIIAERIGEEYISDLDQLR KLLSYVDDEAFIRDVAKVKQENKLKFAAYLEREYKVHINPNSLFDVQVKRIH EYKRQLLNCLHVITLYNRIKKEPNKFVVPRTVMIGGKAAPGYHMAKMIIKLIT AIGDVVNHDPVVGDRLRVIFLENYRVSLAEKVIPAADLSEQISTAGTEASGTG NMKFMLNGALTIGTMDGANVEMAEEAGEENFFIFGMRVEDVDRLDQRGYN AQEYYDRIPELRQIIEQLSSGFFSPKQPDLFKDIVNMLMHHDRFKVFADYEEY VKCQERVSALYKNPREWTRMVIRNIATSGKFSSDRTIAQYAREIWGVEPSRQR LPAPDEKIP
57	Rabbit ALDOA (Fructose-Bisphosphate Aldolase A)	HSHPALTPEQKKELSDIAHRIVAPGKGILAADESTGSIAKRLQSIGTENTEENR RFYRQLLLTADDRVNPCIGGVILFHETLYQKADDGRPFPQVIKSKGGVVGIKV DKGVVPLAGTNGETTTQGLDGLSERCAQYKKDGADFAKWRCVLKIGEHTPS ALAIMENANVLARYASICQQNGIVPIVEPEILPDGDHDLKRCQYVTEKVLAAV YKALSDHHIYLEGTLLKPNMVTPGHACTQKYSHEEIAMATVTALRRTVPPAV TGVTFLSGGQSEEEASINLNAINKCPLLKPWALTFSYGRALQASALKAWGGK KENLKAAQEEYVKRALANSLACQGKYTPSGQAGAAASESLFISNHAY
58	Rabbit MYL1 (Myosin Light Chain, Skeletal Muscle Isoform)	APKKDVKKPAAAAAPAPAPAPAPAPAPAKPKEEKIDLSAIKIEFSKEQQDEFK EAFLLYDRTGDSKITLSQVGDVLRALGTNPTNAEVKKVLGNPSNEEMNAKKI EFEQFLPMLQAISNNKDQGTYEDFVEGLRVFDKEGNGTVMGAELRHVLATL GEKMKEEEVEALMAGQEDSNGCINYEAFVKHIMSI
59	Rabbit G3P (Glyceraldehyde-3-Phosphate Dehydrogenase)	MVKVGVNGFGRIGRLVTRAAFNSGKVDVVAINDPFIDLHYMVYMFQYDSTH GKFHGTVKAENGKLVINGKAITIFQERDPANIKWGDAGAEYVVESTGVFTTM EKAGAHLKGGAKRVIISAPSADAPMFVMGVNHEKYDNSLKIVSNASCTTNCL APLAKVIHDHFGIVEGLMTTVHAITATQKTVDGPSGKLWRDGRGAAQNIIPAS TGAAKAVGKVIPELNGKLTGMAFRVPTPNVSVVDLTCRLEKAAKYDDIKKV VKQASEGPLKGILGYTEDQVVSCDFNSATHSSTFDAGAGIALNDHFVKLISWY DNEFGYSNRVVDLMVHMASKE
60	Rabbit TPM1 (Tropomyosin Alpha-1 Chain)	MDAIKKKMQMLKLDKENALDRAEQAEADKKAAEDRSKQLEDELVSLQKKL KGTEDELDKYSEALKDAQEKLELAEKKATDAEADVASLNRRIQLVEEELDRA QERLATALQKLEEAEKAADESERGMKVIESRAQKDEEKMEIQEIQLKEAKHI AEDADRKYEEVARKLVIIESDLERAEERAELSEGKCAELEEELKTVINNLKSL EAQAEKYSQKEDKYEEEIKVLSDKLKEAETRAEFAERSVTKLEKSIDDLEDEL YAQKLKYKAISEELDHALNDMTSI

TABLE 2-continued

Rabbit Proteins		
SEQ ID NO	Description	Amino Acid Sequence
61	Rabbit Myosin-Binding protein C (fast-type)	EAKPAARKAPKGKDAPKEEPPKEAAAEPPKEAPPEDQSPTAEEPTGIFLKKPD SVSVETGKDAVIMAKLNGKELPAKPTIKWFKGKWLELGSKSGARFSFKETHD AASNVYSLELHIGKVVLGDRGDYRLEVKAKDVCDSCSFNIDVEAPRQDSAGQ SLESFKRAGDAKPDGAGELDFSGLLKKREVVEEEKKKKKDDDDLGIPPEIWE LLKGAKKSEYERIAFQYGITDLRGMLKRLKKAKVEVKKSAAFTKKLDPAYQ VDRGNKVKLVVEISDPDLPLKWFKNGQEIKPSLKYVFENVGKKRILTINKCTL ADDAAYEVVVNDEKCFTELFVKEPPVLIVTPLEDQQVFVGDRVEMSVEVSEE GAQVMWMKDGVELTREGSFKSNYRFKKDGKRHILIYSEVNLEDRGHYQVM TNGGQCEAELIVEEKQLEVLQDIADLTVKASEQAMFKCEVSDEKVTGKWYK NGVEVRPSKRITISHVGRFHKLVIDDVRPEDEGDYTFVPDGYALSLSAKLNFL EIKVEYVPKQEPPKIHLDCSGKTSDNAIVVVAGNKLRLDVSITGEPPPVATWM KGDEVFTVTEGRVRIEQRPDSSSFVIESAERGDEGRYTIKVTNPAGEDVATVM LRVVDVPDPPEAVRVTSVGEDWAILVWDPPKYDGGQPVTGYLLERKKKGSQ RWMKLNFEVFTETTYESTKMIEGILYEMRVFAVNAIGVSQPSMNTKPFMPIAP TSAPQHLTVEDVTDTTTTLKWRPPDRIGAGGIDGYLVEYCVEGSDEWIPANK EPVERCGFTVKGLPTGAKILFRVIGVNIAGRSEPATLVQPVTIREIVEQPKIRLP RHLRQTYVRKVGEHINLVIPFQGKPRPQVVWTKGGAPVDTSRVHVRTSDFDT VFFVREAARSDSGEYELSVQIENMKDTATVRIRVVEKAGPAENVMVKEVWG TNALVEWQPPKDNGNSEITGYLVQKADKKTMEWFTVYEHNRHTSCTVSDLI VGNEYYFRVYTENVCGLSDSPGVSKNTAHILKTGITFKPFEYKEHDFRSPPKF LTPLIDRVVVAGYAAALNCAVRGHPKPKVVWMKNKVEIHEDPKFLMTNHQ GVLTLNIRRPSPFDAGTYSCRAFNELGEALAECRLDVRVPQ
62	Rabbit G1T4T6 (Myomesin 2)	MSLVAVPFYQKRHKHFDQSYRNIQTKYLMDEYTAKKRASTQMSTHKSASTQ MATHKSALQKASSQTSSGGTTCRLCAKRLSAAEEEEQESKNRYRSLVASYGE AKRQRFLSELAQLEESVHLARSQARDVVDKYALQQMVEDKMAWERHVFEE RLSRAPEILVQLRSHTVWEKMSVKLCFTVQGFPTPVVQWYKDGSLICQAGEP GKYRIDSKYGVHTLEISRASFDDTATYSAVATNAHGQVSTNAAVVVRRYRG DEEPFYSVGLPIGYVLVKESKWTRMFFGEGQASLSFSHLNKDDEGLYTLRIVS RGGVSDHSAFLFVRDADPLVTGAPGAPMDLQCHDANRDYVIVTWKPPNTTT ESPVIGYFVDRCEVGTNNWVQCNDAPVKICKYPVTGLYEGRSYIFRVRAVNS AGISRPSRVSDAVAALDPADLRRLQAIHLEGEKEVVIYQDDLEGDVQVPGPPT NVHASEISRNYVVLSWEPPVPRGKEPLMYFIEKSVVGSGSWQRVNAQTAVRS PRYAVFDLAEGKPYVFRVLSANKHGLSDPSEITSPIQAQDTIAVPSAPGRVLAS RNTKTSVVVQWDRPKHEEDLLGYYVDCCVAGTNAWEPCNHKPIGYNRFVV HGLTTGEQYIFRVKAVNALGTSENSQESEVIKVQAALTVPSHPYGVTLLNCD GHSMTLGWKVPKFSGGSPILGYYVDKREVHHKNWHEVNSSPIQERILTVEGL TEGSFYEFKIAATNLAGIGQPSDPSEHFKCEAWTAPEPGPAYDLTFCEVRDTSL VMLWKAPVYSGSSPVSGYFVDFKEEDSGEWSTVNEAATPNRYLKVSDLHQG RTYVFRVRAVNASGVGKPSDVSEPVLVEARPGTKEISAGVDEEGNIYLGFDC QEMTDASHFTWCKSYEEIADPERFHVETKGDHSKLYFKNPDKADLGTYSVSV SDTDGVSSSFVLDEEELERLLALSNEIKNPTIPLKSELAYEIFDKGQVRFWLQT EPLSPDASFRFVINDREVSDSETHRMKCDKASGIIEMVMDRFTIENEGTYTVQI HDGKAKNQSSLVLIGDAFRKVLEEAEFQRKEFLRKQGPHPAEYLHWDVTEEC EVRLICKVANTKKETVFKWLKDDVLYETETPPDLEKGVCELLIPKLSKKDHG EYKATLKDERGQDASILEIAGKVYEDMILAMSRVCGMSASPLKILCTPEGIRL QCFMKYFTEEMRVNWYHKDAKISSSEHMRIGGSEDMAWLQICEPTEKDKGK YTFEIFDGKDNHQRSLDLSGQAFDEAFAEFQRLKAAAFAEKNRGKVIGGLPD VVTIMEGKTLNLTCTVFGNPDPEVVWFKNDQDIELSEHFSVKVEQAKYVSLTI KGVTSEDSGKYSINVKNKYGGEKIDVTVSVYKHGEAIPEVAPPQQAKPKLIPA SASATAP
63	Rabbit ENOB (Beta-enolase)	AMQKIFAREILDSRGNPTVEVDLHTAKGRFRAAVPSGASTGIYEALELRDGDK SRYLGKGVLKAVEHINKTLGPALLEKKLSVVDQEKVDKFMIELDGTENKSKF GANAILGVSLAVCKAGAAEKGVPLYRHIADLAGNHDLVLPVPAFNVINGGSH AGNKLAMQEFMILPVGASSFREAMRIGAEVYHHLKGVIKAKYGKDATNVGD EGGFAPNILENNEALELLKTAIQAAGYPDKVVIGMDVAASEFHRNGKYDLDF KSPDDPARHITGQKLGELYKSFIKNYPVVSIEDPFDQDDWATWTSFLSGVDIQI VGDDLTVTNPKRIAQAVEKKACNCLLLKVNQIGSVTESIQACKLAQSNGWGV MVSHRSGETEDTFIADLVVGLCTGQIKTGAPCRSERLAKYNQLMRIEEALGD KAVFAGRKFRNPKAK
64	Rabbit MLRS (Myosin Regulatory Light Chain 2, Skeletal Muscle Isoform Type 2)	APKKAKRRAAAEGGSSNVFSMFDQTQIQEFKEAFTVIDQNRDGIIDKEDLRDT FAAMGRLNVKNEELDAMMKEASGPINFTVFLTMFGEKLKGADPEDVITGAF KVLDPEGKGTIKKQFLEELLTTQCDRFSQEEIKNMWAAFPPDVGGNVDYKNI CYVITHGDAKDQE

TABLE 2-continued

Rabbit Proteins		
SEQ ID NO	Description	Amino Acid Sequence
65	Truncation of Rabbit G1SJN7 (Myosin 4)	IFSISDNAYQFMLTDRENQSILITGESGAGKTVNTKRVIQYFATIAITGDKKKEE PTPGKMQGTLEDQIISANPLLEAFGNAKTVRNDNSSRFGKFIRIHFGATGKLAS ADIETYLLEKSRVTFQLKAERSYHIFYQIMSNKKPELIEMLLITTNPYDFAFVS QGEITVPSIDDQEELMATDSAVDILGFTSDEKVAIYKLTGAVMHYGNLKFKQ KQREEQAEPDGTEVADKAAYLTSLNSADLLKALCYPRVKVGNEYVTKGQTV QQVYNAVGALAKAIYEKMFLWMVTRINQQLDTKQPRQYFIGVLDIAGFEIFD FNSLEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEWEFIDFGMDLAACIE LIEKPMGIFSILEEECMFPKATDTSFKNKLYEQHLGKSNNFQKPKPAKGKAEA HFSLVHYAGTVDYNITGWLDKNKDPLNETVVGLYQKSGMKTLAFLFSGAQT GEEGGGGGKKGGKKKGSSFQTVSALFRENLNKLMTNLRSTHPHFVRCLIPNE TKTPGAMEHELVLHQLRCNGVLEGIRICRKGFPSRILYADFKQRYKVLNASAI PEGQFIDSKKASEKLLGSIDVDHTQYKFGHTKVFFKAGLLGLLEEMRDEKLA QLITRTQAICRGYLTRVEFKKMMERRESIFCIQYNIRAFMNVKHWPWMKLYF KIKPLLKSAETEKEMANMKEEFEKTKEELAKSEAKRKELEEKMVALMQEKN DLQLQVQAEADGLADAEERCDQLIKTKIQLEAKIKEVTERAEDEEEINAELTA KKRKLEDECSELKKDIDDLELTLAKVEKEKHATENKVKNLTEEMAGLDENIA KLTKEKKALQEAHQQTLDDLQAEEDKVNTLTKAKTKLEQQVDDLEGSLEQE KKLRMDLERAKRKLEGDLKLAQESTMDIENDKQQLDEKLKKKEFEMTNLQS KIEDEQALAMQLQKKIKELQARIEELEEEIEAERASRAKAEKQRSDLSRELEEI SERLEEAGGATSAQIEMNKKREAEFQKMRRDLEEATLQHEATAAALRKKHA DSVAELGEQIDNLQRVKQKLEKEKSELKMEIDDLAGNMETVSKAKGNLEKM CRTLEDQLSEVKTKEEEHQRLINELSAQKARLHTESGEFSRQLDEKDAMVSQ LSRGKQAFTQQIEELKRQLEEETKAKSALAHALQSSRHDCDLLREQYEEEQE AKAELQRAMSKANSEVSQWRTKYETDAIQRTEELEEAKKKLAQRLQDAEEH VEAVNSKCASLEKTKQRLQN
66	Truncation of Rabbit G1SJN7 (Myosin 4)	IFYQIMSNKKPELIEMLLITTNPYDFAFVSQGEITVPSIDDQEELMATDSAVDIL GFTSDEKVAIYKLTGAVMHYGNLKFKQKQREEQAEPDGTEVADKAAYLTSL NSADLLKALCYPRVKVGNEYVTKGQTVQQVYNAVGALAKAIYEKMFLWM VTRINQQLDTKQPRQYFIGVLDIAGFEIFDFNSLEQLCINFTNEKLQQFFNHHM FVLEQEEYKKEGIEWEFIDFGMDLAACIELIEKPMGIFSILEEECMFPKATDTSF KNKLYEQHLGKSNNFQKPKPAKGKAEAHFSLVHYAGTVDYNITGWLDKNK DPLNETVVGLYQKSGMKTLAFLFSGAQTGEEGGGGGKKGGKKKGSSFQTVS ALFRENLNKLMTNLRSTHPHFVRCLIPNETKTPGAMEHELVLHQLRCNGVLE GIRICRKGFPSRILYADFKQRYKVLNASAIPEGQFIDSKKASEKLLGSIDVDHT QYKFGHTKVFFKAGLLGLLEEMRDEKLAQLITRTQAICRGYLTRVEFKKMME RRESIFCIQYNIRAFMNVKHWPWMKLYFKIKPLLKSAETEKEMANMKEEFEK TKEELAKSEAKRKELEEKMVALMQEKNDLQLQVQAEADGLADAEERCDQLI KTKIQLEAKIKEVTERAEDEEEINAELTAKKRKLEDECSELKKDIDDLELTLAK VEKEKHATENKVKNLTEEMAGLDENIAKLTKEKKALQEAHQQTLDDLQAEE DKVNTLTKAKTKLEQQVDDLEGSLEQEKKLRMDLERAKRKLEGDLKLAQES TMDIENDKQQLDEKLKKKEFEMTNLQSKIEDEQALAMQLQKKIKELQARIEE LEEEIEAERASRAKAEKQRSDLSRELEEISERLEEAGGATSAQIEMNKKREAEF QKMRRDLEEATLQHEATAAALRKKHADSVAELGEQIDNLQRVKQKLEKEKS ELKMEIDDLAGNMETVSKAKGNLEKMCRTLEDQLSEVKTKEEEHQRLINELS AQKARLHTESGEFSRQLDEKDAMVSQLSRGKQAFTQQIEELKRQLEEETKAK SALAHALQSSRHDCDLLREQYEEEQEAKAELQRAMSKANSEVSQWRTKYET DAIQRTEELEEAKKKLAQRLQDAEEHVEAVNSKCASLEKTKQRLQNEVEDL MIDVERSNAACAALDK
67	Truncation of Rabbit G1SJN7 (Myosin 4)	TLDDLQAEEDKVNTLTKAKTKLEQQVDDLEGSLEQEKKLRMDLERAKRKLE GDLKLAQESTMDIENDKQQLDEKLKKKEFEMTNLQSKIEDEQALAMQLQKKI KELQARIEELEEEIEAERASRAKAEKQRSDLSRELEEISERLEEAGGATSAQIE MNKKREAEFQKMRRDLEEATLQHEATAAALRKKHADSVAELGEQIDNLQRV KQKLEKEKSELKMEIDDLAGNMETVSKAKGNLEKMCRTLEDQLSEVKTKEE EHQRLINELSAQKARLHTESGEFSRQLDEKDAMVSQLSRGKQAFTQQIEELKR QLEEETKAKSALAHALQSSRHDCDLLREQYEEEQEAKAELQRAMSKANSEVS QWRTKYETDAIQRTEELEEAKKKLAQRLQDAEEHVEAVNSKCASLEKTKQR LQNEVEDLMIDVERSNAACAALDKKQRNFDKVLAEWKHKYEETQAELEASQ KESRSLSTEVFKVKNAYEESLDHLETLKRENKNLQQEISDLTEQIAESGKHIHE LEKVKKQIDQEKSELQAALEEAEGSLEHEEGKILRIQLELNQVKSEIDRKIAEK DEEIDQLKRNHLRVVESMQSTLDAEIRSRNDALRIKKKMEGDLNEMEIQLNH ANRQAAEAIKNLRNTQGILKDTQLHLDDAVRGQDDLKEQLA
68	Truncation of Rabbit G1SJN7 (Myosin 4)	LAQESTMDIENDKQQLDEKLKKKEFEMTNLQSKIEDEQALAMQLQKKIKELQ ARIEELEEEIEAERASRAKAEKQRSDLSRELEEISERLEEAGGATSAQIEMNKK REAEFQKMRRDLEEATLQHEATAAALRKKHADSVAELGEQIDNLQRVKQKL EKEKSELKMEIDDLAGNMETVSKAKGNLEKMCRTLEDQLSEVKTKEEEHQR LINELSAQKARLHTESGEFSRQLDEKDAMVSQLSRGKQAFTQQIEELKRQLEE ETKAKSALAHALQSSRHDCDLLREQYEEEQEAKAELQRAMSKANSEVSQWR TKYETDAIQRTEELEEAKKKLAQRLQDAEEHVEAVNSKCASLEKTKQRLQNE

TABLE 2-continued

Rabbit Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		VEDLMIDVERSNAACAALDKKQRNFDKVLAEWKHKYEETQAELEASQKESR SLSTEVFKVKNAYEESLDHLETLKRENKNLQQEISDLTEQIAESGKHIHELEKV KKQIDQEKSELQAALEEAEGSLEHEEGKILRIQLELNQVKSEIDRKIAEKDEEI DQLKRNHLRVVESMQSTLDAEIRSRNDALRIKKKMEGDLNEMEIQLNHANR QAAEAIKNLRNTQGILKDTQLHLDDAVRGQDDLKEQLAMVERRANLMQAEI EELRASLEQTERSRRVAEQELLDASERVQLLHTQNTSLINTKKKLETDISQIQG EMEDIVQEARNAEEKAKKAITDAAMMAEELKKEQDTSAHLERMKKNMEQT VKDLQHRLDEAEQLALKGGKKQIQKLEARVRELENEVESEQKRNVEAVKGL RKHERRVKELTYQTEEDRKNVLRLQDLVDKLQSKVKAYKRQAEEAEEQSN
69	Truncation of Rabbit MYH4 (Myosin 4)	VFVADPKESFVKATVQSREGGKVTAKTEAGATVTVKEDQVFPMNPPKYDKI EDMAMMTHLHEPAVLYNLKERYAAWMIYTYSGLFCVTVNPYKWLPVYNAE VVTAYRGKKRQEAPPHIFSISDNAYQFMLTDRENQSILITGESGAGKTVNTKR VIQYFATIAVTGDKKKEEATSGKMQGTLEDQIISANPLLEAFGNAKTVRNDNS SRFGKFIRIHFGTTGKLASADIETYLLEKSRVTFQLKAERSYHIFYQIMSNKKP DLIEMLLITTNPYDYAFVSQGEITVPSIDDQEELMATDSAIDILGFTSDERVSIY KLTGAVMHYGNMKFKQKQREEQAEPDGTEVADKAAYLQSLNSADLLKALC YPRVKVGNEYVTKGQTVQQVYNAVGALAKAVYEKMFLWMVTRINQQLDT KQPRQYFIGVLDIAGFEIFDFNSLEQLCINFTNEKLQQFFNHHMFVLEQEEYKK EGIEWEFIDFGMDLAACIELIEKPMGIFSILEEECMFPKATDTSFKNKLYEQHL GKSNNFQKPKPAKGKVEAHFSLVHYAGTVDYNITGWLDKNKDPLNETVVGL YQKSAMKTLAFLFTGTAAAEAEGGGKKGGKKKGSSFQTVSALFRENLNKLM TNLRSTHPHFVRCIIPNETKTPGAMEHELVLHQLRCNGVLEGIRICRKGFPSRIL YADFKQRYKVLNASAIPEGQFIDSKKASEKLLGSIDVDHTQYKFGHTKVFFKA GLLGLLEEMRDDKLAQLITRTQAMCRGFLARVEYKKMVERRESIFCIQYNIR AFMNVKHWPWMKLYFKIKPLLKSAETEKEMANMKEEFEKTKESLAKAEAK RKELEEKMVALMQEKNDLQLQVQAEADSLADAEERCDQLIKTKIQLEAKIKE VTERAEDEEEINAELTAKKRKLEDECSELKKDIDDLELTLAKVEKEKHATENK VKNLTEEMAGLDETIAKLTKEKKALQEAHQQTLDDLQAEEDKVNTLTKAKT KLEQQVDDLEGSLEQEKKIRMDLERAKRKLEGDLKLAQESTMDIENDKQQL DEKLKKKEFEMSNLQSKIEDEQALAMQLQKKIKELQARIEELEEEIEAERASR AKAEKQRSDLSRELEEISERLEEAGGATSAQIEMNKKREAEFQKMRRDLEEA TLQHEATAATLRKKHADSVAELGEQIDNLQRVKQKLEKEKSELKMEIDDLAS NMETVSKAKGNLEKMCRTL
70	Truncation of Rabbit Nebulin	TLVSDVDYRNYLHQWTCMPDQNDVIQAKKAYELQSDAVYKADLEWLRGIG WMPSDSVSVNHAKHAADIFSEKKYRTKIETLNFTPVDDRVDYVTAKQSGEIL DDIKYRKDWNSIKSKYTLTETPLLHTVQEAARILDQYLYKESWEKQKATGYI LPPDAVPFVHAHHSSDVQSELKYKAEHVKQRGHYVGVPTMRDDPKLVWFE HAGQIQNDRLYKEDYHKTKAKINIPADMVSVVAAKEGQNLVSDIDYRNYLH QWTCHPDQNDVIQARKAYDLQSDNVYKADLEWLRGIGWIPLDSVDHVRVT KNQEMVNQIKYKKDALDNYPNFTSVVDPPAIVLAKINAVNQSDVKYKETFN KAKGKYTFSPDTPHISHSKDMGKLYSTILYKGAWEGTKAYGYTLDERYIPIVG AKHADFVNSELKYKETYEKQKGHYLAGKVISEFPGVVHCLDFQKMRSVLNY RRDYENTKANVHIPNDMMNHVLAKKCQYILSDLEYRHYFHQWTSLPDEPKV LLARNAQEILSDNVYKDDLNWLKGIGCYVWDTPQILHAKKSYDLQSQLQYT AAGKENLQNYNLVTDTPLYVTALQSGINASEVKYKENYHQIKDKYTTVLETV DYDRTRNLKDLYSSNLYKEAWNKVKATSYILPTSTLSLTHAKNQKHLASHIK YREEYEKFKALYTLPKSVEDDPNTARCLRVGKLNIDRLYRSVYEKNKMKIHI VPDMVEMVTAKDTQKKVSEIDYRLHLHEWICHPDLQVNSHVRKVTDQISDIV YKDDLNWLKGIGCYVWDTPEILHAKHAYDLRNDIKYKAHMMKTRNDYKLV TDTPVYVQAVKSGKQLSDAVYHHDYVHSVRGKVAPTTKTVDLERALHAYK LQSENLYKTSLRTLPTGYRLPGDTPHFKHNKDIRYMSSYFKYKEVYEHIKAN GYTLGPNDVPFVNVRRVNKVTSERLYRELYHKLKDKIHTTPDTPEIRQVKKT QEAVSELIYKSDFFKMQGHMISLPYTPHVLHCRYVGDITSDIKYKEDLQILKG LGCFLYDTPDMVRSRHLRKLWSNYLYTDNARKMRDKYKVVLDTPEYRKVQ ELKTHLSELVYRASGRKQKSIFTSVPDTPDLLRAKRGQKLQSQYLYVELATKE RPHHHAGNQTTALKHARDVKDMVSEKKYKVQYEKMKDKYTPIPDTPILIRA KNAYWNASDLRYKETFQKTKGKYHTVKDALDIVYHRKVTDDISKVKYKEN YMRHLGIWRSIPDRPEHFHHRAVTDAVSDIKYKEDLTWLKGIGCYAYDTPDF TLAEKNKTLYSKYKYKEVFERTKSDFKYVADCPINRHFKYATQLMNEKKYR ADYEQRKDKYQLVVDEPRHLLAKTAGDQISQIKYRKQYEKSKDKFTSIVDTP EHLRTTKVNKQISDILYKLEYNKAKPRGYTTIHDTPMLLHVRKVKDEVSDLK YKEVYQRNKSNCTIKPDAVHIKAAKDAYKVNTNLDYKKQYEANKAHWRWT PDRPDFIQAAKSSLQQSDFEYKLDREFLKGCKLSVTDDKDMVLALKNSLIESD LKYKEKHVKERGSCHAVPDTPQILLAKTVSNLV
71	Truncation of Rabbit Nebulin	EHTRNVNQIQSDNVYKDEYNNFFKGIGWIPIGSLEVEKVKKAGDALNEKKYR QHPDTIKFTSVPDSMGMVLAQHNTKQLSDLNYKVEGEKVKHKYTMDPEVPQ FIQARVNAFNLSDAHYKADWKKTLAKGYDLRPDAIPIVAAKSSRNIASDCKY KEAYEKAKGKQIGFLSLQDDPKLVHYMNVAKIQSDREYRKGYEASKTRYHT PLDMLSVTAAKKSQEVATNANYKQPFHNYTLLPDALPVEHSRNAMQIQSDN

TABLE 2-continued

Rabbit Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		LYKSDFTGWMKGIGWIPIQSLEVEKAKKAGEILSEKRYRQHPEKLKFTYAMD TMEQALNKANKLNMDKRLYTKKWDKDKTTIHVMPDTPDILLSRVNQITMSD KLYRAGWEEEKKKGYDLRPDAISIKAAKASRDIASDYKYKQAYEQAKGKHI GFRSLEDDPKLVHFMQVAKMQSDREYKKAYEKSKTSFHTPVDMLSVVAAK KSQEVATNANYRNVIHTYNMLPDAMGFELAKNMMQIQSDNQYKADYADFM KGIGWLPLGSLEAEKNKKAMEIISEKKYRQHPDTLKYSTLMDSMNMVLAQN NAKIMNEHLYKQAWEADKTKVHIMPDIPQIVLAKANAINMSDKLYKLSMEE AKKKGYDLRTDAIPIKAAKASRDIASDYKYKYNYEKQKGKMVGFRSLEDDP KLVHSMQVAKMQSDREYKKNYEKTKTSYHTPADMLSVTAAKDAQANITNT NYKHLIHKYILLPDAMSLELSRNMNRIQSDNEYKQDYNEWYKGLGWSPAGS LDVEKAKKATEYASDQKYRQHPSNFQFKKLTDSMDMVLAKQNAHTMNKHL YTIDWNKDKTKIHVMPDTPDILQAKQNQTLFSQKFYKLGWEEALKKGYDLP VDAISVQLAKASRDIASDFKYKQGYRKQLGHHVGFRSLQDDPKLVLSMNVA KMQSEREYKKDFEKWKTKFSSPVDMLGVVLAKKCQALVSDVDYKNYLHQ WTCLPDQNDVIHAKKAYELQSENMYKSDLEWLKGIGWSPLGSLEAEKNKRA SEILSEKKYRQPPDTNKFTSIPDAMDIVLAKTNAKNRSDRLYREAWDKDKTQ VHIMPDTPDIILAKANLINTSDKLYRMGYEELKKKGYDLPVDAIPIKAAKASR EIASEYKYKEGFRKQLGHHIGARNIKDDPKMMWSMHVAKIQSDREYKKDFE KWKTKFSSPVDMLGVVLAKKCQTLVSDIDYKNYLHQWTCLPDQNDVTHAR HAYDLQSDNLYKSDLQWLKGIGWIPTGSIEDEKNKRATQILSDHVYRQHPDK FKFSILMDSIPMVLAKNNATNMNHRLYTEAWDKDKTTVHIMPDTPEVLLAK QNQVNCSEKLYKLGLEETKRKGYDMRVDAIPIKAAKASRDIASEFKYKEGYR KQLGHHIGARAIHDDPKMMWSMHVAKIQSDREYKKDFEKWKTKFSSPVDM LGVVLAKKCQALVSDVDYKNYLHQWTCLPDQNDVMHARQAYDLQSDNLY KSDLQWLRGIGWVPIGSLDVEKC
72	Truncation of Rabbit Nebulin	DVTHARHAYDLQSDNLYKSDLQWLKGIGWIPTGSIEDEKNKRATQILSDHVY RQHPDKFKFSILMDSIPMVLAKNNATNMNHRLYTEAWDKDKTTVHIMPDTP EVLLAKQNQVNCSEKLYKLGLEETKRKGYDMRVDAIPIKAAKASRDIASEFK YKEGYRKQLGHHIGARAIHDDPKMMWSMHVAKIQSDREYKKDFEKWKTKF SSPVDMLGVVLAKKCQALVSDVDYKNYLHQWTCLPDQNDVMHARQAYDL QSDNLYKSDLQWLRGIGWVPIGSLDVEKCKKATEILSDKIYRQPPDRFKFTSV TDSLEQVLAKNNAINMNKRLYTEAWDKDKTQIHIMPDTPEIMLARMNKINYS ESLYKLANEEAKKKGYDLRSDAIPIVAAKASRDIASDYKYKDGYRKQLGHHI GARDIKDDPKMMWSMHVAKIQSDREYKKDFEKWKTKFSSPVDMLGVVLAK KCQTLVSDIDYKNYLHQWTCLPDQNDVMHARQAYDLQSDNIYKSDLQWLR GIGWVPIGSLEVAKCKRAAEILSDNIYRQPPDKLKFTSVPDSLEQVLAKNNAIN MNKRLYTEAWDKDKTQIHIMPDTPEIMLARQNKINYSESLYKLANEEAKKKG YDLRSDAIPIVAAKASRDIASDYKYKDGYRRQLGHHIGARDIKDDPKMMWS MHVAKIQSDREYKKDFEKWKTKFSSPVDMLGVVLAKKCQTLVSDIDYKNYL HQWTCLPDQNDVMHARQAYDLQSDNIYKSDLQWLRGIGWVPIGSVDLVKC KRAAEILSDNIYRQSPDKLKFTSVTDSLEQVLAKNNAINMNKRLYTEAWDKD KTQIHMMPDTPEIMLARQNKMNYSENFYRQAMEEAKKEGYDLRSDAIPIVA AKASRDIASDYKYKEAYRKQLGHHIGARAIHDDPKMMWSLHIAKVQSDREY KKEFEKYKTRYSSPVDMLGIVLAKKCQTLVSDVDYKNLLHEWTCLPDQNDV IQARKAYDLQSDNLYKSDLEWLRGIGWVPIGSLEVVKAKRATEILSDNLYRT RPDTLKFTSVTDSLEQVLAKNNAINMNKRLYTEAWENDKKTIHVMPDTPEIM LAKLNKINYSDKLYKLALEESKKEGYDLRVDAIPIQAAKASRDIASDYKYKE GYRKQLGHHIGARNIKDDPKMMWSMHVAKIQSDREYKKDFEKWKTKFSSP VDMLGVVLAKKCQILVSDIDYKHPLHEWICLPDQNDVIQARKAYDLQSDAV YKADLEWLRGIGWVPIGSVEVEKVKRAGEILSDRKYRQPADKLKFTSITDTPE IVLAKSNALTMSKHLYTEAWDADKTSIHVMPDTPEILLAKSNSANISQKLYTK GWDESKMKDYDLKADAIPIKSAKASRDIASDYKYKEAYEKQKGHHIGAQSIE DDPKIMCAIHA
73	Truncation of Rabbit TTN	DKRETSRPNWAQVSANVPITSCSVEKLIEGHEYQFRICAENKYGVGDPVFTEP AIAKNPYDPPGRCDPPVISNVTKDHMTVSWKPPADDGGSPITGYLLEKRETHA VNWTKVNRKPVIERTIKATGLQEGTEYEFRVTAINKAGPGKPSDASKAVYAQ DPLYPPGPPAFPKVYDTTRSSVSLSWGKPAYDGGSPIIGYLVEVKRADTDNW VRCNLPQKLQKTRFEVTGLMENTEYQFRVYAVNKVGYSDPSDVPDKHCPKD ILIPPEGELDADLRKTLILRAGVTMRLYVPVKGRPPPKITWSKPNVNLRERIGL DIKSTDFDTFLRCENVNKYDAGKYILTLENSCGKKEYTIVVKVLDTPGPPVNV TVKEISRDSAYITWDPPIVDGGSPIINYVVEKRDAERKSWSTVTTECSKTSFRV SNLEEGKSYFFRVFAENEYGIGDPGETRDAVKASETPGPVVDLKVLTVTKSSC NIGWKKPRSDGGSRITGYVVDFLTEENKWQRVMKSLSLQYSTKDLNEGKQY TFRVSAENENGEGTPSEITVVAKDDVVAPDLDLKDLPDLCYLAKENSNFRLKI PIKGKPAPSVTWKKGEDPLATDTRVSVESSAVNTTLVVYDCQKSDAGKYTIT LKNVAGTKEGTLSIKVVGKPGIPTGPIKFDEVTAEAITLKWGPPKDDGGSEITN YILEKRDSVNNKWVTCASAVQKTTFRVTRLHEGMEYTFRVSAENKYGVGEG LKSEPIVAKHPFDVPDAPPPPNIVDVRHDSVSLTWTDPRKTGGSPITGYHIEFK ERNSLLWKRANKTPIRMKDFKVTGLTEGLEYEFRVMAINLAGVGKPSLPSEP VVALDPIDPPGKPEVINVTRNSVTLIWTEPKYDGGHKLTGYIVEKRDLPSKTW

TABLE 2-continued

Rabbit Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		MKANHINVPDCAFTVTDLVEGGKYEFRIRAKNTAGAISAPSESTGTIICKDEYE APTIVLDPTIKDGLTIKAGDTIVLNAISILGKPLPKSSWSKAGKDIRPSDITQITS TPTSSMLTVKYASRKDAGEYTITATNPFGTKEEHVRVTVLDVPGPPGPIEISNV SAEKATLTWTPPLEDGGSPIKSYVLEKRETSRLLWTVVAEDIQSCRHVVTKLI QGNEYLFRVSAVNHYGKGEPVQSEPVKMVDRFGPPGPPGKPEVSNVTKNTA TVSWKRPTDDGGSEITGYYVERREKKGLRWVRATKTPVSDLRCKVTGLQEG NTYEFRVSAENRAGIGPPSDASNYVLMKDVAYPPGPPSNARVTDTTKKSASL AWGKPHYDGGLEITGYVVEHQKVGDETWVKDTTGPALRITEFVVPDLHTKE KYNFRISAINDAGVGEPAVIPDVEIVEREMAPDFELDAELRRTLVVRAGLSIRI FVPIKGRPAPEVTWTKDDINLKTRANIENTESFTLLIIPECNRYDTGKFVMTIEN PAGKKSGFVNVRVLDTPGPVLNLRPTDITKDSVTLHWDLPLIDGGSRITNYIV EKREATRKSYSTVTTKCHKCT
74	Truncation of Rabbit KCRM (Creatine Kinase M-Type)	DKETPSGFTLDDVIQTGVDNPGHPFIMTVGCVAGDEESYTVFKDLFDPIIQDR HGGFKPTDKHKTDLNHENLKGGDDLDPHYVLSSRVRTGRSIKGYTLPPHCSR GERRAVEKLSVEALNSLTGEFKGKYYPLKSMTEQEQQQLIDDHFLFDKPVSPL LLASGMARDWPDARGIWHNDNKSFLVWVNEEDHLRVISMEKGGNMKEVFR RFCVGLQKIEEIFKKAGHPFMWNEHLGYVLTCPSNLGTGLRGGVHVKLAHLS KHPKFEEILTRLRLQKRGTGGVDTAAVGSVFDISNADRLGSSEVEQVQLVVD GVKLMVEMEKKLEKGQSIDDMIPAQK
75	Truncation of Rabbit G1U4H8 (Actinin Alpha 3	GGGGGGEYMEQEEDWDRDLLLDPAWEKQQRKTFTAWCNSHLRKAGTQIEN IEEDFRNGLKLMLLLEVISGERLPRPDKGKMRFHKIANVNKALDFIASKGVKL VSIGAEEIVDGNLKMTLGMIWTIILRFAIQDISVEETSAKEGLLLWCQRKTAPY RNVNVQNFHTSWKDGLALCALIHRHRPDLIDYAKLRKDDPIGNLNTAFEVAE KYLDIPKMLDAEDIVNTPKPDEKAIMTYVSCFYHAFAGAEQAETAANRICKV LAVNQENEKLMEEYEKLASELLEWIRRTVPWLENRVGEPSMSAMQRKLEDF RDYRRLHKPPRVQEKCQLEINFNTLQTKLRLSHRPAFMPSEGKLVSDIANAW RGLEQVEKGYEDWLLSEIRRLQRLQHLAEKFQQKASLHETWTRGKEDMLTQ RDYESASLQEVRALLRRHEAFESDLAAHQDRVEHIAALAQELNELDYHEAAS VNSRCQAICDQWDNLGTLTQKRRDALERMEKLLETIDQLQLEFARRAAPENN WLDGAIEDLQDVWLVHSVEETQSLLTAHDQFKATLPEADRERGAILGIQGEIQ KICQTYGLRPSSTNPYINLSPQDINTKWDTVRKLVPSRDQTLQEELARQQMNE RLRRQFAAQANAVGPWIQAKVEEVGRLAAGLAGSLEEQMAGLRQQEQNIIN YKSNIDRLEGDHQLLQESLVFDNKHTVYSMEHIRVGWEQLLTSIARTINEVEN QVLTRDAKGLSQEQLNEFRASFNHFDRKRNGMMEPDDFRACLISMYDLGEV EFARIMTMVDPNAAGVVTFQAFIDFMTR
76	Truncation of Rabbit PYGM (Glycogen Phosphorylase, Muscle Form)	IYYLSLEFYMGRTLQNTMVNLALENACDEATYQLGLDMEELEEIEEDAGLGN GGLGRLAACFLDSMATLGLAAYGYGIRYEFGIFNQKICGGWQMEEADDWLR YGNPWEKARPEFTLPVHFYGRVEHTSQGAKWVDTQVVLAMPYDTPVPGYR NNVVNTMRLWSAKAPNDFNLKDFNVGGYIQAVLDRNLAENISRVLYPNDNF FEGKELRLKQEYFVVAATLQDIIRRFKSSKFGCRDPVRTNFDAFPDKVAIQLN DTHPSLAIPELMRVLVDLERLDWDKAWEVTVKTCAYTNHTVLPEALERWPV HLLETLLPRHLQIIYEINQRFLNRVAAAFPGDVDRLRRMSLVEEGAVKRINMA HLCIAGSHAVNGVARIHSEILKKTIFKDFYELEPHKFQNKTNGITPRRWLVLCN PGLAEIIAERIGEEYISDLDQLRKLLSYVDDEAFIRDVAKVKQENKLKFAAYLE REYKVHINPNSLFDVQVKRIHEYKRQLLNCLHVITLYNRIKKEPNKFVVPRTV MIGGKAAPGYHMAKMIIKLITAIGDVVNHDPVVGDRLRVIFLENYRVSLAEK VIPAADLSEQISTAGTEASGTGNMKFMLNGALTIGTMDGANVEMAEEAGEEN FFIFGMRVEDVDRLDQRGYNAQEYYDRIPELRQIIEQLSSGFFSPKQPDLFKDI VNMLMHHDRFKVFADYEEYVKCQERVSALYKNPREWTRMVIRNIATSGKFS SDRTIAQYAREIWGVEPSRQRLPAPDEKIP
77	Truncation of Rabbit ALDOA (Fructoase-Bisphosphate Aldolase A)	SHPALTPEQKKELSDIAHRIVAPGKGILAADESTGSIAKRLQSIGTENTEENRRF YRQLLLTADDRVNPCIGGVILFHETLYQKADDGRPFPQVIKSKGGVVGIKVDK GVVPLAGTNGETTTQGLDGLSERCAQYKKDGADFAKWRCVLKIGEHTPSAL AIMENANVLARYASICQQNGIVPIVEPEILPDGDHDLKRCQYVTEKVLAAVYK ALSDHHIYLEGTLLKPNMVTPGHACTQKYSHEEIAMATVTALRRTVPPAVTG VTFLSGGQSEEEASINLNAINKCPLLKPWALTFSYGRALQASALKAWGGKKE NLKAAQEEYVKRALANSLACQGKYTPSGQAGAAAS
78	Rabbit ALBU (Albumin)	EAHKSEIAHRENDVGEEHFIGLVLITFSQYLQKCPYEEHAKLVKEVTDLAKAC VADESAANCDKSLHDIFGDKICALPSLRDTYGDVADCCEKKEPERNECFLHH KDDKPDLPPFARPEADVLCKAFHDDEKAFFGHYLYEVARRHPYFYAPELLYY AQKYKAILTECCEAADKGACLTPKLDALEGKSLISAAQERLRCASIQKFGDRA YKAWALVRLSQRFPKADFTDISKIVTDLTKVHKECCHGDLLECADDRADLAK YMCEHQETISSHLKECCDKPILEKAHCIYGLHNDETPAGLPAVAEEFVEDKDV CKNYEEAKDLFLGKFLYEYSRRHPDYSVVLLLRLGKAYEATLKKCCATDDP HACYAKVLDEFQPLVDEPKNLVKQNCELYEQLGDYNFQNALLVRYTKKVPQ VSTPTLVEISRSLGKVGSKCCKHPEAERLPCVEDYLSVVLNRLCVLHEKTPVS EKVTKCCSESLVDRRPCFSALGPDETYVPKEFNAETFTFHADICTLPETERKIK

TABLE 2-continued

Rabbit Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		KQTALVELVKHKPHATNDQLKTVVGEFTALLDKCCSAEDKEACFAVEGPKL VESSKATLG

[0041] An animal protein can comprise a cow protein (Table 3). A cow protein can comprise a cow muscle protein. A cow protein can comprise a cow blood protein. A cow protein can comprise a cow albumin protein. A cow protein can comprise a sequence that has at least or about 70%, at least or about 75%, at least or about 80%, at least or about 90%, at least or about 91%, at least or about 92%, at least or about 93%, at least or about 94%, at least or about 95%, at least or about 96%, at least or about 97% at least or about 98% at least or about 99% or about 100% sequence identity to any one of SEQ ID NO: 79-93.

[0042] A food product may comprise a fragment or a truncation of a cow protein. The food product may include a fragment or truncation of any one of SEQ ID Nos: 79-93. A sequence of the fragment or truncation may be identical or substantially identical to a portion of any one of SEQ ID Nos: 79-93. A sequence of the fragment or truncation may be at least about 70%, 80%, 90%, 95%, 98%, or more identical to a portion of a sequence of any one of SEQ ID Nos: 79-93. A length of the fragment or truncation may be at least about 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of a length of any one of SEQ ID Nos: 79-93. A length of the fragment or truncation may be less than or equal to about 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, or less than a length of any one of SEQ ID Nos: 79-93.

TABLE 3

Cow Proteins		
SEQ ID NO	Description	Amino Acid Sequence
79	Cow KCRM (Creatine kinase M-type)	MPFGNTHNKHKLNFKAEEEYPDLSKHNNHMAKALTLEIYKKLRDKETPSG FTLDDVIQTGVDNPGHPFIMTVGCVAGDEESYTVFKDLFDPIIQDRHGGFKP TDKHKTDLNHENLKGGDDLDPNYVLSSRVRTGRSIKGYALPPHCSRGERR AVEKLSVEALNSLTGEFKGKYYPLKSMTEQEQQQLIDDHFLFDKPVSPLLL ASGMARDWPDARGIWHNDNKSFLVWVNEEDHLRVISMEKGGNMKEVFR RFCVGLQKIEEIFKKAGHPFMWNEHLGYVLTCPSNLGTGLRGGVHVKLAH LSKHPKFEEILTRLRLQKRGTGGVDTAAVGSVFDVSNADRLGSSEVEQVQL VVDGVKLMVEMEKKLEKGQSIDDMIPAQK
80	Cow ACTN3 (Alpha-actinin-3)	MMMVLQPEGLGTGEGPFAGGRGGGEYMEQEEDWDRDLLLDPAWEKQQR KTFTAWCNSHLRKAGTQIENIEEDFRNGLKLMLLLEVISGERLPRPDKGKM RFHKIANVNKALDFIASKGVKLVSIGAEEIVDGNLKMTLGMIWTIILRFAIQ DISVEETSAKEGLLLWCQRKTAPYRNVNVQNFHTSWKDGLALCALIHRHR PDLIDYAKLRKDDPIGNLNTAFEVAEKYLDIPKMLDAEDIVNTPKPDEKAIM TYVSCFYHAFAGAEQAETAANRICKVLAVNQENEKLMEEYEKLASELLEW IRRTVPWLENRVGEPSMSAMQRKLEDFRDYRRLHKPPRVQEKCQLEINFNT LQTKLRLSHRPAFMPSEGKLVSDIANAWRGLEQAEKGYEDWLLSEIRRLQR LQHLAEKFQQKASLHEAWTRGKEDMLSQRDYETASLQEVRALLRRHEAFE SDLAAHQDRVEHIAALAQELNELDYHEAASVNSRCQAICDQWDNLGTLTQ KRRDALERMEKLLETIDQLQLEFARRAAPFNNWLDGAVEDLQDVWLVHS VEETQSLVTAHDQFKATLPEADRERGAILGIQGEIQKICQTYGLRPSSTNPYI TLTPQDINTKWDTVRKLVPSRDQMLQEELTRQQVNERLRRQFAAQANAIG PWIQGKVEEVGRLAAGMAGSLEEQMAGLRQQEQNIINYKSNIDRLEGDHQ LLQESLVFDNKHTVYSMEHIRVGWEQLLTSIARTINEVENQVLTRDAKGLS QEQLNEFRASFNHFDRKRNGMMEPDDFRACLISMGYDLGEVEFARIMTMV DPNAAGVVTFQAFIDFMTRETAETDTAEQVVASFKILAGDKNYITAEELRR ELPAEQAEYCIRRMAPYKGAGAPAGALDYVAFSSALYGESDL
81	Cow PYGM (Glycogen phosphorylase, muscle form)	MSRPLTDQEKRKQISVRGLAGVENVTELKKNFNRHLHFTLVKDRNVATPR DYYFALAYTVRDHLVGRWIRTQQHYYEKDPKRIYYLSLEFYIGRTLQNTM VNLALENACDEATYQLGLDMEELEEIEEDAGLGNGGLGRLAACFLDSMAT LGLAAYGYGIRYEFGIFNQKISGGWQMEEADDWLRYGNPWEKARPEFTLP VHFYGRVEHTSQGAKWVDTQVVLAMPYDTPVPGYRNNVVNTMRLWSAK APNDFNLKDFNVGGYIQAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYF VVAATLQDIIRRFKSSKFGCLDPVRTNFDAFPDKVAIQLNDTHPSLAIPELM RILVDQERLEWEKAWEVTVKTCAYTNHTVLPEALERWPVHLIETLLPRHLQ IIYEINQRFLNRVAAAFPGDVDRLRRMSLVEEGAVKRINMAHLCIAGSHAV NGVARIHSEILKKTIFKDFYELEPHKFQNKTNGITPRRWLVMCNPGLAEIIAE RIGEEYIADLDQLRKLLSYVDDESFIRDVAKVKQENKLKFSAYLEKEYKVHI NPNSLFDIQVKRIHEYKRQLLNCLHVITLYNRIKKEPNKFFVPRTVMIGGKA APGYHMAKMIIKLITAIGDVVNHDPVVGDRLRVIFLENYRVSLAEKVIPAA DLSEQISTAGTEASGTGNMKFMLNGALTIGTMDGANVEMAEEAGEENFFIF GMRVEDVERLDQKGYNAQEYYDRIPELRHVIDQLSSGFFSPKQPDLFKDIV

TABLE 3-continued

Cow Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		NMLMHHDRFKVFADYEEYIKCQERVSALYKNPREWTRMVIRNIATSGKFS SDRTIAQYAREIWGVEPTRQRMPAPDEKI
82	Cow A6QLL8 (Fructose-bisphosphate aldolase)	MPHQYPALTPEQKKELCDIAHRIVAPGKGILAADESTGSIAKRLQSIGTENTE ENRRFYRQLLLTADDRVNPCIGGVILFHETLYQKADDGRPFPQVIKAKGGV VGIKVDKGVVPLAGTNGETTTQGLDGLSERCAQYKKDGADFAKWRCVLKI GEHTPSSLAIMENANVLARYASICQQNGIVPIVEPEILPDGDHDLKRCQYVT EKVLAAVYKALSDHHIYLEGTLLKPNMVTPGHACTQKYSHEEIAMATVTA LRRTVPPAVPGITFLSGGQSEEEASINLNAINKCPLLKPWALTFSYGRALQAS ALKAWGGKKENLKAAQEEYVKRALANSLACQGKYTPSGKAGAAASESLFI SNHAY
83	Cow MYL1 (Myosin light chain 1/3, skeletal muscle isoform)	MAPKKDVKKPAAAAAPAPAPAPAPAPAPAPPKEEKIDLSAIKIEFSKQQQDE FKEAFLLFDRTGECKITLSQVGDVLRALGTNPTNAEVKKVLGNPSNEEMNA KKIEFEQFLPMLQAISNNKDQGTYEDFVEGLRVFDKEGNGTVMGAELRHV LATLGEKMKEEEVEALMAGQEDSNGCINYEAFVKHIMSN
84	Cow G3P (Glyceraldehyde-3-phosphate dehydrogenase)	MVKVGVNGFGRIGRLVTRAAFNSGKVDIVAINDPFIDLHYMVYMFQYDST HGKFNGTVKAENGKLVINGKAITIFQERDPANIKWGDAGAEYVVESTGVFT TMEKAGAHLKGGAKRVIISAPSADAPMFVMGVNHEKYNNTLKIVSNASCT TNCLAPLAKVIHDHFGIVEGLMTTVHAITATQKTVDGPSGKLWRDGRGAA QNIIPASTGAAKAVGKVIPELNGKLTGMAFRVPTPNVSVVDLTCRLEKPAK YDEIKKVVKQASEGPLKGILGYTEDQVVSCDFNSDTHSSTFDAGAGIALND HFVKLISWYDNEFGYSNRVVDLMVHMASKE
85	Cow TPM1 (Tropomyosin alpha-1 chain)	MDAIKKKMQMLKLDKENALDRAEQAEADKKAAEDRSKQLEDELVSLQKK LKATEDELDKYSEALKDAQEKLELAEKKATDAEADVASLNRRIQLVEEEL DRAQERLATALQKLEEAEKAADESERGMKVIESRAQKDEEKMEIQEIQLKE AKHIAEDADRKYEEVARKLVIIESDLERAEERAELSEGKCAELEEELKTVTN NLKSLEAQAEKYSQKEDKYEEEIKVLSDKLKEAETRAEFAERSVTKLEKSID DLEDELYAQKLKYKAISEELDHALNDMTSI
86	Cow E1BNV1 (Myosin binding protein C2)	MPEAKPAAKKAPAGKDAAAKPAPKEATPQEAPAAPPTEAPPEDQSPTAEEP TGVFLKKPDSVSVETGKDTVIVAKLNGKELPAKPAVKWFKGKWLELGSKS GARFSFKESHDAASNVYTIELHITKVVLGDRGDYRIEVKAKDFCDSCAFNID VEAPRQDSAGQSLESFKRSGEAKSDTAGELDFSGLLKKRQVVEEEKKKKK DDDDLGIPPEIWELLKGAKKSEYERIAFQYGITDLRGMLKRLKKAKVEVKK SAAFTKKLDPAYQVDRGNKIKLVVEISDPDLPLKWFKNGQEIKPSSKYVFE NVGKKRILTINKCTLADDAAYEVVVKDEKCFTELFVKEPPVLIVTPLEDQQ VFVGDRVEMAVEVSEEGAQVMWLKDGVELTREDSFKARYRLKKDGKRHI LIYSEVTMEDKGHYQVMTNGGQCEAELIVEEKQLEVLQDIADLTVKASEQ AVFKCEVSDEKVTGKWYKNGVEVRPSKRITISHTGRFHKLVIDDVRPEDEG DYTFVPDGYALSLSAKLKFLEIKVEYVPKQEPPKIHLDCSGQTSENAIVVVA GNKLRMDVSITGEPRPVATWMKEDEVFTGTEGRVRIEQRGDISSFVIESAER GDEGRYTIKVTNPVGEDVASILLKVVDVPDPPEAVRVTSVGEDWAILVWEP PKYDGGRPVTGYLLERKKKGSQRWMKLNFEVFTETTYESTNMIEGILYEM RVFAVNVIGVSQPSMNTKPFMPIAPTSEPQHLTVEDVTDTTTTLKWRPPDKI GAGGIDGYLIEYCVEGSDEWIPANTEPVERCGFTVKNLPTGAKITFRVVGV NIAGRSQPATLAQPVTIREIVEQPKIRLPRHLRQTYVRKVGEHLNLVIPFQGK PRPQVVWTKGGAPIDTSHVHVRTSDFDTVFFVRQAARSDSGEYELTVQIEN MKDTATVCIRVVEKAGPPQNVMVKEVWGTNALVEWQPPKDDGNSEITGY FVQKADKKTMEWFTVYERNRHTSCTVSDLIMGNEYYFRVYSENVCGLSDL PGVSKNTARIVKTGMTLKLPEYKEHDFRTPPKFLTPLPDRVVVAGYAAALN CAVRGHPKPKVVWMKNKMEIREDPKFLMTNQQGVLTLNIRRPSPFDSGTY SCRAVNELGEALAECKLEVRVPQ
87	Cow E1BF23 (Myomesin 2)	MSLVAVPFYQKRHKHFDQSYRNIQTRYLLDEYSAKKKRASAQSSLQRSVT RKSSSQRESSRAALGESTLGETTLGETSLGDTTLGRTTCRLCAKRMSSALEE EAQERKRRYQSMVAAYGEAKRERYLSELAQLEEDVHVARTHARDQLDRL ALQHAGRHGLHHAVDDRLAWERHSFEERISRAPEILVRLRSHTVWERMSV KLCFTVQGFPTPVVQWYKDGSLICQAGEPGKYRIDSKYGVHTLEINRADFD DTATYSAVATNVHGQVSTNAAVVVRRFRGDEEPFHSVGLPIGLPLSSVIPYT HFDVQFIEKFGVTFRREGETLTLKCTLLVTPDLKRVQPRAEWYRDDVLLKE SKWTKMFFGEGQASLSFSHLNKDDEGLYTLRIVSRGGITDHSAFLFVRDAD PLVTGAPGAPMDVHCHDANRDYVIVTWKPPNATKESPVIGYFIDRCEVGTD NWIQCNDAPVKICKYPVTGLFEGRSYIFRVRAVNSAGISRPSRVSEAVAALD PVDLRRLQAVHLEGEKDIVVYQEELEGEVQIPGPPTNVHASETSRTYVVLS WDPPEPRGKEPLMYFIEKSMVGSGSWQRVNAQTAVRSPRYAVFDLAEGKP YVFRVLSANKHGLSDPSEITAPIQAQDTIVVPSAPGRVLASRNTRTSVVVQW DRPKHEEDLLGYYVDCSVAGSNVWEPCNHKPIGYNRFVVHGLTTGEQYIF RVKAVNAVGTSENSQESDVIKVQAALTVPSHPYGITLLNCDGHSMILGWK VPKFSGGSPILGYYVDKREAHHKNWHEVNSSPLKERILTVEGLTEGSLYEF

TABLE 3-continued

Cow Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		KIAAANMAGIGQPSDPSELFKCEAWTMPEPGPAYDLTFCEVRDTSLVLLWK PPVYSGSSPVSGYFVDYKEEDSGEWLTVHETATPHRYLKVCDLHQGKTYV FRVRAVNASGVGRPSDTSEPVLVEARPGTKEVSAGVDEEGNVYLSFDCPEV TDASHFTWCKSYEDIADDDRFKVETAGDHSKLYFKNLDKEDLGTYSVSVS DTDGVSSSFVLDEEELERLKALSNEIKNPTIPLKSELAYEIFDKGQVRFWLQ AEHLSPDSSYRFIINDREVADSETHRIKCDKSTGIIEMVMDRFTIDNEGTYTV QIQDGKAKSQSSLVLIGDAFKAVLKEAEFQRKEFLRKQGPHFAEYLHWDVT EECEVRLVCKVANTKKETVFKWLKDDVLYETEKMPDLEKGVCELLIPKLS KKDHGEYKATLKDDRGQDVSTLEIAGKVYEDMILAMSRVCGASASPLKIL CTPEGIRLQCFMKYFTEEMKVNWYHKDAKITSSEHMRIGGNEQMAWLQIC EPTEKDKGKYTFEILDGKKNHQRSLDLSGQAFDEAFAEFQQLKAAAFAEKN RGKVIGGLPDVVTIMEGKTLNLTCTVFGNPDPEVVWFKNDKDIELSDHFSV KVEQGKYVSMTIKGVTSEDSGKYSIHVKNKYGGEKIDVTVSVYKHGEEIPD VVLPQQAKPKLLPAAAPSPAH
88	Cow ENOB (Beta-enolase)	MAMQKIFAREILDSRGNPTVEVDLHTAKGRFRAAVPSGASTGIYEALELRD GDKSRYLGKGVLKAVEHINKTLGPALLEKKLSVVDQEKVDKFMIELDGTE NKSKFGANAILGVSLAVCKAGAAEKGVPLYRHIADLAGNPELILPVPAFNVI NGGSHAGNKLAMQEFMILPVGASSFREAMRIGAEVYHHLKGVIKAKYGKD ATNVGDEGGFAPNILENNEALELLKTAIQAAGYPDKVVIGMDVAASEFYRN GKYDLDFKSPDDPARHISGEKLGELYKNFIKNYPVVSIEDPFDQDDWATWT SFLSGVNIQIVGDDLTVTNPKRIAQAVEKKACNCLLLKVNQIGSVTESIQAC KLAQSNGWGVMVSHRSGETEDTFIADLVVGLCTGQIKTGAPCRSERLAKY NQLMRIEEALGDKAVFAGRKFRNPKAK
89	Cow MLRS (Myosin regulatory light chain 2)	MAPKKAKRRAAAEGGSSSVFSMFDQTQIQEFKEAFTVIDQNRDGIIDKEDL RDTFAAMGRLNVKNEELDAMMKEASGPINFTVFLNMFGEKLKGADPEDVI TGAFKVLDPEGKGTIKKKFLEELLTTQCDRFSQEEIKNMWAAFPPDVGGNV DYKNICYVITHGDAKDQE
90	Cow MYH1 (Myosin-1)	MSSDQEMAVFGEAAPYLRKSEKERIEAQNKPFDAKTSVFVADPKESFVKAT VQSREGGKVTAKTEAGATVTVKEDQVFPMNPPKFDKIEDMAMMTHLHEP AVLYNLKERYAAWMIYTYSGLFCVTVNPYKWLPVYNAEVVTAYRGKKRQ EAPPHIFSISDNAYQFMLTDRENQSILITGESGAGKTVNTKRVIQYFATIAVT GEKKKEEPTSGKMQGTLEDQIISANPLLEAFGNAKTVRNDNSSRFGKFIRIH FGTTGKLASADIETYLLEKSRVTFQLKAERSYHIFYQIMSNKKPELIEMLLIT TNPYDYAYVSQGEITVPSIDDQEELMATDSAIEILGFTSDERVSIYKLTGAV MHYGNLKFKQKQREEQAEPDGTEVADKAAYLQGLNSADLLKALCYPRVK VGNEFVTKGQTVEQVYNAVGALAKAVYDKMFLWMVARINQQLDTKQPR QYFIGVLDIAGFEIFDFNSLEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGI EWEFIDFGMDLAACIELIEKPMGIFSILEEECMFPKATDMSFKNKLYEQHLG KSNNFQKPKPAKGKAEAHFSLIHYAGTVDYNITGWLDKNKDPLNETVVGL YQKSSVKTLALLFSGPASGEAEGGPKKGGKKKGSSFQTVSALFRENLNKLM TNLRSTHPHFVRCIIPNETKTPGAMEHELVLHQLRCNGVLEGIRICRKGFPSR ILYADFKQRYKVLNASAIPEGQFIDSKKASEKLLASIDVDHTQYKFGHTKVF FKAGLLGLLEEMRDEKLAQLITRTQARCRGFLARVEYQKMVERRESIFCIQ YNVRAFMNVKHWPWMKLYFKIKPLLKSAETEKEMANMKEEFEKTKEELA KSEAKRKELEEKMVTLTQEKNDLQLQVQSEADALADAEERCDQLIKTKIQL EAKIKEVTERAEDEEEINAELTAKKRKLEDECSELKKDIDDLELTLAKVEKE KHATENKVKNLTEEMAGLDETIAKLTKEKKALQEAHQQTLDDLQAEEDK VNTLTKAKTKLEQQVDDLEGSLEQEKKLRMDLERAKRKLEGDLKLAQEST MDIENDKQQLDEKLKKKEFEMSNLQSKIEDEQALAMQLQKKIKELQARIEE LEEEIEAERASRAKAEKQRSDLSRELEEISERLEEAGGATSAQIEMNKKREA EFQKMRRDLEEATLQHEATAAALRKKHADSVAELGEQIDNLQRVKQKLEK EKSEMKMEIDDLASNMETVSKAKGNLEKMCRALEDQLSELKTKEDEQQRL INDLTTQRARLQTESGEFSRQLDEKDALVSQLSRGKQAFTQQIEELKRQLEE EIKAKSALAHALQSARHDCDLLREQYEEEQEGKAELQRAMSKANSEVAQ WRTKYETDAIQRTEELEEAKKKLAQRLQDAEEHVEAVNAKCASLEKTKQR LQNEVEDLMIDVERTNAACAALDKKQRNFDKILSEWKQKYEETHAELEAS QKESRSLSTELFKIKNAYEESLDQLETLKRENKNLQQEISDLTEQIAEGGKRI HELEKVKKQVEQEKSEIQAALEEAEASLEHEEGKILRIQLELNQVKSEIDRKI AEKDEEIDQLKRNHIRIVESMQSTLDAEIRSRNDAIRLKKKMEGDLNEMEIQ LNHANRMAAEALKNYRSTQAILKDTQIHLDDALRGQEDLKEQLAMVERR ANLLQAEIEELRATLEQTERSRKIAEQELLDASERVQLLHTQNTSLINTKKK LETDITQIQGEMEDIIQEARNAEEKAKKAITDAAMMAEELKKEQDTSAHLE RMKKNLEQTVKDLQHRLDEAEQLALKGGKKQIQKLEARVRELEGEVESEQ KRNVEAVKGLRKHERRVKELTYQTEEDRKNILRLQDLVDKLQAKVKSYKR QAEEAEEQSNVNLSKFRKLQHELEEAEERADIAESQVNKLRVKSREVHTKII SEE

TABLE 3-continued

Cow Proteins		
SEQ ID NO	Description	Amino Acid Sequence
91	Cow A0A3Q1M707 (Nebulin)	MPNFQTRTRFYETTTTRTSDYEQSETSRPALAQPVPEKPVERKRVIRKKVDP SKFMTPYIAHSQKMQNLFSTNKYKENYEKAKGKPYAITTDTPELRRIKKVQ DQLSEVKYRVDGDVAKTICHVDEKAKDIEHAKKVSQQVSKVLYKQNWED TKDKYLLPPDAPELVQAIKNTAMFSKKLYTEDWEADKTMFYPYNDSPELR RVAQAQKALSDIAYKKGLAEQQTQFTSLPDPPDIEFAKKVTNQVSKQKYKE DYENKVKGKWSETPCFEIATARMNSNNISTKKYQEDFEHMKDQIYFMQTE TPEYKVNKQAGVAASKVKYKQDYEKNKGKADYNVLPASENPLLRQLKAA GDALSDKLYKENYEKTKAKSINYCETPKFKLDTVLHNFSSDTKYKDSYLKD ILGHYVGSYEDPYHTHCMRVSAQNSDKNYKAEYEEDRGKGFFPQTITQEY EAIKKLDQCKDHTYKVHPDKTKFTQVTDSPVLMQAQVNSKQLSDLNYKA KHENEKFKCHIPPDTPAFIQHKLNAYNLSDNIYKHDWEKTKAKKFDIKVDA IPLLAAKANTKIASDVMYKKDYEKSKGKMIGALSINDDPKMLHSLKTAKN QSDRLYKENYEKTKAKSMNYCETPKYQLDTLLKNFSEAKYKDSYVQNVL GHYIGSFEDPYQVHCLKISAQNSDKNYKAEYEEDKGKCYFPQTITQEYEAIK KLDQCKDHTYKVHPDKTKFTAVTDTPVLLQAQLNTKQLSDLNYKAKHEG EKFKCHIPADAPQFIQHRVNAYNLSDNIYKHDWEKTKAKKFDIKVDAIPLL AAKANTKIASDVMYKKDYEKSKGKMIGALSINDDPKMLHSLKTAKNQSD HEYRKDYEKSKTIYTAPLDMLPLTHAKKSQAIASDVDYKHLLHNYSYPPDS VNVDLAKKAYALQSDVEYKADYNSWMKGCGWMPFGSLEMEKAKRASDI LNEKKYRQHPDTLKFTSIEDAPIIVQSKINQAQRSDVAYKAKGEEVIHKYSL PADLPQFIQAKVNAYNISENLYKADLKDLSKKGYDLRIDAIPIKAAKAARQ AASDVQYKKDYEKAKGKMVGFQSLQDDPKLVHYMNVAKIQSDREYKKA YEKTKTRHNTPHDMVNIVAAKKAQDVASNVNYKHSLHHYTYLPDAMDLE LSKNMMHIQSDNVYKEDYNNWMKGIGWIPIGSLEVEKVKKAGDALNEKK YRQHPDTLKFTSIVDSPVMVQAKQNTQQVSDILYKAKGEDVKHKYTMSPD LPQFLQAKCNAYNLSDVCYKRDWHDLIAKGTNVLGDAIPITAAKASRNIAS DYKYKEAYEKAKGKQVGFRSLQDDPKLVHYMNVAKLQSDREYKKNYEN TKTSYHTPGDMVSITAAKMAQDVATNVNYKQPIHHYTYLPDALSLEHIRN VNQIQSDNVYKDEYNHFFKGMGWIPIGSLEVEKVKKAGDALNEKKYRQHP DTIKFTSVPDSMGMVLAQHNTKQLSDLNYKVEGEKVKHKYTMDPDVPQFI QAKVNAYNMSDSHYKADWKKTLAKGYDLRPDAIPIVAAKSSRNIASDFKY KEAYEKTKGKQIGFRSLQDDPKLVHFMNVAKMQSDREYKKGYEASKTKY HTPLDMLSVTAAKKSQEVATNTNYKQPFHHYTLLPDALNVEHSRNAMQIQ SDNLYKSDFTNWMKGIGWLPLESLEVEKAKKAGEILSEKKYRQHPEKLKFT YAMDTMEQALNKSNKLIMDKRLYTEKWNKDKTTIHVMPDTPDILLSRVNQ ITMSDKLYKAGWEEEKKKGYDLRPDAISIKAARASRDIASDYKYKQAYEQ AKGKQIGFRSLEDDPKLVHFMQVAKMQSDREYKKAYEKSKTSFQTPVDML SVVAAKKSQEVATNANYRNVIHTYNMLPDAMSLELAKNMMQIQSDNQYK ADYADFMKGIGWLPLGSLEAEKNKKAMEIISEKKYRQHPDTLKYSTLMDS MNMVLAKNNAKIMNEHLYKQAWEADKTKVHIMPDIPQIILAKANAINISD KLYKLSLEEAKKKGYDLRTDAIPIKAAKASRDIASDYKYKHSYEKERGKM VGFRSLEDDPKLVHSMQVAKMQSDREYKKNYEKTKTSYHTPADMLSVTA AKDAQANITNTNYKHLIHKYILLPDAMNIQLSKNMNRIQSDNEYKQDYNE WYKGLGWSPAGSLEVEKAKKATEYASDQKYRQHPSNFQFTKLNDSMDMV LAKQNAHTMNKYLYTVDWNKDKTKIHVMPDTPDILQAKQNQTMYSQKL YKLGWEEALKKGYDLPVDAISVQAAKASRDIVSDFKYKQGYRKQIGHHIG FRSLQDDPKLVLSMNVAKMQSEREYKKDFEKWKTKFSSPVDMLGVVLAK KCQALVSDVDYKNYLHGWTCLPDQNDVIHAKKAYDLQSENLYKSDLEWL KGIGWSPLGSLEAEKNKRASEIISEKKYRQPPDRNKFTSIPDAMDIVLAKTN AKNRSDILYREAWDKDKTQIHIMPDTPDIILAKANLINTSDKFYRMGYEELR KKGYDLPVDAIPIKAAKASREIASEYKYKEGFRMQLGHHIGARNIKDDPKM MWSMHVAKIQSDREYKKDFEKWKTKFSSPVDMLGVVLAKKCQTLVSDID YKNYLHQWTCLPDQSDVIHARRAYDLQSDNLYKSDLQWLRGIGWLPSGSL EDEKNKRASQILSDHVYRQHPDQFKFSSLMDSIPMVLAKNNAITMNHRLYT EAWDKDKTTVHIMPDTPEVLLAKQNQINYSEKLYKLGLDEAKRKGYDMRI DAIPIRAAKASRDIASEFKYKEGYRRQLGHHIGARAIHDDPKMMWSMHVA KIQSDREYKKDFEKWKTKFSSPVDMLGVVLAKKCQTLVSDIDYKNYLHQ WTCLPDQSDVIHARQAYDLQSDNVYKSDLQWLRGIGWVPIGSLDVEKSKR ASEILSDKLYRQPPDKFKFTSVTDSLEQVLAKNNAINMNKRLYTEAWDKD KTQVHIMPDTPEITLARTNKVNYSENLYKLAHEEAKKKGYDLRSDAIPIVA AKASRDIISDYKYKDGYRKQLGHHIGARDIKDDPKMMWSMHVAKIQSDRE YKKDFEKWKTKFSSPVDMLGVVLAKKCQTLVSDIDYKNYLHQWTCLPDQ SDVIHARQAYDLQSDNVYKSDLQWLRGIGWVPIGSVDVVKCKRAAEILSD NLYRQPPDTLKFTSVPDSLEQVLAKNNAINMNKRLYTEAWDKDKTQIHIMP DTPEIMLARQNKLNYSETLYKLANEEAKKKGYDLRSDAIPIVAAKASRDIIS DYKYKDGYRKQLGHHIGARDIKDDPKMMWSMHVAKIQSDREYKKDFEK WKTKFSSPVDMLGVVLAKKCQTLVSDIDYKNYLHQWTCLPDQSDVIHARR AYDLQSDNIYKSDLQWLRGIGWVPIGSVDVVKCKRAAEILSDNLYRQRPDT LKFTSVPDSLEQVLAKNNAINMNKRLYTEAWDKDKTQIHMMPDTPEITLA RQNKINYSENLYRQAMEEAKKEGYDLRSDAIPIVAAKASREIASDYKYKEA YRKQLGHHIGARAIHDDPKMMWSVHVAKMQSDREYKKDFEKYKTRFSSP

TABLE 3-continued

Cow Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		VDMLGIVLAKKCQTLVSDVDYKHPLHEWTCLPDQNDIIHARKAYDLQSDN LYKSDLEWLKGIGWVPIGSVEVLKAKRAGEILSDNIYRQRPDTLKFTSVTDS PEQVLAKNNAINMNKRLYTEAWDNDKKTIHVMPDTPEIMLAKLNRINYSD KLYKLALEESRKEGYDLRLDAIPIQAAKASREIASDYKYKEGYRKQLGHHI GARNIKDDPKMMWSIHAGKLQSDLEYKKDFEKWKTKFSSPVDMMGLVQA KKCQILVSDIDYKNLLHEWTCLPDQNDIIQARKAYDLQSDAIYKADLEWLR GIGWVPIDSVGVEHAKRAGEILSERKYRQPAVQLKFTSITDTPEIVLAKNNA LNVSKHLYTEAWDADKTSIHVMPDTPEILLAKSNSANISHKLYTKGWDESK MKDYDLRADAISIKSAKASRDIASDYKYKEAYEKHKGHHIGARSVEDDPRI MCAMNAGRIQSEREYKKEFQKWKTKFSSPVDMLGILLAKKCQTLVSDIDY RNYLHHWTCLPDQNDIIQARKAYDLQSDNVYKSDLEWLRGIGWMPEGSVE MNRVKNAQDLINERRYRTRPEALTFTSIVDTPEVVLAKANSLQMSEKLYQE AWNKDKTNISIPSDTPLILQAQINAMQISNKLYQKDWDEAKQKGYDLRAD AIEIKHAKASREIASEYKYKEGYRKQLGHHVGFRTLQDDPKLVWSIHAAKI QSDREYRKAYEKSKGIHNTPLDMMSIVQAKKCQVLVSDIDYRNYLHQWIC LPDQNDVIQAKKAYELQSDAIYKADLEWLRGIGWMPEGSPEVLRVKNAQH IFRDSVYRTPVVKLKYTSIVDTPEVVLAKSNAENISIPKYREVWDKDKTSIHI MPDTPEINLARTNALNVSNKLYREGWDEMKMSCDVRLDAIPIQAAKASREI ASDYKYKLDHEKQKGHYVGTRTARDDNKIRWALIAGKIQNEREYRLHWA KWKSKFQSPPDMLSIEHSKNSQTLVSDIDYRHYLHQWTCMPDQNDVIQAR KAYDLQSDAIYKADLEWLRGIGWMPADSVSVNHAKHMADIFNEKKYRTKI ETLSFTPVDDRVDYVTAKHSGEILNDIKYRKDWNDTKSKYTLTETPQLHTA QEAARILDQYLYKESWEKQKATGYILPPDAVPFVHAHHSGDVQSELKYKA EHVKQKGHYVGVPTMRDDPKLVWFEHAGQIQNDRLYKENYHKTKAKIHI PPDMVSVLAAKEGQALASDIDYRNYLHQWICHPDQNDVIQARKAYDLQSD NIYKADLEWLRGIGWIPLDSVDHVRVTRNQEMMNQIKYKKDALANYPNFT SVVDPPEIVLAKINSVNQSDVKYKETFNKLIKGKYIFSPDTPYITHSKDMEKL YSTILYKRAWDGTKAYGYTLDERYIPIVGAKHADFVNSELKYKETYEKLK GHYLAGKEISEFPNVVHCLDFQKMRSLLNYRRHYEDTKANVHIPQDMMNH VLAKRCQYILSDLEYRHYFHQWTSLPEEPNVVRVRHAQEILSDNVYKDDL NWLKGIGCYVWDTPQILHAKKSYDLQSQILYTAAGKENLKNYNLVTDTPL YVTALQSGINASEVKYKENYHQTKDKYTTVLETVDYDRIKNLKDLFSSNLY KEAWDKVKATSYILPPNTVSLTHAKNQKYMASHIKYREEYEKFKALYTLP RSVEDDPNTARCLRVGKFNIDRLYRSVYEKNKMKIHIVPDMVEMVTAKDS QKKVSEIDYRLHLHEWICHPDLQVNSHVRKVTDQISDIVYKDDLTWLKGIG CYVWDTPEILHAKHAYDLRNDIKYKAHVLKTRNNYKLVTDTPVYVQAVK SGKQLSDAVYHYDYVHSIRGRVAPTTKTVDLDRALHAYKLQSENLYRKAG LHALPTGYRLPVDTPHFKHTKDTRYMSSYFKYKEVYEHMKAYGYTLGPN DVPFVNVRRVNNITSERLYRQLYHKLKDKIHTTPDTPEIRQVKKTQEAVSEL IYKSDFFKMQGHMISLPYTPQVLHCRYVGDITSDIKYKEDLQVLRGMGCFL YDTPDMVRSRHLRKLWSHYLYTDKARKMRDKYKVVLDTPEYRKVQELKT HLSELVYRAAGRKQKSIFTSVPDTPDLTRAKRGQKLQSQYLYVELATKERP HHHAGNQTTALKHARHVKDMVSENKYKIQYEKMKDKYTPVPDTPILIRAK RAYWNASDLRYKETFQKTKGKYHTVKDALDIVYHRTVTDHISKIKYKENY MSQLGIWRSIPDRPEHFHHRAVTDAVSDVKYKQDLTWLKGIGCYAYDTPD FTLAEKNKTLYSKYKYKEVFERTKSNFKYVADCPINRHFKFATQLMNEKK YRADYEQRKDKYHLVVDEPRHLLAKIAGDQISQIKYRKNYEETKDKFTSIV DTPEHLRTTKVNKQISDILYKLEYNKAKPRGYTTIHDTPMLLHVRKVKDEV SDLKYKEVYQRTKSNCTIEPDAVHIKAAKDAYKVNTNLDYKKKYEATKAY WKWTPDRPDFIQAAKSTLQQSDFEYKLDREYLKGCKLSVTDDKDMVLAL KNSIIESDLKYKEKHVKERGSCHAVPDTPQILLAKTVSSLVSENKYKSYVKK HLAQGSYTTLPETRDTIHVKEVTKNVSDTNYKKKFVKEKGKSNYSIMLEPP DVKHAMDVAKKQSNVAYKKDAKENLHYTTVADRPDIKKATQAAKQASE VEYRAKHRKEGSHGLSMLGRPDIEMAKKAAKLSSQVQYRENFNKEKGKTP KYNPKDSQLYKVMKDANTLASEVKYKADLKKLHKPVTDMKESLIMNHVL NTSHLASSYQYKKNYEKSKGHYHTIPDNLEQLHLKEATELQSIVKYKEKYE KERGKPMLDFETPTYITAKESQQMQSGKEYRKDYEESIKGRNLTGLEVTPA LLHVKYATKIASEKEYRKDLEESIRGKGLSEMEDTPDMLRAKNATQILNEK EYKRDLELEVKGRGLNAMANETPDFLRARNATDIASQIKYKQSAEMEKAN FTSVVDTPEIIHAQQVKNLSSQKKYKEDAEKCMSYYETVLDTPEMQRVREN QKNFSLLQYQYDLKNSKGKITVVQDTPEILRVKENQKNFSSVLYKEDVSPG TAIGKTPEMMRVKQTQDHISSVKYKEVIGQGTPIPDLPEVKRVKQTQKHISS VMYKENLGTGIPTPVTPEIERVKRNQENFSSNLSLKHEVTSKSIILTKYNSLK LPPHTPLPPVLYKENLGKGTPTPITPEMERVKRNQENFSSVLYKENMGKGTP LPVTPEMERIKHNQENISSVLYKENMRKATPTPVTPEMERAKRNQENISSVL YSDSFRKQIQGKAAYVLDTPEMRRVRETQRHISTVKYHEDFEKHKGCFTPV VTDPITERVKKNTQDFSDICYRGIQRKVVEMEQKRNDQDQETITGLRVWRT NPGSVFDYDPAEDNIQSRSLHMINAQAQRRSREQSRSASGLSISGGEEKSEH SEAAHLSTYSDGGVFFSAASTGYKHARTTELPQQRSSSVATQQTTVSSIPSH PSTAGKIFRAMYDYMAADADEVSFKDGDAIVNVQAIDEGWMYGTVQRTG RTGMLPANYVEAI

TABLE 3-continued

Cow Proteins		
SEQ ID NO	Description	Amino Acid Sequence
92	Cow E1BP87 (Myosin heavy chain 4)	MSSDQEMAIFGEAAPYLRKSEKERIEAQNKPFDAKTSVFVVDPKESFVKAT VQSREGGKVTAKTEAGATVTVKEDQVFPMNPPKYDKIEDMAMMTHLHEP AVLYNLKERYAAWMIYTYSGLFCVTVNPYKWLPVYNAEVVTAYRGKKRQ EAPPHIFSISDNAYQFMLTGELDVSIGESGAGKTVNTKRVIQYFATIAVTGEK KKEEPTSGKMQGTLEDQIISANPLLEAFGNAKTVRNDNSSRFGKFIRIHFGTT GKLASADIETYLLEKSRVTFQLKAERSYHIFYQIMSNKKPELIEMLLITTNPY DYAYVSQGEITVPSIDDQEELIATDSAIDILGFTSDERVSIYKLTGAVMHYGN LKFKQKQREEQAEPDGTEVADKAAYLQSLNSADLLKALCYPRVKVGNEFV TKGQTVQQVYNAVGALAKAVYEKMFLWMVTRINQQLDTKQPRQYFIGVL DIAGFEIFDFNSLEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEWEFIDF GMDLAACIELIEKPMGIFSILEEECMFPKATDTSFKNKLYEQHLGKSNNFQK PKPAKGKAEAHFSLVHYAGTVDYNIAGWLDKNKDPLNETVVGLYQKSAM KTLAFLFAGGPSAEEGGGGKKGGKKKGSSFQTVSALFRENLNKLMTNLRS THPHFVRCIIPNETKTPGAMEHELVLHQLRCNGVLEGIRICRKGFPSRILYAD FKQRYKVLNASAIPEGQFIDSKKASEKLLGSIDIDHTQYKFGHTKVFFKAGL LGILEEMRDEKLAQLITRTQAICRGFLMRVEFRKMMERRESIFCIQYNVRAF MNVKHWPWMKLYFKIKPLLKSAETEKEMANMKEEFEKTKEELAKSEAKR KELEEKMVTLMQEKNDLQLQVQSEADGLADAEERCDQLIKTKIQLEAKIKE LTERAEDEEEINAELTAKKRKLEDECSELKKDIDDLELTLAKVEKEKHATE NKVKNLTEEMAGLDENIAKLTKEKKALQEAHQQTLDDLQAEEDKVNTLT KSKTKLEQQVEDLEGSLEQEKKLRMDLERAKRKLEGDLKLAQESTMDIEN DKQQLDEKLKKKEFEMSNLQSKIEDEQALAMQLQKKIKELQARIEELEEEIE AERASRAKAEKLRSDLSRELEEISERLEEAGGATSAQIEMNKKREAEFQKM RRDLEEATLQHEATAAALRKKHADSVAELGEQIDNLQRVKQKLEKEKSEM KMEIDDLASNMETVSKAKGNLEKMCRTLEDQLSEVKSKEEEHQRLINELSV QKAHLHTESGEFSRQLEEKDALVSQLSRGKQAFTQQIEELKRQLEEETKAK SALAHALQSARHDCDLLREQYEEEQEGKAELQRALSKANSEVAQWRTKYE TDAIQRTEELEEAKKKLAQRLQDAEEHVEVVNAKCASLEKTKQRLQNEVE DLMLDVERSNAACAALDKKQRNFDKVLSEWKQKYEETQADLEASQKESR SLSTELFKVKNAYEESLDQLETLKRENKNLQQEISDLTEQIAEGGKHIHELE KVKKQIEQEKGELQAALEEAEASLEHEEGKILRIQLELNQVKSEIDRKIAEK DEEIDQLKRNHLRVVESMQSTLDAEIRSRNDALRIKKKMEGDLNEMEIQLN HANRQATEAIKNLRNTQGVLKDTQLHLDDAIRGQDDLKEQLAMVERRAN LMQAEIEELRASLEQTERSRRVAEQELLDASERVQLLHTQNTSLINTKKKLE TDITQIQGEMEDIVQEARNAEEKAKKAITDAAMMAEELKKEQDTSAHLER MKKNMEQTVKDLQHRLDEAEQLALKGGKKQIQKLEARVRELENEVENEQ KRNAEAVKGLRKHERRVKELTYQTEEDRKNVLRLQDLVDKLQSKVKAYK RQAEEAEEQSNVNLAKFRKLQHELEEAEERADIAESQVNKLRVKSREVHTK VISEE
93	Cow ALBU (Albumin)	MKWVTFISLLLLFSSAYSRGVFRRDTHKSEIAHRFKDLGEEHFKGLVLIAFS QYLQQCPFDEHVKLVNELTEFAKTCVADESHAGCEKSLHTLFGDELCKVA SLRETYGDMADCCEKQEPERNECFLSHKDDSPDLPKLKPDPNTLCDEFKAD EKKFWGKYLYEIARRHPYFYAPELLYYANKYNGVFQECCQAEDKGACLLP KIETMREKVLASSARQRLRCASIQKFGERALKAWSVARLSQKFPKAEFVEV TKLVTDLTKVHKECCHGDLLECADDRADLAKYICDNQDTISSKLKECCDKP LLEKSHCIAEVEKDAIPENLPPLTADFAEDKDVCKNYQEAKDAFLGSFLYE YSRRHPEYAVSVLLRLAKEYEATLEECCAKDDPHACYSTVFDKLKHLVDE PQNLIKQNCDQFEKLGEYGFQNALIVRYTRKVPQVSTPTLVEVSRSLGKVG TRCCTKPESERMPCTEDYLSLILNRLCVLHEKTPVSEKVTKCCTESLVNRRP CFSALTPDETYVPKAFDEKLFTFHADICTLPDTEKQIKKQTALVELLKHKPK ATEEQLKTVMENFVAFVDKCCAADDKEACFAVEGPKLVVSTQTALA

[0043] An animal protein can comprise a sheep protein (Table 4). A sheep protein can comprise a sheep muscle protein. A sheep protein can comprise a sheep blood protein. A sheep protein can comprise a sheep albumin protein. A sheep protein can comprise a sequence that has at least or about 70%, at least or about 75%, at least or about 80%, at least or about 90%, at least or about 91%, at least or about 92%, at least or about 93%, at least or about 94%, at least or about 95%, at least or about 96%, at least or about 97% at least or about 98% at least or about 99% or about 100% sequence identity to any one of SEQ ID NO: 94-108.

[0044] A food product may comprise a fragment or a truncation of a sheep protein. The food product may include a fragment or truncation of any one of SEQ ID Nos: 94-108. A sequence of the fragment or truncation may be identical or substantially identical to a portion of any one of SEQ ID Nos: 94-108. A sequence of the fragment or truncation may be at least about 70%, 80%, 90%, 95%, 98%, or more identical to a portion of a sequence of any one of SEQ ID Nos: 94-108. A length of the fragment or truncation may be at least about 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of a length of any one of SEQ ID Nos: 94-108. A length of the fragment or truncation may be less than or equal to about 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, or less than a length of any one of SEQ ID Nos: 94-108.

TABLE 4

Sheep Proteins

SEQ ID NO	Description	Amino Acid Sequence
94	Sheep W5PJ69 (Creatine kinase)	MPFGNTHNKHKLNFKAEEEYPDLSKHNNHMAKALTLEIYKKLRDKETPSG FTLDDVIQTGVDNPGHPFIMTVGCVAGDEESYTVFKDLFDPIIQDRHGGFKP TDKHKTDLNHENLKGGDDLDPHYVLSSRVRTGRSIKGYALPPHCSRGERR AVEKLSVEALNSLTGEFKGKYYPLKSMTEQEQQQLIDDHFLFDKPVSPLLL ASGMARDWPDARGIWHNDNKSFLVWVNEEDHLRVISMEKGGNMKEVFR RFCVGLQKIEEIFKKAGHPFMWNEHLGYVLTCPSNLGTGLRGGVHVKLAH LSKHPKFEEILTRLRLQKRGTGGVDTAAVGSVFDVSNADRLGSSEVEQVQL VVDGVKLMVEMEKKLEKGQSIDDMIPAQK
95	Sheep W5P9P1 (Actinin alpha 3)	MMMVLQPEGLGTGEGPFAGGGRGGGEYMEQEEDWDRDLLLDPAWEKQQ RKTFTAWCNSHLRKAGTQIENIEEDFRNGLKLMLLLEVISGERLPRPDKGK MRFHKIANVNKALDFIASKGVKLVSIGAEEIVDGNLKMTLGMIWTIILRFAI QDISVEETSAKEGLLLWCQRKTAPYRNVNVQNFHTSWKDGLALCALIHRH RPDLIDYAKLRKDDPIGNLNTAFEVAEKYLDIPKMLDAEDIVNTPKPDEKAI MTYVSCFYHAFAGAEQAETAANRICKVLAVNQENEKLMEEYEKLASELLE WIRRTVPWLENRVGEPSMSAMQRKLEDFRDYRRLHKPPRVQEKCQLEINF NTLQTKLRLSHRPAFMPSEGKLVSDIANAWRGLEQAEKGYEDWLLSEIRRL QRLQHLAEKFQQKASLHEAWTRGKEEMLSQRDYETASLQEVRALLRRHEA FESDLAAHQDRVEHIAALAQELNELDYHEAASVNSRCQAICDQWDNLGTL TQKRRDALERMEKLLETIDQLQLEFARRAAPFNNWLDGAVEDLQDVWLV HSVEETQSLVTAHDQFKATLPEADRERGAILGIQGEIQKICQTYGLRPSSTNP YIALTPQDINTKWDTVRKLVPSRDQMLQEELTRQQMNERLRRQFAAQANA IGPWIQGKVEEGGRLAAGMAGSLEEQMAGLRQQEQNIINYKSNIDRLEGDH QLLQESLVFDNKHTVYSMEHIRVGWEQLLTSIARTINEVENQVLTRDAKGL SQEQLNEFRASFNHFDRKRNGMMEPDDFRACLISMGYDLGEVEFARIMTM VDPNAAGVVTFQAFIDFMTRETAETDTAEQVVASFKILAGDKNYITAEELR RELPAEQARRAPDRGAGAPAGALDYVAFSSALYGESDL
96	Sheep PYGM (Glycogen phosphorylase)	MSRPLTDQEKRKQISVRGLAGVENVTELKKNFNRHLHFTLVKDRNVATPR DYYFALAYTVRDHLVGRWIRTQQHYYEKDPKRIYYLSLEFYIGRTLQNTM VNLALENACDEATYQLGLDMEELEEIEEDAGLGNGGLGRLAACFLDSMAT LGLAAYGYGIRYEFGIFNQKISGGWQMEEADDWLRYGNPWEKARPEFTLP VHFYGRVEHTSQGAKWVDTQVVLAMPYDTPVPGYRNNVVNTMRLWSAK APNDFNLKDFNVGGYIQAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYF VVAATLQDIIRRFKSSKFGCLDPVRTNFDAFPDKVAIQLNDTHPSLAIPELM RILVDQERLEWEKAWEVTVKTCAYTNHTVLPEALERWPVHLIETLLPRHLQ IIYEINQRFLNRVAAAFPGDVDRLRRMSLVEEGAVKRINMAHLCIAGSHAV NGVARIHSEILKKTIFKDFYELEPHKFQNKTNGITPRRWLVMCNPGLAEVIA ERIGEEYIADLDQLRKLLSYVDDESFIRDVAKVKQENKLKFSAYLEKEYKV HINPNSLFDIQVKRIHEYKRQLLNCLHVITLYNRIKKEPNKFFVPRTVMIGGK AAPGYHMAKMIIRLITAIGDVVNHDPVVGDRLRVIFLENYRVSLAEKVIPAA DLSEQISTAGTEASGTGNMKFMLNGALTIGTMDGANVEMAEEAGEENFFIF GMRVEDVERLDQKGYNAQEYYDRIPELRHIIDQLSSGFFSPKQPDLFKDIVN MLMHHDRFKVFADYEEYVKCQERVSALYKNPREWTRMVIRNIATSGKFSS DRTIAQYAREIWGVEPTRQRMPAPDEKI
97	Sheep A0A6P7DGM3 (Fructose-bisphosphate aldolase)	MPHQYPALTPEQKKELCDIAHRIVAPGKGILAADESTGSIAKRLQSIGTENTE ENRRFYRQLLLTADDRVNPCIGGVILFHETLYQKADDGRPFPQVIKAKGGV VGIKVDKGVVPLAGTNGETTTQGLDGLSERCAQYKKDGADFAKWRCVLKI GEHTPSSLAIMENANVLARYASICQQNGIVPIVEPEILPDGDHDLKRCQYVT EKVLAAVYKALSDHHIYLEGTLLKPNMVTPGHACTQKYSHEEIAMATVTA LRRTVPPAVPGITFLSGGQSEEEASINLNAINKCPLLKPWALTFSYGRALQAS ALKAWGGKKENLKAAQEEYVKRALANSLACQGKYTPSGKAGAAASESLFI SNHAY

TABLE 4-continued

Sheep Proteins		
SEQ ID NO	Description	Amino Acid Sequence
98	Sheep W5QD16 (Myosin light chain 1)	ILLNKMAPKKDVKKPAAAAAPAPAPAPAPAPAPPKEEKIDLSAIKIEFSK QQQDEFKEAFLLYDRTGDGKITLSQVGDVLRALGTNPTNAEVKKVMGNPS NEEMNAKKIEFEQFLPMLQAISNNKDQGTYEDFVEGLRVFDKEGNGTVMG AELRHVLATLGEKMKEEEVEALMAGQEDSNGCINYEAFVKHIMSN
99	Fragment of Sheep G3P (Glyceraldehyde-3-phosphate dehydrogenase)	IGRLVTRAAFNTGKVDIVAINDPFIDLHYMVYMFQYDSTHGKFHGTVKAEN GKLVINGKAITIFQERDPANIKWGDAGAEYVVESTGVFTTMEKAGAHLKG GAKKVIISAPSADAPMFVMGVNHEKYNNTLKIVSNASCTTNCLAPLAKVIH DHFGIVEGLMTTVHAITATQKTVDGPSGKLWRDGRGAAQNIIPASTGAAKA VGKVIPELNGKLTGMAFRVPTPNVSVVDLTCRLEKPAKYDEIKKVVKQASE GPLKGILGYTEDQVVSCDFNSDTHSSTFDAGAGIALNDHFVKLISWYDNEF GYSNRVVDLMVHMASKE
100	Sheep W5QI03 (Tropomyosin 1)	MDAIKKKMQMLKLDKENALDRAEQAEADKKAAEDRSKQLEDELVSLQKK LKATEDELDKYSEALKDAQEKLELAEKKATDAEADVASLNRRIQLVEEEL DRAQERLATALQKLEEAEKAADESERGMKVIESRAQKDEEKMEIQEIQLKE AKHIAEDADRKYEEVARKLVIIESDLERAEERAELSEGQVRQLEEQLRIMDQ TLKALMAAEDKYSQKEDKYEEEIKVLSDKLKEAETRAEFAERSVTKLEKSI DDLEDELYAQKLKYKAISEELDHALNDMTSIESGSCQRRKP
101	Sheep W5PX04 (Myosin binding protein C2)	MPEAKPAAKKAPAGKDAAAKPAPKEAPPQEAPAAPSTEAPPEDQSPTAEEP TGVFLKKPDSVSVENGKDTVIVAKLNGKELPAKPAVKWFKGKWLELGSKS GARFSFKESHDAASNVYTVELHITKVVLGDRGDYRIEVKAKDFCDSCTFNI DVESPRQDSAGQSLESFKRSSEAKSDTAGELDFSGLLKKRQVVEEEKKKKK DDDDLGIPPEIWELLKGAKKSEYERIAFQYGITDLRGMLKRLKKAKVEVKK SAAFTKKLDPAYQVDKGNKIKLVVEISDPDLPLKWFKNGQEIKPSSKYVFE NVGKKRILTINKCTLADDAAYEVVVKDEKCFTELFVKVEPPVLIVTPLEDQ QVFVGNRVEMAVEVSEEGAQVMWLKDGVELTREDSFKSRYRLKKDGKRH ILIYAEVALEDKGHYKVMTNGDQCEAELIVEEKQLEVLQDIADLTVKASEQ AVFKCEVSDEKVTGKWYKNGVEVRPSKRITISHTGRIHKLVIDDVRPEDEG DYTFVPDGYALSLSAKLNFLGAGLSPLPPEPPKIHLDCSGQTSENSVVVVAG NKLRLDVSITGEPRPVATWMKEDEVFTGTEGRVRIEQRGDISSFVIESAERG DEGRYTIKVTNPAGEDVASILLKVVDVPDPPEAVRVTSVGEDWAVLVWEP PKYDGGRPVTGYLLERKKKGSQRWMKLNFEVFTETTYESTNMIEGILYEM RVFAVNAIGVSQPSMNTKPFMPIAPTSEPQHLTVEDVTDTTTTLKWRPPEKI GAGGIDGYLIEYCVEGSDEWIPANTEPVERCGFTVKNLPTGAKITFRVVGV NIAGRSQPATLAQPVTIREIVEQPKIRLPRHLRQTYVRKVGEHLNLVIPFQGK PRPQVVWTKGGAPIDTNHVHVRTSDFDTVFFVRQAARSDSGEYELTVQIEN MKDTATIRIRVVEKAGPPENVMVKEVWGTNALVEWQPPRDDGNSEITGYL VQKADKKTMEWFTVYERNRHTSCTVSDLIMGNEYYFRVYSENICGLSDLP GVSKNTARIVKTGMTLKLPEYKEHDFRTPPKFLTPLFDRVVVAGYAAALNC AVRGHPKPKVVWMKNKMEIREDPKFLMTNHQGVLTLNIRRPSPFDSGTYS CRAVNELGEALAECKLEVRGEP
102	Sheep W5P6W4 (Myomesin 2)	MSLVAVPFYQKRHKHFDQSYRNIQTRYLLDEYAAKKKRASAQSSLQRSVT RKSSSQRESSRAALGETTLGETTLGRTTCPLCAKRMSSALEEEAQERKRRY QSMVAAYGEAKRERYLSELAQLEEDVHVARTHAREELDRLALQHAGRHG LHHAVDDRLAWERHWFEDRISRAPEILVRLRSHTVWERMSVKLCFTVQGF PTPVVQWYKDGSLICQAGEPGKYRIDSKYGVHTLEINRADFDDTATYSAVA TNVHGQVSTNAAVVVRRFRGDEEPFHSVGLPIGLPLSSVIPYTHFDVQFIEK FGVTFRREGETLTLKCTLLVTPDLKRVQPRAEWYRDDVLLKESKWTKMFF GEGQASLSFSHLNKDDEGLYTLRIVSRGGVTDHSAFLFVRDADPLVTGAPG APMDVHCHDANRDYVIVTWKPPNTTKESPVIGYFIDRCEVGTDNWVQCND APVKICKYPVTGLFEGRSYIFRVRAVNSAGISRPSRVSEAVAALDPVDLRRL QAVHLEGEKDIVIYQEDLEGEVQIPGPPTNVHASETSRTYVVLSWDPPEPRG KEPLMYFIEKSMVGSGSWQRVNAQTAVRSPRYAVFDLAEGKPYVFRVLSA NKHGLSEPSEITAPIQAQDTIVVPSAPGRVLASRNTRTSVVVQWDRPKHEED LLGYYVDCSVAGSNVWEPCNHKPIGYNRFVVHGLTTGEQYIFRVKAVNAV GTSENSQESDVIKVQAALTVPSHPYGITLLNCDGHSMILGWKVPKFSGGSPI LGYYIDKREAHHKNWHEVNSSPLKERILKVEGLTEGSLYEFKIAAANMAG MGSQQTELFKCEAWTMPEPGPAYDLTFCEVRDTSLVLLWKPPVYSGSSPVS GYFVDYKEEDSGEWLTVNEVTTPHRYLKVCDLHQGKTYVFRVRAVNASG VGRPSDTSEPVLVEARPGTKEVSAGVDEEGNIYLSFDCPEVTDASHFTWCK SYEDIADDDRFKVETAGDHSKLYFKNLDKEDLGTYSVSVSDTDGVSSSFVL DEEELERLKALSNEIKNPTIPLKSELAYEIFDKGQVRFWLQAEHLSPDSSYRF IINDREVADSETHRIKCDKSTGVIEMVMDRFTVDNEGTYTVQIQDGKAKSQ SSLVLIGDAFKAVLEEAEFQRKEFLRKQGPHFAEYLHWDVTEECEVRLVCK VANTKKETVFKWLKDDVLYETEKMPDLEKGVCELLIPKLSKKDHGEYKAT LKDDRGQDVSTLEIAGKVYEDMILAMSRVCGASASPLKILCTPEGIRLQCF MKYFTEEMKVNWYHKDAKIASSEHMRIGGSEEMAWLQICEPTEKDKGKY TFEILDGKKNHQRSLDLSGQAFDEAFAEFQQLKAAAFAEKNRGKVIGGLPD VVTIMEGKTLNLTCTVFGNPDPEVVWFKNDKDIELSDHFSVKVEQAKYVS

TABLE 4-continued

Sheep Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		MTIKGVTSEDSGKYSMHVKNKYGGEKIDVTVSVYKHGEEIPDVVLPQQAK PKLLPAAAPSPAH
103	Sheep W5P663 (Phosphopyruvate hydratase)	MAMQKIFAREILDSRGNPTVEVDLHTAKGRFRAAVPSGASTGIYEALELRD GDKSRYLGKGVLKAVEHINKTLGPALLEKKLSVVDQEKVDKFMIELDGTE NKSKFGANAILGVSLAVCKAGAAEKGVPLYRHIADLAGNPELILPVPAFNVI NGGSHAGNKLAMQEFMILPVGASSFREAMRIGAEVYHHLKGVIKAKYGKD ATNVGDEGGFAPNILENNEALELLKTAIQAAGYPDKVVIGMDVAASEFYRN GKYDLDFKSPDDPARHISGEKLGELYKSFIKNYPVVSIEDPFDQDDWATWT SFLSGVNIQIVGDDLTVTNPKRIAQAVEKKACNCLLLKVNQIGSVTESIQAC KLAQSNGWGVMVSHRSGETEDTFIADLVVGLCTGQIKTGAPCRSERLAKY NQLMRIEEALGDKAVFAGRKFRNPKAK
104	Sheep W5P7V0 (Myosin light chain)	QAPKKAKRRAAAEGGSSNVFSMFDQTQIQEFKEAFTVIDQNRDGIIDKEDL RDTFAAMGRLNVKNEELDAMMKEASGPINFTVFLTMFGEKLKGADPEDVI TGAFKVLDPEGKGTIKKQFLEELLTTQCDRFSQEEIRNMWAAFPPDVGGNV DYKNICYVITHGDAKDQE
105	Sheep W5PPG6 (Myosin heavy chain 1)	MSSDQEMAIFGEAAPYLRKSEKERIEAQNKPFDAKTSVFVADPKESFVKAT VQSREGGKVTAKTEAGATVTVKEDQVFPMNPPKFDKIEDMAMMTHLHEP AVLYNLKERYAAWMIYTYSGLFCVTVNPYKWLPVYNAEVVAAYRGKKR QEAPPHIFSISDNAYQFMLTDRENQSILITGESGAGKTVNTKRVIQYFATIAV TGEKKKEEPTSGKMQGTLEDQIISANPLLEAFGNAKTVRNDNSSRFGKFIRI HFGTTGKLASADIETYLLEKSRVTFQLKAERSYHIFYQIMSNKKPELIEMLLI TTNPYDYAFVSQGEITVPSIDDQEELMATDSAIEILGFTSDERVSIYKLTGAV MHYGNLKFKQKQREEQAEPDGTEVADKAAYLQGLNSADLLKALCYPRVK VGNEFVTKGQTVEQVYNAVGALAKAIYDKMFLWMVTRINQQLDTKQPRQ YFIGVLDIAGFEIFDFNSLEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIE WEFIDFGMDLAACIELIEKPMGIFSILEEECMFPKATDTSFKNKLYEQHLGKS NNFQKPKPAKGKAEAHFSLIHYAGTVDYNITGWLDKNKDPLNETVVGLYQ KSSVKTLAFLFSGAASAEAEGGGAKKGGKKKGSSFQTVSALFRENLNKLM TNLRSTHPHFVRCIIPNETKTPGAMEHELVLHQLRCNGVLEGIRICRKGFPSR ILYADFKQRYKVLNASAIPEGQFIDSKKASEKLLGSIDVDHTQYKFGHTKVF FKAGLLGLLEEMRDEKLAQLITRTQARCRGFLARVEYQRMVERRESIFCIQ YNVRAFMNVKHWPWMKLYFKIKPLLKSAETEKEMANMKEEFEKTKEELA KSEAKRKELEEKMVTLTQEKNDLQLQVQSEADALADAEERCDQLIKTKIQL EAKIKEVTERAEDEEEINAELTAKKRKLEDECSELKKDIDDLELTLAKVEKE KHATENKVKNLTEEMAGLDETIAKLTKEKKALQEAHQQTLDDLQAEEDK VNTLTKAKTKLEQQVDDLEGSLEQEKKLRMDLERAKRKLEGDLKLAQEST MDIENDKQQLDEKLKKKEFEMSNLQSKIEDEQALGMQLQKKIKELQARIEE LEEEIEAERASRAKAEKQRSDLSRELEEISERLEEAGGATSAQIEMNKKREA EFQKMRRDLEEATLQHEATAAALRKKQGGVGEQLDNLQRVKQKLEKEKS EMKMEIDDLASNMETVSKAKGNLEKMCRTLEDQLSELKTKEDEQQRLIND LTTQRARLQTESGEFSRQLDEKDALVSQLSRGKQAFTQQIEELKRQLEEEIK AKSALAHALQSARHDCDLLREQYEEEQEAKAELQRALSKANSEVAQWRT KYETDAIQRTEELEEAKKKLAQRLQDAEEHVEAVNAKCASLEKTKQRLQN EVEDLMIDVERTNAACAALDKKQRNFDKILSEWKQKYEETHAELEASQKE SRSLSTELFKIKNAYEESLDQLETLKRENKNLQPTEEISDLTEQIAEGGKRIH ELEKVKKQVEQEKSEIQAALEEAEASLEHEEGKILRIQLELNQVKSEIDRKIA EKDEEIDQLKRNHIRVVESMQSTLDAEIRSRNDAIRLKKKMEGDLNEMEIQ LNHANRMAAEALKNYRNTQAILKDTQLHLDDALRGQEDLKEQLAMVERR ANLLQAEIEELRATLEQTERSRKIAEQELLDASERVQLLHTQNTSLINTKKK LETDITQIQGEMEDIIQEARNAEEKAKKAITDAAMMAEELKKEQDTSAHLE RMKKNLEQTVKDLQHRLDEAEQLALKGGKKQIQKLEARVRELEAEVESEQ KRNVETVKGLRKHERRVKELTYQTEEDRKNILRLQDLVDKLQAKVKSYKR QAEEAEEQSNVNLSKFRKLQHELEEAEERADIAESQVNKLRVKSREVHTKII SEE
106	Sheep W5PFV9 (Nebulin)	MADDEEYEEVVEYYTEETVYEEVPGETITEFYETTTTRTSDYEQSDTSRPAL AKPVPEKPVERKRVIRKKVDPSKFMTPYIAHSQKMQDLFSTNKYKENYEK AKGQPYAITTDTPELRRIKKVQDQLSEVKYRTDGDVAKTICHVDEKAKDIE HAKKVSQQVSKVLYKQNWEDTKDKYLLPPDAPELVQAIKNTAMFSKKLY TEDWEADKTLFYPYNDSPELRRVAQAQKALSDITYKKGLAEQQTQFTSLPD PPDIEFAKKVTNQVSKQKYKEDYENKVKGKWSETPCFEIATARMNSQNIST KKYQEDFENTKDQIYFMQTETPEYKVNKQAGVAASKVKYKQDYEKNKGK ADYNVLPASENPLLRQLKAAGDALSDKLYKENYEKTKAKSINYCETPKFKL DTVLHNFSSDTKYKDSYLKNILGHYVGSYEDPYHTHCMRVSAQNSDKNYK AEYEEDRGKGFFPQTITQEYEAIKKLDQCKDHTYKVHPDKTKFTQVTDSPV LVQAQVNSKQLSDLNYKAKHENEKFKCHIPPDTPAFIQHKVNAYNLSDNIY KHDWEKTKAKKFDIKVDAIPLLAAKANTKIASDVMYKKDYEKSKGKMIG ALSINDDPKMLHSLKTAKNQSDRLYKENYEKTKAKSMNYCETPKYQLDTL LKNFSEAKYKDSYVQNVLGHYIGSFEDPYREHCMKVAAQNSDKNYKAEY

TABLE 4-continued

Sheep Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		EEDKGKCYFPQTITQEYEAIKKLDQCKDHTYKVHPDKTKFTAVTDTPVLLQ AQLNTKQLSDLNYKAKHEGEKFKCHIPADAPQFIQHRVNAYNLSDNIYKH DWEKTKAKKFDIKVDAIPLLAAKANTKIASDVMYKKDYEKSKGKMIGALS INDDPKMLHSLKTAKNQSDREYRKDYEKSKTIYTAPLDMLPLTHAKKSQEI ASDVDYKHLLHNYSYPPDSVNVDLAKKAYALQSDVEYKADYNSWMKGC GWVPFGSLEMEKAKRASDILNEKKYRQHPDTLKFTSIEDAPIIVQSKINQAQ RSDVAYKAKGEEVIHKYSLPADLPQFIQAKVNAYNISENMYKADLKDLSK KGYDLRIDAIPIKAAKAARQAASDVQYKKDYEKAKGKMVGFQSLQDDPKL VHYMNVAKIQSDREYKKDYEKTKTRHNTPHDMVNIVAAKKAQDVASNV NYKHSLHHYTYLPDAMDLELSKNMMHIQSDNVYKEDYNNWMKGIGWIPI GSLEVEKVKKAGDALNEKKYRQHPDTLKFTSIVDSPVMVQAKQNTQQVSD ILYKAKGEDVKHKYTMSPDLPQFLQAKCNAYNLSDVCYKRDWHDLIAKG TNVLGDAIPITAAKASRNIASDYKYKEAYEKAKGKQVGFRSLQDDPKLVH YMNVAKLQSDREYKKNYENTKTSYHTPGDMVSITAAKMAQDVATNVNY KQPIHHYTYLPDALSLEHIRNVNQIQSDKVYKDEYNHFFKGIGWIPIGSLEV EKVKKAGDALNEKKYRQHPDTIKFTSVPDSMGMVLAQHNTKQLSDLNYK VEGEKVKHKYTLDPDVPQFIQAKVNAFNMSDSHYKADWKKTLAKGYDLR PDAIPIVAAKSSRNIASDFKYKEGYEKAKGKQIGFRSLQDDPKLVHFMNVA KMQSDREYKKGYEASKTKYHTPLDMLSVTAAKKSQEVATNTNYKQPFHH YTLLPDALNVEHSRNAMQIQSDNLYKSDFTSWMKGIGWMPLESLEVEKAK KAGEILSEKKYRQHPEKLKFTYAMDTMEQALNKSNKLIMDKRRYTEKWN KDKTTIHVMPDTPDILLSRVNQITMSDKLYKAGWEEEKKKGYDLRPDAIPI KAARASRDIASDYKYKQAYEQAKGKQIGFRSLEDDPKLVHFMQVAKMQS DREYKKAYEKSKTSFHTPVDMLSVVAAKKSQEVATNANYRNVIHNYNML PDAMSLELAKNMMQIQSDNQYKADYADFMKGIGWLPLGSLEAEKNKKA MEIISEKKYRQHPDTLKYSTLMDSMNMVLAKNNAKIMNEHLYKQAWEAD KTKVHIMPDIPQIILAKANAINMSDKLYKLSLEEAKKKGYDLRTDAIPIKAA KASRDIASDYKYKHNYEKERGKMVGFRSLEDDPKLVHSMQVAKMQSDRE YKKNYEKTKTSYHTPADMLSVTAAKDAQANITNTNYKHLIHKYILLPDAM NIQLSKNMNRLQSDNEYKQDYNEWYKGLGWSPAGSLEVEKAKRATEYAS DQKYRQHPSNFQFSKLNDSMDMVLAKQNAHTMNKYLYTVDWNKDKTKI HVMPDTPDILQAKQNQTIYSQKLYKLGWEEALKKGYDLPVDAISVQAAKA SRDIVSDFKYKQGYRKQIGHHIGFRSLQDDPKLVLSMNVAKMQSEREYKK DFEKWKTKFSSPVDMLGVVLAKKCQALVSDVDYKNYLHGWTCLPDQND VIHAKKAYDLQSENLYKSDLEWLKGIGWSPLGSLEAEKNKRASEIISEKKY RQPPDRNKFTSIPDAMDIVLAKTNAKNRSDILYREAWDKDKTQVHIMPDTP DIILAKANLINASDKFYRMGYEELRKKGYDLPVDAIPIKAAKASREIASEYK YKEGFRKQLGHHIGARNIKDDPKMMWSMHVAKIQSDREYKKDFEKWKTK FSSPVDMLGVVLAKKCQTLVSDIDYKNYLHQWTCLPDQSDVIHARRAYDL QSDNLYKSDLQWLKGIGWLPSGSLEDEKNKRATQILSDHVYRQHPDKFKFS SLMDSIPMVLAKNNAITMNHRLYTEAWDKDKTTVHIMPDTPGVLLAKQNQ INYSEKLYKLGLDEAKRKGYDMRIDAIPIRAAKASRDIASEFKYKEGYRKQ LGHHIGARAIHDDPKMMWSMHVAKIQSDREYKKDFEKWKTKFSSPVDML GVVLAKKCQTLVSDIDYKNYLHQWTCLPDQSDVIHARQAYDLQSDNVYK SDLQWLKGIGWVPIGSLDVEKCKRASEILSDKLYRQPPDKFKFTSVTDSLEQ VLAKNNAINMNKRLYTEAWDKDKTQVHIMPDTPEITLARMNKVNYSESLY KLAHEEAKKKGYDLRSDAIPIVAAKASRDIISDYKYKDGYRKQLGHHIGAR DIQDDPKMMWSMHVAKIQSDREYKKDFEKWKTKFSSPVDMLGVVLAKKC QTLVSDIDYKNYLHQWTCLPDQSDVIHARRAYDLQSDNVYKSDLQWLRGI GWVPIGSVDVVKCKRAAEILSDNLYRQPPDTLKFTSVPDSLEQVLAKNNAI NMNKRLYTEAWDKDKTQVHIMPDTPEIMLARQNKLNYSETLYKLANEEA KKKGYDLRSDAIPIVAAKASRDIISDYKYKDGYRKQLGHHIGARDIKDDPK MMWSMHVAKIQSDREYKKDFEKWKTKFSSPVDMLGVVLAKKCQTLVSDI DYKNYLHQWTCLPDQSDVIHARQAYDLQSDNIYKSDLQWLRGIGWVPIGS VDVVKCKRAAEILSDNLYRQRPDTLKFTSVPDSLEQVLAKNNAINMNKRL YTEAWDKDKTQIHMMPDTPEITLARQNKINYSENLYRQAMEEAKKEGYDL RSDAIPIVAAKASREIASDYKYKEAYRKQLGHHIGARAIHDDPKMMWSVH VAKMQSDREYKKDFEKYKTRFSSPVDMLGIVLAKKCQTLVSDVDYKHPLH EWTCLPDQNDIIHARKAYDLQSDNLYKSDLEWLKGIGWVPIGSFDVVKAK RAGEILSDNIYRQRPDTLKFTSVTDSPEQVLAKNNAINMSKRLYTEAWDND KKTIHVMPDTPEIMLAKLNRINYSDKLYKLALEESKKEGYDLRLDAIPIQAA KASREIASDYKYKEGYRKQLGHHIGARNIKDDPKMMWSIHATKLQSDREY KKDFEKWKTKFSSPVDMMGLVQAKKCQILVSDIDYKNLLHEWICLPDQND VIQARKAYDLQSDAIYKADLEWLRGIGWVPIGSVGVENVKRAGEILSERKY RQPAVQLKFTSITDTPEIVLAKNNALNMSKHLYTEAWDADKTSIHVMPDTP EILLAKSNSANISHKLYTKGWDESKMKDYDLRADAISIKSAKASRDIASDY KYKEAYEKQKGHHIGARSVEDDPKIMCAMNAGKIQSEREYKKEFQKWKT KFSSPVDMLGILQAKKCQILVSDIDYRNYLHHWTCLPDQNDIIQARKAYDL QSDAIYKADLEWLRGIGWMPEGSPEVLRVKNAQHIFRDSVYRTPVVKLKY TSIVDTPEVVLAKSNAENISIPKYREVWDKDKTSIHIMPDTPEINLARTNALN VSNKLYREGWDEMKMSCDVRLDAIPIQAAKASREIASDYKYKLDHEKQKG HYVGTRTARDDNKIRWALIAGKIQNEREYRMHWAKWKSKFQSPPDMLSIE

TABLE 4-continued

Sheep Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		HSKNSQTLVSDIDYRHYLHQWTCMPDQNDVIQARKAYDLQSDAIYKADLE WLRGIGWMPADSVSVNHAKHMGDIFNEKKYRTKIETLNFTPVDDRVDYVT AKHSGEILNDIKYRKDWNDTKSKYTLTETPQLHAAQEAARILDQYLYKES WEKQKATGYILPPDAVPFVHAHHSGDVQSELKYKAEHVKQKGHYVGVPT MRDDPKLVWFEHAGQIQNDRLYKESYHKTKAKIHIPPDMVSVLAAKEGQA LASDIDYRTYLHQWICHPDQNDVIQARKAYDLQSDNIYKADLEWLRGIGW MPLDSVDHVRVTRNQEMVNQIKYKKDALANYPNFTSVVDPPEIVLAKINA VNQSDVKYKETFNKLIKGKYIFSPDTPYITHSKDMEKLYSTILYKRAWEGT KAYGYTLDERYIPIVGAKHADFVNSELKYKETYEKLKGHYLAGKEISEFPN VVHCLDFQKMRSLLNYRRHYEDTKANVHIPQDMMNHVLAKRCQYILSDL EYRHYFHQWTSLPEEPNVVRVRHAQEILSDNVYKDDLNWLKGIGCYVWD TPQILHAKKSYDLQSQILYTAAGKENLKNYNLVTDTPLYVTALQSGINASE VKYKENYHQIKDKYTTVLETVDYDRTRNLKDLFSSNLYKEAWDKVKATS YILPPNTVSLTHAKNQKHMASYVKYREEYDKFKALYTLPRSVEDDPNTAR CLRVGKFNIDRLYRSVYEKNKMNIHIVPDMVEMVTAKDSQKKVSEIDYRL HLHEWICHPDLQVNSHVRKVTDQISDIVYKDDLTWLKGIGCYVWDTPEILH AKHAYDLRNDIKYKAHVLKTRNNYKLVTDTPVYVQAVKSGKQLSDAVYH YDYVHSIRGRVAPTTKTVDLDRALHAYKLQSENLYRKAGLHALPTGYRLP VDTPHFKHTKDTRYMSSYFKYKEVYEHMKAYGYTLGPNDVPFVNVRRVN NITSERLYRQLYHKLKDKIHTTPDTPEIRQVKKTQEAVSELIYKSDFFKMQG LRKLWSHYLYTDKARKMRDKYQVVLDTPEYRTVQELKTHLSELVYRASG RKQKSIFTSVPDTPDLIRAKRGQKLQSQYLYVELATKERPHHHAGNQTTAL KHARHVKDMVSENKYKIQYEKMKDKYTPVPDTPILIRAKRAYWNASDLR YKETFQKTKGRYHTVKDALDIVYHRTVTDHISKIKYKENYMSQLGIWRSIP DRPEHFHHRAVTDAVSDVKYKADLTWLKGIGCYAYDTPDFTLAEKNKTLY SKYKYKEVFERTKSNFKYVADCPINRHFKFATQLMNEKKYRADYEQRKDK YHLVVDEPRHLLAKIAGDQISQIRYRKNYEKTKDKFTSIVDTPEHLRTTKVN KQISDILYKLEYNKAKPRGYTTIHDTPMLLHVRKVKDEVSDLKYKEVYQRT KSNCTIEPDAVHIKAAKDAYKVNTNLDYKKQYEATKAYWKWTPDRPDFIQ AAKSTLQQSDFEYKLDREYLKGCKLSVTDDKDMVLALKNSIIESDLKYKEK HVKERGSCHAVPDTPQILLAKTVSSLVSENKYKQYVKKHLAQGSYTTLPET RDTIHVKEVTKNVSDTNYKKKFVKEKGKSNYSIMLEPPDVKHAMDVAKK QSNMKVAYKKDAKENLHYTTVADRPDIKKATQAAKQASEVEYRAKHRKE GSHGLSMLGRPDIEMAKKAAKLSSQVKYRENFNKEKGKTPKYNPKDSQLY KVMKDANTLASEVKYKADLKKLHKPVTDMKESLIMNHVLNTSHLASSYQ YKKNYEKSKGHYHTIPDNLEQLHLKEATELQSIVKYKEKYEKERGKPMLD FETPTYITAKESQQMQSGKEYRRDYEESIKGRNLTGLEVTPALLHVKYATKI ASEKEYRKDLEESIRGKGLSEMEDTPDMLRAKNATQILNEKEYKRDLELEV KGRGLNAMANETPDFLRARNATDIASQIKYKQSAEMEKANFTSVVDTPEII HAQQVKNLSSQKKYKEDAEKCMSYYETVLDTPEMQRVRENQKNFSLLQY QYDLKNSKGKITVVQDTPEILRVKENQKNFSSVLYKEDVSPGTAIGKTPEM MRVKQTQDHISSVKYKEVIGQGTPIPDLPEVKRVKQTQKHISSVMYKENLG TGIPTPVTPEIERVKRNQENFSSVLYKENLGKGTPTPITPEMERVKRNQENFS SILYKENLSKGTPLPVTPEMERVKRNQENFSSVLYKENVGKGTPTPVTPEM QRVKRNQENISSVLYKENLGKATPTPFTPEMERVKRNQENFSSVLYKENMR KATPTPVTPEMERAKRNQENISSVLYSDSFRKQIQGKAAYVLDTPEMRRVR ETQRHISTVKYHEDFEKHKGCFTPVVTDPITERVKKNMQDFSDISYRGIQRK VVEMEQKRNDQDQETVTGLRVWRTNPGSVFDYDPAEDNIQSRSLHMINAQ AQRRSREQSRSASGLSVSGGEEKSEHSEAAHLSTYSDGGVFFSATSTGYKH ARTTELPQQRSSSVATQQTTVSSIPSHPSTAGKIFRAMYDYMAADADEVSF KDGDAIVNVQAIDEGWMYGTVQRTGRTGMLPANYVEAI
107	Sheep W5PIU1 (Myosin-4)	MSSDQEMAIFGEAAPYLRKSEKERIEAQNKPFDAKTSVFVADPKESFVKAT VQSREGGKVTAKTEAGATVTVKEDQVFPMNPPKFDKIEDMAMMTHLHEP GVLYNLKERYAAWMIYTYSGLFCVTVNPYKWLPVYNAEVVTAYRGKKRQ EAPPHIFSISDNAYQFMLTDRENQSILITGESGAGKTVNTKRVIQYFATIAVT GEKKKEEPTSGKMQGTLEDQIISANPLLEAFGNAKTVRNDNSSRFGKFIRIH FGTTGKLASADIETYLLEKSRVTFQLKAERSYHIFYQIMSNKKPELIEMLLIT TNPYDYAFVSQGEITVPSIDDQEELMATDSAIEILGFTSDERVSIYKLTGAVM HYGNLKFKQKQREEQAEPDGTEVADKAAYLQSLNSADLLKALCYPRVKV GNEFVTKGQTVQQVYNAVGALAKAVYEKMFLWMVTRINQQLDTKQPRQ YFIGVLDIAGFEIFDFNSLEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIE WEFIDFGMDLAACIELIEKPMGIFSILEEECMFPKATDTSFKNKLYEQHLGKS NNFQKPKPAKGKAEAHFSLVHYAGTVDYNIAGWLDKNKDPLNETVVGLY QKSAMKTLAFLFAGGPSAEVHVSEGGGGKKGGKKKGSSFQTVSALFRENL NKLMTNLRSTHPHFVRCIIPNETKTPGAMEHELVLHQLRCNGVLEGIRICRK GFPSRILYADFKQRYKVLNASAIPEGQFIDSKKASEKLLGSIDIDHTQYKFGH TKVFFKAGLLGILEEMRDEKLAQLITRTQAICRGFLTRVEFRKMMERSRESI FCIQYNVRAFMNVKHWPWMKLYFKIKPLLKSAETEKEMANMKEEFEKTK EELAKSEAKRKELEEKMVTLMQEKNDLQLQVQSEADGLADAEERCDQLIK TKIQLEAKIKELTERAEDEEEINAELTAKKRKLEDECSELKKDIDDLELTLAK VEKEKHATENKVKNLTEEMAGLDENIAKLTKEKKALQEAHQQTLDDLQA

TABLE 4-continued

Sheep Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		EEDKVNTLTKSKTKLEQQVDDLEGSLEQEKKLRMDLERAKRKLEGDLKLA QESTMDIENDKQQLDEKLKKKEFEMSNLQSKIEDEQALGMQLQKKIKELQ ARIEELEEEIEAERASRAKAEKQRSDLSRELEEISERLEEAGGATSAQIEMNK KREAEFQKMRRDLEEATLQHEATAAALRKKHADSVAELSEQIDNLQRVKQ KLEEEGSNLTVQALDIISNFKIINRYYYIEGNLEKMCRTLEDQLSEVKSKEEE HQRLINELSAQKARLHTESGEFSRQLDEKDALVSQLSRGKQAFTQQIEELKR QLEEETKAKSALAHALQSARHDCDLLREQYEEEQEAKAELQRSMSKANSE VAQWRTKYETDAIQRTEELEEAKKKLAQRLQEAEEHVEAVNAKCASLEKT KQRLQNEVEDLMLDVERSNAACAALDKKQRNFDKVLSEWKQKYEETQAE LEASQKESRSLSTELFKVKNAYEESLDQLETLKRENKNLQQEISDLTEQIAE GGKHIHELEKVKKQIEQEKGELQAALEEAEASLEHEEGKILRIQLELNQVKS EIDRKIAEKDEEIDQLKRNHLRVVESMQSTLDAEIRSRNDALRIKKKMEGDL NEMEIQLNHANRQATEATKNLRNTQGVLKDTQLHLDDAIRGQDDLKEQLA MVERRANLMQAEIEELRASLEQTERSRRVAEQELLDASERVQLLHTQNTSL INTKKKLETDITRVHGSVVGSGIISKHLTFVYISVGCAFQAAMMAEELKKEQ DTSAHLERMKKNLEQTVKDLQHRLDEAEQLALKGGKKQIQKLEARVRELE NEVENEQKRNAEAVKGLRKHERRVKELTYQTEEDRKNVLRLQDLVDKLQ SKVKAYKRQAEEAEEQANVNLSKFRKLQHELEEAEERADIAESQVNKLRV KSREVHTKVISEE
108	Sheep ALBU (Albumin)	MKWVTFISLLLLFSSAYSRGVFRRDTHKSEIAHRENDLGEENFQGLVLIAFS QYLQQCPFDEHVKLVKELTEFAKTCVADESHAGCDKSLHTLFGDELCKVA TLRETYGDMADCCEKQEPERNECFLNHKDDSPDLPKLKPEPDTLCAEFKAD EKKFWGKYLYEVARRHPYFYAPELLYYANKYNGVFQECCQAEDKGACLL PKIDAMREKVLASSARQRLRCASIQKFGERALKAWSVARLSQKFPKADFTD VTKIVTDLTKVHKECCHGDLLECADDRADLAKYICDHQDALSSKLKECCD KPVLEKSHCIAEVDKDAVPENLPPLTADFAEDKEVCKNYQEAKDVFLGSFL YEYSRRHPEYAVSVLLRLAKEYEATLEDCCAKEDPHACYATVFDKLKHLV DEPQNLIKKNCELFEKHGEYGFQNALIVRYTRKAPQVSTPTLVEISRSLGKV GTKCCAKPESERMPCTEDYLSLILNRLCVLHEKTPVSEKVTKCCTESLVNRR PCFSDLTLDETYVPKPFDEKFFTFHADICTLPDTEKQIKKQTALVELLKHKP KATDEQLKTVMENFVAFVDKCCAADDKEGCFVLEGPKLVASTQAALA

[0045] An animal protein can comprise a pig protein (Table 5). A pig protein can comprise a pig muscle protein. A pig protein can comprise a pig blood protein. A pig protein can comprise a pig albumin protein. A pig protein can comprise a sequence that has at least or about 70%, at least or about 75%, at least or about 80%, at least or about 90%, at least or about 91%, at least or about 92%, at least or about 93%, at least or about 94%, at least or about 95%, at least or about 96%, at least or about 97% at least or about 98% at least or about 99% or about 100% sequence identity to any one of SEQ ID NO: 109-123.

[0046] A food product may comprise a fragment or a truncation of a pig protein. The food product may include a fragment or truncation of any one of SEQ ID Nos: 109-123. A sequence of the fragment or truncation may be identical or substantially identical to a portion of any one of SEQ ID Nos: 109-123. A sequence of the fragment or truncation may be at least about 70%, 80%, 90%, 95%, 98%, or more identical to a portion of a sequence of any one of SEQ ID Nos: 109-123. A length of the fragment or truncation may be at least about 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of a length of any one of SEQ ID Nos: 109-123. A length of the fragment or truncation may be less than or equal to about 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, or less than a length of any one of SEQ ID Nos: 109-123.

TABLE5

Pig Proteins		
SEQ ID NO	Description	Amino Acid Sequence
109	Pig KCRM (Creatine kinase M-type)	MPFGNTHNKYKLNFKAEEEYPDLSKHNNHMAKALTLEIYKKLRDKETPSGF TLDDVIQTGVDNPGHPFIMTVGCVAGDEESYVVFKDLFDPIIQDRHGGYKPT DKHKTDLNHENLKGGDDLDPNYVLSSRVRTGRSIKGYTLPPHCSRGERRAV EKLSVEALNSLTGEFKGKYYPLKSMTEQEQQQLIDDHFLFDKPVSPLLLASG MARDWPDARGIWHNDNKSFLVWVNEEDHLRVISMEKGGNMKEVFRRFCV GLQKIEEIFKKAGHPFMWNEHLGYVLTCPSNLGTGLRGGVHVKLAHLSKHP KFEEILTRLRLQKRGTGGVDTAAVGSVFDVSNADRLGSSEVEQVQLVVDGV KLMVEMEKKLEKGQSIDDMIPAQK

TABLE5-continued

Pig Proteins		
SEQ ID NO	Description	Amino Acid Sequence
110	Pig A0A480REY2 (Actinin alpha 3)	MMMVMQPEGLGTREGPFAGGSGGGGEYMEQEEDWDRDLLLDPAWEKQQ RKTFTAWCNSHLRKAGTQIENIEEDFRNGLKLMLLLEVISGERLPRPDKGKM RFHKIANVNKALDFIASKGVKLVSIGAEEIVDGNLKMTLGMIWTIILRFAIQDI SVEETSAKEGLLLWCQRKTAPYRNVNVQNFHTSWKDGLALCALIHRHRPDL IDYAKLRKDDPIGNLNTAFEVAEKYLDIPKMLDAEDIVNTPKPDEKAIMTYV SCFYHAFAGAEQAETAANRICKVLAVNQENEKLMEEYEKLASELLEWIRRT VPWLENRVGEPSMSAMQRKLEDFRDYRRLHKPPRVQEKCQLEINFNTLQTK LRLSHRPAFMPSEGKLVSDIANAWRGLEQVEKGYEDWLLSEIRRLQRLQHL AEKFQQKASLHEAWTRGKEDMLSQRDYESASLQEVRALLRRHEAFESDLA AHQDRVEHIAALAQELNELDYHEAASVNSRCQAICDQWDNLGTLTQKRRD ALERMEKLLETIDQLQLEFARRAAPFNNWLDGAVEDLQDVWLVHSVEETQS LVTAHEQFKATLPEADRERGAILGIQGEIQKICQTYGLRPSSTNPYIALSPQDI NTKWDTVRKLVPSRDQMLQEELTRQQVNERLRRQFAAQANAIGPWIQGKV EEVGRLAAGMAGSLEEQMAGLRQQEQNIINYKSNIDRLEGDHQLLQESLVF DNKHTVYSMEHIRVGWEQLLTSIARTINEVENQVLTRDAKGLSQEQLNEFR ASFNHFDRKRNGMMEPDDFRACLISMGYDLGEVEFARIMTMVDPNAAGVV TFQAFIDFMTRETAETDTAEQVVASFKILAGDKNYITAEELRRELPAEQAEYC IRRMAPYKGAGAPAGALDYVAFSSALYGESDL
111	Pig A0A5G2QU59 (Alpha-1,4 glucan phosphorylase)	MAMSRPLTDQEKRKQISVRGLAGVENVAELKKNFNRHLHFTLVKDRNVAN ARDYYFALAYTVRDHLVGRWIRTQQHYYEKDPKRIYYLSLEFYMGRTLQN TMVNLALENACDEATYQLGLDMEELEEMEEDAGLGNGGLGRLAACFLDSM ATLGLAAYGYGIRYEFGIFNQKITGGWQMEEADDWLRYGNAWEKARPEFT LPVHFYGRVEHTNQGAKWVDTQVVLAMPYDTPVPGYRNNVVNTMRLWSA KAPNDFNLKDFNVGGYIQAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYF VVAATLQDIIRRFKSSKFGCRDPVRTSFDAFPDKVAIQLNDTHPSLAIPELMRI LVDEERLEWEKAWDVTVRTCAYTNHTVLPEALERWPVHLMETLLPRHLQII YEINQRFLNRVAAAYPGDVDRLRRMSLVEEGAVKRINMAHLCIAGSHAVNG VARIHSEILKKTIFKDFYELEPHKFQNKTNGITPRRWLVMCNPGLAEVIAERI GEEYIADLDQLRKLLSYVDDEAFIRDVAKVKQENKLKFSAYLEREYKVHINP NSLFDIQVKRIHEYKRQLLNCLHVITLYNRIKKEPNKFVVPRTVMIGGKAAP GYHMAKMIIKLITAIGDVVNHDPVVGDRLRVIFLENYRVSLAEKVIPAADLS EQISTAGTEASGTGNMKFMLNGALTIGTMDGANVEMAEEAGEENFFIFGMR VEDVERLDQKGYNAQEYYDRIPELRHVIDQLSSGFFSPKQPDLFKDIVNMLM HHDRFKVFADYEDYIKCQERVSALYKNPREWTRMVIRNIATSGKFSSDRTIA QYAREIWGVEPSRQRLPAPDEKI
112	Pig A0A286ZYX8 (Fructose-bisphosphate aldolase)	MPHQYPALTPEQKKELSDIAHRIVAPGKGILAADESTGSIAKRLQSIGTENTE ENRRFYRQLLLTADDRVNPCIGGVILFHETLYQKADDGRPFPQVIKAKGGVV GIKVDKGVVPLAGTNGETTTQGLDGLSERCAQYKKDGADFAKWRCVLKIG EHTPSSLAIMENANVLARYASICQQVGLQNGIVPIVEPEILPDGDHDLKRCQY VTEKVLAAVYKALSDHHIYLEGTLLKPNMVTPGHACTQKYSHEEIAMATVT ALRRTVPPAVPGITFLSGGQSEEEASINLNAINKCPLLKPWALTFSYGRALQA SALKAWGGKKENLKAAQEEYVKRALANSLACQGKYTPSGKAGAAASESLF ISNHAY
113	Pig A0A287BJF1 (Myosin light chain 1)	MAPKKDVKKPAAAAAPAPAPAPAPAPAPAPPKEEKIDLSAIKSFSADQIAEFK EAFLLFDRTGECKITLSQVGDVLRALGTNPTNAEVKKVLGNPSNEEMNAKKI EFEQFLPMLQAISNNKDQGSYEDFVEGLRVFDKEGNGTVMGAELRHVLATL GEKMKEEEVEALMAGQEDSNGCINYEAFVKHIMSI
114	Pig G3P (Glyceraldehyde-3-phosphate dehydrogenase)	MVKVGVNGFGRIGRLVTRAAFNSGKVDIVAINDPFIDLHYMVYMFQYDSTH GKFHGTVKAENGKLVINGKAITIFQERDPANIKWGDAGATYVVESTGVFTT MEKAGAHLKGGAKRVIISAPSADAPMFVMGVNHEKYDNSLKIVSNASCTTN CLAPLAKVIHDHFGIVEGLMTTVHAITATQKTVDGPSGKLWRDGRGAAQNII PASTGAAKAVGKVIPELNGKLTGMAFRVPTPNVSVVDLTCRLEKPAKYDDI KKVVKQASEGPLKGILGYTEDQVVSCDFNSDTHSSTFDAGAGIALNDHFVK LISWYDNEFGYSNRVVDLMVHMASKE
115	Pig TPM1 (Tropomyosin alpha-1 chain)	MDAIKKKMQMLKLDKENALDRAEQAEADKKAAEDRSKRLEDELVSLQKK LKATEDELDKYSEAPKDAQEKLELAEKKATDAEADVASLNRRIQLVEEELD RAQERLATALQKLEEAEKAADESERGMKVIESRAQKDEEKMEIQEIQLKEA KHIAEDADRKYEEVARKLVIIESDLERAEERAELSEGKCAELEEELKTVINNL KSLEAQAEKYSQKEDKYEEEIKVLSDKLKEAETRAEFAERSVTKLEKSIDDL EDELYAQKLKYKAISEELDHALNDMTSI

TABLE5-continued

Pig Proteins		
SEQ ID NO	Description	Amino Acid Sequence
116	Pig A0A5G2QB13 (Myosin binding protein C2)	MPEAKPATKKAPAGKDAAKPAPKEAPPKGAPEEPPKEAPPEDQSPTAEEPTG IFLKKPDSVSVENGKDALIVAKVNGKELPGKPEIKWFKGKWLELGSKSGARF SFKESHDSASNVYTLELRIGKVVLGDRGDYRLEVKAKDFCDSCAFNIDVEAP RQDSAGQSLESFKRSGEGKPEGAGELDFSGLLKKRQVVEEEKKKKKDDDDL GIPPEIWELLKGAKKSEYEKIAFQYGITDLRGMLKRLKKAKVEVKKSAAFTK KLDPAYQVDQGNKIKLVVEISDPDLPLKWYKNGQEIKPSSKYVFENVGKKRI LCINKCTLADDAAYEVAVKDEKCFTELFVREPPVLIVTPLEDQQVYVGDRVE MTVEVSEEGAQVMWLKDGVELTREDSFKSRYRFKKDGKRHILIYSDVSLED RGRYQVMTNGGQCEAELSVEEKQLEVLQDIADLTVKASEQAVFKCEVSDEK VTGKWYKNGVEVRPSKRITISHVGRFHKLVIDDVRPEDEGDYTFVPDGYALS LSAKLNFLEIKVEYVPKQEPPKIHLDCSGKTSENAVVVVAGNKLRLDVSITG EPPPVATWMKEDEVFSGTEGRVRLEQRADISSFVIESAERGDEGRYTIKVTNP AGEDVASIFLRVVDVPDPPEAVRVTTVGEDWAILVWEPPKYDGGQPVTGYL LERKKKGSQRWMKLNFEVFTETTYESTNMIEGILYEMRVFAVNAIGVSQPSV NTKPFMPIAPTSEPQHLTVEDVTDTTTTLKWRPPDRIGAGGIDGYLVEYCVE GSDEWIPANTEPVERCGFTVKNLPTGARITFRVVGVNIAGRSQPATLAQPVTI REIVGEPPRQSSVLIRASPDSLVATTLSVPLRRVPQPYAEPAPGIVSPEQPKIRL PRHLRQTYIRKVGEHLNLVIPFQGKPRPQVVWTKGGAPIDTSRVHVRTSNFD TVFFVRQAARSDSGEYELSVQIENMKDTATVHIRVVEKAGPPQNVMVKEV WGTNALVEWQPPKDDGNTEITGYFVQKADKKTMEWFNVYEHNRHTSCTV SDLIMGNEYYFRVYSENICGLSDSPGVSKNTARILKTGITFKLPEFKEHDFRTP PKFLTPLPDRVVVAGYTAALNCAVRGHPKPKVVWMKNQMEIGEDPKFLMT NHQGVLTLNIRRPSPFDSGTYSCRASTSWGRPWPSASWKSECRSEAPASTQG SWLLSPDPDLPEPLSLSRVITRFLSRTQ
117	Pig A0A287BSB6 (Uncharacterized protein)	TMQHVYTKMSLVAVPFYQKRHRHFDQSYRNIQTRYLLDEYAAKKRASTQS SSQKSVTQESSSQRESSSERASSQTLLGGTTCRLCAKRVSTALEEEEQERKRR YHSLLAAYGEAKRERYLSELAQLEEDVHLARTHARDKLDRFALQHVDTHA LRHAVGDRLAWERHSFEDRISRAPEILVRLRSHTVWERMSVKLCFTVQGFPT PVVQWYKDGSLICQAGEPGKYRIESKYGVHTLEINRADFDDTATYSAVATN VHGQVSTNAAVVVRRFHGDEEPFHSVGLPIRLPLSSVIPYTHFDVQFLEKFGV TFRREGETLTLKCTLLVTPDLKRVQPRAEWYRDDVLLKESKWTKMFFGEGQ ASLSFSHLNKDDEGLYTLRIVSRGGISDHSAFLFVRDADPLVMGAPGAPMDV HCHDANRDYVIVTWKPPNTTKESPVLGYFIDRCEVGTDNWVQCNDAPVKIC KYPVSGLVEGRSYTFRVRAVNSAGVSRPSRASEAVAALDPLDLKRLQGIHW EGDKDIVIYQDDLEGEVQIPRPPTNVHASEISRTYVVLSWDPPTPRGKEPLMY FIEKSVVGSGTWQRENAQTAVRSPRYAVFDLAEGKSYVFRVLSANKHGLSD PSEITSPIQAQDTIVVPSAPGRVLASRNTKTSVVVQWDQPKHGEDLLGYYVD CCVAGSNVWEPCNHKPIGYNRFVVHGLTTGEQYIFRVKAVNAVGTSENSQE SDVIKVQAALTVPSHPYGITLLNCDGHSMTLGWKVPKFSGGSPIIGYYVDKR EGHHKNWHEVNSSPLKERILTVEGLTEGSTYEFKIAAPTWWGSGCRRIPVSS SSVRPGQRLSQVPTSGWVGTLGFFLVLLSLYRRYLLGSGEHGDS
118	Pig ENOB (Beta-enolase)	MAMQKIFAREILDSRGNPTVEVDLHTAKGRFRAAVPSGASTGIYEALELRDG DKSRYLGKGVLKAVEHINKTLGPALLEKKLSVVDQEKVDKFMIELDGTENK SKFGANAILGVSLAVCKAGAAEKGVPLYRHIADLAGNPDLVLPVPAFNVING GSHAGNKLAMQEFMILPVGASSFREAMRIGAEVYHHLKGVIKGKYGKDAT NVGDEGGFAPNILENNEALELLKTAIQAAGYPDKVVIGMDVAASEFYRNGK YDLDFKSPDDPSRHITGEKLGELYKSFIKNYPVVSIEDPFDQDDWKTWTSFLS GVNIQIVGDDLTVTNLKRIAQAVEKKACNCLLLKVNQIGSVTESIQACKLAQ SNGWGVMVSHRSGETEDTFIADLVVGLCTGQIKTGAPCRSERLAKYNQLMR IEEALGDKAVFAGRKFRNPKAK
119	Pig Q5XLD2 (Myosin light chain)	MAPKKAKRRAAAEGSSNVFSMFDQTQIQEFKEAFTVIDQNRDGIIDKEDLRD TFAAMGRLNVKNEELDAMMKEASGPINFTVFLTMFGEKLKGADPEDVITGA FKVLDPEGKGTIKKHFLEELLTTQCDRFSQEEIKNMWAAFPPDVGGNVDYK NICYVITHGDAKDQE
120	Pig MYH1 (Myosin-1)	MSSDQEMAIFGEAAPYLRKSEKERIEAQNKPFDAKTSVFVAEPKESFVKGTV QSREGGKVTVKTEAGATLTVKEDQVFPMNPPKFDKIEDMAMMTHLHEPAV LYNLKERYAAWMIYTYSGLFCVTVNPYKWLPVYNAEVVTAYRGKKRQEAP PHIFSISDNAYQFMLTDRENQSILITGESGAGKTVNTKRVIQYFATIAVTGEKK KEEPTSGKMQGTLEDQIISANPLLEAFGNAKTVRNDNSSRFGKFIRIHFGTTG KLASADIETYLLEKSRVTFQLKAERSYHIFYQIMSNKKPELIEMLLITTNPYDY AFVSQGEITVPSIDDQEELMATDSAIEILGFTSDERVSIYKLTGAVMHYGNLK FKQKQREEQAEPDGTEVADKAAYLQGLNSADLLKALCYPRVKVGNEFVTK GQTVQQVYNAVGALAKAVYDKMFLWMVTRINQQLDTKQPRQYFIGVLDIA GFEIFDFNSLEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEWEFIDFGMD LAACIELIEKPMGIFSILEEECMFPKATDTSFKNKLYEQHLGKSNNFQKPKPA KGKVEAHFSLIHYAGTVDYNITGWLDKNKDPLNETVVGLYQKSSVKTLAFL FTGAAGADAEAGGGKKGGKKKGSSFQTVSALFRENLNKLMTNLRSTHPHF

TABLE5-continued

Pig Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		VRCIIPNETKTPGAMEHELVLHQLRCNGVLEGIRICRKGFPSRILYADFKQRY KVLNASAIPEGQFIDSKKASEKLLGSIDIDHTQYKFGHTKVFFKAGLLGLLEE MRDEKLAQLITRTQARCRGFLARVEYQKMVERRESIFCIQYNIRAFMNVKH WPWMKLYFKIKPLLKSAETEKEMANMKEEFEKTKESLAKAEAKRKELEEK MVALMQEKNDLQLQVQAEADSLADAEERCDQLIKTKIQLEAKIKEVTERAE DEEEINAELTAKKRKLEDECSELKKDIDDLELTLAKVEKEKHATENKVKNLT EEMAGLDETIAKLTKEKKALQEAHQQTLDDLQAEEDKVNTLTKAKTKLEQ QVDDLEGSLEQEKKLRMDLERAKRKLEGDLKLAQESTMDIENDKQQLDEK LKKKEFEMSNLQSKIEDEQALAMQLQKKIKELQARIEELEEEIEAERASRAK AEKQRSDLSRELEEISERLEEAGGATSAQIEMNKKREAEFQKMRRDLEEATL QHEATAATLRKKHADSVAELGEQIDNLQRVKQKLEKEKSEMKMEIDDLAS NMETVSKAKGNLEKMCRTLEDQLSELKTKEEEQQRLINDLTAQRARLQTES GEYSRQLDEKDTLVSQLSRGKQAFTQQIEELKRQLEEEIKAKSALAHAVQSS RHDCDLLREQYEEEQEAKAELQRAMSKANSEVAQWRTKYETDAIQRTEELE EAKKKLAQRLQDAEEHVEAVNAKCASLEKTKQRLQNEVEDLMIDVERSNA ACAALDKKQRNFDKILAEWKQKYEETHAELEASQKESRSLSTELFKVKNAY EESLDQLETLKRENKNLQQEISDLTEQIAEGGKRIHELEKIKKQVEQEKSEIQ AALEEAEASLEHEEGKILRIQLELNQVKSEVDRKIAEKDEEIDQLKRNHVRV VESMQSMLDAEIRSRNDAIRLKKKMEGDLNEMEIQLNHANRMAAEALRNY RNTQGILKDTQIHLDDALRSQEDLKEQLAMVERRANLLQAEIEELRATLEQT ERSRKVAEQELLDASERVQLLHTQNTSLINTKKKLETDISQIQGEMEDIIQEA RNAEEKAKKAITDAAMMAEELKKEQDTSAHLERMKKNLEQTVKDLQHRL DEAEQLALKGGKKQIQKLEARVRELEGEVESEQKRNVETVKGLRKHERRVK ELTYQTEEDRKNILRLQDLVDKLQAKVKSYKRQAEEAEEQSNVNLSKFRKL QHELEEAEERADIAESQVNKLRVKSREVHTKIISEE
121	Pig F1SHX0 (Nebulin)	MADDEEYEEVVEYYTEETVYEEVPGETITEVYETTTRTSGYEQSDTSQPAV AQPMPSKPVERKKVIRKKVDPSKFMTPYIAHSQKMQDLFSTNKYKKNYEKA KGKPYAITTDTPELRRIKKVQDQLSEVKYRIDGDVAKTICHVDEKAKDIEHA KKVSQQVSKVLYKQNWEDTKDKYLLPPDAPELVQAIKNTAMFSKKLYTED WEADKSLFYPYNDSPELRRVAQAQKALSDIAYKKGHAEQQTQFTSLPDPPDI EFAKKVTNQVSKQKYKEDYENKIKGKWSETPCFEIATARMNAKNISTRKYQ EDFENMKDQIYFMQTETPEYKVNKQAGTAASKVKYKEDYEKNKGKADYN VLPASENPLLRHLKAAGDALSDKRYKENYEKTKAKSINYCETPKFKLDTVL HNFSSDTKYKDSYLKNILGHYVGSFEDPYHTHCMKVAAENSDKNYKAEYE EDRGKGFFPQTITQEYEAIKKLDQCKDHTYKVHPDKTKFTQVTDSPVLVQA QVNSKQLSDLNYKAKHENEKFKCHIPPDTPAFIQHKVNAYNLSDNVYKHD WEKTKAKKFDIKVDAIPLLAAKANTKNTSDVMYKKDYEKSKGKMIGALSIN DDPKMLHSLKTAKNQSDRLYKENYEKTKAKSMNYCETPKYQLDTLLKNFS EAKYKDSYVQNVLGHYIGSFEDPYQLHCMKVSAQNSDKNYKAEYEEDKGK CYFPQTITQEYEAIKKLDQCKDHTYKVHPDKTKFTAVTDTPVLLQAQLNTK QLSDLNYKAKHEGEKFKCNIPADAPQFIQHRVNAYNLSDNVYKHDWEKTK AKKFDIKVDAIPLLAAKANTKNTSDVMYKKDYEKSKGKMIGALSINDDPKM LHSLKTAKNQSDREYRKDYEKSKTIYTAPLDMLQVTQAKKSQAIASDVDYK HLLHNYSYPPDSINVDLAKKAYALQSDVEYKADYNSWMKGCGWIPFGSLE MEKAKRASDILNEKKYRQHPETLKFTSIEDAPIIVQSKINQAQRSDIAYKAKG EEIIHKYSLPADLPFIQAKVNAYNISESMYKADLKDLSKKGYDLRTDAIPIKA AKAARQAASDVQYKKDYEKAKGKMVGFQSLQDDPKLVHYMNVAKIQSDR EYKKDYEKTKTKYNTPHDMFNVVAAKKAQDVASNVNYRHSLHHYTYLPD AMDLELSKNMMHIQSDNAYKEDYNTWMKGIGWIPLGSLEVEKVKKAGDA LNEKKYRQHPDTLKFTSIVDSPVMVQAKQNTRQVSDILYKAKGEDVKHKYT MGPDLPQFLQAKCNAYNISDVCYKRDWHDLIAKGNNVLGDAIPITAAKASR NIASDYKYKEAYEKAKGKHVGFRSLQDDPKLVHYMNVAKLQSDREYKKN YENTKTSYHTPGDMVSITAAKMAQDVATNVNYKQPIHHYTYLPDALSLEHT RNVNQIQSDNVYKDEYNHFFKGIGWIPLGSLEVEKVKKAGDALNERKYRQH PDTIKFTSVPDSMGMVLAQHNTKQLSDLNYKVEGEKLKHKYTMDPDVPQFI QAKVNAINMSDAHYKADWKKTLAKGYDLTPDAIPIVAAKSSRNIASDFKYK EAYEKAKGKQVGFRSLQDDPKLVHYMNVAKIQSDREYKKGYEASKTRYHT PLDMLSVTAAKKSQEVATNTNYKQPFHHYTLLPDALNVEHSRNAMQIQSDN LYKSDFTNWMKGIGWIPMESLEVEKAKKAGEILSEKKYRQHPEKLKFTYAM DTMEQALNKSNKLIMDKRLYTEKWNKDKTTIHVMPDTPDILLSRVNQITLS DKLYKAGWEEEKKKGYDLRPDAISIKAARASRDIASDYKYKQAYEQAKGK HIGFRSLEDDPKLVHFMQVAKMQSDREYKKAYEKSKTSFHTPVDMLSVVA AKKSQEVATNANYRNVIHTYNMLPDAMSFELAKNMMQLQSDNQYKADYA DFMKGIGWLPLGSLEAEKNKKAMEIISEKKYRQHPDTLKYSTLMDSMNMVL AKNNAKIMNEHLYKQAWEADKTKVHIMPDIPQIILAKANAVNMSDKLYKLS LEESKKKGYDLRTDAISIQAAKASRDIASDYKYKHSYEKERGKMVGFRSLED DPKLVHSMQVAKMQSDREYKKSYEKTKTSYHTPADMLSVTAARDAQANIT NTNYKHLIHKYILLPDAMNIQLSKNMNHIQSDNEYKQDYNECYKGLGWSPA GSLEVEKAKKATEYASDQKYRQHPSNFQFQKLNDSMDMVLAKQNAHTMN KYLYTADWNKDKTKIHVMPDTPDILQAKQNQTLFSQKLYKLGWEEALKKG YDLPVDAISVQLAKVSRDNASDFKYKQGYRKQLGHHIGFRNLQDDPKLVLS

TABLE5-continued

Pig Proteins	
SEQ ID NO Description	Amino Acid Sequence
	MNVAKMQSEREYKKDFEKWKTKFTSPVDMLGVVLAKKCQALVSDVDYKN YLHQWTCLPDQNDVIQAKRSYELQSENLYKSDLEWLRGIGWSPLGSLEAEK NKRASEIISEKKYRQPPDRNKFTSIPDAMDIVLAKTNAKNRSDRLYREAWDK DKTQVHIMPDTPDIILAKANLINSSDKLYRMGYEELKKKGYDLPVDAIPIKA AKASREIASEYKYKEGFRKQLGHHVGARNIEDDPKMMWSMHVAKIQSDRE YKKDFEKWKTKFSSPVDMLGVVLAKKCQTLVSDIDYKNYLHQWTCLPDQS DVIHARRAYDLQSDNLYKSDLQWLKGIGWLPSGSLEDEKNKRATQILSDHV YRQHPDKFKFSSLMDSIPMVLAKNNAITMNHRLYTEAWDKDKTTVHIMPDT PEVLLAKQNKINYSEKLYKLGLEEAKRKGYDMRLDAIPIRAAKASRDIASEF KYKEGYRKQLGHHIGARAIHDDPKMMWSMHVAKVQSDREYKKDFEKWKT KFSSPVDMLGVVLAKKCQTLVSDIDYKNYLHQWTCLPDQSDVIHARRAYDL QSDNMYKSDLQWLRGIGWVPIGSLDVEKCKRATEILSDKIYRQPPDTFKFTS VPDSLEQVLAKNNAINMNKRLYTEAWDKDKTQVHIMPDTPEIMLARQNRL NYSETLYKLANEEAKKKGYDLRSDAIPIVAAKASRDIVSDYKYKDGYRKQL GHHIGARNIEDDPKMMWSMHVAKIQSDREYKKDFEKWKTKFSSPVDMLGL VQAKKCQILVSDVDYKNYLHQWTCLPDQNDVIHARQAYDLQSDNIYKSDL QWLKGIGWVPIGSVDVVKCKRAAEILSDHIYRQPPDKLKFTSVPDSLEQVLA KNNAINMNKRLYTEAWDKDKTQIHIMPDTPEIMLARQNKINYSESLYRQAM EEAKKEGYDLRSDAIPIVAAKASRDIASDYKYKEAYRKQLGHHIGARAVHD DPKMMWSLHVAKMQSDREYKKDFEKYKTRFSSPVDMLGIVLAKKCQTLVS DVDYKHPLHEWTCLPDQNDIIHARKAYDLQSDNLYKSDLEWMKGIGWIPLG SVEVVKAKRAGEILSDNIYRQRPDTLKFTSVTDSPELVLAKNNALNMNKRL YTEAWDNDKKTIHVMPDTPEIMLAKLNRINYSDKLYRLALEESKKEGCDLR VDAIPIQAAKASREIASDYKYKEAYRKQLGHHIGARNIEDDPKMMWSIHAA KIQSDREYKKDFEKWKTRFSSPVDMLGLVQAKKCQILVSDIDYKHPLHEWIC LPDQNDIIQARKAYDLQSDAIYKADLEWLRGIGWVPLGSVEVEKVKRAGEIL SERKYRQPADQLKFTCITDTPEIVLAKNNALTMSKHLYTEAWDTDKTSIHV MPDTPEILLAKSNSANISQKLYTKGWDESKMKDYDLKADAISIKSAKASRDI ASDYKYKEAYEKQKGHHIGAQSVEDDPKIMCAIHAGKIQSEREYKKDFQKW KTKFTSPVDMLGILLAKKCQTLVSDIDYRNYLHHWTCLPDQNDIIQARKAYE LQSDALYKADLEWLRGIGWMPQGSPEVLRVKNAQEILRDSVYRTPVVKLK YTSIVDTPEVVLAKSNAENISIPKYREVWDKDKTSVHIMPDTPEINLARANAL NVSNKIYREGWDEMKMSCDVRLDAIPIQAAKASREIASDYKYKLDHEKQKG HYVGTLTARDDNKIRWALIAGKIQNEREYRLHWAKWKSKFQSPADMLSITH SKNCQALVSDIDYRNYLHQWTCLPDQHDVIQARKAYDLQSDAIYKADLEW LRGIGWMPSDSVSVNHAKHAADIFSEKKYRTKIETLNFTPVDDRVDYVTAK QSGEILNDIKYRKDWNDTKSKYTLTETPQLHAAQEAARILDQYLYKESWEK QKATGYILPPDAVPFVHAHHSSDVQSELKYKADHVKQKGHYVGVPTMRDD PKLVWFEHAGQIQNDRLYKEHYHKTKAKINIPPDMVSVVAAKEGQILVSDID YRNYLHQWTCHPDQNDVIQARKAYDLQSDNVYKADLEWLRGIGWIPLDSV DHVRVTRNQEMVNQIKYKKDALANYPNFTSVVDPPAVVLAKINSVNQSDV KYKETFNKLLKGKYIFSPDTPYITHSKDMGKLYSTILYKGAWEGTKAYGYTL DERYIPIVGAKHADFVNSELKYKETYEKQKGHYLAGKEISEFPNVVHCLDFQ KMRSVLNYRRHYEDTKANVHIPNDMMNHVLAKRCQYILSDLEYRHYFHQ WTSLPEEPNVIRARNAQEILSDNVYKDDLNWLKGIGCYVWDTPQILHAKKS YDLQSQILYTADGKENLKNYNMVTDTPLYVTALQSGINASEVKYKENYHQI KDKYTTVLETVDYDRTRNLKDLYSSNVYKEAWDKVKATSYILPSNTVSLTH AKNQKHLASHIKYREEYEKFKALYTLPKSVDDDPNTARCLRVGKFNIDRLY RSVYEKNKMKIHMVPDMVEMVTAKDSQKKVSEIDYRLHLHEWICHPDLQV NSHVRKVTDQISDIVYKDDLNWLKGIGCYVWDTPEILHAKHAYDLRNDIKY QAHVKKTRNNYKLVTDTPLYVQAVKSGKQLSDAVYHYDYVHSVRGKVAP TTKTVDLDRALHAYKLQSENLYRKDGLNSLPTGYRLPGDTPHFKHNKDIRY MSSYFKYKEVYEHFKANGYTLGPKDVPFVHVRRVNNVTSERLYRQLYHKL KDKIHTTPDTPEIRQVKKTQEAVSELIYKSDFFKMQGHMISMPYTPHVLHCR YVGDITSDIKYKEDLQILRGLGCFLYDTPDMVRSRHLRKLWSNYLYTDNAR KMRDKYKVVLDTPEYRKVQELKTHLSELVYRAAGKKQKSIFTSVPDTPDLL RAKRGQKLQSQYLYVELATKERPHHHAGNQTTALKHARDVKDMVSEKKY KVQYEKMKDKYTPVPDTPILIRAKKAYWNASDLRYKETFQKTKGKYHTVK DALDIVYHRKVTDHISKIKYKENYMSQLGIWRSIPDRPEHFHHRAVTDAVSD VKYKEDLTWLKGIGCYAYDTPDFTLAEKNKTLYSKYKYKEVFERTKSDFKY EADCPINRHFKFATQLMNEKKYRADYEQRKDKYQLVVDEPRHLLAKIAGD QISQIKYRKNYEKTKDKFTSVVDTPEHLRTTKVNKQISDILYKLEYNKAKPR GYTTIHDTPMLLHVRKVKDEVSDLKYKEVYQRTKSNCTIEPDAVHIKAAKD AYKVNTNLDYKKQYEANKAHWRWTPDRPDFIQAAKSSLQQSDFEYKLDRE FLKGCKLSVTDDKDMVLALRNSVIESDLKYKEKHVKERGTCHAVPDTPQIL LAKTVSNLVSENKYKRHVKKHLAQGSYTTLPETRDTVHVKEVTKHVSDTN YKKKFVKERGKSNYSIMLEPPEVKHAMEVAKKQSDVAYKKDAKENLHYTT VADRPDIKKATQAAKQASEVEYRAKHRKEGSHGLSMLGRPDIEMAKKAAT LSSQVKYRENFNKEKGKTPKYNPKESQLYKIMKDANNLASEVKYKADLKK LHKPVTDMKESLIMNHVLNTSHLASSYQYKKNYEKSKGHYHTIPDNLEQLH LKEATELQSIVKYKEKYEKERGKPMLDFETPTYITAKESQQMQSGKEYRKD YEESIKGRNLTGLEVTPALLHVKYATKIASEKEYRKDLEESIRGKGLSEMEDT

TABLE5-continued

Pig Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		PDMLRAKNATQILNEKEYKRDLELEVKGRGLNAMANETPDFMRARNATDI ASQIKYKQLAEMEKANFTSVVDTPEIIHAQQVKNLSSQKKYKEDAEKCMSY YETVLDTPEMQRVRENQKNFSLLQYQCDLKNSKGKITVVQDTPEILRVKEN QKNFSSVLYKEDVSPGTAIGKTPEMMRVKQTQDHISSVKYKETIGQGTPIPD LPEVKRVKETQKHISSVMYKENLGTGIPTPVTPEIERVKRNQENFSSVLYKEN LGKGTPTPITPEMERVKRNQENFSSVLYKENMGKGTPLPVTPEMERVKHNQ ENISSILYKENVGKGTPTPVTPEMQRVKRNQENISSVIYKENLGKATPTPFTPE MERVKRNQENFSSVLYKENMRKATPTPVTPEMERAKRNQENISSVLYSDSF RKQIQGKAAYVLDTPEMRRVRETQRHISTVKYHEDFEKHKGCFTPVVTDPIT ERVKKNTQDFSDINYRGIQRKVVEMEQKRNDQDQETITGLRVWRTNPGSVF DYDPAEDNIQSRSLHMLNAQAQRRSREQSRSASALSISGGEEKSEHSEAARL STYSDGGVFFSATSAAYKHAKTTELPQQRSSSVATQQTTVSSIPSHPSTAGKI FRAMYDYMAADADEVSFRDGDAIVNVQAIDEGWMYGTVQRTGRTGMLPA NYVEAI
122	Pig MYH4 (Myosin-4)	MSSDQEMAIFGEAAPYLRKSEKERIEAQNKPFDAKTSVFVAEPKESFVKGTV QSREGGKVTVKTEAGATLTVKEDQVFPMNPPKFDKIEDMAMMTHLHEPAV LYNLKERYAAWMIYTYSGLFCVTVNPYKWLPVYNAEVVTAYRGKKRQEAP PHIFSISDNAYQFMLTDRENQSILITGESGAGKTVNTKRVIQYFATIAVTGEKK KEEPTPGKMQGTLEDQIISANPLLEAFGNAKTVRNDNSSRFGKFIRIHFGTTG KLASADIETYLLEKSRVTFQLKAERSYHIFYQIMSNKKPELIEMLLITTNPYDY AFVSQGEITVPSIDDQEELMATDSAIEILGFTSDERVSIYKLTGAVMHYGNLK FKQKQREEQAEPDGTEVADKAAYLQGLNSADLLKALCYPRVKVGNEFVTK GQTVQQVYNAVGALAKAVYDKMFLWMVTRINQQLDTKQPRQYFIGVLDIA GFEIFDFNSLEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEWEFIDFGMD LAACIELIEKPMGIFSILEEECMFPKATDTSFKNKLYEQHLGKSNNFQKPKPA KGKAEAHFSLIHYAGTVDYNITGWLDKNKDPLNETVVGLYQKSSVKTLAFL FAERQSSEEGGTKKGGKKKGSSFQTVSALFRENLNKLMTNLRSTHPHFVRCI IPNETKTPGAMEHELVLHQLRCNGVLEGIRICRKGFPSRILYADFKQRYKVLN ASAIPEGQFIDSKKASEKLLGSIDIDHTQYKFGHTKVFFKAGLLGTLEEMRDE KLAQLITRTQAMCRGFLMRVEFRKMMERRESIFCIQYNIRAFMNVKHWPW MKLYFKIKPLLKSAETEKEMANMKEEFEKTKEDLAKSEAKRKELEEKMVAL MQEKNDLQLQVQAEADGLADAEERCDQLIKTKIQLEAKIKEVTERAEDEEEI NAELTAKKRKLEDECSELKKDIDDLELTLAKVEKEKHATENKVKNLTEEMA GLDENIAKLTKEKKALQEAHQQTLDDLQAEEDKVNTLTKAKTKLEQQVDD LEGSLEQEKKLRMDLERAKRKLEGDLKLAQESTMDIENDKQQLDEKLKKKE FEMSNLQSKIEDEQALAMQLQKKIKELQARTEELEEEIEAERASRAKAEKQR SDLSRELEEISERLEEAGGATSAQIEMNKKREAEFQKMRRDLEEATLQHEAT AAALRKKHADSVAELGEQIDNLQRVKQKLEKEKSELKMEIDDLASNMETVS KAKGNLEKMCRTLEDQLSEVKTKEEEHQRLINELSAQKARLQTESGEFSRQL DEKEALVSQLSRGKQAFTQQIEELKRQLEEETKAKSALAHAVQSSRHDCDLL REQYEEEQEAKAELQRAMSKANSEVAQWRTKYETDAIQRTEELEEAKKKL AQRLQDAEEHVEAVNAKCASLEKTKQRLQNEVEDLMLDVERSNAACAALD KKQRNFDKILAEWKHKYEETQAELEASQKESRSLSTELFKVKNAYEESLDQ LETLKRENKNLQQEISDLTEQIAEGGKHIHELEKVKKQIEQEKSELQAALEEA EASLEHEEGKILRIQLELNQVKSEIDRKIAEKDEEIDQMKRNHIRVVESMQST LDAEIRSRNDALRIKKKMEGDLNEMEIQLNHANRQATEAIRNLRNTQGVLK DTQLHLDDAIRGQDDLKEQLAMVERRANLMQAEIEELRASLEQTERSRRVA EQELLDASERVQLLHTQNTSLINTKKKLETDISQIQGEMEDIVQEARNAEEKA KKAITDAAMMAEELKKEQDTSAHLERMKKNMEQTVKDLQHRLDEAEQLA LKGGKKQIQKLEARVRELENEVENEQKRNVEAVKGLRKHERRVKELTYQT EEDRKNVLRLQDLVDKLQSKVKAYKRQAEEAEEQSNVNLSKFRKLQHELEE AEERADIAESQVNKLRVKSREVHTKVISEE
123	Pig ALBU (Albumin)	MKWVTFISLLFLFSSAYSRGVFRRDTYKSEIAHRFKDLGEQYFKGLVLIAFSQ HLQQCPYEEHVKLVREVTEFAKTCVADESAENCDKSIHTLFGDKLCAIPSLR EHYGDLADCCEKEEPERNECFLQHKNDNPDIPKLKPDPVALCADFQEDEQK FWGKYLYEIARRHPYFYAPELLYYAIIYKDVFSECCQAADKAACLLPKIEHL REKVLTSAAKQRLKCASIQKFGERAFKAWSLARLSQRFPKADFTEISKIVTDL AKVHKECCHGDLLECADDRADLAKYICENQDTISTKLKECCDKPLLEKSHCI AEAKRDELPADLNPLEHDFVEDKEVCKNYKEAKHVFLGTFLYEYSRRHPDY SVSLLLRIAKIYEATLEDCCAKEDPPACYATVFDKFQPLVDEPKNLIKQNCEL FEKLGEYGFQNALIVRYTKKVPQVSTPTLVEVARKLGLVGSRCCKRPEEERL SCAEDYLSLVLNRLCVLHEKTPVSEKVTKCCTESLVNRRPCFSALTPDETYK PKEFVEGTFTFHADLCTLPEDEKQIKKQTALVELLKHKPHATEEQLRTVLGN FAAFVQKCCAAPDHEACFAVEGPKFVIEIRGILA

[0047] An animal protein can comprise a bison protein (Table 6). A bison protein can comprise a bison muscle protein. A bison protein can comprise a bison blood protein. A bison protein can comprise a bison albumin protein. A bison protein can comprise a sequence that has at least or about 70%, at least or about 75%, at least or about 80%, at least or about 90%, at least or about 91%, at least or about 92%, at least or about 93%, at least or about 94%, at least or about 95%, at least or about 96%, at least or about 97% at least or about 98% at least or about 99% or about 100% sequence identity to any one of SEQ ID NO: 124-138.

[0048] A food product may comprise a fragment or a truncation of a bison protein. The food product may include a fragment or truncation of any one of SEQ ID Nos: 124-138. A sequence of the fragment or truncation may be identical or substantially identical to a portion of any one of SEQ ID Nos: 124-138. A sequence of the fragment or truncation may be at least about 70%, 80%, 90%, 95%, 98%, or more identical to a portion of a sequence of any one of SEQ ID Nos: 124-138. A length of the fragment or truncation may be at least about 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of a length of any one of SEQ ID Nos: 124-138. A length of the fragment or truncation may be less than or equal to about 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, or less than a length of any one of SEQ ID Nos: 124-138.

TABLE6

Bison Proteins

SEQ ID NO	Description	Amino Acid Sequence
124	Bison A0A6P3IIV2 (Creatine kinase)	MPFGNTHNKHKLNFKAEEEYPDLSKHNNHMAKALTLEIYKKLRDKETPSG FTLDDVIQTGVDNPGHPFIMTVGCVAGDEESYTVFKDLFDPIIQDRHGGFKP TDKHKTDLNHENLKGGDDLDPNYVLSSRVRTGRSIKGYALPPHCSRGERR AVEKLSVEALNSLTGEFKGKYYPLKSMTEQEQQQLIDDHFLFDKPVSPLLL ASGMARDWPDARGIWHNDNKSFLVWVNEEDHLRVISMEKGGNMKEVFR RFCVGLQKIEEIFKKAGHPFMWNEHLGYVLTCPSNLGTGLRGGVHVKLAH LSKHPKFEEILTRLRLQKRGTGGVDTAAVGSVFDVSNADRLGSSEVEQVQL VVDGVKLMVEMEKKLEKGQSIDDMIPAQK
125	Bison A0A6P3J4W6 (Alpha-actinin-2)	MNQIEPGVQYNYVYEDDEYMIQEEEWDRDLLLDPAWEKQQRKTFTAWCN SHLRKAGTQIENIEEDFRNGLKLMLLLEVISGERLPKPDRGKMRFHKIANVN KALDYIASKGVKLVSIGAEEIVDGNVKMTLGMIWTIILRFAIQDISVEETSAK EGLLLWCQRKTAPYRNVNIQNFHTSWKDGLGLCALIHRHRPDLIDYSKLNK DDPIGNINLAMEIAEKHLDIPKMLDAEDIVNTPKPDERAIMTYVSCFYHAFA GAEQAETAANRICKVLAVNQENERLMEEYERLASELLEWIRRTIPWLENRT PEKTMQAMQKKLEDFRDYRRKHKPPKVQEKCQLEINFNTLQTKLRISNRPA FMPSEGKMVSDIAGAWQRLEQAEKGYEEWLLNEIRRLERVEHLAEKFRQK ASTHETWAYGKEQILLQKDYESSTLTEVRALLRKHEAFESDLAAHQDRVE QIAAIAQELNELDYHDAVNVNDRCQKICDQWDRLGTLTQKRREALERTEK LLETIDQLHLEFAKRAAPFNNWMEGAMEDLQDMFIVHSIEEIQSLITAHEQF KATLPEADGERQSILAIQNEVEKVIQSYSIRISSSNPYSTVTVDEIRSKWDKV KQLVPIRDQSLQEELARQHANERLRRQFAAQANAIGPWIQNKMEEIARSSIQ ITGALEDQMNQLKQYEHNIINYKNNIDKLEGDHQLIQEALVFDNKHTNYTM EHIRVGWELLLTTIARTINEVETQILTRDAKGITQEQMNEFRASFNHFDRRK NGLMDHEDFRACLISMGYDLGEAEFARIMTLVDPNGQGTVTFQSFIDFMTR ETADTDTAEQVIASFRILASDKPYILAEELRRELPPDQAQYCIKRMPAYSGP GSVPGALDYTAFSSALYGESDL
126	Bison A0A6P3HAI4 (Alpha-1,4 glucan phosphorylase)	MSRPLTDQEKRKQISVRGLAGVENVTELKKNFNRHLHFTLVKDRNVATPR DYYFALAYTVRDHLVGRWIRTQQHYYEKDPKRIYYLSLEFYIGRTLQNTM VNLALENACDEATYQLGLDMEELEEIEEDAGLGNGGLGRLAACFLDSMAT LGLAAYGYGIRYEFGIFNQKISGGWQMEEADDWLRYGNPWEKARPEFTLP VHFYGRVEHTSQGAKWVDTQVVLAMPYDTPVPGYRNNVVNTMRLWSAK APNDFNLKDFNVGGYIQAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYF VVAATLQDIIRRFKSSKFGCLDPVRTNFDAFPDKVAIQLNDTHPSLAIPELM RILVDQERLEWEKAWEVTVKTCAYTNHTVLPEALERWPVHLIETLLPRHLQ IIYEINQRFLNRVAAAFPGDVDRLRRMSLVEEGAVKRINMAHLCIAGSHAV NGVARIHSEILKKTIFKDFYELEPHKFQNKTNGITPRRWLVMCNPGLAEIIAE RIGEEYIADLDQLRKLLSYVDDESFIRDVAKVKQENKLKFSAYLEKEYKVHI NPNSLFDIQVKRIHEYKRQLLNCLHVITLYNRIKKEPNKFFVPRTVMIGGKA APGYHMAKMIIKLITAIGDVVNHDPVVGDRLRVIFLENYRVSLAEKVIPAA DLSEQISTAGTEASGTGNMKFMLNGALTIGTMDGANVEMAEEAGEENFFIF GMRVEDVERLDQKGYNAQEYYDRIPELRHIIDQLSSGFFSPKQPDLFKDIVN MLMHHDRFKVFADYEEYVKCQERVSALYKNPREWTRMVIRNIATSGKFSS DRTIAQYAREIWGVEPTRQRMPAPDEKI
127	Bison A0A6P3H8G9 (Fructose-bisphosphate aldolase)	MPHQYPALTPEQKKELCDIAHRIVAPGKGILAADESTGSIAKRLQSIGTENTE ENRRFYRQLLLTADDRVNPCIGGVILFHETLYQKADDGRPFPQVIKAKGGV VGIKVDKGVVPLAGTNGETTTQGLDGLSERCAQYKKDGADFAKWRCVLKI GEHTPSSLAIMENANVLARYASICQQNGIVPIVEPEILPDGDHDLKRCQYVT EKVLAAVYKALSDHHIYLEGTLLKPNMVTPGHACTQKYSHEEIAMATVTA LRRTVPPAVPGITFLSGGQSEEEASINLNAINKCPLLKPWALTFSYGRALQAS ALKAWGGKKENLKAAQEEYVKRALANSLACQGKYTPSGKAGAAASESLFI SNHAY

TABLE6-continued

Bison Proteins		
SEQ ID NO	Description	Amino Acid Sequence
128	Bison A0A6P3J7N1 (Myosin light chain 1/3)	MAPKKDVKNPAAAAAPAPAPAPAPAPAPPKEEKIDLSAIKIEFSKQQQDE FKEAFLLFDRTGECKITLSQVGDVLRALGTNPTNAEVKKVLGNPSNEEMNA KKIEFEQFLPMLQAISNNKDQGTYEDFVEGLRVFDKEGNGTVMGAELRHV LATLGEKMKEEEVEALMAGQEDSNGCINYEAFVKHIMSN
129	Bison A0A6P3J4R7 (Glyceraldehyde-3-phosphate dehydrogenase)	MVKVGVNGFGHIGRLVTRAAFNSGKVDVIAINDPFIDLHYIVYMFQYDTTH GKFHGTVKAENGKLIINGKAVTIFQERDPASIKWGDAGAEYVVESTGVFTT MEKAGAHLKGGAKSVIISAPSADAPMFVLGMNHEKYNNNLKIVSNASCTT NCLAPXAKVIHDHFGIVEGLMTTFHAITVTQKTVDGPSGKLWHDGGGAAQ NIIPASTGAAKAVGKVIPELNRKLTGMAFCVPTPNVSVMDLTYHLEKPAKY DEIKKVVKQTSEGPLEGILGYTQDQVVSCDFNSDSYXSTFDAGAGIALNDH FVKLISWYDNEFGYSNRVVDLMVHMASKE
130	Bison A0A6P3GY04 (Tropomyosin alpha-1 chain isoform X2)	MDAIKKKMQMLKLDKENALDRAEQAEADKXPVPHGHSQLEDELVSLQKK LKATEDELDKYSEALKDAQEKLELAEKKATDAEADVASLNRRIQLVEEEL DRAQERLATALQKLEEAEKAADESERGMKVIESRAQKDEEKMEIQEIQLKE AKHIAEDADRKYEEVARKLVIIESDLERAEERAELSEGKCAELEEELKTVTN NLKSLEAQAEKYSQKEDKYEEEIKVLSDKLKEAETRAEFAERSVTKLEKSID DLEDELYAQKLKYKAISEELDHALNDMTSM
131	Bison A0A6P3IB01 (Myosin-binding protein C, fast type)	MPEAKPAAKKAPAGKDAAAKPAPKEATPQEAPAAPPTGEIDGAPAGRAEP GGSSAASRVVPPRPLLRSVFPTAPGLTALPYSQPCRAPQRFQGGGLTVPLPP HSPLLPGKDTVIVAKLNGKELPAKPAVKWFKGKWLELGSKSGARFSFKESH DAASNVYTIELHITKVVLGDRGDYRIEVKAKDFCDSCAFNIDVEAPRQDSA GQSLESFKRSGEAKSDTAGELDFSGLLKKRQVVEEEKKKKKDDDDLGIPPEI WELLKGAKKSEYERIAFQYGITDLRGMLKRLKKAKVEVKKSAAFTKKLDP AYQVDRGNKIKLVVEISDPDLPLKWFKNGQEIKPSSKYVFENVGKKRILTIN KCTLADDAAYEVVVKDEKCFTELFVKEPPVLIVTPLEDQQVFVGDRVEMA VEVSEEGAQVMWLKDGVELTREDSFKARYRLKKDGKRHILIYSEVTMEDK GHYQVMTNGGQCEAELIVEEKQLEVLQDIADLTVKASEQAVFKCEVSDEK VTGKWYKNGVEVRPSKRITISHTGRFHKLVIDDVRPEDEGDYTFVPDGYAL SLSAKLKFLEIKVEYVPKQEPPKIHLDCSGQTSENAIVVVAGNKLRMDVSIT GEPRPVATWMKEDEVFTGTEGRVHIEQRGDISSFVIESAERGDEGRYTIKVT NPVGEDVASILLKVVDVPDPPEAVRVTSVGEDWAVLVWEPPKYDGGRPVT GYLLERKKKGSQRWMKLNFEVFTETTYESTNMIEGILYEMRVFAVNVIGVS QPSMNTKPFMPIAPTSEPRHLTVEDVTDTTTTLKWRPPDKIGAGGIDGYLIE YCVEGSDEWIPANTEPVERCGFTVKNLPTGAKITFRVVGVNIAGRSQPATL AQPVTIREIVEQPKIRLPRHLRQTYVRKVGEHLNLVIPFQGKPRPQVVWTKG GAPIDTSHVHVRTSDFDTVFFVRQAARSDSGEYELTVQIENMKDTATVCIR VVEKAGPPQNVMVKEVWGTNALVEWQPPKDDGNSEITGYFVQKADKKT MEWFTVYERNRHTSCTVSDLIMGNEYYFRVYSENVCGLSDLPGVSKNTARI VKTGMTLKLPEYKEHDFRTPPKFLTPLPDRVVVAGYAAALNCAVRGHPKP KVVWMKNKMEIREDPKFLMTNQQGVLTLNIRRPSPFDSGTYSCRAVNELG EALAECKLEVRVPQ
132	Bison A0A6P3GV86 (Myomesin-2)	MSLVAVPFYQKRHKHFDQSYRNIQTRYLLDEYSAKKKRASAQSSLQRSVT RKSSSQRESSRAALGESTLGETTLGETSLGDTTLGRTTCRLCAKRMSSALEE EAQERKRRYQSMVAAYGEILVRLRSHTVWERMSVKLCFTVQGFPTPVVQ WYKDGSLICQAGEPGKYRIDSKYGVHTLEINRADFDDTATYSAVATNVHG QVSTNAAVVVRRFRGDEEPFHSVGLPIGLPLSSVIPYTHFDVQFIEKFGVTFR REGETLTLKCTLLVTPDLKRVQPRAEWYRDDVLLKESKWTKMFFGEGQAS LSFSHLNKDDEGLYTLRIVSRGGITDHSAFLFVRDADPLVTGAPGAPMDVH CHDANRDYVIVTWKPPNTTKESPVIGYFIDRCEVGTDNWIQCNDAPVKICK YPVTGLFEGRSYIFRMRAVNSAGISRPSRVSEAVAALDPVDLRRLQAVHLE GEKDIVIYQEELEGEVQIPGPPTNVHASETSRTYVVLSWDPPEPRGKEPLMY FIEKSMVGSGSWQRVNAQTAVRSPRYAVFDLAEGKPYVFRVLSANKHGLS DPSEITAPIQAQDTIVVPSAPGRVLASRNTRTSVVVQWDRPKHEEDLLGYY VDCSVAGSNVGEPCNHKPIGYNRFVVHGLTTGEQYIFRVKAVNAVGTSENS QESDVIKVQAALTVPSHPYGITLLNCDGHSMILGWKVPKFSGGSPILGYYV DKREAHHKNWHEVNSSPLKERILTVEGLTEGSLYEFKIAAANMAGIGQPSD PSELFKCEAWTMPEPGPAYDLTFCEVRDTSLVLLWKPPVYSGSSPVSGYFV DYKEEDSGEWLTVHETATPHRYLKVCDLHQGKTYVFRVRAVNASGVGRP SDTSEPVLVEARPGTKEVSAGVDEEGNVYLSFDCPEVTDASHFTWCKSYED IADDDRFKVETAGDHSKLYFKNLDKEDLGTYSVSVSDTDGVSSSFVLDEEE LERLKALSNEIKNPTIPLKSELAYEIFDKGQVRFWLQAEHLSPDSSYRFIIND REVADSETHRIKCDKSTGIIEMVMDRFTIDNEGTYTVQIQDGKAKSQSSLVL IGDAFKAVLKEAEFQRKEFLRKQGPHFAEYLHWDVTEECEVRLVCKVANT KKETVFKWLKDDVLYETEKMPDLEKGVCELLIPKLSKKDHGEYKATLKDD RGQDVSTLEIAGKVYEDMILAMSRVCGASASPLKILCTPEGIRLQCFMKYFT EEMKVNWYHKDAKITSSEHMRIGGSEQMAWLQICEPTEKDKGKYTFEILD DKKNHQRSLDLSGQAFDEAFAEFQQLKAAAFAEKNRGKVIGGLPDVVTIM EGKTLNLTCTVFGNPDPEVVWFKNDKDIELSDHFSVKVEQAKYVSMTIKG

TABLE6-continued

Bison Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		VTSEDSGKYSIHVKNKYGGEKIDVTVSVYKHGEEIPDVVLPQQAKPKLLPA AAPSPAH
133	Bison A0A6P3HSX8 (Phosphopyruvate hydratase)	MRTLRAMAMQKIFAREILDSRGNPTVEVDLHTAKGRFRAAVPSGASTGIYE ALELRDGDKSRYLGKGVLKAVEHINKTLGPALLEKKLSVVDQEKVDKFMI ELDGTENKSKFGANAILGVSLAVCKAGAAEKGVPLYRHIADLAGNPELILP VPAFNVINGGSHAGNKLAMQEFMILPVGASSFREAMRIGAEVYHHLKGVIK AKYGKDATNVGDEGGFAPNILENNEALELLKTAIQAAGYPDKVVIGMDVA ASEFYRNGKYDLDFKSPDDPARHISGEKLGELYKNFIKNYPVVSIEDPFDQD DWATWTSFLSGVNIQIVGDDLTVTNPKRIAQAVEKKACNCLLLKVNQIGSV TESIQACKLAQSNGWGVMVSHRSGETEDTFIADLVVGLCTGQIKTGAPCRS ERLAKYNQLMRIEEALGDKAVFAGRKFRNPKAK
134	Bison A0A6P3H0Z6 (Myosin regulatory light chain 2)	MAPKKAKRRAAAEGGSSSVFSMFDQTQIQEFKEAFTVIDQNRDGIIDKEDL RDTFAAMGRLNVKNEELDAMMKEASGPINFTVFLNMFGEKLKGADPEDVI TGAFKVLDPEGKGTIKKQFLEELLTTQCDRFSHEEIKNMWAAFPPDVGGNV DYKNICYVITHGDAKDQE
135	Bison A0A6P3HQ42 (Myosin-1)	MSSDQEMAIFGEAAPYLRKSEKERIEAQNKPFDAKTSVFVADPKESFVKAT VQSREGGKVTAKTEAGATVTVKEDQVFPMNPPKFDKIEDMAMMTHLHEP AVLYNLKERYAAWMIYTYSGLFCVTVNPYKWLPVYNAEVVTAYRGKKRQ EAPPHIFSISDNAYQFMLTDRENQSILITGESGAGKTVNTKRVIQYFATIAVT GEKKKEEPTSGKMQGTLEDQIISANPLLEAFGNAKTVRNDNSSRFGKFIRIH FGTTGKLASADIETYLLEKSRVTFQLKAERSYHIFYQIMSNKKPELIEMLLIT TNPYDYAYVSQGEITVPSIDDQEELMATDSAIEILGFTSDERVSIYKLTGAV MHYGNLKFKQKQREEQAEPDGTEVADKAAYLQGLNSADLLKALCYPRVK VGNEFVTKGQTVEQVYNAVGALAKAVYDKMFLWMVARINQQLDTKQPR QYFIGVLDIAGFEIFDFNSLEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGI EWEFIDFGMDLAACIELIEKPMGIFSILEEECMFPKATDMSFKNKLYEQHLG KSNNFQKPKPAKGKAEAHFSLIHYAGTVDYNITGWLDKNKDPLNETVVGL YQKSSVKTLALLFSGPASGEAEGGPKKGGKKKGSSFQTVSALFRENLNKLM TNLRSTHPHFVRCIIPNETKTPGAMEHELVLHQLRCNGVLEGIRICRKGFPSR ILYADFKQRYKVLNASAIPEGQFIDSKKASEKLLASIDVDHTQYKFGHTKVF FKAGLLGLLEEMRDEKLAQLITRTQARCRGFLARVEYQKMVERRESIFCIQ YNVRAFMNVKHWPWMKLYFKIKPLLKSAETEKEMANMKEEFEKTKEELA KSEAKRKELEEKMVTLTQEKNDLQLQVQSEADALADAEERCDQLIKTKIQL EAKIKEVTERAEDEEEINAELTAKKRKLEDECSELKKDIDDLELTLAKVEKE KHATENKVKNLTEEMAGLDETIAKLTKEKKALQEAHQQTLDDLQAEEDK VNTLTKAKTKLEQQVDDLEGSLEQEKKLRMDLERAKRKLEGDLKLAQEST MDIENDKQQLDEKLKKKEFEMSNLQSKIEDEQALAMQLQKKIKELQARIEE LEEEIEAERASRAKAEKQRSDLSRELEEISERLEEAGGATSAQIEMNKKREA EFQKMRRDLEEATLQHEATAAALRKKHADSVAELGEQIDNLQRVKQKLEK EKSEMKMEIDDLASNMETVSKAKGNLEKMCRALEDQLSELKTKEDEQQRL INDLTTQRARLQTESGEFSRQLDEKDALVSQLSRGKQAFTQQIEELKRQLEE EIKAKSALAHALQSARHDCDLLREQYEEEQEGKAELQRAMSKANSEVAQ WRTKYETDAIQRTEELEEAKKKLAQRLQDAEEHVEAVNAKCASLEKTKQR LQNEVEDLMIDVERTNAACAALDKKQRNFDKILSEWKQKYEETHAELEAS QKESRSLSTELFKIKNAYEESLDQLETLKRENKNLQQEISDLTEQIAEGGKRI HELEKVKKQVEQEKSEIQAALEEAEASLEHEEGKILRIQLELNQVKSEIDRKI AEKDEEIDQLKRNHIRVVESMQSTLDAEIRSRNDAIRLKKKMEGDLNEMEI QLNHANRMAAEALKNYRNTQAILKDTQIHLDDALRGQEDLKEQLAMVER RANLLQAEIEELRATLEQTERSRKIAEQELLDASERVQLLHTQNTSLINTKK KLETDITQIQGEMEDIIQEARNAEEKAKKAITDAAMMAEELKKEQDTSAHL ERMKKNLEQTVKDLQHRLDEAEQLALKGGKKQIQKLEARVRELEGEVESE QKRNVEAVKGLRKHERRVKELTYQTEEDRKNILRLQDLVDKLQAKVKSY KRQAEEAEEQSNVNLSKFRKLQHELEEAEERADIAESQVNKLRVKSREVHT KIISEE
136	Bison A0A6P3H861 (Nebulin isoform X11)	MADDEEYEEVVEYYTEETVYKEVPGETITGFYETTTTRTSDYEQSETSRPAL AQPVPEKPVERKRVIRKKVDPSKFMTPYIAHSQKMQNLFSTNKYKENYEK AKGKPYAITTDTPELRRIKKVQDQLSEVKYRVDGDVAKTICHVDEKAKDIE HAKKVSQQVSKVLYKQNWEDTKDKYLLPPDAPELVQAIKNTAMFSKKLY TEDWEADKTMFYPYNDSPELRRVAQAQKALSDIAYKKGLAEQQTQFTSLP DPPDIEFAKKVTNQVSKQKYKEDYENKVKGKWSETPCFEIATARMNSNNIS TKKYQEDFEHMKDQIYFMQTETPEYKVNKQAGVAASKVKYKQDYEKNK GKADYNVLPASENPLLRQLKAAGDALSDKLYKENYEKTKAKSINYCETPK FKLDTVLHNFSSDTKYKDSYLKDILGHYVGSYEDPYHTHCMRVSAQNSDK NYKAEYEEDRGKGFFPQTITQEYEAIKKLDQCKDHTYKVHPDKTKFTQVT DSPVLMQAQVNSKQLSDLNYKAKHENEKFKCHIPPDTPAFIQHKLNAYNLS DNIYKHDWEKTKAKKFDIKVDAIPLLAAKANTKIASDVMYKKDYEKSKGK MIGALSINDDPKMLHSLKTAKNQSDRLYKENYEKTKAKSMNYCETPKYQL

TABLE6-continued

Bison Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		DTLLKNFSEAKYKDSYVQNVLGHYIGSFEDPYQVHCLKISAQNSDKNYKA EYEEDKGKCYFPQTITQEYEAIKKLDQCKDHTYKVHPDKTKFTAVTDTPVL LQAQLNTKQLSDLNYKAKHEGEKFKCHIPADAPQFIQHRVNAYNLSDNIYK HDWEKTKAKKFDIKVDAIPLLAAKANTKIASDVMYKKDYEKSKGKMIGAL SINDDPKMLHSLKTAKNQSDHEYRKDYEKSKTIYTAPLDMLPLTHAKKSQ AIASDVDYKHLLHNYSYPPDSVNVDLAKKAYALQSDVEYKADYNSWMKG CGWMPFGSLEMEKAKRASDILNEKKYRQHPDTLKFTSIEDAPIIVQSKINQA QRSDVAYKAKGEEVIHKYSLPADLPQFIQAKVNAYNISENLYKADLKDLSK KGYDLRIDAIPIKAAKAARQAASDVQYKKDYEKAKGKMVGFQSLQDDPKL VHYMNVAKIQSDREYKKAYEKTKTRHNTPHDMVNIVAAKKAQDVASNV NYKHSLHHYTYLPDAMDLELSKNMMHIRSDNVYKEDYNNWMKGIGWIPI GSLEVEKVKKAGDALNEKKYRQHPDTLKFTSIVDSPVMVQAKQNTQQVSD ILYKAKGEDVKHKYTMSPDLPQFLQAKCNAYNLSDVCYKRDWHDLIAKG TNVLGDAIPITAAKASRNIASDYKYKEAYEKAKGKQVGFRSLQDDPKLVH YMNVAKLQSDREYKKNYENTKTSYHTPGDMVSITAAKMAQDVATNVNY KQPIHHYTYLPDALSLEHIRNVNQIQSDNVYKDEYNHFFKGMGWIPIGSLEV EKVKKAGDALNEKKYRQHPDTIKFTSVPDSMGMVLAQHNTKQLSDLNYK VEGEKVKHKYTMDPDVPQFIQAKVNAYNMSDSHYKADWKKTLAKGYDL RPDAIPIVAAKSSRNIASDFKYKEAYEKTKGKQIGFRSLQDDPKLVHFMNV AKMQSDREYKKGYEARKTKYHTPLDMLSVTAAKKSQEVATNTNYKQPFH HYTLLPDALNVEHSRNAMQIQSDNLYKSDFTNWMKGIGWLPLESLEVEKA KKAGEILSEKKYRQHPEKLKFTYAMDTMEQALNKSNKLIMDKRLYTEKW NKDKTTIHVMPDTPDILLSRVNQITMSDKLYKAGWEEEKKKGYDLRPDAIS IKAARASRDIASDYKYKQAYEQAKGRQIGFRSLEDDPKLVHFMQVAKMQS DREYKKAYEKSKTSFHTPVDMLSVVAAKKSQEVATNANYRNVIHTYNML PDAMSLELAKNMMQIQSDNQYKADYADFMKGIGWLPLGSLEAEKNKKA MEIISEKKYRQHPDTLKYSTLMDSMNMVLAKNNAKIMNEHLYKQAWEAD KTKVHIMPDIPQIILAKANAINISDKLYKLSLEEAKKKGYDLRTDAIPIKAAK ASRDIASDYKYKHSYEKERGKMVGFRSLEDDPKLVHSMQVAKMQSDREY KKNYEKTKTSYHTPADMLSVTAAKDAQANITNTNYKHLIHKYILLPDAMNI QLSKNMNRIQSDNEYKQDYNEWYKGLGWSPAGSLEVEKAKKATEYASDQ KYRQHPSNFQFTKLNDSMDMVLAKQNAHTMNKYLYTIDWNKDKTKIHV MPDTPDILQAKQNQTMYSQKLYKLGWEEALKKGYDLPVDAISVQAAKAS RDIVSDFKYKQGYRKQIGHHIGFRSLQDDPKLVLSMNVAKMQSEREYKKD FEKWKTKFSSPVDMLGVVLAKKCQALVSDVDYKNYLHGWTCLPDQNDVI HAKKAYDLQSENLYKSDLEWLKGIGWSPLGSLEAEKNKRASEIISEKKYRQ PPDRNKFTSIPDAMDIVLAKTNAKNRSDILYREAWDKDKTQIHIMPDTPDIIL AKANLINTSDKFYRMGYEELRKKGYDLPVDAIPIKAAKASREIASEYKYKE GFRMQLGHHIGARNIKDDPKMMWSMHVAKIQSDREYKKDFEKWKTKFSS PVDMLGVVLAKKCQTLVSDIDYKNYLHQWTCLPDQSDVIHARRAYDLQS DNLYKSDLQWLRGIGWLPSGSLEDEKNKRASQILSDHVYRQHPDQFKFSSL MDSIPMVLAKNNAITMNHRLYTEAWDKDKTTVHIMPDTPEVLLAKQNQIN YSEKLYKLGLDEAKRKGYDMRIDAIPIRAAKASRDIASEFKYKEGYRRQLG HHIGARAIHDDPKMMWSMHVAKIQSDREYKKDFEKWKTKFSSPVDMLGV VLAKKCQTLVSDIDYKNYLHQWTCLPDQSDVIHARRAYDLQSDNVYKSDL QWLRGIGWVPIGSLDVEKSKRASEILSDKLYRQPPDKFKFTSVTDSLEQVLA KNNAINMNKRLYTEAWDKDKTQVHIMPDTPEITLARTNKVNYSENLYKLA HEEAKKKGYDLRSDAIPIVAAKASRDIISDYKYKDGYRKQLGHHIGARDIK DDPKMMWSMHVAKIQSDREYKKDFEKWKTKFSSPVDMLGVVLAKKCQT LVSDIDYKNYLHQWTCLPDQSDVIHARRAYDLQSDNVYKSDLQWLRGIG WVPIESVDVVKCKRAAEILSDNLYRQPPDTLKFTSVPDSLEQVLAKNNAIN MNKRLYTEAWDKDKTQIHIMPDTPEIMLARQNKLNYSETLYKLANEEAKK KGYDLRSDAIPIVAAKASRDIISDYKYKDGYRKQLGHHIGARDIKDDPKMM WSMHVAKIQSDREYKKDFEKWKTKFSSPVDMLGVVLAKKCQTLVSDIDY RNYLHQWTCLPDQSDVIHARQAYDLQSDNIYKSDLQWLRGIGWVPIGSVD VVKCKRAAEILSDNLYRQRPDTLKFTSVPDSLEQVLAKNNAINMNKRLYTE AWDKDKTQIHMMPDTPEITLARQNKINYSENLYRQAMEEAKKEGYDLRSD AIPIVAAKASREIASDYKYKEAYRKQLGHHIGARAIHDDPKMMWSVHVAK MQSDREYKKDFEKYKTRFSSPVDMLGIVLAKKCQTLVSDVDYKHPLHEWT CLPDQNDIIHARKAYDLQSDNLYKSDLEWLKGIGWVPIGSVEVLKAKRAGE ILSDNIYRQRPDTLKFTSVTDSPEQVLAKNNAINMNKRLYTEAWDNDKKTI HVMPDTPEIMLAKLNRINYSDKLYKLALEESRKEGYDLRLDAIPIQAAKAS REIASDYKYKEGYRKQLGHHIGARNIKDDPKMMWSIHAGKIQSDLEYKKD FEKWKTKFSSPVDMMGLVQAKKCQILVSDIDYKNLLHEWTCLPDQNDIIQ ARKAYDLQSDAIYKADLEWLRGIGWVPIDSVGVEHAKRAGEILSERKYRQP AVQLKFTSITDTPEIVLAKNNALNVSKHLYTEAWDADKTSIHVMPDTPEILL AKSNSANISHKLYTKGWDESKMKDYDLRADAISIKSAKASRDIASDYKYKE AYEKHKGHHIGARSVEDDPRIMCAMNAGRIQSEREYKKEFQKWKTKFSSP VDMLGILLAKKCQTLVSDIDYRNYLHHWTCLPDQNDIIQARKAYDLQSDAI YKADLEWLRGIGWMPEGSPEVLRVKNAQHIFRDSVYRTPVVKLKYTSIVD TPEVVLAKSNAENISIPKYREVWDKDKTSIHIMPDTPEINLARTNALNVSNK LYREGWDEMKMSCDVRLDAIPIQAAKASREIASDYKYKLDHEKQKGHYV

TABLE6-continued

Bison Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		GTRTARDDNKIRWALIAGKIQNEREYRLHWAKWKSKFQSPPDMLSIEHSK NSQTLVSDIDYRHYLHQWTCMPDQNDVIQARKAYDLQSDAIYKADLEWL RGIGWMPADSVSVNHAKHMADIFNEKKYRTKIETLSFTPVDDRVDYVTAK HSGEILNDIKYRKDWNDTKSKYTLTETPQLHTAQEAARILDQYLYKESWEK QKATGYILPPDAVPFVHAHHSGDVQSELKYKAEHVKQKGHYVGVPTMRD DPKLVWFEHAGQIQNDRLYKENYHKTKAKIHIPPDMVSVLAAKEGQALAS DIDYRNYLHQWICHPDQNDVIQARKAYDLQSDNIYKADLEWLRGIGWIPL DSVDHVRVTRNQEMMNQIKYKKDALANYPNFTSVVDPPEIVLAKINSVNQ SDVKYKETFNKLIKGKYIFSPDTPYITHSKDMEKLYSTILYKRAWDGTKAY GYTLDERYIPIVGAKHADFVNSELKYKETYEKLKGHYLAGKEISEFPNVVH CLDFQKMRSLLNYRRHYEDTKANVHIPQDMMNHVLAKRCQYILSDLEYR HYFHQWTSLPEEPNVVRVRHAQEILSDNVYKDDLNWLKGIGCYVWDTPQI LHAKKSYDLQSQILYTAAGKENLKNYNLVTDTPLYVTALQSGINASEVKY KENYHQIKDKYTTVLETVDYDRIKNLKDLFSSNLYKEAWDKVKATSYILPP NTVSLTHAKNQKYMASHIKYREEYEKFKALYTLPRSVEDDPNTARCLRVG KFNIDRLYRSVYEKNKMKIHIVPDMVEMVTAKDSQKKVSEIDYRLHLHEWI CHPDLQVNSHVRKVTDQISDIVYKDDLTWLKGIGCYVWDTPEILHAKHAY DLRNDIKYKAHVLKTRNNYKLVTDTPVYVQAVKSGKQLSDAVYHYDYVH SIRGRVAPTTKTVDLDRALHAYKLQSENLYRKAGLHALPTGYRLPVDTPHF KHTKDTRYMSSYFKYKEVYEHMKAYGYTLGPNDVPFVNVRRVNNITSERL YRQLYHKLKDKIHTTPDTPEIRQVKKTQEAVSELIYKSDFFKMQGHMISLPY TPQVLHCRYVGDITSDIKYKEDLQVLRGMGCFLYDTPDMVRSRHLRKLWS HYLYTDKARKMRDKYKVVLDTPEYRKVQELKTHLSELVYRAAGRKQKSIF TSVPDTPDLTRAKRGQKLQSQYLYVELATKERPHHHAGNQTTALKHARHV KDMVSENKYKIQYEKMKDKYTPVPDTPILIRAKRAYWNASDLRYKETFQK TKGKYHTVKDALDIVYHRTVTDHVSKIKYKENYMSQLGIWRSIPDRPEHFH HRAVTDAVSDVKYKQDLTWLKGIGCYAYDTPDFTLAEKNKTLYSKYKYK EVFERTKSNFKYVADCPINRHFKFATQLMNEKKYRADYEQRKDKYHLVVD EPRHLLAKIAGDQISQIKYRKNYEETKDKFTSIVDTPEHLRTTKVNKQISDIL YKLEYNKAKPRGYTTIHDTPMLLHVRKVKDEVSDLKYKEVYQRTKSNCTI EPDAVHIKAAKDAYKVNTNLDYKKKYEATKAYWKWTPDRPDFIQAAKST LQQSDFEYKLDREYLKGCKLSVTDDKDMVLALKNSIIESDLKYKEKHVKE RGSCHAVPDTPQILLAKTVSSLVSENKYKSYVKKHLAQGSYTTLPETRDTIH VKEVTKNVSDTNYKKKFVKEKGKSNYSIMLEPPDVKHAMDVAKKQSNVA YKKDAKENLHYTTVADRPDIKKATQAAKQASEVEYRAKHRKEGSHGLSM LGRPDIEMAKKAAKLSSQVQYRENFNKEKGKTPKYNPKDSQLYKVMKDA NTLASEVKYKADLKKLHKPVTDMKESLIMNHVLNTSHLASSYQYKKNYEK SKGHYHTIPDNLEQLHLKEATELQSIVKYKEKYEKERGKPMLDFETPTYITA KESQQMQSGKEYRKDYEESIKGRNLTGLEVTPALLHVKYATKIASEKEYRK DLEESIRGKGLSEMEDTPDMLRAKNATQILNEKEYKRDLELEVKGRGLNA MANETPDFLRARNATDIASQIKYKQSAEMEKANFTSVVDTPEIIHAQQVKN LSSQKKYKEDAEKCMSYYETVLDTPEMQRVRENQKNFSLLQYQYDLKNS KGKITVVQDTPEILRVKENQKNFSSVLYKEDVSPGTAIGKTPEMMRVKQTQ DHISSVKYKEVIGQGTPIPDLPEVKRVKQTQKHISSVMYKENLGTGIPTPVTP EIERVKRNQENFSSVLYKENLGKGTPTPITPEMERVKRNQENFSSVLYKEN MGKGTPLPVTPEMERIKHNQENISSVLYKENVGKGTPTPVTPEMQRVKRNQ ENISSVLYKENLGKATPTPFTPEMERVKRNQENFSSVLYKENMRKATPTPV TPEMERAKRNQENISSVLYSDSFRKQIQGKAAYVLDTPEMRRVRETQRHIS TVKYHEDFEKHKGCFTPVVTDPITERVKKNTQDFSDICYRGIQRKVVEMEQ KRNDQDQETITGLRVWRTNPGSVFDYDPAEDNIQSRSLHMINAQAQRRSRE QSRSASGLSISGGEEKSEHSEAAHLSTYSDGGVFFSAASTGYKHARTTELPQ QRSSSVATQQTTVSSIPSHPSTAGKIFRAMYDYMAADADEVSFKDGDAIVN VQAIDEGWMYGTVQRTGRTGMLPANYVEAI
137	Bison A0A6P3HPK5 (Myosin-4)	MSSDQEMAIFGEAAPYLRKSEKERIEAQNKPFDAKTSVFVVDPKESFVKAT VQSREGGKVTAKTEAGATVTVKEDQIFPMNPPKYDKIEDMAMMTHLHEPA VLYNLKERYAAWMIYTYSGLFCVTVNPYKWLPVYNAEVVTAYRGKKRQE APPHIFSISDNAYQFMLTDRENQSILITGESGAGKTVNTKRVIQYFATIAVTG EKKKEEPTSGKMQGTLEDQIISANPLLEAFGNAKTVRNDNSSRFGKFIRIHF GTTGKLASADIETYLLEKSRVTFQLKAERSYHIFYQIMSNKKPELIEMLLITT NPYDYAFVSQGEITVPSIDDQEELIATDSAIDILGFTSDERVSIYKLTGAVMH YGNLKFKQKQREEQAEPDGTEVADKAAYLQSLNSADLLKALCYPRVKVG NEFVTKGQTVQQVYNAVGALAKAVYEKMFLWMVTRINQQLDTKQPRQY FIGVLDIAGFEIFDFNSLEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEW EFIDFGMDLAACIELIEKPMGIFSILEEECMFPKATDTSFKNKLYEQHLGKSN NFQKPKPAKGKAEAHFSLVHYAGTVDYNIAGWLDKNKDPLNETVVGLYQ KSAMKTLAFLFAGGPSAEEGGGGKKGGKKKGSSFQTVSALFRENLNKLMT NLRSTHPHFVRCIIPNETKTPGAMEHELVLHQLRCNGVLEGIRICRKGFPSRI LYADFKQRYKVLNASAIPEGQFIDSKKASEKLLGSIDIDHTQYKFGHTKVFF KAGLLGILEEMRDEKLAQLITRTQAICRGFLMRVEFRKMMERRESIFCIQYN VRAFMNVKHWPWMKLYFKIKPLLKSAETEKEMANMKEEFEKTKEELAKS EAKRKELEEKMVTLMQEKNDLQLQVQSEADGLADAEERCDQLIKTKIQLE

TABLE6-continued

Bison Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		AKIKELTERAEDEEEINAELTAKKRKLEDECSELKKDIDDLELTLAKVEKEK HATENKVKNLTEEMAGLDENIAKLTKEKKALQEAHQQTLDDLQAEEDKV NTLTKSKTKLEQQVDDLEGSLEQEKKLRMDLERAKRKLEGDLKLAQESTM DIENDKQQLDEKLKKKEFEMSNLQSKIEDEQALAMQLQKKIKELQARIEEL EEEIEAERASRAKAEKQRSDLSRELEEISERLEEAGGATSAQIEMNKKREAE FQKMRRDLEEATLQHEATAAALRKKHADSVAELGEQIDNLQRVKQKLEKE KSEMKMEIDDLASNMETVSKAKGNLEKMCRTLEDQLSEVKSKEEEHQRLI NELSAQKARLHAESGEFSRQLEEKDALVSQLSRGKQAFTQQIEELKRQLEE ETKAKSALAHALQSARHDCDLLREQYEEEQEGKAELQRALSKANSEVAQ WRTKYETDAIQRTEELEEAKKKLAQRLQDAEEHVEAVNAKCASLEKTKQR LQNEVEDLMLDVERSNAACAALDKKQRNFDKVLSEWKQKYEETQAELEA SQKESRSLSTELFKVKNACEESLDQLETLKRENKNLQQEISDLTEQIAEGGK HIHELEKVKKQIEQEKGELQAALEEAEASLEHEEGKILRIQLELNQVKSEIDR KIAEKDEEIDQLKRNHLRVMESTQSTLDAEIRSRNDALRIKKKMEGDLNEM EIQLNHANRQATEAIKNLRNTQGVLKDTQLHLDDAIRGQDDLKEQLAMVE RRANLMQAEIEELRASLEQTERSRRVAEQELLDASERVQLLHTQNTSLINTK KKLETDITQIQGEMEDILQEARNAEEKAKKAITDAAMMAEELKKEQDTSA HLERMKKNMEQTVKDLQHRLDEAEQLALKGGKKQIQKLEARVRELENEV ENEQKRNAEAVKGLRKHERRVKELTYQTEEYRKNVLRLQDLVDKLQSKV KAYKRQAEEAEEQSNVNLAKFRKLQHELEEAEERADIAESQVNKLRVKSR EVHTKVISEE
138	Bison A0A6P3HBK1 (Serum albumin)	MKWVTFISLLLLFSSAYSRGVFRRDTHKSEIAHRFKDLGEEHFKGLVLIAFS QYLQQCPFDEHVKLVNELTEFAKTCVADESHAGCEKSLHTLFGDELCKVA SLRETYGNMADCCEKQEPERNECFLSHKDDSPDLPKLKPDPNTLCDEFKAD EKKFWGKYLYEIARRHPYFYAPELLYYANKYNGVFQECCQAEDKGACLLP KIETMREKVLTSSARQRLRCASIQKFGERALKAWSVARLSQKFPKAEFVEV TKLVTDLTKVHKECCHGDLLECADDRADLAKYICDNQDTISSKLKECCDKP LLEKSHCIAEVEKDAVPENLPPLTADFAEDKDVCKNYQEAKDAFLGSFLYE YSRRHPEYAVSVLLRLAKEYEATLEECCAKDDPHACYSTVFDKLKHLVDE PQNLIKQNCDQFEKLGEYGFQNALIVRYTRKVPQVSTPTLVEVSRSLGKVG TRCCTKPESERMPCTEDYLSLILNRLCVLHEKTPVSEKVTKCCTESLVNRRP CFSALTPDETYVPKAFDEKLFTFHADICTLPDTEKQIKKQTALVELLKHKPK ATEEQLKTVMENFVAFVDKCCAADDKEACFAVEGPKLVVSTQTALA

[0049] An animal protein can comprise a deer protein (Table 7). A deer protein can comprise a deer muscle protein. A deer protein can comprise a deer blood protein. A deer protein can comprise a deer albumin protein. A deer protein can comprise a sequence that has at least or about 70%, at least or about 75%, at least or about 80%, at least or about 90%, at least or about 91%, at least or about 92%, at least or about 93%, at least or about 94%, at least or about 95%, at least or about 96%, at least or about 97% at least or about 98% at least or about 99% or about 100% sequence identity to any one of SEQ ID NO: 139-147.

[0050] A food product may comprise a fragment or a truncation of a deer protein. The food product may include a fragment or truncation of any one of SEQ ID Nos: 139-147. A sequence of the fragment or truncation may be identical or substantially identical to a portion of any one of SEQ ID Nos: 139-147. A sequence of the fragment or truncation may be at least about 70%, 80%, 90%, 95%, 98%, or more identical to a portion of a sequence of any one of SEQ ID Nos: 139-147. A length of the fragment or truncation may be at least about 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of a length of any one of SEQ ID Nos: 139-147. A length of the fragment or truncation may be less than or equal to about 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, or less than a length of any one of SEQ ID Nos: 139-147.

TABLE7

Deer Proteins		
SEQ ID NO	Description	Amino Acid Sequence
139	Deer A0A212DH10 (ACTN3)	MLDAEDIVNTPKPDEKAIMTYVSCFYHAFAGAEQAETAANRICKVLAVNQ ENEKLMEEYEKLASELLEWIRRTVPWLENRVGEPSMSAMQRKLEDFRDYR RLHKPPRVQEKCQLEINFNTLQTKLRLSHRPAFMPSEGKLVSDIAKAWRGL EQVEKGYEDWLLSEIRRLQRLQHLAEKFQQKASLHEAWTRGKEEMLSQRD YETASLQEVRALLRRHEAFESDLAAHQDRVEHIAALAQELNELDYHEAAS VNSRCQAICDQWDNLGTLTQKRRDALERMEKLLETIDQLQLEFARRAAPF NNWLDGAVEDLQDVWLVHSVEETQSLVTAHDQFKATLPEADRERGAILGI QGEIQKICQTYGLRPSSTNPYIALTPQDINTKWDTVRKLVPSRDQMLQEELT RQQVNERLRRQFAAQANAIGPWIQGKVEEVGRLAAGMAGSLEEQMAGLR QQEQNIINYKSNIDRLEGDHQLLQESLVFDNKHTVYSMEHIRVGWEQLLTSI ARTINEVENQVLTRDAKGLSQEQLNEFRASFNHFDRKRNGMMEPDDFRAC

TABLE7-continued

Deer Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		LISMGYDLGEVEFARIMTMVDPNAAGVVTFQAFIDFMTRETAETDTAEQV VASFKILAGDKNYITAEELRRELPAEQAEYCIRRMAPYKGAGAPAGALDYV AFSSALYGESDL
140	Deer A0A212DGZ3 (Alpha-1,4 glucan phosphorylase)	MSRPLTDQEKRKQISVRGLAGVENVTELKKNFNRHLHFTLVKDRNVATPR DYYFALAYTVRDHLVGRWIRTQQHYYEKDPKRIYYLSLEFYMGRTLQNTM VNLALENACDEATYQLGLDMEELEEIEEDAGLGNGGLGRLAACFLDSMAT LGLAAYGYGIRYEFGIFNQKISGGWQMEEADDWLRYGNPWEKARPEFTLP VHFYGRVEHTSQGAKWVDTQVVLAMPYDTPVPGYRNNVVNTMRLWSAK APNDFNLKDFNVGGYIQAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYF VVAATLQDIIRRFKSSKFGCLDPVAIQLNDTHPSLAIPELMRILVDQERLEWE KAWEVTVKTCAYTNHTVLPEALERWPVHLIETLLPRHLQIIYEINQRFLNRV AAAFPGDVDRLRRMSLVEEGAVKRINMAHLCIAGSHAVNGVARIHSEILKK TMTSKSWSLINSRIRPMGSPLGAGWLCRIGEEYIADLDQLRKLLSYVDDESF IRDVAKVKQENKLKFSAYLEKEYKVHINPNSLFDIQVKRIHEYKRQLLNCL HAAPGYHMAKMIIKLITAIGDVVNHDPVVGDRLRVIFLENYRVSLAEKVIP AADLSEQISTAGTEASGTGNMKFMLNGALTIGTMDGANVEMAEEAGEENF FIFGMRVEDVERLDQRGYNAQEYYDRIPELRHVIDQLSSGFFSPKQPDLFKD IVNMLMHHDRFKVFADYEEYVKCQERVSALYKNPREWTRMVIRNIATSGK FSSDRTIAQYAREIWGVEPNRQRMPAPDEKI
141	Deer A0A212CZ49 (Fructose-bisphosphate aldolase)	MAQRKPEGSGFHMTCLSMALAYSFRPDANIQPHPQLGNTQKQTELGKELT GTMPHQYPALTPEQKKELCDIAHRIVAPGKGILAADESTGSIAKRLQSIGTE NTEENRRFYRQLLLTADDRVNPCIGGVILFHETLYQKADDGRPFPQVIKAK GGVVDKGVVPLAGTNGETTTQGLDGLSERCAQYKKDGADFAKWRCVLKI GEHTPSSLAIMENANVLARYASICQQNGIVPIVEPEILPDGDHDLKRCQYVT EKVLAAVYKALSDHHIYLEGTLLKPNMVTPGHACTQKYSHEEIAMATVTA LRRTVPPAVPGITFLSGGQSEEEASINLNAINKCPLLKPWALTFSYGRALQAS ALKAWGGKKENLKAAQEEYVKRALANSLACQGKYTPSGKAGAAASESLFI SNHAY
142	Fragment of Deer A0A212D3R8 (MYL1)	PAPAPPKEEKIDLSAIKIEFSKQQQDEFKEAFLLFDRTGECKITLNQVGDVLR ALGTNPTNAEVKKVLGNPSNEEMNAKKIEFEQFLPMLQAISNNKDQGTYE DFVEGLRVFDKEGNGTVMGAELRHVLATLGEKMKEEEVEALMAGQEDSN GCINYEAFVKHIMSN
143	Deer A0A212CCI7 (GAPDH)	MFVMGVNHEKYNNTLKIVSNASCTTNCLAPLAKVIHDHFGIVEGLMTTVH AITATQKTVDGPSGKLWRDGRGAAQNIIPASTGAAKAVGKVIPELNGKLTG MAFRVPTPNVSVVDLTCRLEKPAKYDEIKKVVKQASEGPLKGILGYTEDQV VSCDFNSDTHSSTFDAGAGIALNDHFVKLISWYDNEFGYSNRVVDLMVHM ASKE
144	Deer A0A212D927 (Phosphopyruvate hydratase)	MRTLRAMAMQKIFAREILDSRGNPTVEVDLHTAKGRFRAAVPSGASTGIYE ALELRDGDKSRYLGKGVLKAVEHINKTLGPALLEKASLGRLYLLPPAKFGA NAILGVSLAVCKAGAAEKGVPLYRHIADLAGNPELILPVPAFNVINGGSHA GNKLAMQEFMILPVGASSFKEAMRIGAEVYHHLKGVIKAKYGKDATNVG DEGGFAPNILENNEALELLKTAIQAAGYPDKVVIGMDVAASEFYRNGKYDL DFKSPDDPARHISGEKLGELYKNFIKNYPVVSIEDPFDQDDWATWTSFLSGV NIQIVGDDLTVTNPKRIAQAVEKKACNCLLLKVNQIGSVTESIQACKLAQSN GWGVMVSHRSGETEDTFIADLVVGLCTGQIKTGAPCRSERLAKYNQLMRIE EALGDKAVFAGRKFRNPKAK
145	Deer A0A212CY74 (MYLPF)	MFDQTQIQEFKEAPKKAKRRAAAEGGSSSVFSMFDQTQIQEFKEAFTVIDQ NRDGIIDKEDLRDTFAAMGRLNVKNEELDAMMKEASGPINFTVFLNMFGE KLKGADPEDVITGAFKVLDPEGKGTIKKQFLEELLTTQCDRFSQEEMRNM WAAFPPDVGGNVDYKNICYVITHGDAKDQE
146	Deer A0A212CNJ3 (Actin)	MCEEEDSTALVCDNGSGLCKAGFAGDDAPRAVFPSIVGRPRHQGVMVGM GQKDSYVGDEAQSKRGILTLKYPIEHGIITNWDDMEKIMFETFNVPAMYVA IQAVLSLYASGRTTGIVLDSGDGVTHNVPIYEGYALPHAIMRLDLAGRDLT DYLMKILTERGYSFVTTAEREIVRDIKEKLCYVALDFENEMATAASSSSLEK SYELPDGQVITIGNERFRCPETLFQPSFIGMESAGIHETTYNSIMKCDIDIRKD LYANNVLSGGTTMYPGIADRMQKEITALAPSTMKIKIIAPPERKYSVWIGGS ILASLSTFQQMWISKQEYDEAGPSIVHRKCF
147	Deer A0A212D5P0 (ALB)	MKWVTFISLLLLFSSAYSRGVFRRDTHKSEIAHRFKDLGEDNFQGLVLIAFS QYLQQCPFDEHVKLVKELTDFAKECVADESHAGCEKSLHTLFGDELCKVA SLRETYGDMADCCEKQEPERNECFLKHKDDNPDLPKLKPDPDTLCAEFKA DEKKFWGKYLYEVARRHPYFYAPELLYYANKYNGVFQECCQADDKAACL LPKIDTMREKVLTSSARQRLKCASIQKFGERALKAWTVALLSQKFPKADFA EVTKIVTDLTKVHKECCHGDLLECADDRADLAKYICDNQDTLSSKLKECC DKPVLEKSHCIAEIEKDAVPENLPPLTADFAEDKEVCKNYQEAKDVFLGTF

TABLE7-continued

Deer Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		LYEYARRHPEYAVSVLLRVAKEYEATLEDCCAKDDPHACYATVFDKLQHL VDEPQNLIKQNCDLFEKVGEYGFQNALIVRYTKKAPQVSTPTLVEVSRNLG RVGTKCCTKPESERMSCTEDYLSLVLNRLCVLHEKTPVSEKVTKCCTESLV NRRPCFSALTLDETYVPKAFDEKLFTFHADICTVPDAEKQIKKQTALVELLK HKPKATDEQLKTVMENFVAFVDKCCAAVDKEACFAVEPTLRVRGKKEK

[0051] An animal protein can comprise a chicken protein (Table 8). A chicken protein can comprise a chicken muscle protein. A chicken protein can comprise a chicken blood protein. A chicken protein can comprise a chicken albumin protein. A chicken protein can comprise a sequence that has at least or about 70%, at least or about 75%, at least or about 80%, at least or about 90%, at least or about 91%, at least or about 92%, at least or about 93%, at least or about 94%, at least or about 95%, at least or about 96%, at least or about 97% at least or about 98% at least or about 99% or about 100% sequence identity to any one of SEQ ID NO: 148-161.

[0052] A food product may comprise a fragment or a truncation of a chicken protein. The food product may include a fragment or truncation of any one of SEQ ID Nos: 148-161. A sequence of the fragment or truncation may be identical or substantially identical to a portion of any one of SEQ ID Nos: 148-161. A sequence of the fragment or truncation may be at least about 70%, 80%, 90%, 95%, 98%, or more identical to a portion of a sequence of any one of SEQ ID Nos: 148-161. A length of the fragment or truncation may be at least about 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of a length of any one of SEQ ID Nos: 148-161. A length of the fragment or truncation may be less than or equal to about 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, or less than a length of any one of SEQ ID Nos: 148-161.

TABLE8

Chicken Proteins		
SEQ ID NO	Description	Amino Acid Sequence
148	Chicken KCRM (Creatine kinase M-type)	MPFSSTHNKHKLKFSAEEEFPDLSKHNNHMAKVLTPELYKRLRDKETPSGF TLDDVIQTGVDNPGHPFIMTVGCVAGDEESYEVFKDLFDPVIQDRHGGYKP TDKHRTDLNHENLKGGDDLDPKYVLSSRVRTGRSIKGYSLPPHCSRGERRA VEKLSVEALNSLEGEFKGRYYPLKAMTEQEQQQLIDDHFLFDKPVSPLLLA SGMARDWPDARGIWHNDNKTFLVWVNEEDHLRVISMEKGGNMKEVFRRF CVGLKKIEEIFKKAGHPFMWTEHLGYILTCPSNLGTGLRGGVHVKLPKLSQ HPKFEEILHRLRLQKRGTGGVDTAAVGAVFDISNADRLGFSEVEQVQMVV DGVKLMVEMEKKLEQNQPIDDMIPAQK
149	Chicken (ACTN2 Alpha-actinin-2)	MNSMNQIETNMQYTYNYEEDEYMTQEEEWDRDLLLDPAWEKQQRKTFTA WCNSHLRKAGTQIENIEEDFRNGLKLMLLLEVISGERLPKPDRGKMRFHKI ANVNKALDYIASKGVKLVSIGAEEIVDGNVKMTLGMIWTIILRFAIQDISVE ETSAKEGLLLWCQRKTAPYRNVNIQNFHLSWKDGLGLCALIHRHRPDLIDY SKLNKDDPIGNINLAMEIAEKHLDIPKMLDAEDIVNTPKPDERAIMTYVSCF YHAFAGAEQAETAANRICKVLAVNQENERLMEEYERLASELLEWIRRTIPW LENRTPEKTMQAMQKKLEDFRDYRRKHKPPKVQEKCQLEINFNTLQTKLRI SNRPAFMPSEGKMVSDIAGAWQRLEQAEKGYEEWLLNEIRRLERLEHLAE KFRQKASTHEQWAYGKEQILLQKDYESASLTEVRAMLRKHEAFESDLAAH QDRVEQIAAIAQELNELDYHDAASVNDRCQKICDQWDSLGTLTQKRREAL ERTEKLLETIDQLHLEFAKRAAPFNNWMEGAMEDLQDMFIVHSIEEIQSLIS AHDQFKATLPEADGERQAILSIQNEVEKVIQSYSMRISASNPYSTVTVEEIRT KWEKVKQLVPQRDQSLQEELARQHANERLRRQFAAQANVIGPWIQTKME EIARSSIEMTGPLEDQMNQLKQYEQNIINYKHNIDKLEGDHQLIQEALVFDN KHTNYTMEHIRVGWELLLTTIARTINEVETQILTRDAKGITQEQMNDFRASF NHFDRRKNGLMDHDDFRACLISMGYDLGEAEFARIMSLVDPNGQGTVTFQ SFIDFMTRETADTDTAEQVIASFRILASDKPYILADELRRELPPEQAQYCIKR MPQYTGPGSVPGALDYTSFSSALYGESDL
150	Chicken E1BSN7 (Alpha-1,4 glucan phosphorylase)	MAAPLSDGERRKQISVRGIAGLGDVAEVRKSFNRHLHFTLVKDRNVATPR DYFFALAHTVRDHLVGRWIRTQQHYYERDPKRIYYLSLEFYMGRTLQNTM VNLGLQNACDEAIYQLGLDLEELEEIEEDAGLGNGGLGRLAACFLDSMATL GLAAYGYGIRYEFGIFNQKIIDGWQVEEADDWLRYGNPWEKARPEYMLPV HFYGRVEHSPEGVKWVDTQVVLAMPYDTPVPGYKNNTVNTMRLWSAKA PNDFNLQEFNVGDYIEAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYFV VAATLQDIIRRFKSSKFGCQDPVRTSFETFPDKVAIQLNDTHPALSIPELMRIL VDVEKVDWDKAWEITKQTCAYTNHTVLPEALERWPVSMFEKLLPRHLEII YALNQMHLDRVAALYPGDVDRLRRMSVIEEGDCKRINMAHLCVIGSHAV NGVARIHSDIVKNTVFKDFYELEPEKFQNKTNGITPRRWLLLCNPGLADVIA EKIGEDFVTDLSQLKKLLDFINNETFIRDVAKVKQENKLKFSAYLEERYKVK INPSSMFDIQVKRIHEYKRQLLNCLHAITLYNRIRSDPSKSFVPRTIMIGGKA

TABLE8-continued

Chicken Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		APGYHMAKMIIKLITSIGEVVNNDPYVGDKLKVIFLENYRVSLAEKVIPAAD LSQQISTAGTEASGTGNMKFMLNGALTIGTMDGANVEMAEEAGEENLFIFG MRVEEVEALDRRGYNAREYYDRIPELRQAIDQISSGFFSPRDPGCFRDVVN MLMHHDRFKVFADYEAYIKCQGQVDQLFMDPREWTKKVIRNIACSGKFSS DRTITEYAREIWGVEPSATKIPPPNLPRD
151	Chicken R4GM10 (Fructose-bisphosphate aldolase)	MTHQYPALTAEQKKELSDIALRIVAPGKGILAADESVGSMAKRLNQIGVEN TEENRRLYRQILFSADSRVKKCIGGVIFFHETMYQKADDGTPFVQMIKDKGI VVGIKVDKGVVPLAGTDGETTTQGLDGLSERCAQYKKDGADFAKWRCVL KISDNTPSALAIMENANVLARYASICQQNGIVPIVEPEILPDGDHDLKRCQY VTEKVLAAVYKALSDHHVYLEGTLLKPNMVTPGHSCPTKYSPEEIAMATV TALRRTVPPAVPGVTFLSGGQSEEEASINLNAINTCPLVRPWALTFSYGRAL QASALSAWRGQRDNANAATEEFVKRAEVNGLAALGKYEGSGDDSGAAGQ SLYVANHAY
152	Chicken MLE1 (Myosin light chain 1)	MAPKKDVKKPAAAAAPAPAPAPAPAPAPAKPKEPAIDLKSIKIEFSKEQQD DFKEAFLLFDRTGDAKITLSQVGDIVRALGQNPTNAEINKILGNPSKEEMNA KKITFEEFLPMLQAAANNKDQGTFEDFVEGLRVFDKEGNGTVMGAELRHV LATLGEKMTEEEVEELMKGQEDSNGCINYEAFVKHIMSV
153	Chicken G3P (Glyceraldehyde-3-phosphate dehydrogenase)	MVKVGVNGFGRIGRLVTRAAVLSGKVQVVAINDPFIDLNYMVYMFKYDST HGHFKGTVKAENGKLVINGHAITIFQERDPSNIKWADAGAEYVVESTGVFT TMEKAGAHLKGGAKRVIISAPSADAPMFVMGVNHEKYDKSLKIVSNASCT TNCLAPLAKVIHDNFGIVEGLMTTVHAITATQKTVDGPSGKLWRDGRGAA QNIIPASTGAAKAVGKVIPELNGKLTGMAFRVPTPNVSVVDLTCRLEKPAK YDDIKRVVKAAADGPLKGILGYTEDQVVSCDFNGDSHSSTFDAGAGIALND HFVKLVSWYDNEFGYSNRVVDLMVHMASKE
154	Chicken TPM1 (Tropomyosin alpha-1 chain)	MDAIKKKMQMLKLDKENALDRAEQAEADKKAAEERSKQLEDELVALQK KLKGTEDELDKYSESLKDAQEKLELADKKATDAESEVASLNRRIQLVEEEL DRAQERLATALQKLEEAEKAADESERGMKVIENRAQKDEEKMEIQEIQLKE AKHIAEEADRKYEEVARKLVIIEGDLERAEERAELSESKCAELEEELKTVTN NLKSLEAQAEKYSQKEDKYEEEIKVLTDKLKEAETRAEFAERSVTKLEKSID DLEDELYAQKLKYKAISEELDHALNDMTSI;
155	Chicken MYPC2 (Myosin-binding protein C)	MPEPSKAAPKKEAKKKEEKKEEKKEAPPPQEHKDEAPDDVHPPETPDPEGL FLSKPQNVMVESGRDVTVSARVAGAALPCAPAVKWFKGKWAELGDKSAR CRLRHSVDDDKVHTFELTITKVAMGDRGDYRCEVTAKEQKDSCSFSIDVEA PRSSEGNVLQAFKRTGEGKDDTAGELDFSGLLKKREVQVEEKKKKKDEDD QFPPEIWELLKGVTKKSEYERIAFQYGITDLRGMLKRLKKVHVEPKKSEAFI RKLDPAYQVDKGNKIKLVVELSDPDLPLKWYKNGQLLKPSTKYVFENVGL KRILTIHKCSLADDAAYECRVNDEKCFTEVFVKEPPVTVVRGLEDQQVVVG DRVVLEAEVSEEGAQVMWLKDGVDVTRDDAFKYRFKKDGKKHFLIINEA ELSDSAHYKIMTNGGESEAELSVEEKQLEVLQDMADLTVKASEQAVFKCE VSDEKVTGRWFRNGVEVKPSKRIHISHNGRFHKLVIDDVRPEDEGDYTFIPD GYALSLSAKLNFLEIKVEYVPKQEPPKIHLDCSGKAAENTIVVVAGNKVRL DVPISGEPAPTVTWKRGDQLFTATEGRVHIDSQADLSSFVIESAERSDEGRY CITVTNPVGEDSATLHVRVVDVPDPPQSVRVTSVGEDWAVLSWEAPPFDG GMPITGYLMERKKKGSMRWMKLNFEVFPDTTYESTKMIEGVFYEMRVFA VNAIGVSQPSLNTQPFMPIAPTSEPTHVVLEDVTDTTATIKWRPPERIGAGG VDGYLVEWCREGSNEWVAANTELVERCGLTARGLPTGERLLFRVISVNMA GKSPPATMAQPVTIREIVERPKIRLPRHLRQTYIRRVGEQVNLVIPFQGKPRP QVTWSREGGALPAEVQTRTSDVDSVFFIRSAARPLSGNYEMRVRIDNMEDC ATLRLRVVERPGPPQAVRVMEVWGSNALLQWEPPKDDGNAEISGYTVQK ADTRTMEWFTVLEHSRPTRCTVSELVMGNEYRFRVYSENVCGTSQEPATS HNTARIAKEGLTLKMVPYKERDLRAAPQFLTPLVDRSVVAGYTVTLNCAV RGHPKPKVTWLKNSVEIGADPKFLSRHGLGVLSLLIRRPGPFDGGTYGCRA VNEMGEATTECRLDVRVPQ
156	Chicken MPSF (M-protein)	MSSVAVPFYQRRHKHFDQSYRNIQTRYVLEEYAARKAASRQAAHYESTGL GKTTCRLCARRARSLAHEAMQESRKRTHEQKSHASDEKRIKFASELSSLER EIHMARHHAREQLDRLAIQRMVEENMALERHVVEEKISRAPEILVRLRSHT VWEKMSVRLCFTVQGFPSPVVQWYKNEELITPASDPAKYSVENKYGVHVL HINRADFDDSATYSAVATNIHGQASTNCAVVVRRFRESEEPHPAGIMPFHLP LSYDVCFTHFDVQFLEKFGVTFATEGETLTLKCSVLVTPELKRLRPRAEWY RDDVLIKDSKWTKLYFGEGQAALSFTHLNKDDEGLYTLRMVTKGGVNECS AFLFVRDADALIAGAPGAPMDVKCHDANRDYVIVTWKPPNTTSQNPVIGY FVDKCEVGLENWVQCNDAPVKICKYPVTGLYEGRSYIFRVRAVNSAGISRP SRVSEPVAALDPVDLERTQTVHVDEGRKIVISKDDLEGDIQIPGPPTNVHAS EISKTYVVLSWDPPVPRGREPLTYFIEKSMVGSGSWQRVNAQVAVKSPRYA VFDLAEGKPYVFRVLSANKHGISDPSEITEPIQPQDIVVVPSAPGRVVATRNT KTSVVVQWDKPKHEENLYGYYIDYSVVGSNQWEPANHKPINYNRFVVHG

TABLE8-continued

Chicken Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		LETGEQYIFRVKAVNAVGFSENSQESEAIKVQAALTCPSYPHGITLLNCDGH SMTLGWKAPKYSGGSPILGYYIDKREANHKNWHEVNSSVISRTIYTVEDLT EDAFYEFKIAAANVVGIGHPSDPSEHFKCKAWTMPEPGPAYDLTVCEVRNT SLVLLWKAPVYEGKSPITGYLVDYKEVDTEDWITANEKPTSHRYFKVTDLH QGHTYVFKVRAVNDAGVGKSSEISEPVFVEASPGTKEIFSGVDEEGNIYLGF ECKEATDASHFLWGKSYEEIEDSDKFKIETKGDHSKLYFKHPDKSDLGTYCI SVSDTDGVSSSFVLDEEELERLMTLSNEIKNPTIPLKSELAYEVLDKGEVRF WIQAESLSPNSTYRFVINDKEVENGDRHKISCDHSNGIIEMVMDKFTIDNEG TYTVQIQDGKAKNQSSLVLIGDAFKAILAESELQRKEFLRKQGPHFSEFLYW EVTEECEVLLACKIANTKKETVFKWYRNGSGIDVDEAPDLQKGECHLTVPK LSRKDEGVYKATLSDDRGHDVSTLELSGKVYNDIILALSRVSGKTASPLKIL CTEEGIRLQCFLKYYNEEMKVTWSHRESKISSGEKMKIGGGEDVAWLQITE PTEKDKGNYTFEIFSDKESFKRTLDLSGQAFDDALTEFQRLKAAAFAEKNR GKVIGGLPDVVTIMDGKTLNLTCTVFGNPDPEVVWFKNDKALELNEHYLV SLEQGKYASLTIKGVTSEDSGKYSIYVKNKYGGETVDVTVSVYRHGEKIPE VNQGQLAKPRLIPPSSST
157	Chicken ENOB (Beta-enolase)	MSIQKIHAREILDSRGEPTVEVDLHTAKGHFRAAVPSGASTGIHEALEPRDG DKKRFLGKGVLKAVEHINKTIGPALIEKKISVVEQEKIDKVMIEMDGTENKS KFGANAILGVSLAVCKAGAAEKGVPLYRHIADLAGNTELILPVPAFNVING GSHAGNKLAMQEFMVLPVGAASFHDAMRVGAEVYHSLKGVIKAKYGKD ATNVGGEGGFAPNILDNHEALELLKAAIAQAGYTDKVVIGMDVAASEFCR DGRYHLDFKSPPHTKRYITGEQLGEIYRGFIKDYPVVSIEDPFDQDDWEAW KRFVFHVDIQVVGDDLTVTNPKRIAHGAEQHACNCLLLKVNQIGSVTESIQ ACKLAQSHGWGVMVSHRSGETEDTFIADLVVGLCTGQIKTGAPCRSERLA KYNQLMRIEEALGDKAKFAGRKFRNPKAK
158	Chicken MYL11 (Myosin regulatory light chain 11)	MAPKKAKRRAAEGSSNVFSMFDQTQIQEFKEAFTVIDQNRDGIIDKDDLRE TFAAMGRLNVKNEELDAMIKEASGPINFTVFLTMFGEKLKGADPEDVIMG AFKVLDPDGKGSIKKSFLEELLTTQCDRFTPEEIKNMWAAFPPDVAGNVDY KNICYVITHGEDKEGE
159	Chicken MYH1B (Myosin-1B)	MATDADMAIFGEAAPYLRKSEKERIEAQNKPFDAKSSVFVVHAKESYVKST IQSKESGKVTVKTEGGETLTVKEDQIFSMNPPKYDKIEDMAMMTHLHEPAV LYNLKERYAAWMIYTYSGLFCVTVNPYKWLPVYNPEVVLAYRGKKRQEA PPHIFSISDNAYQFMLTDRENQSILITGESGAGKTVNTKRVIQYFATIAASGD KKKEEQPAGKMQGTLEDQIISANPLLEAFGNAKTVRNDNSSRFGKFIRIHFG ATGKLASADIETYLLEKSRVTFQLKAERSYHIFYQIMSNKKPELIEMLLITTN PYDYQYVSQGEITVPSINDQEELMATDSAIDILGFTPDEKTAIYKLTGAVMH YGNLKFKQKQREEQAEPGGTEVADKAAYLMGLNSADLLKALCYPRVKVG NEYVTKGQTVQQVYNSVGALAKSVFEKMFLWMVVRINQQLDTKQPRQYF IGVLDIAGFEIFDFNSLEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEWE FIDFGMDLAACIELIEKPMGIFSILEEECMFPKATDTSFKNKLYDQHLGKSNN FQKPKPGKGKAEAHFSLVHYAGTVDYNITGWLEKNKDPLNETVVGLYQKS SLKTLALLFASVGGAEAESGAGGKKGGKKKGSSFQTVSALFRENLNKLMS NLRSTHPHFVRCLIPNETKTPGAMEHELVLHQLRCNGVLEGIRICRKGFPIRI LYADFKQRYKVLNASAIPEGQFIDSKKASEKLLGSIDVDHTQYKFGHTKVF FKAGLLGLLEEMRDEKLAQLITRTQARCRGFLMRVEFKKMMERRESIFCIQ YNVRAFMNVKHWPWMKLFFKIKPLLKSAESEKEMANMKEEFEKTKEELA KSEAKRKELEEKMVSLLQEKNDLQLQVQAEADGLADAEERCDQLIKTKIQ LEAKIKELTERAEDEEEMNAELTAKKRKLEDECSELKKDIDDLELTLAKVE KEKHATENKVKNLTEEMAALDETIAKLTKEKKALQEAHQQTLDDLQAEED KVNTLTKAKTKLEQQVDDLEGSLEQEKKLRMDLERAKRKLEGDLKMTQE STMDLENDKQQLDEKLKKKDFEISQIQSKIEDEQALGMQLQKKIKELQARIE ELEEEIEAERTSRAKAEKHRADLSRELEEISERLEEAGGATAAQIDMNKKRE AEFQKMRRDLEEATLQHEATAAALRKKHADSTADVGEQIDNLQRVKQKL EKEKSELKMEIDDLASNMESVSKAKANLEKMCRSLEDQLSEIKTKEEEQQR TINDISAQKARLQTESGEYSRQVEEKDALISQLSRGKQAFTQQIEELKRHLEE EIKAKKCPAHALQSARHDCDLLREQYEEEQEAKGELQRALSKANSEVAQW RTKYETDAIQRTEELEEAKKKLAQRLQDAEEHVEAVNSKCASLEKTKQRL QNEVEDLMIDVERSNAACAALDKKQKNFDKILSEWKQKYEETQAELEASQ KESRSLSTELFKMKNAYEESLDHLETLKRENKNLQQEISDLTEQIAEGGKAI HELEKVKKQIEQEKSELQTALEEAEASLEHEEGKILRVQLELNQVKSDIDRK IAEKDEEIDQLKRNHLRVVDSMQSTLDAEIRSRNEALRLKKKMEGDLNEIEI QLSHANRQAAEAQKNLRNTQGVLKDTQIHLDDALRSQEDLKEQVAMVER RANLLQAEIEELRAALEQTERSRKVAEQELLDASERVQLLHTQNTSLINTKK KLESDISQIQSEMEDTIQEARNAEEKAKKAITDAAMMAEELKKEQDTSAHL ERMKKNLDQTVKDLQHRLDEAEQLALKGGKKQIQKLEARVRELEGEVDA EQKRSAEAVKGVRKYERRVKELTYQSEEDRKNVLRLQDLVDKLQMKVKS YKRQAEEAEELSNVNLSKFRKIQHELEEAEERADIAESQVNKLRAKSREIGK KAESEE

TABLE8-continued

Chicken Proteins		
SEQ ID NO	Description	Amino Acid Sequence
160	Chicken R4GIG1 (Uncharacterized protein)	MATDADMAIFGEAAPYLRKSEKERIEAQNKPFDAKSSVFVVHAKESYVKST IQSKESGKVTVKTEGGETLTVKEDQIFSMNPPKYDKIEDMAMMTHLHEPAV LYNLKERYAAWMIYTYSGLFCVTVNPYKWLPVYNPEVVLAYRGKKRQEA PPHIFSISDNAYQFMLTDRENQSILITGESGAGKTVNTKRVIQYFATIAASGD KKKEEQPAGKMQGTLEDQIISANPLLEAFGNAKTVRNDNSSRFGKFIRIHFG ATGKLASADIETYLLEKSRVTFQLKAERSYHIFYQIMSNKKPELIEMLLITTN PYDYQYVSQGEITVPSINDQEELMATDSAIDILGFTPDEKTAIYKLTGAVMH YGNLKFKQKQREEQAEPDGTEVADKAAYLMGLNSADLLKALCYPRVKVG NEYVTKGQTVQQVYNSVGALAKSVFEKMFLWMVVRINQQLDTKQPRQYF IGVLDIAGFEIFDFNSLEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEWE FIDFGMDLAACIELIEKPMGIFSILEEECMFPKATDTSFKNKLYDQHLGKSNN FQKPKPGKGKAEAHFSLVHYAGTVDYNITGWLEKNKDPLNETVVGLYQKS SLKTLALLFASVGGAEAESGAGGKKGGKKKGSSFQTVSALFRENLNKLMS NLRSTHPHFVRCLIPNETKTPGAMEHELVLHQLRCNGVLEGIRICRKGFPSRI LYADFKQRYKVLNASAIPEGQFIDSKKASEKLLGSIDVDHTQYKFGHTKVF FKAGLLGLLEEMRDEKLAQLITRTQARCRGFLMRVEFKKMMERRESIFCIQ YNVRAFMNVKHWPWMKLFFKIKPLLKSAESEKEMANMKEEFEKTKEELA KSEAKRKELEEKMVSLLQEKNDLQLQVQAEADGLADAEERCDQLIKTKIQ LEAKIKELTERAEDEEEMNAELTAKKRKLEDECSELKKDIDDLELTLAKVE KEKHATENKVKNLTEEMAALDETIAKLTKEKKALQEAHQQTLDDLQAEED KVNTLTKAKTKLEQQVDDLEGSLEQEKKLRMDLERAKRKLEGDLKMTQE STMDLENDKQQLDEKLKKKDFEISQIQSKIEDEQALGMQLQKKIKELQARIE ELEEEIEAERTSRAKAEKHRADLSRELEEISERLEEAGGATAAQIDMNKKRE AEFQKMRRDLEEATLQHEATAAALRKKHADSTAELGEQIDNLQRVKQKLE KEKSELKMEIDDLASNMESVSKAKANLEKMCRSLEDQLSEIKTKEEEQQRT INDISAQKARLQTESGEYSRQVEEKDALISQLSRGKQAFTQQIEELKRHLEEE IKAKNALAHALQSARHDCDLLREQYEEEQEAKGELQRALSKANSEVAQWR TKYETDAIQRTEELEEAKKKLAQRLQDAEEHVEAVNSKCASLEKTKQRLQ NEVEDLMIDVERSNAACAALDKKQKNFDKILSEWKQKYEETQAELEASQK ESRSLSTELFKMKNAYEESLDHLETLKRENKNLQQEISDLTEQIAEGGKAIH ELEKVKKQIEQEKSELQAALEEAEASLEHEEGKILRVQLELNQVKSDIDRKI AEKDEEIDQLKRNHLRVVDSMQSTLDAEIRSRNEALRLKKKMEGDLNEIEI QLSHANRQAAEAQKNLRNTQGVLKDTQIHLDDALRSQEDLKEQVAMVER RANLLQAEIEELRAALEQTERSRKVAEQELLDASERVQLLHTQNTSLINTKK KLESDISQIQSEMEDTIQEARNAEEKAKKAITDAAMMAEELKKEQDTSAHL ERMKKNLDQTVKDLQHRLDEAEQLALKGGKKQIQKLEARVRELEGEVDA EQKRSAEAVKGVRKYERRVKELTYQSEEDRKNVLRLQDLVDKLQMKVKS YKRQAEEAEELSNVNLSKFRKIQHELEEAEERADIAESQVNKLRAKSREIGK KAESEE
161	Chicken ALBU (Albumin)	MKWVTLISFIFLFSSATSRNLQRFARDAEHKSEIAHRYNDLKEETFKAVAMI TFAQYLQRCSYEGLSKLVKDVVDLAQKCVANEDAPECSKPLPSIILDEICQV EKLRDSYGAMADCCSKADPERNECFLSFKVSQPDFVQPYQRPASDVICQEY QDNRVSFLGHFIYSVARRHPFLYAPAILSFAVDFEHALQSCCKESDVGACLD TKEIVMREKAKGVSVKQQYFCGILKQFGDRVFQARQLIYLSQKYPKAPFSE VSKFVHDSIGVHKECCEGDMVECMDDMARMMSNLCSQQDVFSGKIKDCC EKPIVERSQCIMEAEFDEKPADLPSLVEKYIEDKEVCKSFEAGHDAFMAEFV YEYSRRHPEFSIQLIMRIAKGYESLLEKCCKTDNPAECYANAQEQLNQHIKE TQDVVKTNCDLLHDHGEADFLKSILIRYTKKMPQVPTDLLLETGKKMTTIG TKCCQLGEDRRMACSEGYLSIVIHDTCRKQETTPINDNVSQCCSQLYANRR PCFTAMGVDTKYVPPPFNPDMFSFDEKLCSAPAEEREVGQMKLLINLIKRK PQMTEEQIKTIADGFTAMVDKCCKQSDINTCFGEEGANLIVQSRATLGIGA

[0053] An animal protein can comprise a turkey protein (Table 9). A turkey protein can comprise a turkey muscle protein. A turkey protein can comprise a turkey blood protein. A turkey protein can comprise a turkey albumin protein. A turkey protein can comprise a sequence that has at least or about 70%, at least or about 75%, at least or about 80%, at least or about 90%, at least or about 91%, at least or about 92%, at least or about 93%, at least or about 94%, at least or about 95%, at least or about 96%, at least or about 97% at least or about 98% at least or about 99% or about 100% sequence identity to any one of SEQ ID NO: 162-174.

[0054] A food product may comprise a fragment or a truncation of a turkey protein. The food product may include a fragment or truncation of any one of SEQ ID Nos: 162-174. A sequence of the fragment or truncation may be identical or substantially identical to a portion of any one of SEQ ID Nos: 162-174. A sequence of the fragment or truncation may be at least about 70%, 80%, 90%, 95%, 98%, or more identical to a portion of a sequence of any one of SEQ ID Nos: 162-174. A length of the fragment or truncation may be at least about 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of a length of any one of SEQ ID Nos: 162-174. A length of the fragment or truncation may be less than or equal to about 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, or less than a length of any one of SEQ ID Nos: 162-174.

TABLE9

Turkey Proteins

SEQ ID NO	Description	Amino Acid Sequence
162	Turkey G1NL16 (Creatine kinase)	MPFSNSHNLLKMKHSADDEFPDLSAHNNHMAKVLTLDLYKKLRDRQTSSG FTLDDVIQTGVDNPGHPFIMTVGCVAGDEESYEVFKELFDPVIEDRHGGYK PTDEHKTDLNADNLQGGDDLDPNYVLSSRVRTGRSIRGFCFPHCSRGERRA IEKLSVEALGSLGGDLKGKYYALRNMTDAEQQQLIDDHFLFDKPVSPLLLA SGMARDWPDARGIWHNDNKTFLVWINEEDHLRVISMQKGGNMKEVFTRF CTGLTQIETLFKSKNYEFMWNPHLGYILTCPSNLGTGLRAGVHIKLPNLGK HEKFGEVLKRLRLQKRGTGGVDTAAVGGVFDVSNADRLGFSEVELVQMV VDGVKLLIEMEKRLEKGQSIDDLMPAQK
163	Turkey A0A803Y9V3 (Actinin alpha 2)	MNSMNQIETNMQYTYNYEEDEYMTQEEEWDRDLLLDPAWEKQQRKTFTA WCNSHLRKAGTQIENIEEDFRNGLKLMLLLEVISGERLPKPDRGKMRFHKI ANVNKALDYIASKGVKLVSIGAEEIVDGNVKMTLGMIWTIILRFAIQDISVE ETSAKEGLLLWCQRKTAPYRNVNIQNFHLSWKDGLGLCALIHRHRPDLIDY SKLNKDDPIGNINLAMEIAEKHLDIPKMLDAEDIVNTPKPDERAIMTYVSCF YHAFAGAEQAETAANRICKVLAVNQENERLMEEYERLASELLEWIRRTIPW LENRTPEKTMQAMQKKLEDFRDYRRKHKPPKVQEKCQLEINFNTLQTKLRI SNRPAFMPSEGKMVSDIAGAWQRLEQAEKGYEEWLLNEIRRLERLEHLAE KFRQKASTHEQWAYGKEQILLQKDYESASLTEVRAMLRKHEAFESDLAAH QDRVEQIAAIAQELNELDYHDAASVNDRCQKICDQWDSLGTLTQKRREAL ERTEKLLETIDQLHLEFAKRAAPFNNWMEGAMEDLQDMFIVHSIEEIQSLIS AHDQFKATLPEADGERQAILSIQNEVEKVIQSYSMRISASNPYSTVTVEEIRT KWEKVKQLVPQRDQSLQEELARQHANERLRRQFAAQANVIGPWIQTKME EIARSSIEMTGPLEDQMNQLKQYEQNIINYKHNIDKLEGDHQLIQEALVEDN KHTNYTMEHIRVGWELLLTTIARTINEVETQILTRDAKGITQEQMNDFRASF NHFDRRKNGLMDHDDFRACLISMGYDLGEAEFARIMSLVDPNGQGTVTFQ SFIDFMTRETADTDTAEQVIASFRILASDKPYILADELRRELPPEQAQYCIKR MPQYTGPGSVPGALDYTSFSSALYGESDL
164	Turkey A0A803Y8V3 (Alpha-1,4 glucan phosphorylase)	MAAPLSDGERRKQISVRGIAGLGDVAEVRKSFNRHLHFTLVKDRNVATPR DYFFALAHTVRDHLVGRWIRTQQHYYERDPKRIYYLSLEFYMGRTLQNTM VNLGLQNACDEAIYQLGLDLEELEEIEEDAGLGNGGLGRLAACFLDSMATL GLAAYGYGIRYEFGIFNQKIVDGWQVEEADDWLRYGNPWEKARPEYMLP VHFYGRVEHSPEGVKWVDTQVVLAMPYDTPVPGYKNNTVNTMRLWSAK APNDFNLQEFNVGDYIEAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYF VVAATLQDIIRRFKSSKFGCQDPVRTSFETFPDKVAIQLNDTHPALSIPELMR ILVDVEKVDWDKAWEITKQTCAYTNHTVLPEALERWPVSMFEKLLPRHLEI IYALNQMHLDRVAALYPGDIDRLRRMSVIEEGDCKRINMAHLCVIGSHAVN GVARIHSDIVKNTVFKDFYELDPEKFQNKTNGITPRRWLLLCNPGLADVIAE KIGEDFVTDLSQLKKLLDFIDDETFIRDVAKVKQENKLKFSAYLEEHYKVKI NPSSMFDIQVKRIHEYKRQLLNCLHAITLYNRKSDPSKSFVPRTIMIGGKAA PGYHMAKMIIKLITSIGEVVNNDPYVGDKLKVIFLENYRVSLAEKVIPAADL SQQISTAGTEASGTGNMKFMLNGALTIGTMDGANVEMAEEAGEENLFIFG MRVEEVEALDRQGYNAREYYDRIPELRQAIDQISSGFFSPRDPGCFRDVVN MLMHHDRFKVFADYEAYIKCQGQVDQLFMDPREWTKKVIRNIACSGKFSS DRTITEYAREIWGVEPSGDVFPAWEA
165	Turkey GIN2A4 (Fructose-bisphosphate aldolase)	MTHQYPALTAEQKKELSDIALRIVAPGKGILAADESVGSMAKRLNQIGVEN TEENRRLYRQILFSADSRVKKCIGGVIFFHETMYQKADDGTPFVQMIKDKGI VVGIKVDKGVVPLAGTDGETTTQGLDGLSERCAQYKKDGADFAKWRCVL KISDNTPSALAIMENANVLARYASICQQNGIVPIVEPEILPDGDHDLKRCQY VTEKVLAAVYKALSDHHVYLEGTLLKPNMVTPGHSCPTKYSPEEIAMATV TALRRTVPPAVPGVTFLSGGQSEEEASINLNAINTCPLVRPWALTFSYGRAL QASALSAWRGQRDNANAATEEFVKRAEVNGLAALGKYEGSGDDSGAAGQ SLYVANHAY

TABLE9-continued

Turkey Proteins		
SEQ ID NO	Description	Amino Acid Sequence
166	Turkey Q6W5H1 (Myosin essential light chain isoform 1)	MAPKKDVKKPAAAAAPAPAPAPAPAPAPAKPKEPAIDLKSIKIEFSKEQQD DFKEAFLLFDKTGDAKITLSQVGDIVRALGQNPTNAEMNKILGNPSKEEMN AKKITFEEFLPMLQAAANNKDQGTYEDFVEGLRVFDKEGNGTVMGAELRH VLATLGEKMTEEEVEELMKGQEDSNGCINYEAFVKHIMSV
167	Turkey G1NMR6 Glyceraldehyde-(3-phosphate dehydrogenase)	MVKVGVNGFGRIGRLVTRAAVLSGKVQVVAINDPFIDLNYMVYMFKYDST HGHFKGTVKAENGKLVINGHAITIFQERDPSNIKWADAGAEYVVESTGVFT TMEKAGAHLKGGAKRVIISAPSADAPMFVMGVNHEKYDKSLKIVSNASCT TNCLAPLAKVIHDNFGIVEGLMTTVHAITATQKTVDGPSGKLWRDGRGAA QNIIPASTGAAKAVGKVIPELNGKLTGMAFRVPTPNVSVVDLTCRLEKPAK YDDIKRVVKAAADGPLKGILGYTEDQVVSCDFNGDSHSSTFDAGAGIALND HFVKLVSWYDNEFGYSNRVVDLMVHMASKE
168	Turkey A0A803XNQ2 (Tropomyosin 3)	MEAIKKKMQMLKLDKENALDRAEQAEAEQKQAEERSKQLEDELAAMQK KLKGTEDELDKYSEALKDAQEKLELAEKKAADAEAEVASLNRRIQLVEEE LDRAQERLATALQKLEEAEKAADESERGMKVIENRALKDEEKMELQEIQL KEAKHIAEEADRKYEEVARKLVIIEGDLERTEERAELAESKCSELEEELKNV TNNLKSLEAQAEKYSQKEDKYEEEIKILTDKLKEAETRAEFAERSVAKLEK TIDDLEDELYAQKLKYKAISEELDHALNDMTSI
169	Turkey A0A803YPL8 (Myosin-binding protein C)	MCLQLFVFSELPPPFSFLKNKELTFVYCLHKALMVECWCHSICSSPTLADQI REVVGLESLVSSGKCILERIKHWLIIDIFLFLCISVWSIGESPPEETEKHDDSQ RSTLFIEKPQSGSVSVGGNITFIAKVEAKDLLRKPNVKWFKGKWMDLASKA GKHLQLKESFERHTKIHTFEMHIIQAKENYAGNYRCEVSYKDKFDSCSFDL EVTESSQAAPSIDIRSAFKRSGDGQDDAGELDFSGLLKRREIKQQEEEPEVD VWELLKNANPSEYEKIAFQYGITDLRGMLKRLKRMRREVKKSAAFSKGLD PAYQVDKGGKVRFMVELADPTVELKWYKNGQEIRPSAKYIFEHKGNQRIL FINNCTMTDDARYYVTAGDEKCSTELFVRDPPILVTKGLEDTNTYVGERVE LSCEVSEDDANVKWFKNGVELTNEPKSRYRIKVEGKKHTLIIEEAAKNDNA TYSVMTTGGQSEAKLSVDLRPLKISLALEDQTVRLGQEIHLKCEISENVEGK WYKNGQLIEASDRVKLYHKGRIHRFIIASAAVDDEGEYMFVPDAYNINIPC KVHVVDPPKLHLDGLGENNTVTVVAGTKLRLEIPITGEPTPKVMWSKGDK WITDSGRIRAETYSDSSCLVIDTAEREDSGPFRITLKNEAGEDSALINIKVVD VPDPPQAPNVTEVGEDWCVMTWEPPANDGGSPILGYFIERKKKQSSRWMR LNFELCKETTFEPKKMIEGVAYEVRVFAVNAIGTSKPSMPSKSFVPLAVTSP PTLLAVDSVTDTSVTMKWRPPDHIGAAGLDGYVVEYCFEGTDEWIVANPE LTDKTKFTINGLPTGSKILVRVKAVNAAGESEPRYHPQPILVKEVIEPPKIRL PRHLKQTYTRRVGETVNLVIPFQVRTKNTDRKTLNTHFSYATASYLLFLYV NLGYWTAFVNILHPEGKSHDKASGIIHSYLGTKIGHYHMSSQSN
170	Turkey G3URS1 (Myomesin 1)	MNVKGEISAYASLVVKRYKGEIDASHLHAGTFSMPLSLGVTPHGFASRFEI KFIDKFEVMFGREGETMSLGCTVVISPDIKRFKPDIEWYRNGVLITPSKWVQ MHWTGERASLTLTHANKEDEGLYTLRVVMGEYYEEYSGYVFVRDADAEI PGAPGAPLDVQCLDANKDYVIVSWKQPAVDGGNPILGYFIDRCEVGTTHW TQCNETPVKFARFPVTGLIEGRSYIFRVRAVNRAGVSLPSRVSEAVAALDPA DRARLRSHPSAPWTGQIIVTEEEPAEGVIPGPPTDLQVIEATKNYVVLSWKP PGQRGHEGIMYFVEKCIAGTEDWQRVNTEIPVKSPRFAVFDLSEGKSYCFR VRCCNSAGIGEPSEATEATVVDDKLDIPKPPGRIMPTRNTDTSVVVTWTEPP DAKELVGYYIESSVVGSGHWEPCNNNPVKGTRFICHGLVNGEKYIFRVRAV NAAGLSEFSQDSEPIEVQAAIAAPSPPYDITVLESVRDSMVLGWKQPKATGG TEITGYYVNYREVIGGVPGKWREANIKAISERAYRIENLKENMVYQFQVAA ANLAGVGTPSLPSKPFKCEEWTIAVPGPPHDVTYTEVRKDSLVLLWKEPVY TGRSPIAGYYIDMKETEAKEEHWRSVNEKPLQKKFLKIAGLTEGVSYVFRV RATNQAGVGRPSDITDPVVAETRPGTKEVVVNVDDDGVISLNFECDQMSP NSKFVWSKNYEPIEDDSRLDIDSKGGKSKAIFKNLGEDDLGIYSCVVTDTDG VSSGYTIDEEEMKRLLALSHERKFPTIPLASELAVEILEKGEVRFWLQAEKLS GNAKANFVFNDKEIFNGEKYKMKVDQRTGLVEMIMDKLEDKDEGTYTFQ LQDGKATNQSSLVLIGDVFKKLRDEAEFQRKEWHRKQGPHFVDQLGWEV TDECNVMLKCKVANIKKETHIVWYKDDREIMVDEEHDFKDGVCTLLISEFS KKDTGIYEVILKDDRGKDVSELKLTDTAFADLMNEVCRRIALSATDLKIQST AEGIRLYSFVTYYVDDLRVGWVHNDTQIRYTDRVKTGVTGEQIWLQINEPT PQDKGKYTMELFDGKTGHKKTVDLSGQAFDEAFAEFQRLKQAAIAEKNRA RVLGGLPDVVTIQEGKALNLSCTVWGDPTPEVSWLKNEKSFVSDANCILKF EAGKNVSFTISAVSTHDSGKYSLVVKNKYGTETSDVTVIVITVSGQTDTAK HTVSISGQSQNNITVSKKKKATTMSKPQQEKEKKDIKPHVDEIIATEETSDS KEIPRVEKEKNTENNQQTAKKTHGKAAAARTVP

TABLE9-continued

Turkey Proteins		
SEQ ID NO	Description	Amino Acid Sequence
171	Turkey A0A803YQS2 (Phosphopyruvate hydratase)	MSILKIQAREIFDSRGNPTVEVDLYTNKGLFRAAVPSGASTGIYEALELRDN DKTRYLGKGVSKAVEHVNKTIAPALISKNVNVVEQEKIDKLMLEMDGTEN KSKFGANAILGVSLAVCKAGAAEKGVPLYRHIADLAGNPEVILPVPAFNVI NGGSHAGNKLAMQEFMILPVGADNFKEAMRIGAEVYHNLKNVIKEKYGK DATNVGDEGGFAPNILENKEALELLKTAIGKAGYSDKVVIGMDVAASEFYR DGKYDLDFKSPDDPSRYISPDQLADLYKGFVKNYPLVSIEDPFDQDDWAA WKKFTASVGIQVVGDDLTVTNPKRIAKAVEEKSCNCLLLKVNQIGSVTESL QACKLAQSNGWGVMVSHRSGETEDTFIADLVVGLCTGQIKTGAPCRSERL AKYNQLLRIEEELGSKARFAGRNFRNPRIN
172	Turkey G1N0W1 (Myosin light chain 2)	MAPKKAKKRIEGANSNVFSMFEQAQIQEFKEAFTIMDQNRDGFIDKADLRD TFAALGRLNVKNEEIDEMIKEAPGPINFTVFLTMFGEKLKGADPEETILNAF KVFDPEGKGLKSAYIKEMLMTQGERFSQEEIDQMFAAFPPDVSGNLDYKNL VHVITHGEEKD
173	Turkey GIMSQ0 (Myosin heavy chain 2)	MSSDAEMAIFGEAAPYLRKSEKERIEAQNRPFDAKAACFVVDQKQMYVKG TIQSREGGKVTVKTYDDTTVTVKDDEVFPMNPPKYDKIEDMAMMTHLHEP AVLYNLKERYAAWMIYTYSGLFCVTVNPYKWLPVYNPEVVSGYRGKKRQ EAPPHIFSISDNAYQFMLTDRDNQSILITGESGAGKTVNTKRVIQYFATIAVT GEKKKDQQPSKMQGTLEDQIISANPLLEAFGNAKTVRNDNSSRFGKFIRIHF GTTGKLASADIETYLLEKSRVTFQLASERSYHIFYQIMSNKKPELIGEYLLLI TTNPYDFTYVSQGEITVASIDDSEELLATDNAVDILGFSPDEKVGMYKLTGA VMHYGNMKFKQKQREEQAEPDGTEVADKAAYLMGLNSAELLKALCYPR VKVGNEYVTKGQNVQQVYNSVGALAKSVYEKMFLWMVTRINQQLDTKQ PRQHFIGVLDIAGFEIFDFNSLEQLCINFTNEKLQQFFNHHMFVLEQEEYKKE GIDWEFIDFGMDLAACIELIEKPMGIFSILEEECMFPKATDTSFKNKLYDQHL GKSNNFQKPKPAKGKAEAHFSLIHYAGTVDYNITGWLEKNKDPLNETVVG LYQKSSLKLLSFLFSNYAGAETSDSGSGKKGGKKKGGSFQTVSAVFRENLN KLMTNLRSTHPHFVRCLIPNETKTPGEMDHYLVMHQLRCNGVLEGIRICRK GFPSRILYADFKQRYKILNASAIPEGQFIDSKKASEKLLSSIDVDHNQYRFGH TKVFFKAGLLGLLEEMRDEKLVSLITHMQAMCRGYLMRTEFKKMNARRES IYIIQYNVRAFMNVKHWPWMKLYFKMKPMLKSVESEKEMANMKEEFEKT KEELAKSEAKRKELEEKMVSLLQEKNDLQLQVQSESENLADAEERCEGLIK SKIQLEARIKELNERMEDEEEMNAELTAKKRKLEDECSELKKDIDDLELTLA KVEKEKHATENKVKNLTEEMAALDENISKLTKEKKALQEAHQQTLEYLQA EEDKVSILTKIKTKLEQQVDDLEGSLEQERKQRMDLERAKRKLDGDLKMS QDSIMDLKNDKQQTDERLKKKDFEISHLQSKIEDEQAQSSQLQKKIKELQA RTEELEEEIEAERANCAKTEKQRADLSRELEEISERLEEAGGATAAQIEMNK KREAEFQKMRRDLEEATLQHEATAAALRKKHADSTAELGEQIDNLQRVKQ KLEKEKSELKMEIDDLASNMESVSKAKSNLEKMCRALEDQYSEVRTKDDE HMRLINDMNTQKTRLQTENGELTRQLEEKESLISQLTRGKQAITQQTEELK RQLEEENKAKNALAHGLQSARHDCDLLREQYEEELEAKGELQRALSKANS EVAQWRTKYETDAIQCMEELEDAKKKLAQRLQDSEEQIEAVNSKCASLEK TKQRLQGEVDDLMIAVERSNAACAAFDKKQKNFDKVMAEWKQKYQESQ AELEAAQKESRTLSTEIFRMKNAYEEVLDQVETVRRENKNLQQEISDLTEQ LAEAGKANHGLEKAKKQTEQEKYDLQAALEEAEVSLEHEEGKILRVQLEL NQVKSGVDRRSAEKDEEIQQLKKNHQRVLESMQTMLDAEIRSRNDALRLK KKMEGDLNDMEIQLSRANCQIAEIQKHLKGMQGQLKDSQLHLDDALREND DLKEQLAVVERRNNLMMTELEQMRAALEQTERARKVSEQELTDTSERVQI LHSQNTSLLNTKKKLEVDITHLQNEVEDSIQEASNAEEKAKKAITDAAMMA EELKKEQDTSAHLERMKRNLEQMVKDLQHRLDEAEQLTLKGGKKQLQKL EARIHELENELDIEQKRGTESLKGARKYERRLKELTYQSEEDKKNILRLQNL VDKLQLKVKAYKKQAEEAEEQASTNLSRCRKTQHELEEAEERADIAECQV NKLRAKSRDTGRQKSGLSSELGDNTFLPVIWHNMCNIDRPETTPQVILLHQ
174	Turkey GINCR2 (Albumin)	MKWVTLISFIFLFSSATSRNLQRVARDAEHKSEIAHRYNDLKEETFKAVAMI TFAQYLQRCSYDGLSKLVKDVVDLAQKCVANEDAPECSKSLPSIILDEICQV EKLRDSYGAMADCCSKADPERNECFLSFKVSQPDFIQPYQRPASDVICKEY QDNRVSLLGHFIYSVARRNPYMYAPTILSFAADYEHALQSCCQESDVGACL DAKETAMREKAKKISVKQQYSCGILKKFGDRVFQAEKLVRLSQKYPKAPFS EVAKLVRESKDVHKECCEGDMVECMDDMAEIINNMCSKQDVFSSKIKGCC EKPIMERSQCIMEAEFDEKPADLPSLVEKYIQDKEVCKSFEAGHDAFMSEFV YEYSRRHPEFSTQLILRIAKGYESLLEKCCKTDNPAECYANAQDQLNQHIKE TQDVVKTNCDLLSAHGEPDFLKSILIRYTKKMPQVSTETLLEIGKKMTAIGT KCCQLSEDRRMACSEGYLSIVIHDMCRRQETTPINDNVSHCCSDSYAHRRP CFTAMGVDTKYVPPPFNPDMFSFDEKLCSAPAEEREVGQMKLLINLIKRKP QMTEEQIKTIADGFTAMVDKCCKQSDINTCFGEEGANLIVQSRATLGIGV

[0055] An animal protein can comprise a duck protein (Table 10 and Table 11). A duck protein can comprise a duck muscle protein. A duck protein can comprise a duck blood protein. A duck protein can comprise a duck albumin protein. A duck protein can comprise a protein from a mallard duck (*Anas platyrhynchos*). A duck protein can comprise a protein from a Northern mallard duck (*Anas platyrhynchos platyrhynchos*). A duck protein can comprise a protein from a Muscovy duck (*Cairina moschata*). A duck protein can comprise a protein from a domestic Muscovy duck (*Cairina moschata domestica*). A duck protein can comprise a sequence that has at least or about 70%, at least or about 75%, at least or about 80%, at least or about 90%, at least or about 91%, at least or about 92%, at least or about 93%, at least or about 94%, at least or about 95%, at least or about 96%, at least or about 97% at least or about 98% at least or about 99% or about 100% sequence identity to any one of SEQ ID NO: 175-200.

[0056] A food product may comprise a fragment or a truncation of a duck protein. The food product may include a fragment or truncation of any one of SEQ ID Nos: 175-200. A sequence of the fragment or truncation may be identical or substantially identical to a portion of any one of SEQ ID Nos: 175-200. A sequence of the fragment or truncation may be at least about 70%, 80%, 90%, 95%, 98%, or more identical to a portion of a sequence of any one of SEQ ID Nos: 175-200. A length of the fragment or truncation that is at least about 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of a length of any one of SEQ ID Nos: 175-200. A length of the fragment or truncation may be less than or equal to about 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, or less than a length of any one of SEQ ID Nos: 175-200.

TABLE 10

Northern Mallard Duck Proteins

SEQ ID NO	Description	Amino Acid Sequence
175	Northern Mallard Duck A0A493TFK1 (Creatine kinase)	MPFSNSHNLLKMKHSATDEFPDLSAHNNHMAKVLTLDLYKKLRDKQTPSG FTLDDVIQTGVDNPGHPFIMTVGCVAGDEESYEVFKELFDPVIEDRHGGYK PTDQHKTDLNAGNLQGGDDLDPNYVLSSRVRTGRSIRGFCLPPHCSRGERR AIEKLSVEALGSLGGDLKGKYYALRNMTDAEQQQLIDDHFLFDKPVSPLLL ASGMARDWPDARGIWHNDNKTFLVWINEEDHLRVISMQKGGNMKEVFTR FCTGLTQIENLFKSKNHEFMWNPHLGYILTCPSNLGTGLRAGVHIKLPHLGK HEKFGEVLKKLRLQKRGTGGVDTAAVGGVFDVSNADRLGFSEVELVQMV VDGVKLLIEMEKRLEKGQSIDDLIPAQK
176	Northern Mallard Duck A0A493TCE5 (Actinin alpha 2)	MNSMNQIETNMQYTYNYEEDEYMTQEEEWDRDLLLDPAWEKQQRKTFTA WCNSHLRKAGTQIENIEEDFRNGLKLMLLLEVISGERLPKPDRGKMRFHKI ANVNKALDYIASKGVKLVSIGAEEIVDGNVKMTLGMIWTIILRFAIQDISVE ETSAKEGLLLWCQRKTAPYRNVNIQNFHLSWKDGLGLCALIHRHRPDLIDY SKLNKDDPIGNINLAMEIAEKHLDIPKMLDAEDIVNTPKPDERAIMTYVSCF YHAFAGAEQAETAANRICKVLAVNQENERLMEEYERLASELLEWIRRTIPW LENRTPEKTMQAMQKKLEDFRDYRRKHKPPKVQEKCQLEINFNTLQTKLRI SNRPAFMPSEGKMVSDIAGAWQRLEQAEKGYEEWLLNEIRRLERLEHLAE KFRQKASTHEQWAYGKEQILLQKDYESASLTEVRAMLRKHEAFESDLAAH QDRVEQIAAIAQELNELDYHDAASVNDRCQKICDQWDSLGTLTQKRREAL ERTEKLLETIDQLHLEFAKRAAPFNNWMEGAMEDLQDMFIVHSIEEIQSLIS AHDQFKATLPEADGERQAILSIQNEVEKVIQSYSMRISASNPYSTVTVEEIRT KWEKVKQLVPQRDQSLQEELARQHANERLRRQFAAQANVIGPWIQTKME EIARSSIEMTGPLEDQMNQLKQYEQNIINYKHNIDKLEGDHQLIQEALVEDN KHTNYTMEHIRVGWELLLTTIARTINEVETQILTRDAKGITQEQMNDFRASF NHFDRRKNGLMDHDDFRACLISMGYDLGEAEFARIMSLVDPNGQGTVTFQ SFIDFMTRETADTDTAEQVIASFRILASDKPYILADELRRELPPEQAQYCIKR MPPYTGPGSVPGALDYTSFSSALYGESDL
177	Northern Mallard Duck U3ISG6 (Alpha-1,4 glucan phosphorylase)	MAAPLSDGEKRKQISVRGIAGLGDVAEVRKSFNRHLHFTLVKDRNVATPR DYFFALAHTVRDHLVGRWIRTQQHYYERDPKRIYYLSLEFYMGRTLQNTM VNLGLQNACDEAIYQLGLDLEELEEIEEDAGLGNGGLGRLAACFLDSMATL GLAAYGYGIRYEFGIFNQKIVDGWQVEEADDWLRYGNPWEKARPEYMLP VHFYGRVEHTPEGVKWVDTQVVLAMPYDTPVPGYKNNTVNTMRLWSAK APNDFNLQEFNVGDYIEAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYF VVAATLQDIIRRFKSSKFGCRDPVRTSFETFPDKVAIQLNDTHPALSIPELMR ILVDVEKVDWDKAWEITKRTCAYTNHTVLPEALERWPVSMFEKLLPRHLEI IYALNQMHLDRVAALYPGDVDRLRRMSVIEEGDCKRINMAHLCVIGSHAV NGVARIHSDIVKHSVFKDFYELEPEKFQNKTNGITPRRWLLLCNPGLADVIA EKIGEGFVTDLSQLKKLLDFIDNEMFIRDVAKVKQENKLKFSAYLEEQYKV KINPSSMFDVQVKRIHEYKRQLLNCLHAITLYNRIRSDPSKSFVPRTIMIGGK AAPGYHMAKMIIKLITSIGEVINNDPYVGDKLKVIFLENYRVSLAEKVIPAA DLSQQISTAGTEASGTGNMKFMLNGALTIGTMDGANVEMAEEAGEENLFIF GMRVEDVEALDRKGYNALEYYERLPELRQAIDQISSGFFSPRDPGCFRDVV NMLMHHDRFKVFADYEAYIKCQGQVDQLFMDPREWTKKVIRNIACSGKFS SDRTITEYAQEIWGVEPSSTKIPPPNIPRD

TABLE 10-continued

Northern Mallard Duck Proteins		
SEQ ID NO	Description	Amino Acid Sequence
178	Northern Mallard Duck U3IHG8 (Fructose-bisphosphate aldolase)	MTHQFPALSPEQKKALSDIAQRIVASGKGILAADESVGTMGNRLQRINVEN TEENRRAFREILFSSDTSINQSIGGVILFHETLYQKDSSGKPFPAIIKEKGIVVG IKLDKGTSPLAGTNGETTIEGLDGLSERCAQYKKDGVDFGKWRAVLKITST TPSQLAIQENANTLARYASICQQNGLVPIVEPEVLPDGDHDLQRCQYVTEK VLAAVYKALNDHHVYLEGTLLKPNMVTAGHACTKKYTPQDVAAATVTTL LRTVPAAVPGICFLSGGQSEEEASLNLNAMNQSSLPKPWKLTFSYGRALQA SALAAWLGKSENKKAAQEAFRKRAQINSLACRGQYVVSGKTDAAATQSLF TASYTY
179	Northern Mallard Duck U3IA79 (Myosin light chain 1)	MAPKKDVKKPAAAAAPAPAPAPAPAPAPAKPKEPAIDLKSIKIEFSKEQQD DFKEAFLLFDRTGDAKITLSQVGDIVRALGQNPTNAEINKILGNPSKEEMNA KKITFEEFLPMLQAAANNKDQGTYEDFVEGLRVFDKEGNGTVMGAELRHV LATLGEKMTEEEVDELMKGQEDSNGCINYEAFVKHILSV
180	Northern Mallard Duck A0A493TQC9 (Glyceraldehyde-3-phosphate dehydrogenase)	MVKVGVNGFGRIGRLVTRAAILSGKVQVVAINDPFIDLNYMVYMFKYDST HGHFRGTVKAENGKLVINGNAITIFQERDPSNIKWADAGAEYVVESTGVFT TMEKAGAHLKGGAKRVIISAPSADAPMFVMGVNHEKYDKSLKIVSNASCT TNCLAPLAKVIHDNFGIVEGLMTTVHAITATQKTVDGPSGKLWRDGRGAA QNIIPASTGAAKAVGKVIPELNGKLTGMAFRVPTPNVSVVDLTCRLEKPAK YDDIKRVVKAAADGPLKGILAYTEDQVVSCDFNGDSHSSTFDAGAGIALND HFVKLVSWYDNEYGYSNRVLALSTLLHAANIFPSSR
181	Northern Mallard Duck A0A493TRM1 (Tropomyosin 3)	MEAIKKKMQMLKLDKENALDRAEQAEAEQKQAEERSKQLEDELAAMQK KLKGTEDELDKYSEALKDAQEKLELAEKKAADAEAEVASLNRRIQLVEEE LDRAQERLATALQKLEEAEKAADESERGMKVIENRALKDEEKMELQEIQL KEAKHIAEEADRKYEEVARKLVIIEGDLERTEERAELAESKCSELEEELKNV TNNLKSLEAQAEKYSQKEDKYEEEIKILTDKLKEAETRAEFAERSVAKLEK TIDDLEDELYAQKLKYKAISEELDHALNDMTSI
182	Northern Mallard Duck U3J9W2 (Myosin binding protein C3)	MPEPVKKTVSAFTKKPKTTEVTAGSTAVFEAETEKASVKVKWQRAGTEIT DSEKYVIKAEGNKHSLTINNVVKEDDVTYAVIAGSSKVKFELKVKEPEKSE TVTPAEAAPAPAASELPAPPAESSQNTEVPSAGTQPEEPADPIGLFVTRPQDG EVTVGGNITFTAKVAGESLLKKPSVKWFKGKWMDLASKVGQHLQLHDSY DRNNKVYTFELEIIEANMTFAGGYRCEVSTKDKFDSSNFNLTVNEAPLSGD LDIRAAFRRTTEGHDEAGELNFSALLKKRGDGKTESQPDVDVWEILRKAPP SEYEKIAFQYGITDLRGMLKRLKRIKKEEKKSTAFLKKLDPAYQVDKGQKI KLMVEVANPDADVKWLKNGQEIQVSGSRYIFEAVGNKRILTINHCSLADD AAYECVVGEEKSFTELFVKEPPILITRPLEDQMVMVGERVEFECEVSEEGAT VKWEKDGVELTREETFKYRFKKDGKKQYLIINESTKEDSGHYTVKTNGGV SVAELIVQEKKLEVYQSIADLTVKARDQAVFKCEVSDENVKGIWLKNGKE VVPDERIKISHIGRIHKLTIEDVTPGDEADYSFIPQGFAYNLSAKLQFLEVKID FVPREEPPKIHLDCLGQSPDTIVVVAGNKLRLDVPISGDPTPTVIWQKVNKK TDLLRSNDDSMTPSENSSDLNTDSKLLFESEGRVRVEKHEDHCVFIIEGAEK EDEGVYRVIVKNPVGEDKADITVKVIDVPDPPEAPKISNIGEDYCTVQWQPP KYDGGQPVLGYILERKKKKSYRWMRLNFDLLKELTYEAKRMIEGVVYEM RIYAVNSIGMSRPSPASQPFMPIAPPSEPTHFTVEDVSDSTVALKWRPPERIG AGGLDGYIVEYCKEGSTEWIPALPGLTERTSALIKDLVTGDKLHFRIKAINL AGESGAAIIKEPITVQEILQRPKIWLPRHLRQTLVKKVGETINIVIPFQGKPRP KITWTKDGQSLDAKDVGIRNSNTDTILFIRKAELHHSGAYEITLQIENMTDK VAITIQIIDKPGPPQNIKLTDVWGFNAALEWTPPQDDGNAQILGYTVQKAD KKTMEWYTVFDHYRRTNCVVSELIMGNEYFFRVFSENLCGLSETAATTKN PAYIQKTGTTYKPPCYKEHDFSEPPKFTHPLVNRSVIAGYNTTLSCAVRGIP KPKIYWYKNKVDLSGDAKYRMFSKQGVLTLEIRKPTPFDGGFYTCKAVNE CGEAEIECKLDVRVPQ
183	Northern Mallard Duck A0A493TRW8 (Myomesin 2)	MSSLAVPFYQRRHKHFDQSYRNIQTRYVLEEYAARKAASRQAAYYESSGL GKTTCRLCARRARTLAHETMQESRKRTHEQKSHASEEKRIRFASELSALEK EIHTARHRAREQLDRLAIQRMVEENMAMERHAVEERISRAPEILVRLRSHT VWEKMSVSLCFTVQGFPTPVVQWYKNEELITPASDPAKYRVESKYGVHVL HINRADFDDSATYSAVATNVHGQASTNCAVVVRRFREGEEPHPAGIMPFHL PLSYDVCFTHFDVQFLEKFGVTFATEGETLTLKCSVLVTPELKRLRPRAEW YRDDVLIKDSKWTKLYFGEGQAALSFTHLNKDDEGLYTLRMITKGGVNEC SAYLFVRDADALIAGAPGAPMDVKCHDANRDYVIVTWKPPNTTSQSPVIG YFVDKCEVGLENWVQCNDAPVKICKYPVTGLYEGRSYIFRVRAVNNAGIS RPSRVSEPVAALDPVDLERTQTIHVDEGRKIVISKDDLEGDIQIPGPPTNVHA SEISKTYVVLSWDPPVPRGREPLTYFIEKSMVGSGSWQRVNAQVAVKSPRY AVFDLAEGKPYVFRVVSANKHGISDPSEITEPIQPQDTVVVPSAPGRVVATR NTKTSVVVQWDKPKHEEDLYGYYIDYSVVGSNHWEPSNHKPIKYNRFVVH GLETGEQYIFRVKAVNAVGFSENSQESEAIKVQAALTCPSYPHGITLLNCDG HSMTLGWKAPKYSGGSPILGYYMDKREANHQNWHEVNSSLITRTIYTVED

TABLE 10-continued

Northern Mallard Duck Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		LTEDAFYEFKIAAANIVGIGQPSDPSEHFKCKAWTMPEPGPAYDLTVCEVR NTSLVLLWKAPVYEGKSPITGYLVDYKEVDKEDWITANEKPTAQRYFKVT DLHEGHTYVFKVRAVNDAGVGKSSEVSEPVLVEAKPGTKEIFSGVDDEGNI YLAFECKEATDASHFAWGKSYEEIEDSDKYKIETEGDHSKLYFKNPDKSDL GTYSISVSDTDGVSSSFVMDDEELERLMAISNEIKNPTIPLKSELAYEVLEKG EVRFWIQAESLSPDSTFRFVINDKEVENSDRHKINCDRPNGIIEMVMDRFTID NEGTYTVQIQDGKAKNQSSLVLIGDAFKAVLAESELQRKEFLRKQGPHFSE FLYWEVTEECEVLLACKVANTKKETVFKWYKDGSGIDVDEMPDLQKGEC HLSIPRLSRKDEGVYKATLSDDRGHDVSTLELSGKVYDDIILALSRVSGKTA SPLKILCTEEGIRLQCFLKYYNEEMKISWSHRESKISSGEKMKIGGGENIAWL QITEPTEKDKGNYTFEIFGDKESFKRTLDLSGQAFDDAYAEFQRLKAAAFAE KNRGKVIGGLPDVVTIMDGKTLNLTCTVFGNPDPEVVWFKNDKALELNEH YVVSLEQGKYASLTIKGVTSEDSGKFSIHVKNKYGGETVDVTVSVYRHGEK VPEVKPGQLAKPRPIPPSEDSS
184	Northern Mallard Duck A0A493T6U6 (Phosphopyruvate hydratase)	MPERSWIPVGTPLLRWTCTHTKACFEQRSPAVPPLVSTKHWSSETTTSPVSL EKGLSVVDQEKIDNLMLEMDGTENKSKFGANAILGVSLAVCKAGAAEKDV PLYRHIADLAGNSDLILPVPAFNVINGGSHAGNKLAMQEFMILPVGAESFRD AMRIGAEVYHNLKSVIKEKYGKDATNVGDEGGFAPNILENSEALELLKEAI DKAGYTDKIVIGMDVAASEFYRDGKYDLDFKSPDDPSRYISADELGDLYQS FVRDYPVVSIEDPFDQDDWEAWSKFTANVGIQIVGDDLTVTNPKRIERAVE EKACNCLLLKVNQIGSVTEAIQACKLAQENGWGVMVSHRSGETEDTFIADL VVGLCTGQIKTGAPCRSERLAKYNQLMRIEEELGDEARFAGHNFRNPSVL
185	Northern Mallard Duck U3I9B1 (Myosin light chain 2)	MQGENHRAPKKAKKRIEGANSNVFSMFEQAQIQEFKEAFTIMDQNRDGFID KADLRDTFAALGRLNVKNEEIDEMIKEAPGPINFTVFLTMFGEKLKGADPE ETILNAFKVFDPEGKGLKSAYIKEMLMTQGERFSQEEVDQMFAAFPPDVSG NLDYKNLVHVITHGEEKD
186	Northern Mallard Duck U3J7T0 (Myosin motor domain-containing protein)	FSISDNAYQFMLTDRENQSILITGESGAGKTVNTKRVIQYFATIAASGEKKK EEQSGKMQGTLEDQIISANPLLEAFGNAKTVRNDNSSRFGKFIRIHFGATGK LASADIETYLLEKSRVTFQLKAERSYHIFYQIMSNKKPELIDMLLITTNPYDY HFVSQGEITVASINDQEELMATDSAIDILGFTADEKVAIYKLTGAVMHYGN LKFKQKQREEQAEPDGTEVADKAAYLMGLNSADLLKALCYPRVKVGNEY VTKGQTVQQVNNSVGALAKAVYEKMFLWMVVRINQQLDTKQPRQYFIGV LDIAGFEIFDFNSLEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEWEFID FGMDLAACIELIEKPMGIFSILEEECMFPKATDTSFKNKLYDQHLGKSSNFQ KPKPAKGKVEAHFSLVHYAGTVDYNITGWLEKNKDPLNETVIGLYQKSSV KTLALLFANYGGAEINPQKIIIKDTFKGSSFQTVSALFRENLNKLMTNLRSTH PHFVRCIIPNETKTPGAMEHELVLHQLRCNGVLEGIRICRKGFPSRVLYADF KQRYKVLNASAIPEGQFIDSKKASEKLLGSIDVDHTQYKFGHTKVFFKAGL LGLLEEMRDEKLAQLITRTQARCRGFLMRVEYQRMVERRESIFCIQYNVRA FMNVKHWPWMKLFFKIKPLLKSAESEKEMANMKEEFEKTKEELAKSEAK KKELEEKMVKLVQEKNDLQLQVQAEADALADAEERCDQLIKTKIQLEAKI KEVTERAEDEEEINAELTAKKRKLEDECSELKKDIDDLELTLAKVEKEKHA TENKVKNLTEEMAALDETIAKLTKEKKALQEAHQQTLDDLQAEEDKVNTL TKAKTKLEQQVDDLEGSLEQEKKLRMDLERAKRKLEGDLKLSQDSIMDLE NDKQQLDEKLKKKDFEISQIQSKIEDEQALGMQLQKKIKELQARIEELEEEIE AERTSRAKAEKHRADLSRELEEISERLEEAGGATAAQIEMNKKREAEFQKM RRDLEEATLQHEATAAALRKKHADSTAELGEQIDNLQRVKQKLEKEKSEL KMEIDDLASNMESVSKAKANLEKMCRTLEDQLSEIKTKEEEHQRMINDLTA QRARLQTESGEYSRQVEEKDSQISQLSRSKQAFTQQIEELKRHLEEEIKAKN ALAHALQSARHDCDLLREQYEEEQEAKGELQRALSKANSEVAQWRTKYE TDAIQRTEELEEAKKKLAQRLQDAEEHVEAVNAKCASLEKTKQRLQNEVE DLMIDVERSNAACAALDKKQKNEDKILAEWKQKYEETQAELEASQKESRS LSTELFKMKNAYEESLDHLETLKRENKNLQQEISDLTEQIAEGGKAIHELEK VKKQIEQEKSEIQAALEEAEASLEHEEGKILRLQLELNQVKSEIDRKIAEKDE EIDQLKRNHLRIVESMQSTLDAEIRSRNEALRLKKKMEGDLNEMEIQLSHA NRQAAEAQKNLRNTQGVLKDTQIHLDDALRTQEDLKEQVAMVERRANLL QAEIEELRAALEQTERSRKVAEQELMDASERVQLLHTQNTSLINTKKKLET DITQIQSEMEDAIQEARNAEEKAKKAITDASMMAEELKKEQDTSAH

TABLE 10-continued

Northern Mallard Duck Proteins		
SEQ ID NO	Description	Amino Acid Sequence
187	Northern Mallard Duck U3IU92 (Albumin)	MITFAQYLQRCSYDGLSKLVKDVVDLAQKCVANEDAPECSKSLPSIFLDEIC QVEKLRDSYGAMADCCAKSDPERNECFLSFKIPQPDFVQPYQRPASDVICK EYEDNRVSFLGHFIYSVARRNPFLYAPTILGVAADYEHALKSCCKESDVGA CLDGKETGIREKVKKISVKQQYSCGILKKFGDRIFQADKLALLSQKYPKASF AEISKLIHDVKDVYKECCEGDMVECMDDRAELVNYMCSKHEAFSSKIKDC CEKPIVERSQCIMEAEFDEKPADLPSLVEKYIQDKEVCKSFEAGHDAFMSEF VYEYSRRHPEFSTQLILRIAKGYETLLEKCCKTDNPAECYANAQDQLNEHIK ETQDVVKTNCDLLTTHGEADFLKALLIRYTKKMPQVSTETLLEIGKKMTGV GTKCCKLPEDRRLPCSEGYLSMVIQDMCRRQETSPINDNVSHCCSDSYAYR RPCFSAMGVDTKYVPPPFNPDVFNFDEKLCSAPAEERAEGQMKLLINLIKR KPQMTEEQIKTIADGFTAMVDKCCKQSDIETCFGEEGANLIVQSRATLGIGA

TABLE 11

Domestic Muscovy Duck Proteins		
SEQ ID NO	Description	Amino Acid Sequence
188	Domestic Muscovy Duck A0A8C3BBL6 (Creatine kinase)	MCCLKCQQGFWDFCAPAVLARGPCFGCGALVTAVGSAMPFSNSHNLLKM KHSATDEFPDLSAHNNHMAKVLTLDLYKKLRDKQTPSGFTLDDVIQTGVD NPGHPFIMTVGCVAGDEESYEVFKELFDPVIEDRHGGYKPTDQHKTDLNAG NLQGGDDLDPNYVLSSRVRTGRSIRGFCLPPHCSRGERRAIEKLSVEALGSL GGDLKGKYYALRNMTDAEQQQLIDDHFLFDKPVSPLLLASGMARDWPDA RGIWHNDNKTFLVWINEEDHLRVISMQKGGNMKEVFTRFCTGLTQIENLFK SKNHEFMWNPHLGYILTCPSNLGTGLRAGVHIKLPHLGKHEKFGEVLKKLR LQKRGTGGVDTAAVGGVFDVSNADRLGFSEVELVQMVVDGVKLLIEMEK RLEKGQSIDDLIPAQK
189	Domestic Muscovy Duck A0A8C3BEA7 (Actinin alpha 2)	RAGDKPQDEYMTQEEEWDRDLLLDPAWEKQQRKTFTAWCNSHLRKAGT QIENIEEDFRNGLKLMLLLEVISGERLPKPDRGKMRFHKIANVNKALDYIAS KGVKLVSIGAEEIVDGNVKMTLGMIWTIILRFAIQDISVEETSAKEGLLLWC QRKTAPYRNVNIQNFHLSWKDGLGLCALIHRHRPDLIDYSKLNKDDPIGNI NLAMEIAEKHLDIPKMLDAEDIVNTPKPDERAIMTYVSCFYHAFAGAEQAE TAANRICKVLAVNQENERLMEEYERLASELLEWIRRTIPWLENRTPEKTMQ AMQKKLEDFRDYRRKHKPPKVQEKCQLEINFNTLQTKLRISNRPAFMPSEG KMVSDIAGAWQRLEQAEKGYEEWLLNEIRRLERLEHLAEKFRQKASTHEQ WAYGKEQILLQKDYESASLTEVRAMLRKHEAFESDLAAHQDRVEQIAAIA QELNELDYHDAASVNDRCQKICDQWDSLGTLTQKRREALERTEKLLETIDQ LHLEFAKRAAPFNNWMEGAMEDLQDMFIVHSIEEIQSLITAHDQFKATLPE ADGERQAILSIQNEVEKVIQSYSMRISASNPYSTVTVEEIRTKWEKVKQLVP QRDQSLQEELARQHANERLRRQFAAQANVIGPWIQTKMEEIARSSIEMTGP LEDQMNQLKQYEQNIINYKHNIDKLEGDHQLIQEALVFDNKHTNYTMEHIR VGWELLLTTIARTINEVETQILTRDAKGITQEQMNDFRASFNHFDRRKNGL MDHDDFRACLISMGYDLGEAEFARIMSLVDPNGQGTVTFQSFIDFMTRETA DTDTAEQVIASFRILASDKPYILADELRRELPPEQAQYCIKRMPPYTGPGSVP GALDYTSFSSALYGESDL
190	Domestic Muscovy Duck A0A8C3BW27 (Alpha-1,4 glucan phosphorylase)	MAAPLSDGEKRKQISVRGIAGLGDVAEVRKSFNRHLHFTLVKDRNVATPR DYFFALAHTVRDHLVGRWIRTQQHYYERDPKRIYYLSLEFYMGRTLQNTM VNLGLQNACDEAIYQLGLDLEELEEIEEDAGLGNGGLGRLAACFLDSMATL GLAAYGYGIRYEFGIFNQKIVDGWQVEEADDWLRYGNPWEKARPEYMLP VHFYGRVEHNPEGVKWVDTQVVLAMPYDTPVPGYKNNTVNTMRLWSAK APNDFNLQEFNVGDYIEAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYF VVAATLQDIIRRFKSSKFGCRDPVRTSFETFPDKVAIQLNDTHPALSIPELMR ILVDVEKVDWDKAWEITKRTCAYTNHTVLPEALERWPVSMFEKLLPRHLEI IYALNQMHLDRVAALYPGDVDRLRRMSVIEEGDCKRINMAHLCVIGSHAV NGVARIHSDIVKHSVFKDFYELEPEKFQNKTNGITPRRWLLLCNPGLADVIA EKIGEGFVTDLSQLKKLLDFIDNKMFIRDVAKVKQENKLKFSAYLEEQYKV KINPSSMFDVQVKRIHEYKRQLLNCLHAITLYNRIRSDPSKSFVPRTIMIGGK AAPGYHMAKMIIKLITSIGEVINNDPYVGDKLKVIFLENYRVSLAEKVIPAA DLSQQISTAGTEASGTGNMKFMLNGALTIGTMDGANVEMAEEAGEENLFIF GMRVEDVEALDRKGYNALEYYERLPELRQAIDQISSGFFSPGDPGCFRDVV NMLMHHDRFKVFADYEAYIKCQGQVDQLFMDPWEWTKKVIRNIACSGKF SSDRTITEYAQEIWGVEPSSTKIPPPNIPRD

TABLE 11-continued

Domestic Muscovy Duck Proteins		
SEQ ID NO	Description	Amino Acid Sequence
191	Domestic Muscovy Duck A0A8C3CD90 (Fructose-bisphosphate aldolase)	MTHQFPALSPEQKKALSDIAQRIVASGKGILAADESVGTMGNRLQRINVEN TEENRRAFREILFSSDTSINQSIGGVILFHETLYQKDSSGKPFPAIIKEKGIVVG IKLDKGTSPLAGTNGETTIEGLDGLSERCAQYKKDGVDFGKWRAVLKITST TPSQLAIQENANTLARYASICQQNGLVPIVEPEVLPDGDHDLQRCQYVTEK VLAAVYKALNDHHVYLEGTLLKPNMVTAGHACTKKYTPQDVAAATVTTL LRTVPAAVPGICFLSGGQSEEEASLNLNAMNQSSLPKPWKLTFSYGRALQA SALAAWLGKPENKKAAQEAFRKRAQINSLACRGQYVVSGKTDAAATQSLF TASYTY
192	Domestic Muscovy Duck A0A8C3CJK8 (Glyceraldehyde-3-phosphate dehydrogenase)	MVKVGVNGFGRIGRLVTRAAILSGKVQVVAINDPFIDLNYMVYMFKYDST HGHFRGTVKAENGKLVINGNAITIFQERDPSNIKWADAGAEYVVESTGVFT TMEKAGAHLKGGAKRVIISAPSADAPMFVMGVNHEKYDKSLKIVSNASCT TNCLAPLAKVIHDNFGIVEGLMTTVHAITATQKTVDGPSGKLWRDGRGAA QNIIPASTGAAKAVGKVIPELNGKLTGMAFRVPTPNVSVVDLTCRLEKPAK YDDIKRVVKAAADGPLKGILAYTEDQVVSCDFNGDSHSSTFDAGAGIALND HFVKLVSWYDNEYGYSNRVVDLMVHMASKE
193	Domestic Muscovy Duck A0A8C3GR18 (Tropomyosin 1)	MDAIKKKMQMLKLDKENALDRAEQAEADKKAAEERSKQLEDELVALQK KLKGTEDELDKYSESLKDAQEKLELADKKATDAESEVASLNRRIQLVEEEL DRAQERLATALQKLEEAEKAADESERGMKVIENRAQKDEEKMEIQEIQLKE AKHIAEEADRKYEEVARKLVIIESDLERAEERAELSESKCAELEEELKTVTN NLKSLEAQAEKYSQKEDKYEEEIKVLTDKLKEAETRAEFAERSVTKLEKSID DLEDELYAQKLKYKAISEELDHALNDMTSM
194	Domestic Muscovy Duck A0A8C3CSW2 (Myosin binding protein C3)	YPTTKPKTTEVTAGSTAVFEAETEKAGVKVKWQRAGTEITDSEKYVIKAEG NKHSLTINNVVKEDDVTYAVIAGSSKVKFELKVKEPEKSETVTPAEAAPAP AASELPAPPAESSQNTEVPSAGTQPEEPVDPIGLFVTRPQDGEVTVGGNITFT AKVAGESLLKKPSVKWFKGKWMDLASKVGQHLQLHDSYDRNNKVYTFE LEIIEANMTFAGGYRCEVSTKDKFDSSNFNLTVNEIYHLCHDIVSFPFPSKTN YIFLLLRSTEGHDEAGELNFSALLKKRGDGKTESQPDVDVWEILRKAPPSEY EKIAFQYGITDLRGMLKRLKRIKKEEKKSTAFLKKLDPAYQVDKGQKIKLM VEVANPDADVKWLKNGQEIQVSGRYIFEAVGNKRILTINHCSLADDAAYEC VVGEEKSFTELFVKEPPILITHPLEDQMVMVGERVEFECEVSEEGATVKWE KDGVELTREETFKYRFKKDGKKQYLIINESTKEDSGHYTVKTNGGVSVAEL IVQEKKLEVYQSIADLTVKARDQAVFKCEVSDENVKGIWLKNGKEVVPDE RIKISHIGRIHKLTIEDVTPGDEADYSFIPQGFAYNLSAKLQFLEVKIDFVPRE EPPKIHLDCLGQSPDTIVVVAGNKLRLDVPISGDPTPTVIWQKVNKKTDLLR SNDDSMTPSENSSDLNTDSKLLFESEGRVRVEKHEDHCVFIIEGAEKEDEGV YRVIVKNPVGEDKADITVKVIDVPDPPEAPKISNIGEDYCTVQWQPPKYDG GQPVLGYILERKKKKSYRWMRLNFDLLKELTYEAKRMIEGVVYEMRIYAV NSIGMSRPSPASQPFMPIAPPSEPTHFTVEDVSDSTVALKWRPPERIGAGGLD GYIVEYCKEGSTEWIPALPGLTERTSALIKDLVTGDKLHFRIKAINLAGESGA AIIKEPITVQEILQRPKIWLPRHLRQTLVKKVGETINIVIPFQGKPRPKITWTK DGQSLDAKDVGIRNSNTDTILFIRKAELHHSGAYEITLQIENMTDKVAITIQII DKPGPPQNIKLTDVWGFNAALEWTPPQDDGNAQILGYTVQKADKKTMEW YTVFDHYRRTNCVVSELIMGNEYFFRVFSENLCGLSETAATTKNPAYIQKT GTTYKPPCYKEHDFSEPPKFTHPLVNRSVIAGYNTTLSCAVRGIPKPKIYWY KNKMDLSGDAKYRMFSKQGVLTLEIRKPTPFDGGFYTCKAVNERGEAEIEC KLDVRGNHMSIMHNTNTAVK
195	Domestic Muscovy Duck A0A8C3CXN6 (Myomesin 2)	MSSLAVPFYQRRHKHFDQSYRNIQTRYVLEEYAARKAASRQAAYYESSGL GKTTCRLCARRARTLAHETMQESRKRTHEQKSHASEEKRIRFASELSALEK EIHTARHRAREQLDRLAIQRMVEENMAMERHAVEERISRAPEILVRLRSHT VWEKMSVSLCFTVQGFPTPVVQWYKNEELITPASDPAKYRVESKYGVHVL HINRADFDDSATYSAVATNVHGQASTNCAVVVRRFREGEEPHPAGIMPFHL PLSYDVCFTHFDVQFLEKFGVTFATEGETLTLKCSVLVTPELKRLRPRAEW YRDDVLIKDSKWTKLYFGEGQAALSFTHLNKDDEGLYTLRMITKGGVNEC SAYLFVRDADALIAGAPGAPMDVKCHDANRDYVIVTWKPPNTTSQSPVIG YFVDKCEVGLENWVQCNDAPVKICKYPVTGLYEGRSYIFRVRAVNNAGIS RPSRVSEPVAALDPVDLERTQTIHVDEGRKIVISKDDLEGDIQIPGPPTNVHA SEISKTYVVLSWDPPVPRGREPLTYFIEKSMVGSGSWQRVNAQVAVKSPRY AVFDLAEGKPYVFRVVSANKHGISDPSEITEPIQPQDTVVVPSAPGRVVATR NTKTSVVVQWDKPKHEEDLYGYYIDYSVVGSNHWEPSNHKPIKYNRFVVH GLETGEQYIFRVKAVNAVGFSENSQESEAIKVQAALTCPSYPHGITLLNCDG HSMTLGWKAPKYSGGSPILGYYMDKREANHQNWHEVNSSLITRTIYTVED LTEDAFYEFKIAAANIVGIGQPSDPSEHFKCKAWTMPEPGPAYDLTVCEVR NTSLVLLWKAPVYEGKSPITGYLVDYKEVDKEDWITANEKPTAQRYFKVT DLHEGHTYVFKVRAVNDAGVGKSSEVSEPVLVEAKPGTKEIFSGVDDEGNI YLAFECKEATDASHFAWGKSYEEIEDSDKYKIETEGDRSKLYFKNPDKSDL GTYSISVSDTDGVSSSFVMDDEELERLMAISNEIKNPTIPLKSELAYEVLEKG EVRFWIQAESLSPDSTFRFVINDKEVENSDRHKINCDRPNGIIEMVMDRFTID

TABLE 11-continued

Domestic Muscovy Duck Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		NEGTYTVQIQDGKAKNQSSLVLIGDAFKAVLAESELQRKEFLRKQGPHFSE FLYWEVTEECEVLLACKVANTKKETVFKWYKDGSGIDVDEMPDLQKGEC HLSIPRLSRKDEGVYKATLSDDRGHDVSTLELSGKVYDDIILALSRVSGKTA SPLKILCTEEGIRLQCFLKYYNEEMKISWSHRESKISSGEKMKIGGGENIAWL QITEPTEKDKGNYTFEIFGDKESFKRTLDLSGQAFDDAYAEFQRLKAAAFAE KNRGKVIGGLPDVVTIMDGKTLNLTCTVFGNPDPEVVWFKNDKALELNEH YVVSLEQGKYASLTIKGVTSEDSGKFSIHVKNKYGGETVDVTVSVYRHGEK VPEVKPGQLAKPRPIPPSEGSS
196	Domestic Muscovy Duck A0A8C3CT67 (Phosphopyruvate hydratase)	MVFKMSILKIHAREIFDSRGNPTVEVDLYTNKGLFRAAVPSGASTGIYEALE LRDNDKTRYMGKGVSKAVEHVNKTIAPALISKNVNVVEQEKIDKLMLDM DGSENKSKFGANAILGVSLAVCKAGAAEKGVPLYRHIADLAGNPEVILPVP AFNVINGGSHAGNKLAMQEFMILPVGADSFKEAMRIGAEVYHNLKNVIKE KYGKDATNVGDEGGFAPNILENKEALELLKTAIGKAGYSDKVVIGMDVAA SEFYRDGKYDLDFKSPDDPSRYISPDQLADLYKGFVKNYPVVSIEDPFDQD DWAAWKKFTGSVGIQVVGDDLTVTNPKRIAKAVEEKSCNCLLLKVNQIGS VTESLQACKLAQSNGWGVMVSHRSGETEDTFIADLVVGLCTGQIKTGAPC RSERLAKYNQLLRIEEELGSKARFAGRNFRNPRIN
197	Domestic Muscovy Duck A0A8C3BXQ0 (Myosin light chain 2)	MAPKKAKKRIEGANSNVFSMFEQAQIQEFKEAFTIMDQNRDGFIDKADLRD TFAALGRLNVKNEEIDEMIKEAPGPINFTVFLTMFGEKLKGADPEETILNAF KVFDPEGKGLKSAYIKEMLMTQGERFSQEEVDQMFAAFPPDVSGNLDYKN LVHVITHGEEKD
198	Domestic Muscovy Duck A0A8C3BJT4 (Uncharacterized protein)	MTSPDAEMAAFGEAAPYLRKSEKERIEAQNKPFDAKSSVFVAHPKESFVKG TIQSKESGKVTVKTEGGETLTVKEDQIFSMNPPKYDKIEDMAMMTHLHEPA VLYNLKERYAAWMIYTYSGLFCVTVNPYKWLPVYNPEVVLAYRGKKRQE APPHIFSISDNAYQFMLTDRENQSILITGESGAGKTVNTKRVIQYFATIAASG EKKKEEQSGKMQGTLEDQIISANPLLEAFGNAKTVRNDNSSRFGKFIRIHFG ATGKLASADIETYLLEKSRVTFQLKAERSYHIFYQIMSNKKPELIDMLLITTN PYDYHFVSQGEITVASINDQEELMATDSAIDILGFTADEKVAIYKLTGAVMH YGNLKFKQKQREEQAEPDGTEVADKAAYLMGLNSADLLKALCYPRVKVG NEYVTKGQTVQQVNNSVGALAKAVYEKMFLWMVVRINQQLDTKQPRQY FIGVLDIAGFEIFDFNSLEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEW EFIDFGMDLAACIELIEKPMGIFSILEEECMFPKATDTSFKNKLYDQHLGKSS NFQKPKPAKGKAEAHFSLVHYAGTVDYNITGWLEKNKDPLNETVIGLYQK SSVKTLALLFANYGGAEAEASGGGASGGKKGGKKKGSSFQTVSALFRENL NKLMTNLRSTHPHFVRCIIPNETKTPGAMEHELVLHQLRCNGVLEGIRICRK GFPSRVLYADFKQRYKVLNASAIPEGQFIDSKKASEKLLGSIDVDHTQYKFG HTKVFFKAGLLGLLEEMRDEKLAQLITRTQARCRGFLMRVEYQRMVERRE SIFCIQYNVRAFMNVKHWPWMKLFFKIKPLLKSAESEKEMANMKEEFEKT KEELAKSEAKRKELEEKMVKLVQEKNDLQLQVQAEADALADAEERCDQLI KTKIQLEAKIKEVTERAEDEEEINAELTAKKRKLEDECSELKKDIDDLELTL AKVEKEKHATENKVKNLTEEMASLDETIAKLTKEKKALQEAHQQTLDDLQ AEEDKVNTLTKAKTKLEQQVDDLEGSLEQEKKLRMDLERAKRKLEGDLKL SQDTIMDLENDKQQLDEKLKKKDFEISQIQSKIEDEQALGMQLQKKIKELQ ARIEELEEEIEAERTSRAKAEKHRADLSRELEEISERLEEAGGATAAQVEMN KKREAEFQKMRRDLEEATLQHEATAATLRKKHADSTAELGEQIDNLQRVK QKLEKEKSELKMEIDDLASNMESVSKAKANLEKMCRTLEDQLSEIKTKEEE HQRMINDLTAQRARLQTESGEYSRQVEEKDSQISQLSRSKQAFTQQIEELKR HLEEEIKAKNALAHALQSARHDCDLLREQYEEEQEAKGELQRALSKANSE VAQWRTKYETDAIQRTEELEEAKKKLAQRLQDAEEHVEAVNAKCASLEKT KQRLQNEVEDLMIDVERTNAACAALDKKQKNFDKILAEWKQKYEETQAE LEASQKESRSLSTELFKMKNAYEESLDHLETLKRENKNLQQEISDLTEQIAE GGKAIHELEKVKKQIEQEKSEIQAALEEAEASLEHEEGKILRLQLELNQVKS EIDRKIAEKDEEIDQLKRNHLRIVESMQSTLDAEIRSRNEALRLKKKMEGDL NEMEIQLSHANRQAAEAQKNLRNTQAVLKDTQIHLDDALRTQEDLKEQVA

TABLE 11-continued

Domestic Muscovy Duck Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		MVERRANLLQAEIEELRAALEQTERSRKVAEQELMDASERVQLLHTQNTSL INTKKKLETDIVQIQGEMEETIQEARNAEEKAKKAITDASMMAEELKKEQD TSAHLERMKKNLDQTVKDLQHRLDEAEQLALKGGKKQIQKLEARVRELEG DVDAEQKRSAEAVKGVRKYERRVKELTYQSEEDRKNILRLQDLVDKLQM KVKSYKRQAEEAEELSNVNLSKFRKIQHELEEAEERADIAESQVNKLRAKS REFHGKKIEEEE
199	Domestic Muscovy Duck A0A8C3BJT4 (Uncharacterized protein)	MTSPDAEMAAFGEAAPYLRKSEKERIEAQNKPFDAKSSVFVAHPKESFVKG TIQSKESGKVTVKTEGGETLTVKEDQIFSMNPPKYDKIEDMAMMTHLHEPA VLYNLKERYAAWMIYTYSGLFCVTVNPYKWLPVYNPEVVLAYRGKKRQE APPHIFSISDNAYQFMLTDRENQSILITGESGAGKTVNTKRVIQYFATIAASG EKKKEEQSGKMQGTLEDQIISANPLLEAFGNAKTVRNDNSSRFGKFIRIHFG ATGKLASADIETYLLEKSRVTFQLKAERSYHIFYQIMSNKKPELIDMLLITTN PYDYHFVSQGEITVASINDQEELMATDSAIDILGFTADEKVAIYKLTGAVMH YGNLKFKQKQREEQAEPDGTEVADKAAYLMGLNSADLLKALCYPRVKVG NEYVTKGQTVQQVNNSVGALAKAVYEKMFLWMVVRINQQLDTKQPRQY FIGVLDIAGFEIFDFNSLEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEW EFIDFGMDLAACIELIEKPMGIFSILEEECMFPKATDTSFKNKLYDQHLGKSS NFQKPKPAKGKAEAHFSLVHYAGTVDYNITGWLEKNKDPLNETVIGLYQK SSVKTLALLFANYGGAEAEASGGGASGGKKGGKKKGSSFQTVSALFRENL NKLMTNLRSTHPHFVRCIIPNETKTPGAMEHELVLHQLRCNGVLEGIRICRK GFPSRVLYADFKQRYKVLNASAIPEGQFIDSKKASEKLLGSIDVDHTQYKFG HTKVFFKAGLLGLLEEMRDEKLAQLITRTQARCRGFLMRVEYQRMVERRE SIFCIQYNVRAFMNVKHWPWMKLFFKIKPLLKSAESEKEMANMKEEFEKT KEELAKSEAKRKELEEKMVKLVQEKNDLQLQVQAEADALADAEERCDQLI KTKIQLEAKIKEVTERAEDEEEINAELTAKKRKLEDECSELKKDIDDLELTL AKVEKEKHATENKVKNLTEEMASLDETIAKLTKEKKALQEAHQQTLDDLQ AEEDKVNTLTKAKTKLEQQVDDLEGSLEQEKKLRMDLERAKRKLEGDLKL SQDTIMDLENDKQQLDEKLKKKDFEISQIQSKIEDEQALGMQLQKKIKELQ ARIEELEEEIEAERTSRAKAEKHRADLSRELEEISERLEEAGGATAAQVEMN KKREAEFQKMRRDLEEATLQHEATAATLRKKHADSTAELGEQIDNLQRVK QKLEKEKSELKMEIDDLASNMESVSKAKANLEKMCRTLEDQLSEIKTKEEE HQRMINDLTAQRARLQTESGEYSRQVEEKDSQISQLSRSKQAFTQQIEELKR HLEEEIKAKNALAHALQSARHDCDLLREQYEEEQEAKGELQRALSKANSE VAQWRTKYETDAIQRTEELEEAKKKLAQRLQDAEEHVEAVNAKCASLEKT KQRLQNEVEDLMIDVERTNAACAALDKKQKNFDKILAEWKQKYEETQAE LEASQKESRSLSTELFKMKNAYEESLDHLETLKRENKNLQQEISDLTEQIAE GGKAIHELEKVKKQIEQEKSEIQAALEEAEASLEHEEGKILRLQLELNQVKS EIDRKIAEKDEEIDQLKRNHLRIVESMQSTLDAEIRSRNEALRLKKKMEGDL NEMEIQLSHANRQAAEAQKNLRNTQAVLKDTQIHLDDALRTQEDLKEQVA MVERRANLLQAEIEELRAALEQTERSRKVAEQELMDASERVQLLHTQNTSL INTKKKLETDIVQIQGEMEETIQEARNAEEKAKKAITDASMMAEELKKEQD TSAHLERMKKNLDQTVKDLQHRLDEAEQLALKGGKKQIQKLEARVRELEG DVDAEQKRSAEAVKGVRKYERRVKELTYQSEEDRKNILRLQDLVDKLQM KVKSYKRQAEEAEELSNVNLSKFRKIQHELEEAEERADIAESQVNKLRAKS REFHGKKIEEEE
200	Domestic Muscovy Duck A0A8C3CNB8 (Albumin)	MKWVTLISFIFLFSSATSRNLQRVARDADHKSEIAHRYNDLKEETFKAVAMI TFAQYLQRCSYDGLSKLVKDVVDLAQKCVANEDAPECSKSLPSIFLDEICQ VEKLRDSYGAMADCCAKSDPERNECFLSFKIPQPDFVQPYQRPASDVICKE YEDNRVSFLGHFIYSVARRNPFLYAPTILGVAADYEHALKSCCKESDVGAC LDGKETAIREKVKKISVKQQYSCGILQKFGDRIFQADKLAFLSQKYPKASFT EISKLIHDVKDVYKECCEGDMVECMDDRAELVNYICSKHESFSSKIKDCCE KPIVERSQCIMEAEFDEKPADLPSLVEKYIQDKEVCKSFEAGHDAFMSEFVY EYSRRHPEFSTQLILRIAKGYETLLEKCCKTDNPAECYANAQDQLNEHIKET QDVVKTNCDLLTTHGEADFLKALLIRYTKKMPQVSTETLLEIGKKMTGVG TKCCKLPEDRRLPCSEGYLSMVIQDMCRRQETSPINDNVSHCCSDSYAYRR PCFSAMGVDTKYVPPPFNPEMFNFDEKLCSAPAEERAAGQMKLLVNLIKRK PQMTEEQIKTIADGFTAMVDKCCKQSDIETCFGEEGANLIVQSRATLGIGA

[0057] An animal protein can comprise a salmon protein (Table 12). A salmon protein can comprise a salmon muscle protein. A salmon protein can comprise a salmon blood protein. A salmon protein can comprise a salmon albumin protein. A salmon protein can comprise a sequence that has at least or about 70%, at least or about 75%, at least or about 80%, at least or about 90%, at least or about 91%, at least or about 92%, at least or about 93%, at least or about 94%, at least or about 95%, at least or about 96%, at least or about 97% at least or about 98% at least or about 99% or about 100% sequence identity to any one of SEQ ID NO: 201-216.

[0058] A food product may comprise a fragment or a truncation of a salmon protein. The food product may include a fragment or truncation of any one of SEQ ID Nos: 201-216. A sequence of the fragment or truncation may be identical or substantially identical to a portion of any one of SEQ ID Nos: 201-216. A sequence of the fragment or truncation may be at least about 70%, 80%, 90%, 95%, 98%,

or more identical to a portion of a sequence of any one of SEQ ID Nos: 201-216. A length of the fragment or truncation may be at least about 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of a length of any one of SEQ ID Nos: 201-216. A length of the fragment or truncation may be less than or equal to about 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, or less than a length of any one of SEQ ID Nos: 201-216.

TABLE 12

Salmon Proteins

SEQ ID NO	Description	Amino Acid Sequence
201	Salmon B5DGP0 (Creatine kinase)	MPFGNTHNNFKLNFKVEEEYPDLTKHNNHMAKVLTKDMYAKLRDKQTPSG FTLDDVIQTGVDNPGHPFIMTVGCVAGDEESYEIFKDLLDPIISDRHSGYKPT DKHKTDLNFENLKGGDDLDPNYVLSSRVRTGRSIKGYTLPPHNSRGERRAV EKLSVEALNTLDGEFKGKYYPLNKMTDAEQEQLIADHFLFDKPVSPLLLGA GMARDWPDARGIWHNDAKSFLVWVNEEDHLRVISMEKGGNMKEVFRRFC VGLKRIEETFKKHNHGFMWNEHLGYVLTCPSNLGTGLRGGVHVKLPKLSTH PKFEEILTRLRLQKRGTGGVDTASVGGVFDISNADRLGSSEVDQVQMVVDG VKLMVEMEKKLEKGEAIDGMIPAQK
202	Salmon A0A1S3RGX6 (Alpha-actinin-3)	MTAIETQVQYGSYMTSEQVYMTQEDDWDRDLLLDPAWEKQQRKTFTAWC NSHLRKAGTQIENIEEDFRNGLKLMLLLEVISGERLPKPDKGKMRFHKIANV NKALDFICSKGVKLVSIGAEEIVDGNVKMTLGMIWTIILRFAIQDISVEETSAK EGLLLWCQRKTAPYRNVNVQNFHISWKDGLALCALIHRHRPDLIDYSKLRK DDPMGNLNTAFEVAEKYLDIPKMLDAEDIVNTPKPDEKAIMTYVSCFYHAF AGAEQAETAANRICKVLAVNQENEKLMEEYEKLASELLEWIRRTIPWLENR VAEQTMRAMQQKLEDFRDYRRVHKPPKVQEKCQLEINFNTLQTKLRLSNRP AFMPSEGKMVSDIANAWKGLEQVEKGYEEWLLTEIRRLERLDHLAEKFKQK SSLHESWTSGKVELLSMKDYESASLMEIRALMRKHEAFESDLAAHQDRVEQ IAAIAQELNELDYHDAVTINARCQGICDQWDNLGTLTQKRRDSLERVEKLW ETIDQLYLEFAKRAAPFNNWMDGAMEDLQDMFIVHSIEEIQSLITAHDQFKA TLPEADKERVATMGIQNEIVKIAQTYGIKLSGVNPYTNLSPQDITDKWDAVK HLVPLRDQMLQEEVARQQSNERLRRQFAAQANIIGPWIQTKMEEISHVSVDI AGSLEEQMSNLKQYEQNIINYKCNIDKLEGDHQLSQESLIFDNKHTNYTMEH VRVGWEQLLTTIARTINEVENQILTRDAKGISQEQLNEFRASFNHFDRKRNG MMDPDDFRACLISMGYDLGEVEFARIMTLVDSNNTGVVTFQAFIDFMTRET AETDTADQVMASFKILASDKTYITVEELRRELPPEQADYCISRMTSYIGSGAP PGALDYISFSSALYGESDL
203	Salmon C0H8W5 (Alpha-1,4 glucan phosphorylase)	MPKPLTDQDKRKQISVRGLAGVENVSDLKTNFNRHLHFTLVKDRNVATKR DYYFALANTVRDHLVGRWIRTQQHYYEKDPKRVYYLSLEFYMGRTLQNTM VNLALENACDEATYQLGLDMEELQDIEEDAGLGNGGLGRLAACFLDSMAS LGLAAYGYGIRYEFGIFNQKIVNGWQVEEADDWLRYGNPWEKARPEYMRP VHFYGRVEHTPEGVKWVDTQVVLALPYDTPVPGYRNNVVNTMRLWSAKA PCDFNLKDFNVGGYIQAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYFVV AATLQDIVRRFKSSKFGSTEVVRVDLSTLPDKVAIQLNDTHPAMAIPELMRIL LDTEHQTWEKAWDICTRTCAYTNHTVLPEALERWPTDLLQNLLPRHLEIIYE INRRHLERIAKLYPGDHDRLRRMSLVEEGDQKKINMAHLCIVGSHAVNGVA RIHSDILKATLFKDFYEVDPHKFQNKTNGITPRRWLVMCNPGLAEVIAERIGE DYIRDLDQLKKLLTFVDDDSLIHDIAKVKQENKLKFSAYLEEHYKVKINPNS MFDVQVKRIHEYKRQLLNCLHIITLYNRIKKEPNKNWTPRTIMIGGKAAPGY HTAKLIIRLITAIGEIVNHDTVVGDRLKVIFLENYKVTLAEKIIPASDLSEQIST AGTEASGTGNMKFMLNGALTIGTMDGANVEMAEEAGEENLFIFGMRVDDV DAMDKSGYDAMDYYNRIPELKQAMDQIAGGFFSPDQHDLFKDIVNMLLHH DRFKVFADYEAYIKCQEKVSALYKNPKEWTKMVIHNIAGCGKFSSDRTISQY AREIWGMEPSLERIPAPDDPRQ
204	Salmon C0H911 (Fructose-bisphosphate aldolase)	MPHSFPFLSTEQKKELSDIAQKIVAPGKGILAADESTGSVAKRFQSINAENTE ENRRLYRQLLFTADRVEPCIGGVIFFHETLYQKTDDGKLFPQLIKERDMVVGI KVDKGVVPLAGTNGETTTQGLDGLYERCAQYKKDGADFAKWRCVLKITDT TPSELAIKENANVLARYASICQMHGIVPIVEPEILPDGDHDLKRCQYVTEKVL AAVYKALSDHHVYLEGTLLKPNMVTPGHSCSQKYSNQEIAMATVTALRRT VPPAVPGITFLSGGQSEEEASINLNAMNQCPLHRPWAITFSYGRALQASALK AWGGKPENGKACQEEFIKRALNNSKACQGKYVSSGDKGAAGGESLYVANH AY
205	Salmon B5DH12 (Myosin light chain 1-1)	MAPKKDAKAPAKKAEPAKAAAPAPAPEPDVVAAPPPLDLSTVKVEFTPDQQ EDYKEAFGLFDRVGDAKVAYNQVADIMRALGQNPTNKEVRKVLGNPSDED MAAKRLEFEAFLPMLQHIVNDPNKGTEDDYVEGLRVFDKEGNGTVMGAEL RIVLGTLGEKMSEAEIDALMQGQEDENGSINFEAFVKHIMSI

TABLE 12-continued

Salmon Proteins		
SEQ ID NO	Description	Amino Acid Sequence
206	Salmon B5DGR3 (Glyceraldehyde-3-phosphate dehydrogenase)	MVKVGVNGFGRIGRLVTRAAFHSKKGVEIVAINDPFIDLDYMVYMFKYDST HGRFHGEVKAEGGKLVIDGHKITVFHERDPANIKWGDAGATYVVESTGVFT TIEKASTHLKGGAKRVVISAPSADAPMFVMGVNHEKYDNSLKVVSNASCTT NCLAPLAKVIHDNYHIIEGLMSTVHAVTATQKTVDGPSGKLWRDGRGASQN IIPASTGAAKAVGKVIPELNGKITGMAFRVPTPNVSVVDLTVRLEKPASYDAI KKVVKAAADGPMKGILGYTEQQVVSSDFNGDTHSSIFDAGAGIALNDHFVK LVTWYDNEFGYSNRVIDLMAHMATKE
207	Salmon Q91472 (Fast myotomal muscle tropomyosin)	MDAIKKKMQMLKLDKENALDRAEGAEGDKKAAEDKSKQLEDDLVALQKK LKGTEDELDKYSESLKDAQEKLEVAEKTATDAEADVASLNRRIQLVEEELD RAQERLATALTKLEEAEKAADESERGMKVIENRASKDEEKMELQDIQLKEA KHIAEEADRKYEEVARKLVIIESDLERTEERAELSEGKCSELEEELKTVTNNL KSLEAQAEKYSQKEDKYEEEIKVLTDKLKEAETRAEFAERSVAKLEKTIDDL EDELYAQKLKYKAISEELDNALNDMTSI
208	Salmon A0A1S3RC27 (Myosin-binding protein C)	MPEAKKPAAKTDKAAAAPADKEARSAIEEGSECRDELPDDETVPGGQSELT GLFVERPESTTVIKGKNVTFVAKVDSSDLLRKPNMKWLKGKWLDLGSKAG KHVQFKEAYDRNSKVYTYEMSIIKVVEGDAGGYRCEVTSKDKCDSCTFEVT VEAVQEEQPANILDAFKRSGDAGEDAGDLDFSALLKKREKKARDEPKEDVD VWEILKDAKPCDYEKIAFDYGITDLRGLLKRLKKMKKVEPKKSDAFLKKLE QCYSVDKGKRTQMHVELHDPNVQVKWLKNGVEIKPSAKYVFECVGNKRTL TINKCNLSDDAAYECVVGEEKSFTEVFVKEPPVTITKLLDDVHVVVGEKVEF ECEVSEEGANVKWMKDGVELTKDGKIRIKKDGKKHTLVINEATIEDIGMYY VYTNGGESKGELEVEAKELEVLQSIADLSVKACEQAMFKCEVSDEKVVGK WFKDGVEVKPGNRIKMSHIGRIHKLTIDDVKPQDEGNYTFVPEGYALSLSAK LNFIEIKIEYVPRQDPPKIHLDTSSTGSKNTIVVVAGNKLRLDVEITGEPVPTV CWMKGDTVISEAEGRVRVETRTTLSSFVIEGAERPDEGQYSIIVTNPAGEDRA ELTVKIVDVPNPPENVRCMGVGEDTATITWDPPKFDGGAPIKGYLMERKKQ GSSRWTKLNFEVFESTTYEAKKMIEGVFYEMRVFAVNGIGISQPSGNSKPFM PIAPTSEPTRLTVEDVTDSTCALKWRPPERVGAGGVDGYIIEWCKEGEDNWV AANKEPVDKNTYRVKGLPTGEKLLFRVVAMNIAGRSPPCTMKQSVTIREIM EYPKIRLPRQLRTKFIRKVGEKINLVIPFQGKPRPVVNWLKDGEPLENKSVGI RTSDFDTILFIRSAERDHSGTYTLSVQIDNMQDKADIHIQIVDKPGPPINVMVT DVWGFNAALEWKPPKDTGNTDITGYTIQKADKKTQGWFTVYEHNRRPSCT ASDLVMGNEYSFRVFSENICGLSDEVAFSKNTAIIGKTDLEYNKPAFKERDM SSSPKFTAPLVDRVVVAGYSTAISCAVRAYPKAKIVWMKNKMIIGEDPKFLM QNNQGVLTLNIRKPGQFDGGKYSCKAINDLGEDEVECRLEVRVLKEKKEGD EEKK
209	Salmon A0A1S3L6A0 (Myomesin-2 isoform X1)	MATKVMHFNQKKHVSHVSHHSSHTTKHVVKEQSSRKSSSKSVNQVQHTHA TEAMAEPAYTVPAFRQRSADEIEEYQRVSSHVGKGLASIQKELHRMRVVTK TQVENIAIKREVQEMMHKKMTLSTETDKAPDFMVALRPHTVWEKTPVKLF CTVDGHPRPIVKWYKGGKPVDPLSAPGKYKIESKYGVHSLIISRCMVSDTAE YSAVATNSHGSTTSKASVIVKRPAGGAYGSCQLFGQVPHLTVIHHSKLEITM LDRFGVSFGTEGGSISLVCTMVVVPDLPNVPPLAQWYRDDKLLKAGKLAEI KVGGGAATLTLPHLAKDDEGLYTLRMWTKDGTTEHSAYLFVKDAAPSVAG APGAPISVKAFDINSDYVLVAWKPPNTTNEAAITGYFVDKRESGSATWSQCN DAPVQICKYPVHGLNVGHAYHFRVRAVNSAGISRPSRESDKVTALDPAERE RLQVIKLDGKHEVVIKDDDLEGYVTLPGEPTDMHASEIFKSYIVLSWKPPSPR GRAPLWYIIEKSVSGTGAWQRINTAVKLQSPRYPVFDLEAGKKYQFRVYSE NLYGASEASKPTDPIEKVDEHGVPSAPGQVVATRNTKSSVFVQWDPPKHPN HLMGYYIDASLVGSKTWAPCNHKPYKNTRFVVHGLTPGETYVFRVQAVNV YGLSDESQESTPIAVEPALTTPSAPHGITMLSCDGASMIIAWNSPKHRGGSKI NAYYVDKRDADSLAWKEVNSAPTNTRTYTVDGLTEGTFYEFKVQAGNLAG VGVPSAPSTPLKCEAWTSAVPGPAYDLAFREVRGDSLVILWKAPVYTGASA VTGYIVEMAKKGHHYHPLNEEAIGHCYLQVTGLEAGAEYTFKVKAVNAEG VGKASQTSEPVFAKALPGTQEIQCGVDEDTGDIFLSFEACEISETSNFAWSKN YKEISDCTRVSIETKGKHSKLTFVNPDVSDLGAFSVHVTDTDGVSASHTVTE DALNTMLELSYNIRHPIVPLKHQLNYEVLEKGHVRFWLQCLKMSPAVNYRF IVNDKEVTSSDSHKISHDVNTGVIQMTVDHFSKASEATYTVQIHDGRAKNQS SLVLVGDVFKAALKEAEFQRSEHIRKQGPHFAEYLSYHVDDDCTVLLVCKV ANVKKETVFHWFKDEEEIVPETPPNVMSGSVALPISLFSRKDQGHYKAKLSD DRGKDTSVFDISGQVFDDIINAIAHIAGSSASELVLQCTPEGIRLQCYMKYYT EEIRTAWLHKDSKISSSEKMRIAGTAEMAWMQIREPSEKEKGHYSIQIMDAT KSHTRTFDLSGQAYTDAYEEYLRLKAAAFAEKNRGRVVGGLPDVVTIMEQ KTLSLTCTVWGDPSPEVTWFKNENEVVSTEHAKITLEANKFASLTITAVTSE DSGKYSINVRNKYGGEFVEITVSVYKQGESPPEPKMGGRGATPAPLASKTPA KTPTPAQTPTPSLKSPTPSLKSPTPSLRSPAKSPTPTPKSPTPPRRVKSPSPARSL KSPTPPRK

TABLE 12-continued

Salmon Proteins		
SEQ ID NO	Description	Amino Acid Sequence
210	Salmon ENOB (Beta-enolase)	MSITKIHAREILDSRGNPTVEVDLYTAKGRFRAAVPSGASTGVHEALELRDG DKSRYLGKGTVKAVDHVNKDIAAKLIEKKFSVVDQEKIDHFMLELDGTENK SKFGANAILGVSLAVCKAGAAEKGVPLYRHIADLAGHKDVILPCPAFNVING GSHAGNKLAMQEFMILPIGASNFHEAMRIGAEVYHNLKNVIKAKYGKDATN VGDEGGFAPNILENNEALELLKTAIEKAGYPDKIIIGMDVAASEFYKAGKYD LDFKSPDDPARYITGDQLGDLYKSFIKGYPVQSIEDPFDQDDWAAWTKFTAA VDIQVVGDDLTVTNPKRIQQAVEKKACNCLLLKVNQIGSVTESIKACKLAQS NGWGVMVSHRSGETEDTFIADLVVGLCTGQIKTGAPCRSERLAKYNQLMRI EEELGAKAKFAGKDYRHPKIN
211	Salmon Q7ZZN0 (Myosin regulatory light chain 2)	MAPKKAKRRGAAAEGGSSNVFSMFEQSQIQEYKEAFTIIDQNRDGIISKDDL RDVLASMGQLNVKNEELEAMVKEASGPINFTVFLTMFGEKLKGADPEDVIV SAFKVLDPEATGFIKKEFLQELLTTQCDRFSAEEMKNLWAAFPPDVAGNVD YKQICYVITHGEEKEE
212	Salmon A0A1S3LA68 (Myosin heavy chain)	MTTLGLTEGHPAKYQHLTNLVPAISYELGQKKERVIMSTDAEMAAFGPAAI YLRKPERERIASQAAPFDAKTAFFVVEPKEMYLKGVLQSKEGGKATVKTLC NKVLTVKEDDIHPMNPPKYDKIEDMAMMTHLNEATVLYNLKERYAAWMI YTYSGLFCVTVNPYKWLPVYDSVVVNAYRGKKRIEAPPHIFSISDNAYQFML TDHENQSILITGESGAGKTVNTKRVIQYFATIAVAGGKKEQTAAATSGKIKG SLEDQIIAANPLLEAYGNAKTVRNDNSSRFGKFIRIHFGTTGKLASADIETYLL EKSRVTFQLSAERSYHIFYQLMTGHQPQLLEALLITTNPYDYPIISHGEIAVKS IDDTEEFIATDTAIDILGFTAEEKIGIYKLTGSVMHHGAMKFKQKQREEQAEP DGTEEADKISYLMGLNSADLLKALCYPRVKVGNEMVTKGQTVPQVNNSTM ALCKSVYEKMFLWMVVRINEMLDTKQSRQFFIGVLDIAGFEIFDYNSLEQLC INFTNEKLQQFFNHHMFVLEQEEYKKEGIEWAFIDFGMDLAACIELIEKPMGI FSILEEECMFPKASDTTFKSKLYDQHIGKTKAFEKPKPGKGKAEAHFALVHY AGTVDYNVTGWLDKNKDPLNDSVVQLYQKSSVKLLAMLYVGAAASQADD AARGGGGKKKKGGSFQTVSALFRENLGKLMTNLRSTHPHFVRCLIPNESKTP GLMENFLVIHQLRCNGVLEGIRICRKGFPSRIQYGDFKQRYRVLNASVIPDGQ FIDNKKASEKLLGSIDVDKSQYKFGHTKVFFKAGLLGTLEEMRDEKLATLVT MTQALCRGYVMRKEFVKMMARREAIYSIQYNIRSFMNVKHWPWMKLYFKI KPLLKSAETEKEMAAMKENFEKMKEDLAKALAKIKGLEEKIVSLLQENNDL QLQIAAEENTLCDAEERCEGLIKSKIQMEAKLKETFERLEDEEELNAELTAKK RKLEDECSELKKDIDDLELTLAKVEKEKHATENKVKNLTEEMSSQDESLIKL TKEKKALQEVHQQVLDDLQAEEDKVNNLTKAKTKLEQQVDDLEGSLEQEK KVRMDLERAKRKLEGDLKLALESLMDLENDKQQSDEKIKKKDFETSQLLSK IEDEQALGAQLQKKIKELQARIEELEEEIEAERAARAKVEKQRSDLSRELEEIS ERLEEAGGATSVQIEMSKKREAEFQKLRRDLEESTLQHEATAAALRRKQAD TVAELGEQIDNLQRVKQKLEKEKSEYKMEIDDLSSNMEAIAKAKGNLEKMC RSLEDQLSELKTKNDEHVRQLNDINVQRARLLTENGELGRQMEEKESLVSQ LTRAKQTFTQQIEDLKRQIEEEVKAKNALAHGLQSARHDCDLLREQYEEEQE AKAELQRALSKANSEVAQWRSRYETDAIQRTEELEDAKKKLTQRLQDAQE AVEATNAKCASLEKTKQRLLGEVEDLMIDVERANASAAQLDKKQRNFDKV LSEWKQKYEEGQAELEGSLKEARSLSTELFKMKNSYEEALDHLETLKRENK NLQHEISDLTELVGESGKSIHELDKAKKTVENEKAEIQAALEEAEGTLEHEES KILRIQLELNQIKGEVDRKLAEKDEETEQIKRNSQRMMDSMQSTLDSEIRSRN DAMRVKKKMEGDLNEMEIQLSHANRQAAESQKQLRNVQGALKDAQLQLD DAIRVADDMKEQVAMVERRNNLMMAEIEELRAALEQTERGRKVAEQELVD ASERVGLLHSQNTSLINTKKKLEVDLTQVQGEMEDTVQEARNAEEKAKKAI TDAAMMAEELKKEQDTSSHLERMKKNLEVTVKDLQHRLDEAENLAMKGG KKQLQKLEARVRELESEVDAEQKRGVEAIKGVRKYERRVKELTYQTEEDNK NVSRLQDLVDRLQLKMKAYKRAAEDAEEQSNIHLSRYRKVQHELEEAQER ADISESQVNKLRAKSREIGSKGNVAEE
213	Salmon A0A1S3NZV0 (Nebulin isoform X4)	MDENEEFEEEFEEEEIIEEEYFEETVTSTTSHQTTSRAAAVTQQPGVPVRRVV RKKVKVDTSKFMTPYLAHSQKMLDLFSYNKYRRAYDKAKGKPYAITADTP EMMRIKKAQEQLSEVKYRMEGNKVRCTSLYDGEAREIAHVKHVSELISKVL YRAKWDETKDRYLLPPDAPELVLAVKNAANYSKKLYTEQWDEDKTMCYP YSDSPELRRVAKAQEVLSDILYKKGHAERTAKYTSLAVPNDMELAKKNFEH RSDLKYHEDYNKNVKGHWCETPYIEVAVARVAMDNLSNKMYTAAYEDIK DQINFMQTETPEYKQHKKAGHAASSIKYKKDYENSKSVADYNVLPATDNPL LRQLRYAGTILSDKVYKASYEKSRGSSINYCDTPKFKMDSVVQQFSDAHYK DKYENEVKGHYIGSYEDIYMQHCKKMEEIKSEQNYKADYEDIKTRCFYPQT MSPEYEAHKKVAQCKDKVYRQHPDTVKFTQVTDSPVQLQAAINAKQLSDL NYHAAYELDKFKCALPPDYPFFIQSRANAYNLSDSCYRYDWEKVKAKQFEV KGDAISILAARAHTNIASDVKYKKEFEKNKGKMVGALSIHDDPKILHSMHV AKIQSDREYKKNYEKTKTKYITTLDMMSNTLAKKSQGIASMTGYRTIHNRY LLPHDAMSLDLAKKANIIQSDNEYHSDYNNSTKGTGWVAFGSMEVEKAKK AGDILNERKYRQHPDTVPFTSIADHPVMMQAQVNQIMRSDLHYKKGLEEIH QKYSLPHNDPQFLQAKCNAYNISKNYYKQSWEETIAKGYDMRPDAISVIHA

TABLE 12-continued

Salmon Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		KHGRHIASDIQYKKDYVKGKGKQIGCRNLQDDPLLVHYMEVARLQSDKNY KKDYHKSKLKYHSPVDMMSVVHAKHASTVQTFAGYRKIQHHYSLLPGAM NLELARTMNTNASDVNYKADWENSVKGTGWIPIGSIPVELAKQAGAVLDEH KYRQHPENFKFSSKMDGMNMELCKHNQKIMNVQEYKASGEKFLHTYNLPA DAPEFMQAKCNAVNISKNYYTYAHRQDILKGHQVKADAIPILAARACTNIAS NYKYKLAYEKGRGKHVGFRNVKDDPLLVHYMEVAKMQSEQNYKKDYHK SKLKYHSPVDMMSLVHAKHASAAQTFAGYRKSLHNYTLLPEAMPVQLART MMEIQSNNAYKSDYNNMCRGSGWVPIGSMQVELAKTAMAARDEQGYRVH PSTLKFTSLTGQMNMALAMTNTKLLDDASYRASGEKFKHTYNLPADSPQFL QAKCNAQTMSESHYKHKWVEDIAKGYDIRPDAISVIHAKHGRHIASDIQYK KDYVKGKGKQIGCRNLQDDPLLVHYMEVAKMQSDKNYKKDYHKSKLKYH SPVDMMSLEHAKHASTVQTFAGYRKIQHHYTLTADSMNLALARTMNASSS DYEYKSDWNNSVRGTGWIPIGSLPVELAKVGGDIQSEQKYRTHPSKFQFKKL MDSMDIVLATDNNKIMNKQEYTKQWHNDKLTVHVMPDAPEIVLAKANAIN MSQKTYKAGLAKMYKNGYDLKVDAIAIQAAKQGTKIASDYKYKLSHRMN RGKQIGCRNLQDDPLLVHYMEVARLQSDKNYKKDYHKSKLKYHSPVDMM SVVHAKHASTVQTFAGYRKIQHHYTVLADAMNVELAKQMQTIASDNLYKS DYETSVKGTGWIPIGSLDVEKAKTAGLIGSEKLYRQHPSSFKFTKDMHSMDV VLANANNAIMNKQTYVDAWNKDKTSVHIMPDSMEVVLARANKINCSATLY KLANEQDKKKGYDLRNDAIAILAAKQGTSIASDFKYRAGYRKQVGHHIGAR CVEEDPLLMLAINSAKIASDALYKKDFHQSKTKFHLPLDMIAFELAKKNQIQ VNHAKYVTRLHNWTCLPDSTDVVHARHVYDIASDAVYKADLKCLQGIGW VPIGSLDVEKAKKAGEVLSEKLYRQHPSNFKFTSSVQDIGMVLAKQNNDNA NKQAYTEAWETDKTKNHIMPDAMDIVLANANKSNYSWNQYRKAMVEDK KKGYDMRNDAISIIAAKHGREIISDHKYKSSYRKQVGHHIGARCVEDDPLLV LALNSAKIASDALYKKDFHQSKTKFHLPLDMIAFELCKKNQIQVNDFNYRTH LHNWTCLPDSTDVVHARHVYDIASEAVYRADMKWLRGIGWVPTGSLDVEK AKKAGEILSEKKYRQHPSTLDFKSPLDGMDIVLAKHNGQIMNKRQYVEQW DADKIKIHVGNELPELVLAKANAITMSNKLYKQGFDETIKKGYYLPVDAISV KAAKHGREIISDYKYKAGYSKQVGHHIGARCVEEDPLLMLAINSAKIASDAL YKKDFHQSKTKFHLPLDMIAFELCKKNQIQVNDFNYRTHLHNWTCLPDSTD VVHARRVYDLHSDAVYRADMEFIRGTGWVPIGSLDVLKAKKAAEILSDRLY KQKPDSLKYTTDMHSMPMVLAKANAEIINKGNYVAQWEADKIKIHVGNEL PELVLAKANAINNSYKLYKAELEELYKKGYNLKSDCIAIKAAKHGREIISDY KYKAGYSEQVGHHIGARCVEEDPLIMLALNSAKIASDALYKKDFHQSKTKF HLPVDMLTIELAKKCQMQVNDFNYRTRLHNWTCLPDSSDVVQARKVYALH SDAVYRADMEWLKGTGWVPIGSIDVEKAKKAGEILSERKYREHPSNFKFTA TTDDMPYALATANAQIMDKRQYVQQWEADKIQIHVGNELPELVLAKANAI TMSNKLYREGFNEMIGQGYHLDKDAIAVQHAKKSREIISDYKYKAGFRNQC GHHIGAHCVEDDPLLVLAANAARISSDINYKKDFHLTKTKFSIPLDMMSFDL AKKSQFQIQEFTYRTKLHNWTCMPDNTDVVRARKVYSLHSDAVYRADLEW LRGCGWLPQGSVDVIKAKNAQTILNERLYRQPPSSIKFTSVVDSPEVVLSKQ NADILSDRKYREVWDTEKTNIHLNNDLPSLVLSKANAIAISNKLYTKDWDA DKAKGYHIKEDAIAVLKARKSRDIISDYKYHEGYRKQVGHHIGARCVEEDPL IMLALNSAKIASDALYKKDFNKSKTKFHLPVDMLTIELAKKCQVQVNDFNY RTHLHNWTCLPDSTDVVQARKTYDLQSDAVYKADLEWLRGCGWNPHESV EVLKVKHAQAVLADRGYRVKLDQQKFTAPVDRIDMVCAKNAAEILNEANY RALWHTEKTKYTITDTPGLATAREVAKNLHPKLYTKDWDKVKATGFFMPG DAVPIKQCIHNNKVQSDHKYKADYRKQAGHHIGARCVEDDPLIMLAINSAKI ASDALYKKDFNKSKTKFHLPVDMLTIELAKKCQVQVNDDHYRTRLHNWTC LPDSTDVTQARKAYDLQSDVVYKADLEWLRGCGWVAADCVDHVKVRNA QKIMNERLYKKDAKDNFAKYTHIVDRPEIILATINAFNLSDLKYKETFNAEK GKYIGSEDTPQLAHSREVAKVISEKYYKLAWDESKATGYQLDAHYIPVVGA KKTREIVSDVKYHDTHEKAKGHYLAGTLVDFPEVMRCGEQEKQKGLRDYH KEYNTNKAKNHIPADIIPNMVAKRCQDMLSDVLYRTLLHNWTCHPDQQDII RARKNNEILSDVYYKDDLNWIKGIGCYAWDTPELVRAKNAHLLQSDLKYK EEGKKEFNNYSIVTDTPVYVTALLGHTWASDLNYREAYHKDKHMYTTILDT YNYAHNATLKDIYSTKTYTSAWDKIKAKDYTLPSDSHALKHAKQQKVILSN VKYREDYEKFKSLYTLPKSLEDDPQTLRCLKAGKMGLDRLYKADYEKTKA KNHIPPDMLDVLSARQTQNYVSEIGYRKHLHQWICMPDMQVYVQARKVNE QLSDIFYKDDLNWLRGIGCYAWDTPEILRVKAADKLQSDNKYRAKGIEAFK DYSVVIDTPTYETAKLSSQNLSDLHYRNDYATNVKGTNSSPAVTVDTERAR LANFNQSDNLYKEANKNFMPTGYSLPHDNPLMKQAKQNSIVASSVKYKND YEKSKATNYTMDPDGVSFVNIRNANKVTNPRLYRELYEKGKGKIHSTYDTP EIRQVKMTQEAVSDLCYKEKYFANRGTMIPMGVTPQMIHCNHVNEITSDLR YKEDLLWLRGIGCFLYDTPEMSTVRNITKFRSNYEVEAKKNHANFTVVLDT PEYQRVTELKTHMSDLMYRAQGNVEKTKCTTTLDSLDVKRVKWAQQLSN KYMYTDLASKERAHFTPEKDTPIMSHARAMKVVFSEIKYKDQYEKMKHKY TSIHDTPILVRAKKAYLQSSDLRYKESFERSKGHYHSVKDALDIVCHRRVTD DISEVKYREKYINSLGTWRSMPDRPEFFHSRIIGENVSDIKYKEDLDWLKGIG CYVWDRPEMVQAEKNKVLFSERHYKETFEKNRHRFKYTCDTPFFQAAKKA SVIINDRHYRAAYEKTKDKYSSVLDDPRNLLAKENAKNSQWKYREQYERA

TABLE 12-continued

Salmon Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		KNKFTSILETPEHESHKRSKAIGDNVYKMDYNKNKAKGYTLGHDTPHSTHM KKVKEITSVLKYKELYEKSKAQINMDPESFEIRAAKEAYKNISNLDYRKKYQ ATKNKWIWTVDRPDFIQAAKANFQQSDIEYTYDKEMMKGCVISIVDDKLAV LAKMNADLSNVKYREKFEKAKGQYKTVLDTPTNLHAKSVRILASESKYKLA GKMEQQTGQFTTLPTTRDTIHAKDMGKLVSEKQYKQKFEKEKGKSKYNSM RVPPDVQHAMDVAKNQSNFAYRQEAKAKLNYTSVAMRPDIMRATQASKL ASNIGYREKARQEAAHGGSLVHRPDIALATAVSKLNSQVQYKAKYDKEKGL RPKYDMKESVVYKTQKDAHQLASEVKYKGDLKKLHKPVTDMAESLSMQH NLSTSKLSSDYQYKKKHEDSKGHYLMIADSPEQLHHKEASELQSHVKYKEK YEKAKGKAMLDFETPTYVTAKEAQHMQSAKDYHKEYEESIKGKNLSGLEV TPAMLHVRHATKIASERDYRRDLEEGVKGKGLTVFEETPELLRARNATAILS QKEYKKTLETEIKGKGMLALATDTPDFMRARNATDILSQQKYKQTAEADRA SYTSVIDTPDILHAQNIRNMVSQKKYREEAEKCMSHYVPVLDTPEMMRVRE NQKNFSTIKYKDLLCRCTAISDLPEVKRVREMQKNISSVLYKDSSAKGTPQV FTPEMERAKRNQEQISSVLYSDSFRKQVQGKAAFVLDTPEMRRVKETQRNIS GVKYHQDFEKSKGSFTPTTSDPVMDRVKKNTQEFSDINYRGIQRRVVEMEK RRVVEHEQETVTDLRVWRTNPGSVFDYDPAEDNIQSRSLHMMSVQAQRRS KEHSRSTSAMSGVGDEKSEVSEDPAHHVSLYSNGFATSSMGYTHTKTVELQ QRSSSVATQQTTVSSVPSHPSTTGKTVRAMYDYSSADMDEVSFKDGDVIVN VQSIDEGWMYGTVQRSGKTGMLPANYVEAI
214	Salmon Q78BU2 (Actin alpha 1-1)	MCDDDETTALVCDNGSGLVKAGFAGDDAPRAVFPSIVGRPRHQGVMVGM GQKDSYVGDEAQSKRGILTLKYPIEHGIITNWDDMEKIWHHTFYNELRVAPE EHPTLLTEAPLNPKANREKMTQIMFETFNVPAMYVAIQAVLSLYASGRTTGI VLDSGDGVTHNVPIYEGYALPHAIMRLDLAGRDLTDYLMKILTERGYSFVTT AEREIVRDIKEKLCYVALDFENEMATAASSSSLEKSYELPDGQVITIGNERFR CPETLFQPSFIGMESAGIHETAYNSIMKCDIDIRKDLYANNVLSGGTTMYPGI ADRMQKEITALAPSTMKIKIIAPPERKYSVWIGGSILASLSTFQAMWITKQEY DEAGPSIVHRKCF
215	Salmon ALBU2 (Albumin 2)	MQWLSVCSLLVLLSVLSRSQAQNQICTIFTEAKEDGFKSLILVGLAQNLPDST LGDLVPLIAEALAMGVKCCSDTPPEDCERDVADLFQSAVCSSETLVEKNDLK MCCEKTAAERTHCFVDHKAKIPRDLSLKAELPAADQCEDFKKDHKAFVGRF IFKFSKSNPMLPPHVVLAIAKGYGEVLTTCCGEAEAQTCFDTKKATFQHAIA KRVAELKSLCIVHKKYGDRVVKAKKLVQYSQKMPQASFQEMAGMVDKIV ATVAPCCSGDMVTCMKERKTLVDEVCADESVLSRAAGLSACCKEDAVHRG SCVEAMKPDPKPDGLSEHYDVHADIAAVCQTFTKTPDVAMGKLVYEISVRH PESSQQVILRFAKEAEQALLQCCDMEDHAECVKTALAGSDIDKKITDETDYY KKMCAAEAAVSDDNFEKSMMVYYTRIMPQASFDQLHMVSETVHDVLHAC CKDEPGHFVLPCAEEKLTDAIDATCDDYDPSSINPHIAHCCNQSYSMRRHCIL AIQPDTEFTPPELDASSFHMGPELCTKDSKDLLLSGKKLLYGVVRHKTTITED HLKTISTKYHTMKDKCCAAEDQAACFTEEAPKLVSESAELVKV
216	Salmon ALBU1 (Albumin 1)	MQWLSVCSLLVLLSVLSRSQAQNQICTIFTEAKEDGFKSLILVGLAQNLPDST LGDLVPLIAEALAMGVKCCSDTPPEDCERDVADLFQSAVCSSETLVEKNDLK MCCEKTAAERTHCFVDHKAKIPRDLSLKAELPAADQCEDFKKDHKAFVGRF IFKFSKSNPMLPPHVVLAIAKGYGEVLTTCCGEAEAQTCFDTKKATFQHAV MKRVAELRSLCIVHKKYGDRVVKAKKLVQYSQKMPQASFQEMGGMVDKI VATVAPCCSGDMVTCMKERKTLVDEVCADESVLSRAAGLSACCKEDAVHR GSCVEAMKPDPKPDGLSEHYDIHADIAAVCQTFTKTPDVAMGKLVYEISVR HPESSQQVILRFAKEAEQALLQCCDMEDHAECVKTALAGSDIDKKITDETDY YKKMCAAEAAVSDDSFEKSMMVYYTRIMPQASFDQLHMVSETVHDVLHA CCKDEQGHFVLPCAEEKLTDAIDATCDDYDPSSINPHIAHCCNQSYSMRRHC ILAIQPDTEFTPPELDASSFHMGPELCTKDSKDLLLSGKKLLYGVVRHKTTIT EDHLKTISTKYHTMKEKCCAAEDQAACFTEEAPKLVSESAELVKV

[0059] An animal protein can comprise a cod protein (Table 13). A cod protein can comprise a cod muscle protein. A cod protein can comprise a cod blood protein. A cod protein can comprise a cod albumin protein. A cod protein can comprise a sequence that has at least or about 70%, at least or about 75%, at least or about 80%, at least or about 90%, at least or about 91%, at least or about 92%, at least or about 93%, at least or about 94%, at least or about 95%, at least or about 96%, at least or about 97% at least or about 98% at least or about 99% or about 100% sequence identity to any one of SEQ ID NO: 217-231.

[0060] A food product may comprise a fragment or a truncation of a cod protein. The food product may include a fragment or truncation of any one of SEQ ID Nos: 217-231. A sequence of the fragment or truncation may be identical or substantially identical to a portion of any one of SEQ ID Nos: 217-231. A sequence of the fragment or truncation may be at least about 70%, 80%, 90%, 95%, 98%, or more identical to a portion of a sequence of any one of SEQ ID Nos: 217-231. A length of the fragment or truncation may be at least about 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of a length of any one of SEQ ID Nos: 217-231. A length of the fragment or truncation may be less than or equal to about 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, or less than a length of any one of SEQ ID Nos: 217-231.

TABLE 13

Cod Proteins

SEQ ID NO	Description	Amino Acid Sequence
217	Cod A0A8C4Z9R4 (Creatine kinase)	MAKNCHNDYKMKFSGDEEFPDLSLHNNHMAKVLTKDIYNKLRGKSTPSGF TVDDIIQTGVDNPGHPFIMTVGCVAGDEESYEVFKDLLDPIISDRHSGYKPT DKHRTDLNFENLKGGDDLDPAYVLSSRVRTGRSIKGFTLPPHNSRGERRTIE KLSIEALATLSGEFKGKYYPLNGMTDKEQDQLINDHFLFDKPVSPLLTCAG MARDWPDARGIWHNDAKTFLVWVNEEDHLRVISMQQGGNMKEVFRRFC TGLLKIEETFKKHNHGFMWNEHLGYVLTCPSNLGTGLRGGVHVKLPKLST HAKFEEILTRLRLQKRGTGGVDTASVGGVFDISNADRLGSSEVDQVQMVV DGVKLMVAMEKKLEKSESIDDMIPAQK
218	Cod A0A8C5D617 (Actinin alpha 3b)	FHKLLQTNGKGMIMLAYMTQEEGWDRDLLLDPAWEKQQRKTFTAWCNS HLRKAGTQIENIEEDFRNGLKLMLLLEVISGERLPKPDKGKMRFHKIANVN KALDYICSKGVKLVSIGAEEIVDGNAKMTLGMIWTIILRFAIQDISVEETSAK EGLLLWCQRKTAPYRNVNVQNFHISWKDGLALCALIHRHRPDLIDYSKLR KDDPIGNLNTAFEVAEKFLDIPKMLDAEDIVNTPKPDEKAIMTYVSCFYHAF AGAEQAETAANRICKVLAVNQENEKLMEEYEKLASELLEWIRRTIPWLENR AAEQTMRAMQQKLEDFRDYRRVHKPPRVQEKCQLEINFNTLQTKLRLSNR PAFMPSEGKMVSDIANAWKGLEQVEKGYEEWLLTEIRRLERLDHLAEKFK QKCSLHETWTSGKEELLSQKDFESASLMEIRALMRKHEAFESDLAAHQDR VEQIAAIAQELNELDYHDAVTVNARCQGICDQWDNLGTLTQKRRDSLERV EKLWETIDQLYLEFAKRAAPFNNWMDGAMEDLQDMFIVHSIEEIQSLITAH DQFKATLPEADKERMATMGIQSEILKIAQTYGIKLSGINPYTNLSPQDISNK WDAVKHLVPLRDQMLQEEVARQQANERLRRQFAAQANIIGPWIQTKMEEI GHVSADITGSLEDQMNNLKQYEQNIINYKSNIDKLEGDHQLSQESLIFDNKH TNYTMEHIRVGWEQLLTTIARTINEVENQILTRDAKGISQEQLNEFRASFNH FDRKRNGMMDPDDFRACLISMGYDLGEVEFARIMTLVDANNTGVVTFQAF IDFMTRETAETDTAEQVMASFKILASDKSYITVDELRRELPLEQAEYCISRM TRYVGADGTTGALDYISFSSALYGESDL
219	Cod A0A8C5D074 (Alpha-1,4 glucan phosphorylase)	MAKPLTDQEKRKQISVRGLAGVENVTDIKSNFNRHLHYTLVKDRNVATKR DYYFALANTVRDHLVGRWIRTQQHYYEKDPKRVYYLSLEFYMGRTLQNT MVNLALENACDEATYQLGLDMEELQDIEEDAGLGNGGLGRLAACFLDSM ASLGLAAYGYGIRYEFGIFNQKIVDGWQVEEADDWLRYGNPWEKARPEY MRPVHYYGRVEHTPGGVKWVDTQVVLALPYDTPVPGYRNNVVNTMRLW SAKAPCDFNLKDFNVGGYIQAVLDRNLAENISRVLYPNDNFFEGKELRLKQ EYFVVAATLQDIIRRFKSSKFGSTEFVRMDLSTLPDKVAIQLNDTHPALAIPE LMRILMDLEELPWEKAWDIVVRTCAYTNHTVLPEALERWPIDLFQNLLPRH LDIVYEINRRHLERVAQQFPGDEGRLRRMSLIEEGDHKKINMAHLCIVGSH AVNGVARIHSEIIKDTVFKDFYEMDPEKFQNKTNGITPRRWLVMSNPGLAE VIAERIGDDYIRDLDQLKNLLEFVDDDSLIRDVAKVKQENKMKFSAYLEEH YKVKINPESMFDVQVKRIHEYKRQLLNCLHIITFYNRIKKQPNKAWTPRTIM IGGKAAPGYHTAKMIIKLITAIGDIVNHDAVVGDRLKVIFLENYKVTLAEKII PASDLSEQISTAGTEASGTGNMKFMLNGALTIGTMDGANVEMAEEAGEEN LFIFGMRVDDVKAMDQTGYDAKSYYDRIPELKQAMDQISGGFFSPGQPDLF KDIVNMLMNHDRFKVFADYEEYIKCQDKVSALYKDPKEWTKMVIRNIGSS GKFSSDRTIAQYAREIWGMEPSLEKIPAPDDPRD
220	Cod A0A8C5F6L1 (Fructose-bisphosphate aldolase)	MPHAYPFLSTEQKKELSDIAKKIVAPGKGILAADESTGSVAKRFQGINAENT EENRRLYRQLLFTADDRVAPCIGGVIFFHETLYQKTDDGKLFPQLIRERGQV VGIKVDKGVMPLAGTDGETTTQGLDGLYERCAQYKKDGADFAKWRCVLK ITENTPSELAIMENANVLARYASICQMVRHHGIVPIVEPEILPDGDHDLQRTQ YVTEKVLAAVYKALSDHHVYLEGTLLKPNMVTAGHSCTRKYSAQEVAMA TVTALRRTVPPAVPGVTFLSGGQSEEQASINLNAINQCPLGRPWALTFSYGR ALQASALKAWGGKKENGKACQEEYIKRALNNNLASQGKYVSSGDKGAAG GESLYVANHAY
221	Cod A0A8C4YT70 (Myosin, light chain 1, alkali; skeletal, fast)	MAPKKDVKAPAAAAKKAEPAKKAEPAPEPVAVPAPKTVDLSAVKVEFTPD QMEDYREAFGLFDRVGDNKVCYNQIADIMRALGQNPTNKEVKAILGNPSD EDMNSKRVDFEGFLPMMQTIVNSPNKGTLDDYVEGLRVFDKEGNGTVMG AELRIVLSTLGEKMTEAEIDALMQGQEDENGCINYESFVKHIMSI
222	Cod Q8AWX8 (Glyceraldehyde-3-phosphate dehydrogenase)	MVKVGINGFGRIGRLVTRAGFQSKKVEIVAINDPFIDLEYMVYMFKYDSTH GRFHGEVKIEGDKLVIDGHKISVFHERDPANIKWGDAGAQYVVESTGVFTT IEKASAHLKGGAKRVIISAPSADAPMFVMGVNHEKYENSLQVVSNASCTTN CLAPLAKVIHDNFGIVEGLMSTVHAITATQKTVDGPSGKLWRDGRGASQNI IPASTGAAKAVGKVIPELNGKITGMAFRVPTPNVSVVDLTVRLEKPAKYDDI KKVIKAASEGPMKGILGYTEDQVVSTDFNGDIRSSIFDAGAGIALNDHFVKL VTWYDNEFGYSNRVIDLMAHMSCKE

TABLE 13-continued

Cod Proteins		
SEQ ID NO	Description	Amino Acid Sequence
223	Cod A0A8C5CXA2 (Tropomyosin 1, alpha)	MDAIKKKMQMLKLDKENALDRAEQAEGDKKAAEDRSKQLEDDLVALQK KLKGTEDELDKYSEALKDAQEKLELAEKKAADAESDVASLNRRIQLVEEEL DRAQERLATALTKLEEAEKAADESERGMKVIENRAMKDEEKMELQEIQLK EAKHIAEEADRKYEEVARKLVIIESDLERTEERAELSEGKCSELEEELKTVT NNLKSLEAQAEKYSQKEDKYEEEIKVLTDKLKEAETRAEFAERSVAKLEKT IDDLEDELYAQKLKYKAISEELDHALNDMTSI
224	Cod A0A8C5B6F3 (Myosin binding protein Cb)	MPEAADAGELPTAEEDSGDEEGQGTELTGLFVERPQENVVIIAGADVVITA KVDSTTLTRKPTVKWLKGKWLDLASKAGKHLQFKETYDRETKIYTYEMK VIKAKVGAAGAAGAYRCEVTSKDKFDSASFEITVEAAQQSDDVDILSSFKR ADAGEDDGNLDFSALLKAAKKKKRPEDDKPVIDVWELLKSAHPSEYEKIA FEYGITDLRGMLKRLKKMKAEPKPSEAFLKRMESCYSVEKGKKIVLTCEVF DPNVQVKWLKNGQEIKGSAKYIMETIGNIRKLTIMKTTLADDAAYECVVG EDKCFTEVFVKEPPVTITKLMDDYHVVVGERVEFEVEVSEEGANAMWFFE DIEIFKDTNPAKYRIKKDGRKHMLIILEAGLEDIGMYHCWTNGGHTKGELE VEEKELLVLQDIADLTVKATDQAIFKCEVSDEKVTGKWFKDGVEVLPDER HKMTHIGRFHKLLIDNVKPQDAGDYTFIPNGYALSLSAKLNFLEIKIDYVPR QDPPKIHLDTTGNMVSQNTIIVVAGNKLRLDVEITGEPAPTVVWSKGEEVIT ATEGRVRVDNRKTLSSFVIEGAEREDEANYTICVTNPAGEDKAKLFVKIVD VPDPPQNVKCLSVGEDSSSMVWEPPAFDGGAPLKGYLMERKKKGSSRWT KLNFDVYEATTYEAKRMIEGVYYEMRVFAINSIGMSQPSITSKPFMPLAPTS EPTRLLVHDVTDGTCTLKWMAPERIGAGGLDGYIIEYCKEGETEWVVANK ELCERQGYVVRGLPVGEKINFRVFAKNLAGVSPPATLAQPVTVREIVDLPKI RLPRDLRTKYIRKVGDKINLVIPFQGKPRPVVKWYKDGAPVDEKIVGVRNS TVDSILFIRSAERDHSGKYELVVQIENMEDRATIDLRVVERPGPPTKVKVTE VWGFNAALEWEAPKDDGNCEITGYTIQKADKKTMEWYTIYEHNRRPSCTS SDLIMGNEYMFRVYSENLVGPSEEPGQSKNTAVIGKTGLAYNPPEFKEKDV SCIPTFTQPLLDRSVVAGYSTAMSCSVKGFPKPKIVWMKNKMILGGDPKYI MLNNQGVLTLNIRKPSTFDGGRYSCMAVNDLGQALVECKLEIRIPTDPTAE VK
225	Cod A0A8C5F718 (Myomesin 2b)	MAMRSGNYNHEYYHKQSKYVVKEYSSSESVVDDYQHKAAVRIQGVVRK KMVEKYVREEPMIRGPKFVVRLRSHTVFENSPIKLFCTVEGFPVPIVTWYKD GVILEHGSAGRYIVDDNVGMRSLEIPRCGTDDTAQYSAVASNIHGQASSQA SIIVKRFRQEDESCLYAWLPYTVPMLPSIQYTKINITFTEKFGPVFASEGGSAS LTATMHLEPNLANLQPDSQWYRDDTRLVDSKWVKINTGRGTSTLTIPNLFK DDEGLYTLRLVTKGGTAEHSAFISVSDGPPPVVGAPGAPMDIRIHDANRDY VIVSWKPPNTTTEGPILGYFVDKCEVGTENWAQCNDHPIKICKYPVPGLFEG HSYYFRVRAVNARGISKPSRMSDPIAALDPTEFERLHALKLGGKLDVVTYD DDMEAEGKPPGVSSGIYASETDRTYVVLSWKAPAYSSKAPMWYYIEKCQV GTAAWQRVNTQVPIRSPRYAVFDLAEGKEYQFRVLSANMYGTSEPSEPSKL IQTQALKGVPSAPGQVISTRETDTSLLIQWAPPKEPNNLIGYYIDQCVKGSK DWTSANHKPHSSTKFVVSGLTTGETYVFRVQAINELGLSDESQESSPLSVKA ALTSPTAPYDVALLGCDGQSMVLNWKKPLHTGGTAINAYFVDKRRSGAT MWREVHIPAVTERVYQVEGLTQGAIYQFRVYAANLAGLGKSSKASADFSC EAWTMPEPGPAYDLTASEVRDTSLVLEWKKPVYYGSGPITGYHVEYAKKG SSDWTRANETAVSHRYLKVEGLEEGAKYVFRVRGVNEAGVGMASLPSDS VVAKALAGAQEVSCSVDEKSGDIVFAFESCQMNVGSEFLWKKNYKEITEFS SSIVVHTEGNLCKLMFKNAVKEDFGTYSVSVSHTDGVSSSCSISAEELQKM LELSYSIRHPIIPLKSELAYKILERGRMRFWLQAEEISSKVTYKFFANNKELS GADPNKMAHDVSTGVIEFILDHFTEDSEGTYTCQITDGGGKAQSSLVLIGDA FKAALAEADYQRREYIRVKEGPHFSEFLTVKVGEDCSVTLVCKVANLKKES EFHWYKEDTEIIPDVKPDLGSGVCKLPIHLFSRKTTGVYKATISDDRGKDMS QVDISGKVFDDVINKINQLAGASAAELVINCTATGIQLQCHMKYYTSEMNI HWMHGESKVGHSEKCVIGGTPAMATMEVIEPNDKDKGVYSIVITDLENTN KRTLDLSGDVYDKAYAQFQKLKAEAYAEMNRGKVMGGLPDVVTIMEKK TLSLTCTVCGDPKPLVSWLKNGNDVEPDNQYVVSLDQGKFACLTINGIGM DDSGKYTMIVQNKYGGECVDIVVSVYRHGDTIPVAKPNLTAKTIIPPKMPIE LPQPKTAAPSPAPPASPSPAPPKAVGGRGMKSPTPTRRK
226	Cod A0A8C5C5F8 (Phosphopyruvate hydratase)	MKSNPLTSNTGEHASYLLHVNGVMSITKIHAREILDSRGNPTVEVDLWTAK GLFRAAVPSGASTGIHEALELRDGDKTRYLGKGTQKAVDHVNKDISPKLIE KNFSVVDQEKIDKFMLELDGTENKSKFGANAILGVSLAVCKAGAAEKGVP LYRHIADLAGNKDVILPVPAFNVINGGSHAGNKLAMQEFMILPVGAANFHE AMRIGAEVYHNLKNVIKAKYGKDATNVGDEGGFAPNILENNEALELLKSAI EKAGYPDKIIIGMDVAASEFYRSGKYDLDFKSPDDPARHISGKELGDLYKTF IANYPVQSIEDPFDQDDWANWADFTAAVDIQVVGDDLTVTNPKRIQQAVE KKACNCLLLKVNQIGSVTESIQACKLAQSNGWGVMVSHRSGETEDTFISDL VVGLCTGQIKTGAPCRSERLSKYNQLMRIEEELGDKAKFAGKDYRHPKIN

TABLE 13-continued

Cod Proteins		
SEQ ID NO	Description	Amino Acid Sequence
227	Cod A0A8C4ZUF5 (Myosin light chain)	MAPKKAKRRAGGGDSGSSNVFSMFEQSQIQEYKEAFTIIDQNRDGIISKDDL RDVLASMGQINVKNEELEAMVKEASGPINFTVFLTMFGEKLKGADPEDVIV AAFKILDPEATGSIKKEFLQELLTTQCDRFSPEEIKNMWSAFPPDVAGNVDY KNICYVITHGEEKEE
228	Cod A0A8C5FJN9 (Myosin, heavy chain b)	IMGDGEMQCFGEAAVYLRKPERERLEAQTSPFDAKTAFFVSEPKEMYLKG VLQSRDGGKATVKTLCDKTLTVKEEEIYPMNPPKYDKIEDMAMMTHLNEP TVLYNLKERYAAWMIYTYSGLFCVTVNPYKWLPVYDSSVVNAYRGKKRIE APPHIFSISDNAYQAMLTDRENQSILITGESGAGKTVNTKRVIQYFATIAVSG PKKTEQVAGKMQGSLEDQIIAANPLLEAYGNAKTVRNDNSSRFGKFIRIHF ATSGKLASADIETYLLEKSRVTFQLSAERSYHIFYQLMTGHKPELIDALLITK NPFDYHMISQGEITVKSIDDVEEFIATDDAIDILGFTADEKIGIYKLTGAVLH HGNMKFKQKQREEQAEPDGNEEADKIGYLMGLNSADVLKALCYPRVKVG NEYVTKGQTCPQVHNSVMALCKSVYEKMFLWMVVRINEMLDTKQPRQFF IGVLDIAGFEIFNYNSLEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEW EFIDFGMDLAACIELIEKPMGIFSILEEECMFPKASDTTFKNKLHDQHLGKTK AFEKPKPAKGKAEAHFSLVHYAGTVDYNITGWLDKNKDPLNDSVIQLYQK SSVKLLALLYASHALPDADAASKKGAKKKGGSFQTVSALFRENLGKLMTN LRSTHPHFVRCLIPNESKTPGLMENFLVIHQLRCNGVLEGIRICRKGFPSRVL YGDFKQRYKVLNASVIPEGQFIDNKKASEKLLGSIDVDSTQYKFGHTKVFF KAGLLGALEEMRDEKLAALVTLTQALCRGFNMRKEFVKMMERREAVFTI QYNIRSFMNVKTWPWMKVYFKIKPLLKSAETEKEMAQMKETFGKTKEDL AKALSKKKELEEKMVSLLQEKNDLLLQVQSEAETLSDSEERCEGLIKSKIQL EARLKETSERLEDEEEINSELTAKKRKLEDECSELKKDIDDLELTLAKVEKE KHATENKVKNLTEEMSSQDEIIAKLTKEKKALQEAHQQTLDDLQAEEDKV NTLSKAKCKLEQQVDDLEGSLEQEKKLRMDLERAKRKLEGDLKLAQESIM DLENDKQQSDEKIKKRDFETSQLLNKIEDEQTLSAQFQKKIKELQARIEELE EEIEAERAARAKVEKQRADLSRELEEISERLEEAGGATSLQIEMNKKREAEF QKLRRDLEESTLQHEATSAALRKKQADSVAELGEQIDNLQRVKQKLEKEK SECKMEIDDLSSNMEAVAKSKANLEKLARTLEDQLSELKSKNDEHVRQLN DLSVORARLQTENGEFGRQQEEKEALISQLTRGKQAYTQQIEELKRHIEEEV KAKNALAHSVQSARHDCDLLREQFEEEQEAKAELQRGMSKANSEVAQWR SKYETDAIQRTEELEEAKKKLAQRLQDAEEAVEATNSKCASLEKTKQRLLG EVEDLMIDVERANALAANLDKKQRNFDKVLAEWKQKYEESQAELEGAQK EARSLSTEMFKLKNSYEESLDQLETLKRENKNLQQEISDLSEQMGETGKTIH ELEKGKKTVEVEKSELQTSLEEAEATLEHEESKILRVQLELTQIKGEVDRKM AEKDEEMEQIKRNSLRVIDSMQSNLDSEIRSRNDALRIKKKMEGDLNEMEI QLSHANRQASEAQKQLRNVQGQLKDAQLHLDDALRGQEDMKEQVAMVE RRNNLMLAEIEELRAALEQTERGRKVAEQELVDASERLGLLHSQNTSLINT KKKLETDLIQIQGEVEDTVQEARNAEEKAKKAITDAAMMAEELKKEQDTS SHLERMKKNLELTVKDLQHRLDEAESLAMKGGKKQLQKLEARVRELETEC EAEQRRGSEAIKGVRKYERKVKELSYQCDEDKKNICRLQDLVDKLQLKMK AYKRQAEESEEQANTHLSRYRKVQHELEEAQERADVSESQVNKLRTKSRN GITKIIN
229	Cod A0A8C5F7Y0 (Nebulin)	MATAGMPTRKVRKRVAVDTSKFMTPYLAHSMKMLEQYSHNMYRKAYDL SRGTPPVHIADTPEMVRIRKAQEQLSEVKYRMEGNKARAGSLYDGEAKEIA HVKHMSELLSKVLYRQKWDETKDRYLLPPDAPELVLAVKNAANYSKKLY TEDWDLEKASFLPYSDSPELRRVAKAQEALSDIQYKKGYDQRKLKFTSLAD RPDVELAKKVFTQRSDLKYKEDYNKNVKGQWCETPYFDVATARVAMENL SSRRYTQDWEAAKDQINFMQTETPVYDTNKKAGTSASSVKYRKDYEKTK DKSDYNTLPATENPLLRQLRYAGTILSDKVYKKNYEKSRGFSINYCDTPKF KMDSVLKKFTDTHYQDKYKNEVQGHYIGSYEDIYMLHCQKVEGFKSEQN YKAEYEDIKTRCFFPQTITPEYEVDKKLAQCKDKVYRQHPDKVKFTQVTDS PVQVQAALNAIQLSDINYKAKYEQTKTKNSLPADYPFFIQSRVNAFNLSENC YKYNWEKEKAKRFEMKGDAISILAARAHTNIASDVKYKKEYEKNRGRMV GALSIHDDPKILHSVHVAKIVSNREYTKDYEKVKTKYHTSLDMMSVTLAK KSQAIASMTGYKKISNRYLMPYDAILVDLAKKANSIQSDNEYKSDYNQYTK GSPWITFGSMDVELAKKAGEILSEKKYRKHPDTVPFTAIDDHPVMLQARVN QLQRSDLFYKKGLEEIHQKYSLPIDMPELQLARANSYNISKNYYKAAWEET ISKGYDLKLDAISIKAAKAARHAASDIQYKKAYEKARGHHVGFRNIKDDPL LVHYMQVAKMQSEKNYKKDYHKSKLKYHSPVDMMSLVHAKQASKAQT YAGYRKIQHHYSLMPDAMNLVLARKMNTDASDCEYKSEWNNYVRGTGW VPIGSIGVETAKIGGQIQSENKYRTHPSTFKFSKLMDSMDLELATQAYTSAW EKDKRTIHVMPDSPEIVLAKANAITMSDVREYYYKLHQNGYHLKGDAIPIV AAKLGTKIASNYKYKLDYEKARGHHVGFRNIKDDPLLVHYMQVAKMQSE KNYKKDYEKAKLKYHSPVDMMSLVHAKEASKAQTYAGYRKIPHTYFLLP DNMQLQLCRKMQEIQSDNEYKSEYNNLYKGTGWVPIGSIDVELAKTANSA LREHGYRQHPSSFKFTSKTDSMDMVLALANSKIIDQAAYKASGEKVKHTY HLPADAPEFIQAKFNAQTMSKSHYKQKWVEHIAKGYDMRPDAISILHAKQ GRHIASDIQYKKAYEKARGHHVGFRSIKDDPLLVHYMQVAKMQSEKNYK

TABLE 13-continued

Cod Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		KDYEKAKLKYHSPVDMMSLVHAKEASKAQTYAGYRKIQHHYNLMPDAM NLVLARKMNTDASDCEYKSEWNNYVRGTGWVPIGSIGVETAKIGGEIQSE NKYRTHPSTFKFSKLMDSMDLELATANNKIMNKDKRTIHVMPDSPEIVLAK ANAITMSDKLYKAGVMEMAKKGYHLKGDAIPIVAAKLGTKIASNYKYKL DYEKARGHHVGFRNIKDDPLLVHYMQVAKMQSERNYTKDYQKSKLKYHS PVDMMSLVHAKEASKAQTYAGYRKIEHHYNLMPDAMNLVLARKMNTDA SDCEYKSEWNNYVRGTGWVPIGSIGVETAKIGGEIQSENKYRTHPSTFKFSK LMDSMDLELATANNKIMNKQAYTSAWEKDKRTIHVMPDSPEIVLAKANAI TMSDKLYKAGVMEMAKKGYHLKGDAIPIVAAKLGTKIASNYKYKLDYEK ARGHHVGFRCIQDDPLLVHYMQLAKQQSEKNYKKDYEKAKLKYHSPVDM MSLVHAKDATKAQTYAGYRKIQHHYSLLPDAMNLVLARNMQVNISDNQY KSDYESYCKGIGWIPIGSLDVEKAKIAGKIQSEKLYRQHPSNLKFTKDMFSM DLTLATANNQIMNKVLLALYCVDYTLMNILSSGNNLTNNFVVSFSKLYKL DNELSKKKGHNVRKDAIAILAAKNCTTIASDYKYKAGYRKQVGHHIGARC VEDDPLLVLALNSARIASDALYKKDFNKSKTKFNLPVDMMAFELAKKNQI QVNHANYITRLHNWTCLPDSTDVVNARRAYEIQSDNYYKEDQKLMQGIG WVPIGSLDVEKAKKAGEILSEHLYRKHPSNYKFTSKCDAMPLVLAKANAQI TNKKAYHDAWDKEKSNIHIMPDAMDVVLAKANRANYSIKRYKQANEDAK KKGYDLRNDAIGIIAARAQRNIISDIKYKAGYRKQVGHHIGARCVEDDPLL MLALNSARIASDYLYKKDFNKSKTKFNLPVDMMAFELAKKNQIQVNHAN YITRLHNWTCLPDSTDVVNARRAYEIQSDNYYKEDQKLMQGIGWVPIGSL DVEKAKKAGEILSEHLYRKHPSNYKFTSKCDAMPLVLAKANAQITNKKAY HDAWDKEKSNIHIMPDAMDVVLAKANRANYSIKGYKQANEDAKKKGYD LRNDAIGIIAARAQRNIISDIKYKAGYRKQVGHHIGARCVEDDPLLVLALNS ARIASDYLYKKDFNKSKTKFNLPVDMMAFELAKKNQIQVNHANYITRLHN WTCLPDSTDVVNARRAYEIQSDNYYKEDQKLMQGIGWVPIGSLDVEKAKK AGQILSEHLYRKHPSNYKFTSKCDAMPLVLAKANAQITNKKAYHEAWDKE KSNIHIMPDAMDVVLAKANRANYSIKKYKEANEESKKKGYDLRNDAISIIA ARAQRNIVSDYKYKAGYRKQTGHHIGARCVEDDPLLVLALNSARIASDAL YKKDFNKSKTKFNLPVDMMAFELAKKNQIQVNHANYITRLHNWTCLPDST DVVHARHVYEIQSDAVYKADLKWLQGMGWVPIGSLDVEKTKKAGEILSE RKYRQPASSHKFTSPLDAMEIVLAKNNALTMNQRLYTEAWEKDKTKLHIT SETPEIILSQQNAINMSNKKYKQGYEDTIKKGYYLPGDAISVKSAKHSRNIIS DYKYKAGYRTQVGHHIGARCVEDDPLLMLALNSARIASDYLYKKDFNKSK TKFNLPVDMMAFELAKKNQIQVNHANYITRLHNWTCLPDSTDVVNARKA YDLQSDAVYKADMEFMKGIGWVPIGSLDVVKAKHAGKILSDHLYKQKTD KLKYTSDTKAMPMLLAKANAEVMNKKSYVAAWEKDKIKTHLIPDLPEILL SRSNAYNLSNKMYKEDVVKMFRKGYHLKGDAIAIVAAKKGREIISDIKYK AGYRKQVGHHIGALSVRDDPLIMLALNSARIASDALYKKDFNQSKTKFHLP VDMLTIELAKKCQVQVNDFNYRTRLHNWTCLPDSTDVVQARKAYDLQSD AVYKADMEWIKGQGWVPIGSLDVEKAKKAAEILSERKYREHPSKFKFTAK TDDMPYALAKANAEIMNKQSYVAAWESDKTKLHIMPDTPEIILGQQNKLN NSRKYYREAYEETLKKGYYLPKDAISVLAARASRNIISDYKYKAGFRKQCG HHIGALSVQDDPLLMMAANAAKIFSANLYKQDFHKTKTKYKLPVDMVAF DLAKTAQSHMKEVTYRTRLHNWTCMPDNTDVVHARKVYDLHSDAVYRA DMEWIRGCGWSPQDSVDVVKARNAQKILNENLYRQAPSTVKFTSVVDLPA IVLSKQNAENISEMKYTAAWNKDKLTFHLKPDIPSVQLAKANAINISNKLYT KAWDNQKAKAYHIKEDTIAVLKARASRDIISDYKYKTGYRKQVGHHIGAL SVRDDPLIMLALNSARIASDALYKKDFNQSKTKFHLPVDMLTIELAKKCQIQ VNDFNYRTRLHNWTCLPDSTDVVQARKAYDLQSDAVYKADMEWIRGCG WVPHECVDVVKVKHAQKILADRGYRVKLDTQKFTAPVDRMDFVCAKNA AEVLNDAKYREAWHTDKSKFSLVDTPGLATSREVAKIMHPKLYTKDWDK VKATGYFMPGDAVPIKQCTQTTKVQSSHKYLEGYRKQVGHHIGALSVFDD PLIMLAMNSAKIASEALYKKDFNKSKTKFHLPVDMLTIELAKKCQIQVNDF NYRTRLHNWTCLPDSTDVVNARKAYDLQSDFVYKADLEWIRGCGWVPAE SVHHVKVRKAQEILNERLYKKGAIENFSKYTHVVDRPELILARVAAANLSD LKYKETFNNEKGHYIGSEDTPQLAHSRECAKNISEAKYHESHEKAKGHYM ANTLVDFPEVVRCGEQEKAKALRVYQKEYHESKTNNHIPSDIIANQVAKRC QDMLSDVIYRTHLHQWTCPPDQLDTVRARKGNEILSDVFYKEDLTWMKG MGCYVWDTPEIVRAKKSYELQSELKYKEAGKKEFNNYSIVTDTPVYVTAIL GSTWASELNYREAYHKEKHLYTTILDTHDYARCHDFKHFFSNKIYTAAWD KIKAKGYKIPHDSHAMMHAKQQKVILSGVKYKEDYEKFKSLYSLPKCLQD DPATARCIKAGKLNLDRLYKEDYEKTKAKNHIPADMLDIVSGREVQRNVS EINYRKRLHQWVCLPDMQVFVQARKVNDQLSDIFYKDDLNYLRGIGCYA WDTPEILRVKTAGDQQSENKYKAKAVEAYKDYTVVLDTPTYETAKQNAV NLSDLHYRLDYNANVKGTNTAPAVTMETERARLANYNQSDNLYKEANKN SMATGYLMTNDTPVVKQVKNNSLVASSVKYKAAYEMMKATGYTLHPEG VTFVNTRKVNQVINPRLYREQYEKDKDKIHSTLDTPDIRQVKMNQEHLSEL CYREKYNNTKGQLINMPMTPQLMHCQHVNQITSDLRYKEDLLWLRGLGCF MYDTPEMVNVRNITKFKNNYQVEAKKNHSKFSVVLDTPEYQRVTELKSH MSMVSPTSNTSLNTLDSVDIQRSKWAQQLSNKFLYTDLASKERAHFSQEVE TPAMGLARTMKVTNSYNKYKEQYEKMKHRYTSIADTPILIRAKKAYLQSS

TABLE 13-continued

Cod Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		DLRYKETFERSKGHYHTVKDALDILCHRRVTDDISEVKYREKYINSLGTWK SIPDRPEFFFNRIVGENVSSVKYKEDLEWMKGIGCFVWDTPEMVQADINKT LYSGRLYKASFEKNRYNFKYTCDTPFFHAAKNASSLINERSYRSSYEKNKD KFTITTDDPRYQLARENNKMSQAKYKSRAKELLKSGCNELHRPDILTALYQ SVLGSKWKYREQYERVKSKFTSILETPEYESHKRSKKIGDIVYRMEYNKNK AKGYSLPYDTPHQNHMKKVQALTSNLKYREVYEKAKSQINMDPDAFEIRA AKEAYKNTTNLDYKKKYEATKNKWIWTVDRPDFISHSRNTLQQSDVEYKY DKEMMKGCVIPIVDDKITLLAKKGCEMASDIKYKEKYEKAKGHYNPVADS PQILHAKAVRTLASESKYKEASKKEMQSGSFTTLPETRDTVHSKEVNKLISG KLYKSKFEKEKGQSKYNRMIVPPDVQHAMDVAKSQSSYAYKRDAKDNLH YTTTADRPDIRKATQAAKLISDVGYREKARAEASRGGSLTYRPDIALATEVS KLTSLVDAKSALHGAASYDTLQMRHSKKMSALTSNVKYHEKFDKEMKGK KPKYDLKDSKIYQTLKDATNLASEVKYKGDLKKLHKPVTDMAESLSMQH NLSTSKLSSDYRYKKKFKDSKGHYLVIADTPEQLHHKEATELQSSVSGPSLC GVHVCKDYKKDYEEGMRGKNLCGLEVTPAMIHVRHATKIASEREYRKDL EEGVKGKGLSAPEETPEHLRAKNATHILSEREYKKSLELEIKGKGMLALAT DTPDFMRARNATDILSQTKYKHTAELDRASFTSVIDTPDIIHAQQMRNIISQN KYKAEAGKTMSHYVPVFDTPEMQRVRENQKNFSTLQYQADHKIIKGKVLT VKDTPEMLRVKENTKNFSSVLYKDSKAKGTPVVFTPEMERVKRNQEHISSV LYSDSFRKQVQGKAAFVLDTPEMRRVKESQLNISGVRYHQDHEKSKGSFT AMASYPEMERVKKNTQEFSDINYRGIQRRVVEMERRRVVEHEQETTTDLR VWRTNPGSVFDYDPAEDNIQSRSLHMMSVHAQRRSKEHSRSTSAMSGPAD EKSEMSHEGDHHMSTYSNGFGASSVGYQTTKTVELQQRSSSVATQQTTVS SIPSHPSTTGKTVRAMYDYGAADNDEVTFKDGDVIVNVQSIDEGWMYGTV QRTGKTGMLPANYVEAV
230	Cod A0A8C5ARQ5 (Myosin heavy chain, fast skeletal muscle-like)	MSTDAEMAVYGQAAIYLRKPEKERLEAQSKPFDAKSAAYVTDVKELYLKC TILKKEGGKVTVKVHDTEEEKTVKDDDVTPMNPPKYDKIEDMAMMTHLN EASVLYNLKERYAAWMIYTYSGLFCATVNPYKWLPVYDQHVVCAYRGKK RMEAPPHIFSVSDNAYQFMLTDRENQSVLITGESGAGKTVNTKRVIQYFATI SVGGDKKKDIGAGKIQGSLEDQIIAANPLLEAYGNAKTVRNDNSSRFGKFIR IHFGTSGKLSSADIETYLLEKSRVTFQLPAERGYHIFYQMMTNHKPEIIEMTL ITTNPYDFPTCSMGQIAVASIDDKEELVATDTAIDILGFNNDEKMGIYKLTG AVLHHGNMKFKQKQREEQAEPEGTEDADKVAYLLGLNSADMLKGLCYPR VKVGNEYVTKGQTVPQVLNSVSALAKSIYERMFLWMVIRINQMLDTKQPR SSFIGVLDIAGFEIFDYNSMEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGI IWEFIDFGMDLAACIELIEKPMGIFSILEEECMFPKASDVTFKAKLYDQHLGK TKAFEKPKPAKGKAEAHFSLVHYAGTVDYNITGWLDKNKDPLNDSVVQL YQKSGVKLLPVLYPPVASEDSTKKGGKKKGGSMQTVSSQFRENLGKLMTN LRSTHPHFVRCLIPNESKTPGLMENFLVIHQLRCNGVLEGIRICRKGFPSRILY ADFKQRYKVLNASVIPEGQFIDNKKASEKLLGSIDVNHEEYKFGHTKVFFK AGLLGTLEEMRDEKLATLVGMIQALCRGFQMRKEFKKMMERRHAVYTIQ YNIRSFMNVKTWPWMKVYYKIKPLLKSAETEKELSAMKENYEKMKKDLA TALAKKKELEEKMVSLLQEKNDLSLQVASEGENLNDAEERCEGLIKSKIQL EARLKETTERLEDEEEMNADLTAKKRKLEDECSELKKDIDDLELTLAKVEK EKHATENKVKNLTEEMASQDESVAKLSKEKKALQEAHQQTLDDLQAEED KVNTLSKAKTKLEQQVDDLEGSLEQEKKLRMDLERSKRKLEGDLKLAQES IMDLENDKQQSDEKIKKKDFETSQLLSKIEDEQSLAIQLQKKIKELQARIEEL EEEIEAERAARAKVEKQRADLSRELEEISERLEEAGGATSAQIEMNKKREAE FQKLRRDLEESTLQHEATAAALRKKQADSVAELGEQIDNLQRVKQKLEKE KSEYKMEIDDLSSNMETVAKAKGNLEKICRTLEDQLSEIKAKSDENSRQIND IGAQRARLLTENGEFGRQLEEKEALVSQLTRGKQAYTQQIEELKRQIEEEVK AKNALAHGVQSARHDCDLLREQFEEEQEAKAELQRGMSKANGEVAQWRS KYETDAIQRTEELEESKKKLAQRLQEAEEQIEAVNSKCASLEKTKQRLQGE VEDLMVDVERANGLAANLDKKQRNFDKVLAEWKQKYEEGQAELEGTQK EARSLSTELFKMKNSYEEALDQLETLKRENKNLQQEISDLSEQIGETGKSIH ELEKSKKQVETEKTEIQTALEEAEGTLEHEESKILRVQLELNQIKGEVDRKL AEKDEEMEQIKRNSLRITDSMQSTLDSEVRSRNDALRIKKKMEGDLNEMEI QLSHANRQAAEAQKQLRNVQGQLKDAQLHLDDAVRGGEDLKEQAAMVD RRNGLMVAEIEELRAALEQTERGRKVAEQELIDASERVGLLHSQNTSLLNT KKKLETDLVQVQGEVDDTVQEARNAEEKAKKAITDAAMMAEELKKEQDT SSHLERMKKNLEVTVKDLQHRLDEAENLAMKGGKKQLQKLESRVRELETE VEGEQRRGVDAVKGVRKYERRVKELTYQTEEDKKNGARLQDLVDKLQLK VKAYKRQAEEAEEQANSYLSKCRKVQHELEEAEERADIAETQVNKLRSKS RDGGKGKDPAE
231	Cod Q78AY8 (Actin alpha 1)	MCDDEETTALVCDNGSGLVKAGFAGDDAPRAVFPSIVGRPRHQGVMVGM GQKDSYVGDEAQSKRGILTLKYPIEHGIITNWDDMEKIWHHTFYNELRVAP EEHPTLLTEAPLNPKANREKMTQIMFETFNVPAMYVAIQAVLSLYASGRTT GIVLDSGDGVTHNVPVYEGYALPHAIMRLDLAGRDLTDYLMKILTERGYSF VTTAEREIVRDIKEKLCYVALDFENEMATAASSSSLEKSYELPDGQVITIGN ERFRCPETLFQPSFIGMESAGIHETAYNSIMKCDIDIRKDLYANNVLSGGTT

TABLE 13-continued

Cod Proteins		
SEQ ID NO	Description	Amino Acid Sequence
		MYPGIADRMQKEITALAPSTMKIKIIAPPERKYSVWIGGSILASLSTFQQMWI SKQEYDEAGPSIVHRKCF

[0061] An animal protein can comprise a tuna protein (Table 14). A tuna protein can comprise a tuna muscle protein. A tuna protein can comprise a tuna blood protein. A tuna protein can comprise a tuna albumin protein. A tuna protein can comprise a sequence that has at least or about 70%, at least or about 75%, at least or about 80%, at least or about 90%, at least or about 91%, at least or about 92%, at least or about 93%, at least or about 94%, at least or about 95%, at least or about 96%, at least or about 97% at least or about 98% at least or about 99% or about 100% sequence identity to SEQ ID NO: 232.

[0062] A food product may comprise a fragment or a truncation of a tuna protein. The food product may include a fragment or truncation of any one of SEQ ID Nos: 232. A sequence of the fragment or truncation may be identical or substantially identical to a portion of any one of SEQ ID Nos: 232. A sequence of the fragment or truncation may be at least about 70%, 80%, 90%, 95%, 98%, or more identical to a portion of a sequence of any one of SEQ ID Nos: 232. A length of the fragment or truncation that is at least about 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of a length of any one of SEQ ID Nos: 232. A length of the fragment or truncation may be less than or equal to about 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, or less than a length of any one of SEQ ID Nos: 232.

TABLE 14

Tuna Proteins		
SEQ ID NO	Description	Amino Acid Sequence
232	Tuna XP_042264497.1 (actin alpha 1, skeletal muscle b)	MCDDDETTALVCDNGSGLVKAGFAGDDA PRAVFPSIVGRPRHQGVMVGMGQKDSYV GDEAQSKRGILTLKYPIEHGIITNWDDM EKIWHHTFYNELRVAPEEHPTLLTEAPL NPKANREKMTQIMFETFNVPAMYVAIQA VLSLYASGRTTGIVLDAGDGVTHNVPVY EGYALPHAIMRLDLAGRDLTDYLMKILT ERGYSFVTTAEREIVRDIKEKLCYVALD FENEMATAASSSSLEKSYELPDGQVITI GNERFRCPETLFQPSFIGMESAGIHETA YNSIMKCDIDIRKDLYANNVLSGGTTMY PGIADRMQKEITALAPSTMKIKIIAPPE RKYSVWIGGSILASLSTFQQMWISKQEY DEAGPSIVHRKCF

[0063] An additional gene that is not a gene of interest can be a gene that modulates protein expression. A gene that modulates protein expression can enhance protein expression, decrease protein expression, stop protein expression, or maintain protein expression.

[0064] A gene that modulates protein expression can increase protein expression by at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 100%, at least about 150%, at least about 200%, at least about 300%, at least about 400%, at least about 500%, at least about 600%, at least about 700%, at least about 800%, at least about 900% at least about 1,000%, at least about 5,000%, at least about 10,000% or more. A gene that modulates protein expression can increase protein expression by at most about 10,000%, at most about 5,000%, at most about 1,000%, at most about 900%, at most about 800%, at most about 700%, at most about 600%, at most about 500%, at most about 400%, at most about 300%, at most about 200%, at most about 150%, at most about 100%, at most about 90%, at most about 80%, at most about 70%, at most about 60%, at most about 50%, at most about 40%, at most about 30%, at most about 20%, at most about 10%, at most about 5%, or less.

[0065] A gene that modulates protein expression can increase protein expression by at least or up to about 1-fold, at least or up to about 2-fold, at least or up to about 3-fold, at least or up to about 4-fold, at least or up to about 5-fold, at least or up to about 6-fold, at least or up to about 7-fold, at least or up to about 8-fold, at least or up to about 9-fold, at least or up to about 10-fold, at least or up to about 20-fold, at least or up to about 30-fold, at least or up to about 40-fold, at least or up to about 50-fold, at least or up to about 60-fold, at least or up to about 70-fold, at least or up to about 80-fold, at least or up to about 90-fold, at least or up to about 100-fold, at least or up to about 500-fold, at least or up to about 1,000-fold, at least or up to about 5,000-fold, or at least or up to about 10,000-fold. A gene that modulates protein expression can increase protein expression by at most or less than about 10,000-fold, at most or less than about 5,000-fold, at most or less than about 1,000-fold, at most or less than about 500-fold, at most or less than about 100-fold, at most or less than about 90-fold, at most or less than about 80-fold, at most or less than about 70-fold, at most or less than about 60-fold, at most or less than about 50-fold, at most or less than about 40-fold, at most or less than about 30-fold, at most or less than about 20-fold, at most or less than about 10-fold, at most or less than about 9-fold, at most or less than about 8-fold, at most or less than about 7-fold, at most or less than about 6-fold, at most or less than about 5-fold, at most or less than about 4-fold, at most or less than about 3-fold, at most or less than about 2-fold, or at most or less than about 1-fold.

[0066] A gene that modulates protein expression can decrease protein expression by at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 100%, at least about 150%, at least about 200%, at least about 300%, at least about 400%, at least about 500%, at least about 600%, at least about 700%, at least about 800%, at least about 900% at least about 1,000%, at least about 5,000%, at least about 10,000% or more. A gene that

modulates protein expression can decrease protein expression by at most about 10,000%, at most about 5,000%, at most about 1,000%, at most about 900%, at most about 800%, at most about 700%, at most about 600%, at most about 500%, at most about 400%, at most about 300%, at most about 200%, at most about 150%, at most about 100%, at most about 90%, at most about 80%, at most about 70%, at most about 60%, at most about 50%, at most about 40%, at most about 30%, at most about 20%, at most about 10%, at most about 5%, or less.

[0067] A gene that modulates protein expression can decrease protein expression by at least or up to about 1-fold, at least or up to about 2-fold, at least or up to about 3-fold, at least or up to about 4-fold, at least or up to about 5-fold, at least or up to about 6-fold, at least or up to about 7-fold, at least or up to about 8-fold, at least or up to about 9-fold, at least or up to about 10-fold, at least or up to about 20-fold, at least or up to about 30-fold, at least or up to about 40-fold, at least or up to about 50-fold, at least or up to about 60-fold, at least or up to about 70-fold, at least or up to about 80-fold, at least or up to about 90-fold, at least or up to about 100-fold, at least or up to about 500-fold, at least or up to about 1,000-fold, at least or up to about 5,000-fold, or at least or up to about 10,000-fold. A gene that modulates protein expression can decrease protein expression by at most or less than about 10,000-fold, at most or less than about 5,000-fold, at most or less than about 1,000-fold, at most or less than about 500-fold, at most or less than about 100-fold, at most or less than about 90-fold, at most or less than about 80-fold, at most or less than about 70-fold, at most or less than about 60-fold, at most or less than about 50-fold, at most or less than about 40-fold, at most or less than about 30-fold, at most or less than about 20-fold, at most or less than about 10-fold, at most or less than about 9-fold, at most or less than about 8-fold, at most or less than about 7-fold, at most or less than about 6-fold, at most or less than about 5-fold, at most or less than about 4-fold, at most or less than about 3-fold, at most or less than about 2-fold, or at most or less than about 1-fold.

[0068] An additional gene of interest can be a gene that modulates protein secretion. A gene that modulates protein secretion can enhance protein secretion, decrease protein secretion, stop protein secretion, or maintain protein secretion.

[0069] A gene that modulates protein secretion can increase protein secretion by at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 100%, at least about 150%, at least about 200%, at least about 300%, at least about 400%, at least about 500%, at least about 600%, at least about 700%, at least about 800%, at least about 900% at least about 1,000%, at least about 5,000%, at least about 10,000% or more. A gene that modulates protein secretion can increase protein secretion by at most about 10,000%, at most about 5,000%, at most about 1,000%, at most about 900%, at most about 800%, at most about 700%, at most about 600%, at most about 500%, at most about 400%, at most about 300%, at most about 200%, at most about 150%, at most about 100%, at most about 90%, at most about 80%, at most about 70%, at most about 60%, at most about 50%, at most about 40%, at most about 30%, at most about 20%, at most about 10%, at most about 5%, or less.

[0070] A gene that modulates protein secretion can increase protein secretion by at least or up to about 1-fold, at least or up to about 2-fold, at least or up to about 3-fold, at least or up to about 4-fold, at least or up to about 5-fold, at least or up to about 6-fold, at least or up to about 7-fold, at least or up to about 8-fold, at least or up to about 9-fold, at least or up to about 10-fold, at least or up to about 20-fold, at least or up to about 30-fold, at least or up to about 40-fold, at least or up to about 50-fold, at least or up to about 60-fold, at least or up to about 70-fold, at least or up to about 80-fold, at least or up to about 90-fold, at least or up to about 100-fold, at least or up to about 500-fold, at least or up to about 1,000-fold, at least or up to about 5,000-fold, or at least or up to about 10,000-fold. A gene that modulates protein secretion can increase protein secretion by at most or less than about 10,000-fold, at most or less than about 5,000-fold, at most or less than about 1,000-fold, at most or less than about 500-fold, at most or less than about 100-fold, at most or less than about 90-fold, at most or less than about 80-fold, at most or less than about 70-fold, at most or less than about 60-fold, at most or less than about 50-fold, at most or less than about 40-fold, at most or less than about 30-fold, at most or less than about 20-fold, at most or less than about 10-fold, at most or less than about 9-fold, at most or less than about 8-fold, at most or less than about 7-fold, at most or less than about 6-fold, at most or less than about 5-fold, at most or less than about 4-fold, at most or less than about 3-fold, at most or less than about 2-fold, or at most or less than about 1-fold.

[0071] A gene that modulates protein secretion can decrease protein secretion by at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 100%, at least about 150%, at least about 200%, at least about 300%, at least about 400%, at least about 500%, at least about 600%, at least about 700%, at least about 800%, at least about 900% at least about 1,000%, at least about 5,000%, at least about 10,000% or more. A gene that modulates protein secretion can decrease protein secretion by at most about 10,000%, at most about 5,000%, at most about 1,000%, at most about 900%, at most about 800%, at most about 700%, at most about 600%, at most about 500%, at most about 400%, at most about 300%, at most about 200%, at most about 150%, at most about 100%, at most about 90%, at most about 80%, at most about 70%, at most about 60%, at most about 50%, at most about 40%, at most about 30%, at most about 20%, at most about 10%, at most about 5%, or less.

[0072] A gene that modulates protein secretion can decrease protein secretion by at least or up to about 1-fold, at least or up to about 2-fold, at least or up to about 3-fold, at least or up to about 4-fold, at least or up to about 5-fold, at least or up to about 6-fold, at least or up to about 7-fold, at least or up to about 8-fold, at least or up to about 9-fold, at least or up to about 10-fold, at least or up to about 20-fold, at least or up to about 30-fold, at least or up to about 40-fold, at least or up to about 50-fold, at least or up to about 60-fold, at least or up to about 70-fold, at least or up to about 80-fold, at least or up to about 90-fold, at least or up to about 100-fold, at least or up to about 500-fold, at least or up to about 1,000-fold, at least or up to about 5,000-fold, or at least or up to about 10,000-fold. A gene that modulates protein secretion can decrease protein secretion by at most

or less than about 10,000-fold, at most or less than about 5,000-fold, at most or less than about 1,000-fold, at most or less than about 500-fold, at most or less than about 100-fold, at most or less than about 90-fold, at most or less than about 80-fold, at most or less than about 70-fold, at most or less than about 60-fold, at most or less than about 50-fold, at most or less than about 40-fold, at most or less than about 30-fold, at most or less than about 20-fold, at most or less than about 10-fold, at most or less than about 9-fold, at most or less than about 8-fold, at most or less than about 7-fold, at most or less than about 6-fold, at most or less than about 5-fold, at most or less than about 4-fold, at most or less than about 3-fold, at most or less than about 2-fold, or at most or less than about 1-fold.

[0073] An expression cassette can comprise at least one regulatory sequence. A regulatory sequence can comprise a promoter, a repressor, an enhancer, a terminator, a 5' UTR, a 3' UTR or a combination thereof. A regulatory sequence can be part of a coding sequence. Alternatively, a regulatory sequence can be a non-coding sequence. An expression cassette can comprise at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10 or more regulatory sequences.

[0074] A promoter can be an endogenous *P. pastoris* promoter or variant thereof. Alternatively, a promoter can be a yeast promoter from a different species (e.g., *S. cerevisiae*) or variant thereof. Alternatively, a promoter can be a synthetic or novel promoter. A promoter can be an AOX1 promoter, an ADH3 promoter, a TEF1 promoter, a PGK1 promoter, a TPI1 promoter, a PHO89 promoter, a THI11 promoter, an AOD promoter, or a GAP promoter.

[0075] An enhancer can be an endogenous *P. pastoris* enhancer or variant thereof. Alternatively, an enhancer can be a yeast enhancer from a different species (e.g., *S. cerevisiae*) or variant thereof. Alternatively, an enhancer can be a synthetic or novel enhancer. An enhancer can be an AOX1 enhancer or a GAP enhancer.

[0076] A repressor can be an endogenous *P. pastoris* repressor or variant thereof. Alternatively, a repressor can be a yeast repressor from a different species (e.g., *S. cerevisiae*) or variant thereof. Alternatively, a repressor can be a synthetic or novel repressor. A repressor can be an AOX1 repressor or a GAP repressor.

[0077] A terminator can be an endogenous *P. pastoris* terminator or variant thereof. Alternatively, a terminator can be a yeast terminator from a different species (e.g., *S. cerevisiae*) or variant thereof. Alternatively, a terminator can be a synthetic or novel terminator. A terminator can be an AOX1 terminator, a RPL15A terminator, a RPL41B terminator, a DAS1 terminator, a VMA2 terminator, a SRP54 terminator, a FUR4 terminator, a PLB3 terminator, a GIC1 terminator, a FBP1 terminator, a CYC1 terminator, a IDP1 terminator, a PRC1 terminator, a RPL3 terminator, a PGK1 terminator, a YPT31 terminator, an ASP3-1 terminator, a TPO4 terminator, a RPS28A terminator, a VMA16 terminator, a LSC2 terminator, a DIT1 terminator, a MIG2 terminator, an ADH2 terminator, an ADH3 terminator, an ICY2 terminator, a PHO89 terminator, a THI11 terminator, an AOD terminator, or a GAP terminator.

[0078] A UTR can be a 5' UTR or a 3' UTR. A UTR can be an endogenous *P. pastoris* UTR or variant thereof. Alternatively, a UTR can be a yeast UTR from a different species (e.g., *S. cerevisiae*) or variant thereof. Alternatively, a UTR can be a synthetic or novel UTR. A UTR can be an AOX1 UTR, a RPL15A UTR, a RPL41B UTR, a DAS1 UTR, a VMA2 UTR, a SRP54 UTR, a FUR4 UTR, a PLB3 UTR, a GIC1 UTR, a FBP1 UTR, a CYC1 UTR, an IDP1 UTR, a PRC1 UTR, a RPL3 UTR, a PGK1 UTR, a YPT31 UTR, an ASP3-1 UTR, a TPO4 UTR, a RPS28A UTR, a VMA16 UTR, a LSC2 UTR, a DIT1 UTR, a MIG2 UTR, an ADH2 UTR, an ADH3 UTR, a TEF1 UTR, a PGK1 UTR, a TPI1 UTR, a PHO89 UTR, a THI11 UTR, an AOD UTR, or a GAP UTR.

[0079] An expression cassette can comprise a signal sequence. A signal sequence (e.g., a secretion sequence or a nuclear localization sequence) can aid with and/or induce protein secretion. An expression cassette can comprise at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10 or more signal sequences.

[0080] A signal sequence can be an endogenous *P. pastoris* signal sequence or variant thereof. Alternatively, a signal sequence can be a yeast signal sequence from a different species (e.g., *S. cerevisiae*) or variant thereof. Alternatively, a signal sequence can be a synthetic or novel promoter. A signal sequence can be an amylase signal sequence, an inulinase signal sequence, an invertase signal sequence, a lysozyme signal sequence, an albumin signal sequence, a glucoamylase signal sequence, or a killer protein signal sequence, a *S. cerevisiae* Ost1 signal sequence, a *S. cerevisiae* alpha mating factor, a partial *S. cerevisiae* alpha mating factor, or a *Pichia* Exg1p signal sequence.

[0081] A regulatory sequence can modulate protein yield. A regulatory sequence can enhance protein yield, decrease protein yield, or maintain protein yield.

[0082] A regulatory sequence that modulates protein yield can increase protein yield by at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 100%, at least about 150%, at least about 200%, at least about 300%, at least about 400%, at least about 500%, at least about 600%, at least about 700%, at least about 800%, at least about 900% at least about 1,000%, at least about 5,000%, at least about 10,000% or more. A regulatory sequence that modulates protein yield can increase protein yield by at most about 10,000%, at most about 5,000%, at most about 1,000%, at most about 900%, at most about 800%, at most about 700%, at most about 600%, at most about 500%, at most about 400%, at most about 300%, at most about 200%, at most about 150%, at most about 100%, at most about 90%, at most about 80%, at most about 70%, at most about 60%, at most about 50%, at most about 40%, at most about 30%, at most about 20%, at most about 10%, at most about 5%, or less.

[0083] A regulatory sequence that modulates protein yield can increase protein yield by at least or up to about 1-fold, at least or up to about 2-fold, at least or up to about 3-fold, at least or up to about 4-fold, at least or up to about 5-fold, at least or up to about 6-fold, at least or up to about 7-fold, at least or up to about 8-fold, at least or up to about 9-fold, at least or up to about 10-fold, at least or up to about 20-fold, at least or up to about 30-fold, at least or up to about 40-fold, at least or up to about 50-fold, at least or up to about 60-fold, at least or up to about 70-fold, at least or up to about 80-fold, at least or up to about 90-fold, at least or up to about 100-fold, at least or up to about 500-fold, at least or up to about 1,000-fold, at least or up to about 5,000-fold, or at

least or up to about 10,000-fold. A regulatory sequence that modulates protein yield can increase protein yield by at most or less than about 10,000-fold, at most or less than about 5,000-fold, at most or less than about 1,000-fold, at most or less than about 500-fold, at most or less than about 100-fold, at most or less than about 90-fold, at most or less than about 80-fold, at most or less than about 70-fold, at most or less than about 60-fold, at most or less than about 50-fold, at most or less than about 40-fold, at most or less than about 30-fold, at most or less than about 20-fold, at most or less than about 10-fold, at most or less than about 9-fold, at most or less than about 8-fold, at most or less than about 7-fold, at most or less than about 6-fold, at most or less than about 5-fold, at most or less than about 4-fold, at most or less than about 3-fold, at most or less than about 2-fold, or at most or less than about 1-fold.

[0084] A regulatory sequence that modulates protein yield can decrease protein yield by at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 100%, at least about 150%, at least about 200%, at least about 300%, at least about 400%, at least about 500%, at least about 600%, at least about 700%, at least about 800%, at least about 900% at least about 1,000%, at least about 5,000%, at least about 10,000% or more. A regulatory sequence that modulates protein yield can decrease protein yield by at most about 10,000%, at most about 5,000%, at most about 1,000%, at most about 900%, at most about 800%, at most about 700%, at most about 600%, at most about 500%, at most about 400%, at most about 300%, at most about 200%, at most about 150%, at most about 100%, at most about 90%, at most about 80%, at most about 70%, at most about 60%, at most about 50%, at most about 40%, at most about 30%, at most about 20%, at most about 10%, at most about 5%, or less.

[0085] A regulatory sequence that modulates protein yield can decrease protein yield by at least or up to about 1-fold, at least or up to about 2-fold, at least or up to about 3-fold, at least or up to about 4-fold, at least or up to about 5-fold, at least or up to about 6-fold, at least or up to about 7-fold, at least or up to about 8-fold, at least or up to about 9-fold, at least or up to about 10-fold, at least or up to about 20-fold, at least or up to about 30-fold, at least or up to about 40-fold, at least or up to about 50-fold, at least or up to about 60-fold, at least or up to about 70-fold, at least or up to about 80-fold, at least or up to about 90-fold, at least or up to about 100-fold, at least or up to about 500-fold, at least or up to about 1,000-fold, at least or up to about 5,000-fold, or at least or up to about 10,000-fold. A regulatory sequence that modulates protein yield can decrease protein yield by at most or less than about 10,000-fold, at most or less than about 5,000-fold, at most or less than about 1,000-fold, at most or less than about 500-fold, at most or less than about 100-fold, at most or less than about 90-fold, at most or less than about 80-fold, at most or less than about 70-fold, at most or less than about 60-fold, at most or less than about 50-fold, at most or less than about 40-fold, at most or less than about 30-fold, at most or less than about 20-fold, at most or less than about 10-fold, at most or less than about 9-fold, at most or less than about 8-fold, at most or less than about 7-fold, at most or less than about 6-fold, at most or less than about 5-fold, at most or less than about 4-fold, at most or less than about 3-fold, at most or less than about 2-fold, or at most or less than about 1-fold.

[0086] A regulatory sequence can modulate protein secretion. A regulatory sequence that modulates protein secretion can enhance protein secretion, decrease protein secretion, stop protein secretion, or maintain protein secretion.

[0087] A regulatory sequence that modulates protein secretion can increase protein secretion by at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 100%, at least about 150%, at least about 200%, at least about 300%, at least about 400%, at least about 500%, at least about 600%, at least about 700%, at least about 800%, at least about 900% at least about 1,000%, at least about 5,000%, at least about 10,000% or more. A regulatory sequence that modulates protein secretion can increase protein secretion by at most about 10,000%, at most about 5,000%, at most about 1,000%, at most about 900%, at most about 800%, at most about 700%, at most about 600%, at most about 500%, at most about 400%, at most about 300%, at most about 200%, at most about 150%, at most about 100%, at most about 90%, at most about 80%, at most about 70%, at most about 60%, at most about 50%, at most about 40%, at most about 30%, at most about 20%, at most about 10%, at most about 5%, or less.

[0088] A regulatory sequence that modulates protein secretion can increase protein secretion by at least or up to about 1-fold, at least or up to about 2-fold, at least or up to about 3-fold, at least or up to about 4-fold, at least or up to about 5-fold, at least or up to about 6-fold, at least or up to about 7-fold, at least or up to about 8-fold, at least or up to about 9-fold, at least or up to about 10-fold, at least or up to about 20-fold, at least or up to about 30-fold, at least or up to about 40-fold, at least or up to about 50-fold, at least or up to about 60-fold, at least or up to about 70-fold, at least or up to about 80-fold, at least or up to about 90-fold, at least or up to about 100-fold, at least or up to about 500-fold, at least or up to about 1,000-fold, at least or up to about 5,000-fold, or at least or up to about 10,000-fold. A regulatory sequence that modulates protein secretion can increase protein secretion by at most or less than about 10,000-fold, at most or less than about 5,000-fold, at most or less than about 1,000-fold, at most or less than about 500-fold, at most or less than about 100-fold, at most or less than about 90-fold, at most or less than about 80-fold, at most or less than about 70-fold, at most or less than about 60-fold, at most or less than about 50-fold, at most or less than about 40-fold, at most or less than about 30-fold, at most or less than about 20-fold, at most or less than about 10-fold, at most or less than about 9-fold, at most or less than about 8-fold, at most or less than about 7-fold, at most or less than about 6-fold, at most or less than about 5-fold, at most or less than about 4-fold, at most or less than about 3-fold, at most or less than about 2-fold, or at most or less than about 1-fold.

[0089] A regulatory sequence that modulates protein secretion can decrease protein secretion by at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 100%, at least about 150%, at least about 200%, at least about 300%, at least about 400%, at least about 500%, at least about 600%, at least about 700%, at

least about 800%, at least about 900% at least about 1,000%, at least about 5,000%, at least about 10,000% or more. A regulatory sequence that modulates protein secretion can decrease protein secretion by at most about 10,000%, at most about 5,000%, at most about 1,000%, at most about 900%, at most about 800%, at most about 700%, at most about 600%, at most about 500%, at most about 400%, at most about 300%, at most about 200%, at most about 150%, at most about 100%, at most about 90%, at most about 80%, at most about 70%, at most about 60%, at most about 50%, at most about 40%, at most about 30%, at most about 20%, at most about 10%, at most about 5%, or less.

[0090] A regulatory sequence that modulates protein secretion can decrease protein secretion by at least or up to about 1-fold, at least or up to about 2-fold, at least or up to about 3-fold, at least or up to about 4-fold, at least or up to about 5-fold, at least or up to about 6-fold, at least or up to about 7-fold, at least or up to about 8-fold, at least or up to about 9-fold, at least or up to about 10-fold, at least or up to about 20-fold, at least or up to about 30-fold, at least or up to about 40-fold, at least or up to about 50-fold, at least or up to about 60-fold, at least or up to about 70-fold, at least or up to about 80-fold, at least or up to about 90-fold, at least or up to about 100-fold, at least or up to about 500-fold, at least or up to about 1,000-fold, at least or up to about 5,000-fold, or at least or up to about 10,000-fold. A regulatory sequence that modulates protein secretion can decrease protein secretion by at most or less than about 10,000-fold, at most or less than about 5,000-fold, at most or less than about 1,000-fold, at most or less than about 500-fold, at most or less than about 100-fold, at most or less than about 90-fold, at most or less than about 80-fold, at most or less than about 70-fold, at most or less than about 60-fold, at most or less than about 50-fold, at most or less than about 40-fold, at most or less than about 30-fold, at most or less than about 20-fold, at most or less than about 10-fold, at most or less than about 9-fold, at most or less than about 8-fold, at most or less than about 7-fold, at most or less than about 6-fold, at most or less than about 5-fold, at most or less than about 4-fold, at most or less than about 3-fold, at most or less than about 2-fold, or at most or less than about 1-fold.

[0091] A food product or substance can be comprised of the recombinant proteins described herein. The food product or substance may comprise recombinant proteins described herein provided in an isolated or purified form. Alternatively, or in addition to, the recombinant proteins described herein may be provided to the food product or substance as whole cells comprising the recombinant proteins. A food product or substance can be a food for animals. A food product or substance can be food for a pet (e.g., companion animal), a farm or agricultural animal, a wild animal, or a human. A food product or substance can be food for a obligate carnivore (e.g., a feline), which requires animal protein as a part of the diet.

[0092] A food product can be a wet puree. Alternatively, a food product can be a dry pellet. A food product can replace an animal or human diet. Alternatively, a food product can supplement an animal or human diet.

[0093] A food product can contain additional ingredients such as, but not limited to, animal-based meat, amino acid supplements (e.g., taurine), vitamins (e.g., vitamin E), fats, nutritional yeast, *spirulina*, agar, xanthan gum, kelp powder, squash powder, sweet potato powder, *papaya* powder, flaxseed oil, coconut oil, probiotics, cranberry powder, wheat grass powder, turmeric, coconut flour, chitosan, pumpkin puree, inulin powder, guar gum, green tea extract, koji, stabilizers, vegetables, fruits, microbes, fillers, or flavoring compounds.

EXAMPLES

Example 1: Recombinant Protein Expression

[0094] In this experiment, a *Pichia pastoris* cell was transfected with a DNA sequence, resulting in recombinant protein expression.

[0095] Host cells were chosen based on predetermined characteristics such as, but not limited to, doubling time, ability to properly fold and make appropriate post-translational modifications on recombinant proteins, protein secretion, ability to grow on cost-effective and/or easily available media/feedstock, or genetic tractability. Host cells that do not efficiently secrete proteins may be lysed to release the target recombinant proteins.

[0096] *Pichia* cells were grown overnight in YPD media in baffled flasks at 30° C. to an optical density at 600 nm wavelength (OD 600) of about 1 to 2.6. Cells were collected by centrifugation and resuspended at room temperature for 30 min in 100 mM lithium acetate (LiOAc), 10 mM dithiothreitol (DTT), 0.6M sorbitol, and 10 mM Tris-HCl pH 7.5, at 1e8 cells/ml. Cells were pelleted and washed four times in 1 mL 1M sorbitol. Cells were resuspended in 1M sorbitol at a final concentration of 1e9 cells/ml. 40 μL of cell suspension was mixed with 1 uL (100-1 ug) of DNA, transferred to a 0.2 cm gap cuvette, and incubated on ice for 5 minutes.

[0097] The expression cassette used comprised a promoter, a 5' untranslated region, a signal sequence for protein secretion, an open reading frame (ORF) of an animal muscle protein, a 3' UTR, and a transcriptional termination and 3' end processing sequence. The construct used a *P. pastoris* AOX1 promoter, a *P. pastoris* 5' UTR, a *P. pastoris* 3' UTR, a *S. cerevisiae* alpha mating factor, a rabbit serum albumin ORF, and a *P. pastoris* terminator.

[0098] The expression cassette was transfected into the *P. pastoris* host cell using electroporation at 1.5 kV, 200 ohms, and 25 uF, where the expression cassette was integrated into the genome of *P. pastoris*. Electroporated cells were diluted in 1 mL 1M sorbitol. Cells were transferred to 15 ml round-bottom tubes with 1 ml YPD media and incubated while shaking at 30° C. for 1 hour.

[0099] Cells were plated on YPDS plates and incubated at 30° C. Once the cells' endogenous protein expression machinery expressed the protein of interest, the cells secreted the protein of interest into the media, which was harvested and purified for further use.

Example 2: Analysis of Secreted Proteins

[0100] In this experiment, genes of interest were cloned into plasmids containing the GAP1 promotor, *Saccharomyces cerevisiae* alpha-Mating factor signal sequence, and a zeocin resistance gene. Plasmids were linearized and transformed into *Pichia pastoris* cells using electroporation. Transformed cells were spread onto YPDS plates containing 100 micrograms per milliliter (ug/mL) zeocin and incubated at 30° C. for 3-4 days. Colonies were restruck to single colony twice. To produce secreted protein, cells were resus-

pended in 4 milliliters (mL) of buffered complex glycerol medium (BMGY) (rich glycerol) and cultured shaking at 250 revolutions per minute (rpm) and 30° C. for 4 days. Every 24 hours, 200 microliters (uL) of culture was removed and centrifuged to separate cells from supernatant. A 10 μL aliquot of supernatant was used to analyze secreted proteins on a sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gel.

[0101] FIG. **1** shows a Coomassie stained SDS-PAGE gel of an example media from *Pichia pastoris* cells including a construct for secreted recombinant rabbit albumin or wildtype cells. Wildtype cells lack the construct for secreted recombinant rabbit albumin. Lane 1 **101** shows a band at 66 kilo Daltons (kDa), which corresponds to the recombinant rabbit albumin. Lane 2 **102**, which correspond to the wildtype media, lacks a band at 66 kDa. Lane 3 **103** shows a bovine serum albumin marker with a band at 66 kDa.

[0102] FIG. **2** shows a Coomassie stained sodium SDS-PAGE gel of an example media from *Pichia pastoris* cells including a construct for secreted recombinant rabbit Tropomyosin alpha-1 chain or wildtype cells. The wildtype cells lack the construct. Lane 2 **201** shows the wildtype expressed protein from wildtype cells. Lanes 3-5 **202** show secreted protein expression for Tropomyosin alpha-1 chain. Lanes 3-5 show a band **203** at 33 kDa corresponding to the Tropomyosin alpha-1 chain.

[0103] For proteins produced in a bioreactor, cells were resuspended in YPG medium and grown in shake flasks at 30° C. for 16 hours. The cells were transferred to seed fermenters and grown at 30° C. for another 16 hours. The cultures were transferred to a 100 liter (L) fermenter and batch fed a glycerol-based media for 92 hours at 30° C. The culture was then centrifuged to separate solid and liquid phases, filtered, and spray dried to produce the final protein powder.

[0104] FIG. **3** shows a Coomassie stained SDS-PAGE gel of an example recombinant rabbit albumin. Lane 1 **301** shows an example recombinant rabbit albumin (rALBU-RABIT) grown from a bioreactor. The rabbit albumin peptides were purified and spray dried. The protein powder was resuspended and run on an SDS-PAGE gel. Lane 2 **302** shows a bovine serum albumin marker.

Example 3: Cat Food with Rabbit Proteins

[0105] In this example, cat food is made using rabbit proteins. Rabbits are naturally hunted by felines in the wild, and rabbit proteins are hypoallergenic for cats. A recombinant rabbit protein is produced and used as an ingredient in both wet and dry cat food, snacks, and/or treats. In some formulations, the protein is in dried powder form. In some formulations, the protein is in a solution. For dry food, ingredients are mixed and extruded.

[0106] Twenty-five rabbit proteins are produced using the recombinant protein expression method of Example 1. The twenty-five rabbit proteins tested are found in Table 2.

[0107] The rabbit proteins are harvested from the recombinant protein expression method and produced into a wet pureed cat treat. Additionally, the rabbit proteins are harvested from the recombinant protein expression method and produced into a dry extruded cat food kibble.

Example 4: Pet Food with Cow Proteins

[0108] In this example, pet food is made using cow proteins. A recombinant cow protein is produced and used as an ingredient in both wet and dry pet food, snacks, and/or treats. In some formulations, the protein is in dried powder form. In some formulations, the protein is in a solution. For dry food, ingredients are mixed and extruded.

[0109] The cow proteins are produced using the recombinant protein expression method of Example 1. The cow proteins tested are found in Table 3.

Example 5: Pet Food with Sheep Proteins

[0110] In this example, pet food is made using sheep proteins. A recombinant sheep protein is produced and used as an ingredient in both wet and dry pet food, snacks, and/or treats. In some formulations, the protein is in dried powder form. In some formulations, the protein is in a solution. For dry food, ingredients are mixed and extruded.

[0111] The sheep proteins are produced using the recombinant protein expression method of Example 1. The sheep proteins tested are found in Table 4.

Example 6: Pet Food with Pig Proteins

[0112] In this example, pet food is made using pig proteins. A recombinant pig protein is produced and used as an ingredient in both wet and dry pet food, snacks, and/or treats. In some formulations, the protein is in dried powder form. In some formulations, the protein is in a solution. For dry food, ingredients are mixed and extruded.

[0113] The pig proteins are produced using the recombinant protein expression method of Example 1. The pig proteins tested are found in Table 5.

Example 7: Pet Food with Bison Proteins

[0114] In this example, pet food is made using bison proteins. A recombinant bison protein is produced and used as an ingredient in both wet and dry cat food, snacks, and/or treats. In some formulations, the protein is in dried powder form. In some formulations, the protein is in a solution. For dry food, ingredients are mixed and extruded.

[0115] The bison proteins are produced using the recombinant protein expression method of Example 1. The bison proteins tested are found in Table 6.

Example 8: Pet Food with Deer Proteins

[0116] In this example, pet food is made using deer proteins. A recombinant deer protein is produced and used as an ingredient in both wet and dry pet food, snacks, and/or treats. In some formulations, the protein is in dried powder form. In some formulations, the protein is in a solution. For dry food, ingredients are mixed and extruded.

[0117] The deer proteins are produced using the recombinant protein expression method of Example 1. The deer proteins tested are found in Table 7.

Example 9: Pet Food with Chicken Proteins

[0118] In this example, pet food is made using chicken proteins. A recombinant chicken protein is produced and used as an ingredient in both wet and dry pet food, snacks, and/or treats. In some formulations, the protein is in dried powder form. In some formulations, the protein is in a solution. For dry food, ingredients are mixed and extruded.

[0119] The chicken proteins are produced using the recombinant protein expression method of Example 1. The chicken proteins tested is found in Table 8.

Example 10: Pet Food with Turkey Proteins

[0120] In this example, cat food is made using turkey proteins. A recombinant turkey protein is produced and used as an ingredient in both wet and dry pet food, snacks, and/or treats. In some formulations, the protein is in dried powder form. In some formulations, the protein is in a solution. For dry food, ingredients are mixed and extruded.

[0121] The turkey proteins are produced using the recombinant protein expression method of Example 1. The turkey proteins tested are found in Table 9.

Example 11: Cat Food with Duck Proteins

[0122] In this example, cat food is made using duck proteins. Ducks are naturally hunted by felines in the wild, and duck proteins are hypoallergenic for cats. A recombinant duck protein is produced and used as an ingredient in both wet and dry cat food, snacks, and/or treats. In some formulations, the protein is in dried powder form. In some formulations, the protein is in a solution. For dry food, ingredients are mixed and extruded.

[0123] The duck proteins are produced using the recombinant protein expression method of Example 1. The duck proteins tested are found in Table 10 and Table 11.

Example 12: Pet Food with Salmon Proteins

[0124] In this example, pet food is made using salmon proteins. A recombinant salmon protein is produced and used as an ingredient in both wet and dry pet food, snacks, and/or treats. In some formulations, the protein is in dried powder form. In some formulations, the protein is in a solution. For dry food, ingredients are mixed and extruded.

[0125] The salmon proteins are produced using the recombinant protein expression method of Example 1. The salmon proteins tested are found in Table 12.

Example 13: Pet Food with Cod Proteins

[0126] In this example, pet food is made using cod proteins. A recombinant cod protein is produced and used as an ingredient in both wet and dry pet food, snacks, and/or treats. In some formulations, the protein is in dried powder form. In some formulations, the protein is in a solution. For dry food, ingredients are mixed and extruded.

[0127] The cod proteins are produced using the recombinant protein expression method of Example 1. The cod proteins tested are found in Table 13.

Example 14: Pet Food with Tuna Proteins

[0128] In this example, pet food is made using tuna proteins. A recombinant tuna protein is produced and used as an ingredient in both wet and dry pet food, snacks, and/or treats. In some formulations, the protein is in dried powder form. In some formulations, the protein is in a solution. For dry food, ingredients are mixed and extruded.

[0129] The tuna proteins are produced using the recombinant protein expression method of Example 1. The tuna protein tested is found in Table 14.

Example 15: Food Product for Human Consumption

[0130] In this example, a food product for human consumption is made using any of the proteins described in Tables 2-14. A recombinant protein from Tables 2-14 is produced and provided in a food, beverage, or supplement. In some formulations, the protein is in dried powder form. In some formulations, the protein is in a solution.

[0131] While preferred embodiments of the present invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. It is not intended that the invention be limited by the specific examples provided within the specification. While the invention has been described with reference to the aforementioned specification, the descriptions and illustrations of the embodiments herein are not meant to be construed in a limiting sense. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention. Furthermore, it shall be understood that all aspects of the invention are not limited to the specific depictions, configurations or relative proportions set forth herein which depend upon a variety of conditions and variables. It should be understood that various alternatives to the embodiments of the invention described herein may be employed in practicing the invention. It is therefore contemplated that the invention shall also cover any such alternatives, modifications, variations or equivalents. It is intended that the following claims define the scope of the invention and that methods and structures within the scope of these claims and their equivalents be covered thereby.

SEQUENCE LISTING

```
Sequence total quantity: 232
SEQ ID NO: 1            moltype = AA  length = 35213
FEATURE                 Location/Qualifiers
source                  1..35213
                        mol_type = protein
                        organism = Mus musculus
SEQUENCE: 1
MTTQAPMFTQ PLQSVVVLEG STATFEAHVS GSPVPEVSWF RDGQVISTST LPGVQISFSD  60
GRARLMIPAV TKANSGRYSL RATNGSGQAT STAELLVTAE TAPPNFSQRL QSMTVRQGSQ  120
VRLQVRVTGI PTPVVKFYRD GAEIQSSLDF QISQEGDLYS LLIAEAYPED SGTYSVNATN  180
SVGRATSTAE LVVQGEEVVP AKKTKTIVST AQISETRQTR IEKKIEAHFD ARSIATVEMV  240
IDGATGQLPH KTPPRIPPKP KSRSPTPPSI AAKAQLARQQ SPSPIRHSPS PVRHVRAPTP  300
SPVRSVSPAG RISTSPIRSV KSPLLIRKTQ TTTMATGPEV PPPWKQEGYV ASSTEAEMRE  360
TTMTSSTQIR REERWEGRYG VQEQVTISGA AAAAASASVS SSFTAGAITT GTKEVKQETD  420
KSAAVATVVA AVDMARVREP AISAVEQTAQ RTTTTAVHIQ PAQEQARKEA EKTAVTKVVV  480
AADKAKEQEL KSRTREVMVT TQEQTHISHE QIRKETEKAF VPKVVISATK AKEQETRITG  540
EITTKQEQKR ITQETIRQET EEIAASMVVV ATAKSTKLEA AVGVQEETAI QQDQMHLTHE  600
```

-continued

```
QMMKETRKTV VPKVIVATPK IKEQDLVSRS REAITTKRDQ VQITQEKKRK EVETTALSTI  660
AVATAKAKEQ ETVLRSREAM ATRQEHIQVT HGQVGVGKKA EAVATVVAAV DQARVREPRE  720
PTHVEESHSQ QTTLEYGYKE HISTTKVPEQ PRRPASEPHV VPQAVKPAVI QAPSETHIKT  780
TDQMGMHISS QVKKTTDIST ERLVHVDKRP RTASPHFTVS KISVPKTEHG YEASIAGSAI  840
ATLQKELSAT SSTQKITKSV KAPTVKPGET RVRAEPTPSP QFPFADMPPP DTYKSQAGIE  900
VKKEVGVSIS GSTVREEHFE VLRGREAKVT ETARVPAPAE VPVTPPTLVS GLKNVTVIEG  960
ESVTLECHIS GYPSPKVTWY REDYQIESSI DFQITFQGGI ARLMIREAFA EDSGRFTCSA 1020
VNEAGTVSTS CYLAVQVSEE FDKETTLTEK FATEEKRFVE SRDVVMTDTS ITEEQAGPGE 1080
PAAPFFISKP VVQKLVEGGS VVFECQIGGN PKPHVYWKKS GVPLTTGYRY KVSYNKQTGE 1140
CRLVISMTFA DDAGEYTIVI RNKHGETSAS ASLLEEADYE ALVKTQQEML YQTQMSTFIQ 1200
EPKVGEIAPG FAYSEYEKEY EKEQALIRKK MAKDTVMVRT FVEDQEFHIS SFEERLIKEI 1260
EYRIIKTTLE ELLEEDGEEK MAVDISESEA IESGFDIRIK NYRILEGMGV TFHCKMSGYP 1320
LPKIAWYKDG KRIRHGERYQ MDFLQDGRAS LRIPVVLPED EGIYTAFASN IKGNAICSGK 1380
LYVEPAAPFS APTYMPTPEA VSRIRSVSPR SLSRSPIRMS PAMSPARMSP ARMSPARMSP 1440
ARMSPGRRLE ETDESQLERL YKPVFVLKPA SFKCLEGQTA RFDLKVVGRP MPETFWFHNG 1500
QQIVNDYTHK VVIKEDGTQS LIIVPASPSD SGEWTVVAQN RAGKSTISVT LTVEAVEHQI 1560
KPAFVEKLKN VNIKEGSRLE MKVRATGNPN PDIVWLKNSD IIVPHKYPRI RIEGTRGEAA 1620
LKIDSIISQD SAWYTATAIN KAGRDTTRCK VNVEVEFAEP EPERKLIIPR GTYRAKEIAA 1680
PELEPLHLRY GQEQWEEGDL YDKEKQQKPF FKKKLTSLRL KRFGPAHFEC RLTPIGDPTM 1740
VVEWLHDGKP LEAANRLRMI NEFGYCSLDY GAAYSRDSGV ITCRATNKYG TDHTSATLIV 1800
KDEKSLVEES QLPDGKKGLQ RIEELERMAH EGALTGVTTD QKEKQKPDIV LFPEPVRVLE 1860
GETARFRCRV TGYPQPKVNW YLNGQLIRKS KRFRVRYDGI HYLDIVDCKS YDTGEVKVTA 1920
ENPEGVTEHK VKLEIQQRED FRSVLRRAPE PKPEFHVHEP GKLQFEVQKV DRPVDTSETK 1980
EVVKLKRAER ITHEKVSEES EELRSKFKRR TEEGYYEAIT AVELKSRKKD ESYEELLKKT 2040
KDELLHWTKE LTEEEKKALA EEGKITIPTF KPERIELSPS MEAPKIFERI QSQTVGQGSD 2100
AHFRVRVVGK PDPECEWYKN GVKIERSDRI YWYWPEDNVC ELVIRDVTAE DSASIMVKAI 2160
NIAGETSSHA FLLVQAKQLI TFTQELQDVV AKEKDTMATF ECETSEPFIK VKWYKDGIEV 2220
HAGDKYRMHS DRKVHFLSVL TIDTSDAEDY SCVLVEDENI KTTAKLIVEG AVVEFVKELQ 2280
DIEVPESYSG ELECIISPEN IEGKWYHNDV ELKSNGKYSI TSRRGRQNLT VKDVTKEDQG 2340
EYSFVVDGKK TTCKLKMKPR PIAILQGLSD QKVCEGDIVQ LEVKVSLENV EGVWMKDGQE 2400
VQHSDRVHIV IDKQSHMLLI EDMTKEDAGN YSFTIPALGL STSGNVSVYS VDVITPLKDV 2460
NVIEGTKAVL ECKVSVPDVT SVKWYLNDEQ IKPDDRVQSI VKGTKQRLVI NRTHASDEGP 2520
YKLMVGRVET SCNLSVEKIK IIRGLRDLTC TETQNVVFEV ELSHSGIDVV WNFKGKEIKP 2580
SSKYKIEAHG KIYKLTVLNM MKDDEGEYAF YAGENTTSGK LTVAGGAIST PLTDQTVAES 2640
QEAVFECEVA NPESEGEWLK DGKHLALSNN FRGESDGHKR RLVIAAAKLD DAGEYTYKVA 2700
TSKTSAKLKV EAVKIKKTLR NLTVTETQDA VFSVELTHPD VKGVQWIKNG VVLDSNDKYE 2760
ISVKGTLYSL KIKNCAMADE SVYGFKLGRL GASARLHVET VKIIKKPKDV TALENATVTF 2820
EVSVSHDTVP VKWFHKNVEI KPSDKHRLVS ERKVHKLMLQ SISPSDAGEY TAMVGQLECK 2880
AKLFVETLHI TKTMKSIEVP ETKAASFECE VSHFNVPSMW LKNGVEIEMS EKFKIVVQGK 2940
LHQLIIMNTS TEDSAEYTFV CGNDQVSATL TVTPIMITSM LKDINAEEKD TITFEVTVNY 3000
EGISYKWLKN GVEIKSTDRC QMRTKKLTHS LNIRNVHFGD AADYTFVAGK ATSTATLYVE 3060
ARHIEFRKHI KDIKVLEKKR AMFECEVSEP DITVQWMKDG QELQIADRIK IQKEKYVHRL 3120
LIPSTRMSDA GKYTVVAGGN MSTANLFVEG RDVRIRSIKK EVQVIEKQRA VVEFEVNEDD 3180
VDAHWYKDGI EINFQVQERH QYVVERRIHR MFISETRHSD AGEYTFVAGR NRSSVTLYVN 3240
APEPPQVLQE LQPVTVQSGK PARFCAVISG RPQPKISWYK EEQLLSTGFK CKFLHDGQEY 3300
TLLLIEAFPE DAAVYTCEAK NDYGVATTSA SLSVEVPEVV SPDQEMPVYP PAIVTPLQDT 3360
VTSEGRPARF QCQVSGTDLK VSWYCKDKKI KPSRFFRMTQ FEDTYQLEIA EAYPEDEGTY 3420
AFVANNAVGQ VSSTATLRLE APESILHERI GQQIEMEMKE IASLLSAEED FQTYSSDLRL 3480
PNANETLELL SEPPARSTQF DSRQEGAAPV FIREISDVEI SVEDVAKLSV TVTGCPKPKI 3540
QWFFNGMLLT PSADYKFVFD GDTHSLIILF TRFQDEGEYT CLASNEYGKA VCSAHLRISP 3600
RGERSTEMES GEKKALEKPK GPCPPYFFKE LKPVHCGPGI PAVFEYSVHG EPAPTVLWFK 3660
EDMPLYTSVC YTIIHSPDGS GTFIVNDPQR GDSGLYLCKA QNLWGESTCA AELLVLPEDT 3720
DVPDASCKEE STLGVPGDFL ETSARGPLVQ GVDSRQEITA FAEGTISKAA LIAEETLQLS 3780
YERSVDDSEV GTGVTIGAQK LPPVVLSTPQ GTGELPSIDG AVHTQPGRGP PPTLNLQAVQ 3840
AQTTLPKEAT LQFEEPEGVF PGASSAAQVS PVTIKPLITL TAEPKGNYPQ SSTAAPDHAL 3900
LSSVAAETLQ LGEKKIPEVD KAQRALLLSQ SLAEGCVESL EVPDVAVSNM RSEPQVPFQH 3960
TCTEGKILMA SADTLKSTGQ DVALRTEEGK SLSFPLALEE KQVLLKEEQS EVVAVPTSQT 4020
SKSEKEPEAI KGVKEVREQE LLSKETLFPS MPEEQRLHLK TQVRRALQAA VAREQANLFS 4080
EWLRNIDKVE VTAVNFTQEP KRILCTYLIT SVSSLTEELT VTIEDIDPQM ANLETGLKDA 4140
LCSIVCEERN ILMAEDPRIH EEDKIDVQGG RDHLSDAQKV ETVIEAEADS KYLVSKEEVS 4200
WSKVESQLKD GDTNEVPQAE TLKLAEESGT QKTSTEMSQE EAEGTLADLC PAVLKHLVDT 4260
ISEEGDTVHL TSSISNAKEV HWYFKGNLVP SDGKFKCLKE QNAYTLVIEA VKTEDEGEYV 4320
CEASNDSGKA KTSAKLTVGE RAAPVIKRRI EPLEVALGHL AKFTCEIQGA PNVRFQWFKA 4380
GREIYESDKC SIRSSNYVSS LEILRTQVVD CGEYTCKASN EYGSVSCTAT LTVTEAYPPT 4440
FLSRPKALTT FVGKAAKFLC TVSGTPVIEI IWQKDGAALS PSPDCRVTDA DNKHSLELSN 4500
LTVQDRGIYS CKASNKFGAD ICQAELTIID KPHFIKELEA VQSAINKKIH LECQVDEDRK 4560
VTITWSKDGQ KLPAGKDYKI YFEDKIASLE IPLAKLKDSG TYTCTASNEA GSSSSSAAVA 4620
VREPPSFVKK VDPSYLMLPG ESARLHCKLK GSPVIQVTWF KNNKELSESN TVRMSFVNSE 4680
AILDITDVKV DDSGTYSCEA TNDVGSDSCS TEVVIKEPPS FIKTLEPADI VRGANALLQC 4740
EIAGTGPFEV NWFKDKKQIR SSKKYRLFTQ KTFVYLEISS FNSADVGDYE CVVANEVGKC 4800
GCVATHLLKE PPTFVKKVDD FTALAGQTVT LQAAVRGSEP ISVMWMKGQE VIKEDGKIKM 4860
SFSNGVAVLT IPDVQISLGG KYTCLAENEA GSQTSVGELI VKEPAKIIER AELIQVTAGD 4920
PATLEYTVSG TPELKPKWYK DGRPLVASKK YRISFKNNIA QLKFYSAELH DSGQYTFEIS 4980
NEVGSSSCET TFTVLDRDIA PLFTKPLRNV DSVVGGACRL DCKIAGSLPM RVSWFKDGKE 5040
LTASDRYQIA FVEGTASLEI SRVDMNDAGN FTCRATNSVG SKDSSGALIV QEPPSFVTKP 5100
GSRDVLPGSA VCLKSAFQGS APLTIKWFKG DKELVSGGSC YITKETSESS LELYAVKTSD 5160
```

-continued

```
SGTYTCKVSN VAGSVECSAD LFVKEPATFI EKLEPSQLLK KGDGTQLACK VTGTPPIKIT  5220
WFANDRELRE SSKHKMSFAE STAVLRLTDV AIEDSGEYMC EAQNEAGSDH CTGIVIVKES  5280
PYFTKEFKSI EVLKEYDVML LAEVAGTPPF EITWFKDNTT LRSGRKYKTF LQDQLVSLQV  5340
LKFVAADAGE YQCRVTNEVG SSTCSARVTL REPPSFIKKI EATSSLRGGT AAFQATLKGS  5400
LPITVTWLKD NDEITEDDNI RMTFENNVAS LYLSGIEVKH DGKYVCQAKN DAGIQRCSAL  5460
LSVKEPATIM EEAVSIDVTQ GDPATLQVKF SGTKEISAKW FKDGQELTLG PKYKISVTDT  5520
VSILKIISTE KKDSGEYTFE VQNDVGRSSC KASINVLDLI IPPSFTKKLR KMDSIKGSFI  5580
DLECIVAGSH PISIQWFKDD QEISASDKHK FSFHDNTAFL EISQLEGTDS GTYTCSATNK  5640
AGHSQCSGHL TVKEPPYFVE KPQSQDVNPG TRVQLKALVG GTAPMTIKWF KDNKELHPGA  5700
ARSVWKDDTS TILELFSAKA ADSGTYICQL SNDVGTTSSK ATIFVKEPPQ FIKKPSPVLV  5760
LRNGQSTTFE CQVTGTPEIR VSWYLDGNEI TDLRKYGISF VDGLATFQIS NARVENSGTY  5820
VCEARNDAGT ASCSIELKVK EPPIFIRELE PVEVVKDSDV ELECEVMGTT PFEVTWLKNN  5880
KEIRSGKKYT MSEKMSVFYL HITKCDPSDV GEYQCIIANE GGSCACSARV ALKEPPSFIK  5940
KIENVTTVLK SSATFQSTVA GSPPISITWL KDDQILEEND NVHISFEDSV ATLQVRSVDN  6000
GHSGRYTCQA KNESGIERCY AFLLVQEPAQ IIEKAKSVDV TEKDPVTLEC VVAGTPELKV  6060
KWLKDGKQIV PSRYFSMSFE NNVASFRIQS VMKQDSGQYT FKVENDFGSS SCDAYLRVLD  6120
QDIPPSFTKK LTKMDKVLGS SIHMECKVSG SLPISAQWFK DGKEISTSAK YRLVCHENTV  6180
SLEVSNLELE DTANYTCKVS NVAGDNACSG ILTVKEPPSF LVKPERQQAI PDSTVEFKAV  6240
LKGTPPFKIK WLKDDVELVS GPKCFIGLEG STSFLNLYSV DSSKTGQYTC QVTNDVGSDS  6300
CTTMLLVTEP PKFVKKLEAS KIIKAGDSAR LECKITGSPE IQVVWYRNEH ELTASDKYQM  6360
TFIDSVAVIQ MNSLGTEDSG DFICEAQNPA GSTSCSTKVI VKEPPVFSSF PPIVETLKNT  6420
EVSLECELSG TPPFEVVWYK DKRQLRSSKK YKVASKNFHA SIHILNVEST DIGEYHCKAQ  6480
NEVGSDACVC AVKLKEPPKF ISKLNSLTVV AGEPAELQAS IEGAQPISVQ WLKEKEEVIR  6540
ESENIRISFV NNVATLQFAK VEPANAGKYI CQVKNDGGVR ENMATLTVLE PAVIIEKAGS  6600
MTVTVGETCA LECKVAGTPE LSVEWYKDGK LLTSSQKHKF SFYNKISSLK ILSVEKEDAG  6660
TYTFQVQNNV GKSSCTAVVD VSDRMVPPSF TRRLKDTGGV LGTSCILECK VAGSSPISIA  6720
WFHEKTKIVS GAKYQTTFSD NVCTLQLNSL DSSDMGSYTC VAANVAGSDE CRALLTVQEP  6780
PSFVKEPEPL EVLPGKNITF TSVIRGTPPF KVGWFRGARE LVKGNRCNIY FEDTVAELEL  6840
FNIDISQSGE YTCVVSNNAG QASCTTRLFV KEPATFVKKL SDHSVEPGKS IILEGTYTGT  6900
LPISVTWKKD GVSITPSERC NIVTTEKTCI LEILSSTKGD AGHYSCEIEN EAGRDACDAL  6960
VSTLEPPYFV TELEPLEASV GDSVSLQCQV AGTPEITVSW FKGDTKLRST PEYRTYFTNN  7020
VATLVFNKVS INDSGEYTCM AENSIGTAAS KTIFRIQERQ LPPSFARQLK DIEQTVGLPV  7080
TLTCRLNGSA PIQVCWYRDG VLLRDDENLQ MSFVDNVATL KILQTDLSHS GQYSCSASNP  7140
LGTASSTARL TAREPKKSPF FDIKPVSIDV IAGESADFEC HVTGAQPMRV TWSKDNKEIR  7200
PGGNYTITCV GNTPHLRILK VGKGDSGQYT CQATNDVGKD MCSAQLSVKE PPKFIKKLDA  7260
SKVAKQGESI QLECKISGSP EIKVVWFRND SELHESWKYN MSFVNSVALL TINEASAEDT  7320
GDYICEAHNG VGDASCSTAL KVKAPPVFTQ KPPPVGALKG SDVILQCEIS GTPPFEVVWV  7380
KDRKQVRSSK KFKITSKNFD TSLHIFNLEA PDIGEYHCKA TNEVGSDTCA CTVKFKEPPR  7440
FVKKLSDAST LIGDPVELQA VVEGFQPISV VWLKDKGEVI RESENVRISF VDNIATLQLG  7500
SPEASQSGKY VCQIKNDAGM RECSAVLTVL EPATIVEKPE PMTVTTGNPF TLECVVAGTP  7560
ELSAKWFKDG RELSSGSRHH ITFVRNLASL KIPSAEMNDK GLYTFEVENR VGKSSCTVSV  7620
HVSDRVVPPS FVRRLKDTSA TLGASVVLEC RVSGSAPISV GWFLDGNEII SSPKCQSSFA  7680
DNVCTLTLSS LEPSDTGAYT CVAANVAGQD ESSAVLTVQE PPSFEQTPDS VEVLPGMSLT  7740
FTSVIRGTPP FKVKWFKGSR ELVSGEACTI SLEDFVTELE LLEVEPGQSG DYSCLVTNDA  7800
GSASCTTHLF VKEPATFVKR LADTSVETGS PIVLEATYSG TPPISVSWMK NEYPLSQSPN  7860
CGITTTEKSS ILEILESTIE DYAQYACLIE NEAGQDICEA LVSVLEPPYF IEPLEHVEAA  7920
IGEPITLQCK VDGTPEIRIS WYKEHTKLRS APAYKMQFKN NVASLVINKV DHSDVGEYTC  7980
KAENSVGAVA SSAVLVIKER KLPPSFARKL KDVHETLGFP VAFECRINGS EPLQVSWYKD  8040
GELLKDDANL QMSFVHHVAT LQILQTDQSH VGQYNCSASN PLGTASSSAK LILSEHEVPP  8100
FFDLKPVSVD LALGESGSFK CHVTGTAPIK ITWAKDNREI RPGGNYKMTL VENTATLTVL  8160
KVAKGDAGQY TCYASNVAGK DSCSAQLGVQ EPPRFIKKLD QSRIVKQDEY TRYECKIGGS  8220
PEIKVLWYKD EVEIQESSKF RMSFEDSVAI LEMHSLSVED SGDYTCEARN AAGSASSSTS  8280
LKVKEPPVFR KKPFPVETLK GADVHLECEL QGTPPFQVSW HKDKRELRSG KKYKIMSENL  8340
LTSIHILNVD TADIGEYQCK ATNDVGSDTC VGSVTMKAPP QFVKKLTDIS TIIGKEVQLQ  8400
TTIEGAEPIS VAWFKDKGEI VRESDNIWIS YSENIATLQF SRAEPANAGK YTCQIKNDAG  8460
MQECYATLSV LEPAAIVEKP ESIKVTTGDT CTLECTVSGT PELSTKWFKD GKELTSDNKY  8520
KISFFNKVSG LKIINVVPGD SGVYSFEVQN PVGKDSCKVS IQVSDRIIPP SFTRKLKETN  8580
GLSGSSVVME CKVYGSPPIS VLWFHDGNEI SSGRKYQTTL TDNTCALTVN MLEEADAGDY  8640
TCIATNVAGS DECSAPLTVR EPPSFVQKPD PMDVLTGSNV TFTSIVKGSP PFTVSWFKGS  8700
TELVPGARCN VSLQDSVGEL ELFDVDTSQS GEYTCIVSNE AGRASCTTRL FVKAPAIFVK  8760
RLNDYSIEKG KPLILEGTFS GTPPISVTWK KNGINVIASQ RCNITTTEKS AILEILSSTV  8820
EDSGQYNCYI ENASGKDSCS AQILILEPPY FVKQLEPVKV TVGDSASLQC QLAGTPEIGV  8880
SWYKGDTKLR PTATCKMHFK NNVATLVFTQ VDSSDSGEYI CRAENSVGEV SSSTFLTVQE  8940
QKLPPSFSRQ LRDVQETVGL PVVFECAVSG SEPISVSWYK DGKPLKDSPN IQTSFLDNIA  9000
TLNIFKTDRS LSGQYSCTAT NPIGSASSGA KLILTEGKNP PFFDIPLAPM DAVVGESADL  9060
ECHVTGTQPI KVTWAKDNRE IRSGGNYQIS YLENSAHLTI VKVDKGDSGQ YTCYAVNEVG  9120
KDSCTAQLNI KERLIPPSFT KKLSETVEET EGNSFKLEGR VAGSQPITVA WYKNNVEIHP  9180
TSNCEIMFKN NALLLQVKRA SMADAGLYTC KATNDAGSAL CTSSIVIKEP KKPPVFDQHL  9240
APVTASEGDS VQLSCHVQGS EPIRIQWLKA GREVKPSDRC SFSFASGTAM LELKETAKAD  9300
SGDYVCKASN VAGSDTSKCK VTIKEKPAAA PAAKKAAVDG KLFFVSEPQS IRVVEKTTAT  9360
FIAKVGGDPI PNVKWTKGKW RQLNQGGRIL IHQKGDEAKL EIRDTTKTDS GLYRCVAFNK  9420
HGEIESNVNL QVDERKKQEK IEGDLRAMLK KTPALKKGSG EEEEIDIMEL LKNVDPKEYE  9480
KYARMYGITD FRGLLQAFEL LKQSQEEETH RLEIEELEKS ERDEKEFEEL VAFIQQRLTQ  9540
TEPVTLIKDI ENQTVLKDND AIFEIDIKIN YPEIKLSWYK GTEKLEPSNK YEISIDGDRH  9600
TLRVKNCQPK DQGNYRLVCG PHIASAKLTV IEPAWERHLQ DVTLKEGQTC TMTCQFSVPN  9660
VKSEWFRNGR VLKPQGRVKT EVEHKVHKLT IADVRAEDQG QYTCKHEDLE TSAELRIEAE  9720
```

-continued

```
PIQFTKRIQN IVVSEHQSAT FECEVSFDDA IVTWYKGPTE LTESQKYNFR NDGRCHYMTI  9780
HNVTPDDEGV YSVIARLEPR GEARSTAELY LTTKEIKLEM KPPDIPDSRV PIPTMPIRAV  9840
PPEEIPPAVA PSIPLLLPLP EEKKPPAKRI EVTKKGVKKD TKKVVTKPKE EAPPPPVAKK  9900
PPPPTPMIPA KASEIIDVSS KAEEVKITTI TRKKEVHKEK EAVYEREEAV YEKKVHIEPW  9960
EEPYEELETE PYTEPYEEPY YEEPDEDYEE IKVEAKKQVH EEWEEDFEEG QEYYEREEGY 10020
DEGEEEWEEI YHEREIIQVQ KEVHEELHEK KIPAKVPEKK VPPPKVVKKP VVEKVEKTTR 10080
RMEEEKVQVI KVPEVSKKIV PQKPSRTPVQ EEIIEVKVPA VHTKKMVISE EKMFFASHTE 10140
EEVSVSVPEV QKKTVTEEKI HVAVSKKIEP PPKVPEPPKK PVPEEVVPVP IPKKVEPPAA 10200
KVPEAPKKPV PEEKKPVPIP KKKEPAAPPQ VPEAPKKPAP EEKIPVPVTK KKEAPPAKVP 10260
EVQKKVVTEE KIAIITQREE SPPPAVPEIP KKKVPEEKRP VPRKEEVPPP KVPVPPKKAV 10320
PEAVVPAPIP KKAPPRAEVS KKTVVEEKRF AAEEKLSVAV PQRVELMRHE EEEWTYSEEE 10380
ERVSVSVYRE EERDEEEAEI TEYEVLEEPE EYVVEEKMHF ISKKVEVEPA KVPEKKIIPK 10440
PKVPAKIEEP PPTKVPEPPK KIVPEKKVPA PAPKKVPPAK APEESKRPVP EKRAPAEEVG 10500
IEEPPPTKVA ERHMKITQEE KVLVAVTKKE APPRARVPEE PKKVAPEERF PKLKPRREEE 10560
PPAKVTEVRK RAVKEEKVSI EVPKREPRPT KEVTVTEEKK WSYTREEETV SEHREEEYED 10620
YEDYEEYKEF EEYEPTEEYD QYDEYAEREV EHYEEHEEYV TEPKKPVPVK PAQEPVPAKP 10680
KAPPPKVLKK AVPEEKAPLP IQKKLKPLPP KAPEEPKKVV EEKIQISITK REKQQVTEPV 10740
AKVPMKPKRV VPEAKIPAPT KEVAVPVRVP GVPKKRELEE VVVFKEEVEA HEEYIVEEEE 10800
EYVHEEEYVH KEEYVHEEEY VHKEEYIHEE EEHLHEEEET IAEEEVVPVA PVKVPVVPKK 10860
PVPEEKKPVP VPKKKEAPPA KVPEIPKKPE EKVPVPIPKK EKAPPAKVPE VPKKPVPEEK 10920
PPVPVPKKVE PPPAKVPEVP KKPVPEKKVP APTPKKVEAP PAKVPEVPKK PIPEEKKPTA 10980
LLKKMEAPPP KAPKKREVVP VPVALPREEE EEEVPFEEVP EEEILPEEEV PSEEEAPPEE 11040
VPPEEEEEVLP EEEEVLPEEE EVLPEEEEVQ PEEEALPEIK PKVPKPAPVP KKTVPEKKVP 11100
VPVPKKVEPP PPPKVPEIKK KVPEKKVVVP KKEEAPPTKV PEVSKKVEER RIIPPKEEEV 11160
PPAEVYEEAE EPTPEEIPEE PPSIEEEEIV EEEEEEEEVL PPRAPEVVKK AVPEAPTPVP 11220
KKAEAPPAKV PKKIPEEKVP VPVQKKEAPP AKVPEKKKIP EKKVPVPKKE AVPPAKGKAV 11280
FEEKISVAYQ QEELVQERIE LELVEAKVEE AFEEEEFHEV QEYFEEEEFH EVEEFIRVEE 11340
RRFQEEHKVE EVHRVIEFLE AEEVEVYEKP KIPPKKGPEV SEKVIPPKKP PTKVIPRKEP 11400
PAKVPEVTKK TVVEEKIRAP EEPKVPAPKA PEVPKKITPE EKVREAVPKK PEVPPPKVPE 11460
VPKKIIQEEK LPVVLPEDTE IYMYEASEET VIEEEHVTLP QKARLKVAKV PAPPQTVVTE 11520
EKTYVTIRKT RETLALKESE TTREAPPELK SYKAVPEIPE PPSPEDLEII EDVLPEKRPP 11580
ASKRRKTQLP TAPEAPREMP PEMNTFEEIS VEPEMLPTQV LDTYQEATVE KKTLRISRKK 11640
PELPSDEEVP EAPREVVAKK KVLPPQVPEV VPVKVPGAPK EVVSERKSLE EPPKKPAVRP 11700
VTVPEEPKEV IPEKKVSLVP PKKPAAPPVT VPEAPEEVFS EDEETLAPPQ EPEAPPAKVP 11760
EAPKEVVPEK KVSVVPPKKP EAPPAKVPEA PKEAAPEKKV PVAPKKKPEA PPVKVPEAPK 11820
KVVPEKKLPV AAPKKPEAPA AEVPEVPKTA VPQKKIPEAI PPKPESPPLE VPEVPPKEVT 11880
PEKKVPAAPP TKPEIPPPKV PEAPQAAVVE EKTPEALPKK AEAAPVPVPQ VQETVPEKTR 11940
PVGPPKKPEA TTVPVPKVQK TIPEKTRPEA PPKRPEARTV PETVPEKTRP MAPPKKPEAT 12000
TLPVPEVQET VPEKTRPVGP PKKPEATTVP VPEVQETIPE KTRPPKKPEA TPVPVPEVQE 12060
TVPEKTRPVG PPKKPEATTV SVPEVQETIP EKTRPAAPPK KPEATAVPET IPEKTRPEAP 12120
PKRPEATTVP VPEADQAVVP EKKVPRVPPK KVEAPPITVP EEPKEVIPEK KVSLVPPKKP 12180
AAPPVTVPEA PEEVFSEDEE TLAPPQEPEA PPAKVPEAPK EVVPEKKVSV VPPKKPEAPP 12240
AKVPEAPKEA APEKKVPVAP KKKPEAPPVK VPEAPKKVVP EKKLPVAAPK KPEAPAAEVP 12300
EVPKAAVPQK KIPEAIPPKP ESPPPEVYEE PEEEIVPEEP PEEAVEEPVP APPPKVTEPP 12360
KKPVPEKKAP PAVVKKPEPP PAKVPEVPKE APPEKKVPPK KPEAPPAKVP EVPKEVVTEK 12420
KVAVPKKPEV PPAKVPEVPK KPVIEEKPAI PVVEKVASPP AEVYEEPEEV TAEEEEPAPA 12480
VEEEEYEAPP PPAPVPEEPK KVVPEKKFPV IKKPEAPPPK VPEVPKKAVP VKKVPVVKKP 12540
EPPEAEVPEV PKKLVPVKKE PVPVTKKTEV LPEKVPEAPK KITPEKKESV PVPEEPEAPP 12600
ASVEETPEET IYEEKATITI GRKETPPVEE REIEKFIQPE EPELEPEPEE IPVQEPEPEK 12660
KVIEKPKLKP RPPARPPSPP KEDVKEKMFQ LKAVSKKKVP EKPEVVEKVE PAPLKVPTAE 12720
KKVRKLLPEP KPQPKEEVVL KSVLRKKPEE EEPKVEPKKV EKAKKPEEPQ PPPKAVEVEA 12780
PPEPTPKERK VPEPAKVPEI KPAIPLPGPE PKPKPEPEVK TMKAPPIEPA PTPIAAPVTA 12840
PVVGKKAEAK PKDEAAKPKG PIKGVAKKTP SPIEAERKKL RPGSGGEKPP DEAPFTYQLK 12900
AVPLKFVKEI KDIVLTEAES VGSSAIFECL VSPSTAITTW MKDGSNIRES PKHRFIADGK 12960
DRKLHIIDVQ LSDAGEYTCV LRLGNKEKTS TAKLIVEELP VRFVKTLEEE VTVVKGQPLY 13020
LSCELNKERD VVWRKDGKIV VEKPGRIVPG VIGLMRALTI NDADDTDAGT YTVTVENANN 13080
LECSSCVKVV EIIREWLVKP IRDQHVKPKG TAVFTCDIAK DTPNIKWFKG YDEIPLEPND 13140
KTEILKEGNH LFLKVKNAMP EDIDEYAVEI EGKRYPAKLT LGEREVELLK PIEDVTIYEK 13200
ESASFDAEIS EEDIPGEWKL KGELLRPSPT CEIKAEGGKR FLTLHKVKLD QAGEVLYQAC 13260
NAITTAILTV KEIELDFAVP LKDVTVPEKR QARFECVLTR EANVIWSKGP DIIKASDKFD 13320
IIADGKKHIL VINDSQFDDE GVYTAEVEGK KTSAQLFVTG IRLKFISPLE DQTVKEGQTA 13380
TFVCELSHEK MHVVWFKNDV KLHTTRTVLM SSEGKTYKLE IRETTLDDIS QIKAQVKNLS 13440
STANLKVLEA DPYFTVKLHD KTGVEKDEII LKCEVSKDVP VKWFKDGEEI VPSPKHSVKT 13500
DGLRRILKIK KAELKDKGEY VCDCGTDTTK ANVTVEARLI KVEKPLYGVE VFVGETARFE 13560
IELSEPDVHG QWKLKGEPLT ASPDCEIIED GKKHVLVLYN CQLDMTGEIS FQAANAKSAA 13620
NLKVKELPLI FITPLSDVKV FEKDEAKFEC EVSREPKTFR WLKGTQEITG DDRFELIKDG 13680
TRHSLVIKSA AFEDEAKYMF EAEDKRTSGK LIIEGIRLKF LTPLKDVTAK ERENAVFTVE 13740
LSHDNIPVSW FKNDQRLHTS KRVSMHDEGK THSITFKDLS IDDTSQIRVE AMGISSEAKL 13800
TVLEGDPYFT GKLQDYTGVE KDEVILQCEI SKADAPVKWF KDGKEIKPSK NVVIKADGKK 13860
RMLILKKALK SDIGQYTCDC GTDQTSGKLD IEDREIKLVR PLYSVEVMET ETARFETEIS 13920
EDDIHANWKL KGEALLQTPE CEIKEEGKIH VLILHNCRLD QTGGVDFQAA NVKSSAHLRV 13980
KPRVIGLLRP LKDVTVTAGE TATFDCELSY EDIPVEWYLK GKKLEPNDKV VTRSEGRVHT 14040
LTLRDVKLED AGEVQLTAKD FKTQANLFVK EPPVEFTKPL EDQTVEEEAT AVLECEVSRE 14100
NAKVKWFKNG TEILKSKKYE IVADGRVRKL IIHGCTPEDI KTYTCDAKDF KTSCNLNVVP 14160
PHVEFLRPLT DLQVKEKETA RFECEISKEN EKVQWFKDGA EIKKGKKYDI ISKGAVRILV 14220
INKCLLNDEA EYSCEVRTAR TSGMLTVLEE EAVFTKNLAN LEVSEGDTIK LVCEVSKPGA 14280
```

-continued

```
EVIWYKGDEE IIETGRFEIL TDGRKRILII QNAQLEDAGS YNCRLPSSRT DSKVKVHELA  14340
AEFISKPQNL EILEGEKAEF VCTISKESFE VQWKRDDQTL ESGDKYDIIA DGKKRVLVVK  14400
DATLQDMGTY VVMVGAARAA AHLTVIEKLR IIVPLKDTKV KEQQEVVFNC EVNTEGAKAK  14460
WFRNEEAIFD SSKYIILQKD LVYTLRIRDA RLDDQANFNV SLTNHRGENV KSAANLIVEE  14520
EDLRIVEPLK DIETMEKKSV TFWCKVNRLN VTLKWTKNGE EVAFDNRISY RIDKYKHSLI  14580
IKDCGFPDEG EYVVTAGQDK SVAELLIIEA PTEFVEHLED QTVTEFDDAV FSCQLSREKA  14640
NVKWYRNGRE IKEGKKYKFE KDGSIHRLII KDCRLEDECE YACGVEDRKS RARLFVEEIP  14700
VEIIRPPQDI LEAPGADVIF LAELNKDKVE VQWLRNNMIV VQGDKHQMMS EGKIHRLQIC  14760
DIKPRDQGEY RFIAKDKEAR AKLELAAAPK IKTADQDLVV DAGQPLTMVV PYDAYPKAEA  14820
EWFKENEPLS TKTVDTTAEQ TSFRISEAKK DDKGRYKIVL QNKHGKAEGF INLQVIDVPG  14880
PVRNLEVTET FDGEVSLAWE EPLTDGGSKI IGYVVERRDI KRKTWVLVTD RADSCEFTVT  14940
GLQKGGVEYL FRVSARNRVG TGEPVETDSP VEARSKYDVP GPPLNVTITD VNRFGVSLTW  15000
EPPEYDGGAE ITNYVIELRD KTSIRWDTAM TVRAEDLSAT VTDVVEGQEY SFRVRAQNRI  15060
GVGKPSAATP FVKVADPIER PSPPVNLNAS EQTQSSVQLT WEPPLKDGGS PILGYIIERR  15120
EEGKDNWIRC NMKPVPELTY KVTGLQKGNK YLYRVSAENA AGVSDPSEIL GPLTADDAFV  15180
EPTMDLSAFK DGLEVIVPNP IKILVPSTGY PRPKATWTFG DQVLEEGDRV KMKTISAYAE  15240
LVISPSERTD KGIYTLTLEN PVKSISGEIN VNVIAPPSAP KELKFSDITK DSVHLTWEPP  15300
DDDGGSPLTG YVVEKRDMSR KTWTKVMDFV TDLEFTVPDL VQGKEYLFKV CARNKCGPGE  15360
PAYTDEPVNM SAPATVPDPP ENVKWRDRTA NSIFLTWDPP KNDGGSRIKG YIVEKCPRGS  15420
DKWVACGEPV PDTKMEVTGL EEGKWYAYRV KALNRQGASK PSKPTEEIQA VDTQEAPEIF  15480
LDVKLLAGIT VKAGTKIELP ATVTGKPEPK ITWTKADTLL KPDQRITIEN VPKKSTVTIT  15540
DSKRSDTGTY IIEAVNVCGR ATAVVEVNVL DKPGPPAAFD ITDVTNESCL LTWNPPRDDG  15600
GSKITNYVVE RKATDSDVWH KLSSTVKDTN FKATKLTPNK EYIFRVAAEN MYGVGEPVQA  15660
APIIAKYQFD PPGPPTRLEP SDITKDAVTL TWCEPDDDGG SPITGYWVER LDPDTDKWVR  15720
CNKMPVKDTT YRVKGLTNKK KYRFRVLAEN LAGPGKPSRS TEPILIKDPI DPPWPPGKPT  15780
VKDIGKTSLV LNWTKPEHDG GAKIESYVIE MLKTGTDEWV RVAEGVPTTE HLLTGLMEGQ  15840
EYSFRVRAVN KAGESEPSEP SDPVLCREKL YPPSPPRWLE VINITKNTAD LKWTVPEKDG  15900
GSPITNYIVE KRDVRRKGWQ TVDTTVKDTK CTVTPLTEGS LYVFRVAAEN AIGQSDYTEI  15960
GDSVLAKDTF TTPGPPYALT VVDVTKRHVD LKWEPPKNDG GRPIQRYIIE KKEKLGTRWV  16020
KAGKTSGPDC NFRVTDVIEG TEVQFQVRAE NEAGVGHPSE PTEILSIEDP TSPPSPPLDL  16080
HVTDAGRKHI AIAWKPPEKN GGSPIIGYHV EMCPVGTEKW MRVNSRPIKD LKFKVEEGIV  16140
PDKEYVLRVR AVNAVGVSEP SEISENVVAK DPDCKPTIDL ETHDIVVIEG EKLNIPVPFR  16200
AVPVPTVSWH KDGKEVKASD RLTMKNDHIS AHLEVPKSVH ADAGVYTITL ENKLGSATAS  16260
INVKVIGLPG PCKDIKASDI TKSSCKLTWE PPEFDGGSPI LHYVLERREA GRRTYIPVMS  16320
GENKLSWTVK DLIPNGEYFF RVKAVNKIGG GEYIELKNPV IAQDPKQPPD PPVDVEVHNP  16380
TAKAMTITWK PPLYDGGSKI MGYIIEKITK GEDRWKRCNE HLVPVLTYTA KGLEEGKEYQ  16440
FRVRAENAAG IGEPSRATPP TKAVDPIDAP KVILRTSLEV KRGDEIALDA TISGSPYPTI  16500
TWIKDENVIV PEEIKKRAAP PVRRKKGEAE EEEPFSLPLT ERLSINNSKQ GESQLRIRDS  16560
LRPDHGQYMI KVENDHGVAK APCSVSVLDT PGPPINFVFE DIRKDSVLCK WEPPLDDGGS  16620
EIINYTLEKK DKTKPDSDWI VITSTLRNCK YSVTKLIEGK EYLFRVRAEN RFGPGPPCVS  16680
KPLLAKDPFE PPDAPDKPIV EDVTSNSMLV KWNEPKDNGS PILGYWLEKR EVNSTHWSRV  16740
NKTLLSSLKT KVDGLLEGLT YVFRVCAENA AGPGKFSPPS DPKTARDPIS PPGPPVPRVA  16800
DTSSTTIELE WEPPAFNGGG EIMGYFVDKQ LVGTNEWSRC TEKMIKVRQF TVKEIREGAD  16860
YKLRVSAVNA AGEGPPGETE PVTVAEPQEP PTVELDVSVK GGIQIMAGKT LRIPAEVTGR  16920
PVPTKVWTIE EGELDKERVI IENVGTKSEL IIKNALRKDH GRYVITATNS CGSKFAAVRV  16980
EVFDVPGPVL DLKPVVTNRK MCLLNWSDPA DDGGSDITGF IIERKDAKMH TWRQPIETER  17040
SKCDITGLIE GQEYKFRVIA KNKFGCGPPV EIGPILAVDP LGPPTSPERL TYTERTKSTI  17100
TLDWKEPRSD GGSPIQGYII EKRRHDKPDF ERVNKRLCPT TSFLVENLDE HQMYEFRVKA  17160
VNDVGESEPS LPLNVVIQDD EVPPTIKLRL AVRGDTIKVK AGEPVNIPAD VTGLPMPKIE  17220
WSKNEKVIDK PTDTLNITKE EVSRSEAKTE LSIPKAARED KGTYTITASN RLGSVFRNVH  17280
VEVYDRPSPP RNLAVTDIKA ESCYLTWDAP LDNGGSEITH YIIDKRDASR KKSEWEEVTN  17340
TAVERRYGIW KLIPNGQYEF RVRAVNKYGI SDECKSDKVV IQDPYRLPGP PGKPKVLERT  17400
KGSMLVSWTP PLDNGGSPIT GYWLEKREEG GTYWSRVSRA PITKVGLKGV EFNVPRLIEG  17460
VKYQFRAMAI NAAGIGPPSE PSDPAVAGDP IYPPGPPSCP EVKDKTKSSI SLAWKPPAKD  17520
GGSPIKGYIV EMQEEGTTDW KPVNEPDKLL TACECVVPNL KELRKYRFRV KAVNEAGESE  17580
PSDTTGEIPA TDIQEVPEVF IDIGAQDCLV CKAGSQVKIP AVIKGRPTPK SSWEFDGKAK  17640
KAMKDGVHDI PEDAQLETAE NSSVIIIPEC TRAHSGKYSI TAKNKAGQKT ANCRVKVMDA  17700
PGPPKDLKVS DITRGSCRLS WKMPDDDGGD RIKGYVIEKK TIDGKAWTKV NPNCGSTTFV  17760
VPDLISEQQY FFRVRAENRF GIGPPAETIQ RTTARDPIYP PDPPIKLKIG LITKNTVHLS  17820
WKPPKNDGGS PVTHYIVECL AWDPTGKKKE AWRQCNRRDV EELEFTVEDL IEGGEYEFRV  17880
KAVNEAGVSK PSATVGPVIV KDQTCPPAIE LKEFMEVEEG TDVNIVAKIK GVPFPTLTWF  17940
KAPPKKPDSK EPVVYDTHVN KQVVDDTCTL VIPQSRRSDT GLYSITAVNN LGTASKEMRL  18000
NVLGRPGPPV GPIKFESISA DQMTLSWLPP KDDGGSKITN YVIEKREANR KTWVRVSSEP  18060
KECMYTIPKL LEGHEYVFRI MAQNKYGIGE PLDSEPETAR NLFSVPGAPD KPTVSSVTRN  18120
SMTVNWEEPE YDGGSPVTGY WLEMKDTTSK RWKRVNRDPI KAMTLGVSYK VTGLIEGSDY  18180
QFRVYAINAA GVGPASLPSD PVTARDPVAP PGPPFPKVTD WTKSSVDLEW SPPLKDGGSK  18240
ITGYIVEYKE EGKEEWEKGK DKEVRGTKLV VTGLKEGAFY KFRVRAVNVA GVGEPGEVTD  18300
VIEMKDRIVS PDLQLDASVR DRIVVHAGGV IRIIAYVSGK PPPTVTWSMN ERALPQEAAI  18360
ETTAISSSMV IKNCQRSHQG VYSLLAKNEG GERKKTIIVD VLDVPGPVGI PFLSDNLTND  18420
SCKLTWFSPE DDGGSPITNY VIQKREADRR AWTPVTYTVT RQNATVQGLI QGKSYFFRIA  18480
AENSIGMGPF VETPNALVIR DPITVPERPE DLEVKEVTKN TVSLTWNPPK YDGGSEIINY  18540
VLESRLIGTE KFHKVTNDNL LSRKYTVKGL KEGDTYEYRV SAVNIVGQGK PSFCTKPITC  18600
KDELAPPTLD LDFRDKLTVR VGESFALTGR YSGKPKPKID WFKDEADVLE DDRTHIKTTP  18660
TTLALEKTKA KRSDSGKYCV VVENSTGSRK GFCQVNVVDR PGPPVGPVIF DEVTKEYMVI  18720
SWKPPLDDGG SEITNYIIEK KELGKDIWMP VTSASAKTTC KVPKLLEGKD YIFRIHAENL  18780
YGISDPLVSD SMKAKDRFRV PDAPEQPVVT EVTKDSALVT WNKPNDGGKP ITNYILEKRE  18840
```

-continued

```
TMSKRWVRVT KEPIHPYTKY RVPDLLEGCQ YEFRVSAENE IGIGDPSPPS KPVFARDPIA  18900
KPSPPINPEA IDTTCNSVEL TWQPPRHDGG SKILGYIVEY QKVGDEEWRR ANHTPESCPE  18960
TKYKVTGLRD GQSYKFRVLA VNEAGESDPA HVPEPVLVKD RLEPPELILD ANMAREQHIR  19020
VGDTLRLSAI IKGVPFPKVT WKKEDREAPT KAQIDVTPVG SKLEIRNAAH EDGGIYSLTV  19080
ENPAGTKTVS VKVLVLDKPG PPRDLEVSEI RKDSCYLTWK EPLDDGGSVV TNYVVERKDV  19140
ATAQWSPLST TSKKKSHMAK HLTEGNQYLF RVAAENQYGR GPFVETPKPI KALDPLHPPG  19200
PPKDLHHVDV DKTEVSLVWN KPDRDGGSPI TGYLVEYQEE GDKDWIKFKT VKNLDCVVTG  19260
LQQGKTYRFR VKAENIIGLG LPDTTIPIEC QEKLVPPSVE LDVKLIEGLV VKAGTTVRFP  19320
AIIRGVPVPT AKWTTDGTEI KTDDHYTVET DSFSSVLTIK NCLRKDTGEY QLTVSNAAGT  19380
KTVAVHLTVL DVPGPPTGPI NILDVTPEYM TISWQPPKDD GGSPVINYIV EKQDTRKGTW  19440
GVVSAGSSKL KLKVPHLQKG CEYVFRVKAE NKMGVGPPLD SIPTVAKHKF SPPSPPGKPV  19500
VTDITENAAT VSWTLPKSDG GSPITGYYVE RREITGKWVR VNKTPIADLK FRVTGLYEGN  19560
TYEFRVFAEN LAGLSNPSPS SDPIKACRPI KPPGPPINPK LKDKSKESAD LVWTKPLSDG  19620
GSPILGYVVE YQKPGTAQWD RINKDELIRQ CAFRVPGLIE GNEYRFRIRA ANIVGEGEPR  19680
ELAESVIAKD ILHPPEVELD VTCRDVITVR VGQTIRILAR VKGRPEPDIT WSKEGKVLVK  19740
DKRVDLIHDL PRVELQIKEA VRADHGKYII SAKNSSGHAQ GSAIVNVLDR PGPCQNLKVS  19800
NVTKENCTIS WENPLDNGGS EITNFIVEYR KPNQKGWSIV ASDVTKRLVK ANLLANNEYY  19860
FRVCAENKVG VGPTIETKTP ILAINPIDRP GEPENLHIAD KGKTFVYLKW RRPDYDGGSP  19920
NLSYHVERRL KGSADWERVH KGSIKETHYM VDKCVENQIY EFRVQTKNEG GESDWVRTEE  19980
VVVKEDLQKP VLDLKLSGVL TVKAGDTIRL EAGVRGKPFP EVAWTKDKDA TDLTRSPRVK  20040
IDTSAESSKF SLTKAKRSDG GKYVITATNP AGSFVAYATV NVLDKPGPVR NLKITDVSSD  20100
RCTIRWDPPE DDGGCEIQNY ILEKCESKRM VWSTYSANVL TPSATVTRLI EGNEYIFRVR  20160
AENKIGTGPP TESKPVIAKT KYDRPGRPDP PEVTKVSKEE MTVVWNAPEY DGGKSITGYY  20220
LEKKEKHAVR WVPVNKSAIP ERRLKVQNLL PGHEYQFRVK AENEVGIGEP SLPSRPVVAK  20280
DPIEPPGPPT NFKVVDTTKS SITLAWGKPV YDGGAPIIGY VVEMRPKIAD ASPDEGWKRC  20340
NAAAQLIRME FTVTSLDENQ EYEFRVCAQN QVGIGRPAEL KEAIKPKEIL EPPEIDLDAS  20400
MRKLVVVRAG CPIRLFAIVR GRPAPKVTWR KVGIDNVVRK GQVDLVDTMA FLVIPNSTRD  20460
DSGKYSLTLV NPAGEKAVFV NVKVLDTPGP VADLKVSDVT KTSCHVSWAP PENDGGSQVT  20520
HYIVEKREAE RKTWSTVTPE VKKTSFNVTN LVPGNEYFFR VTAVNEYGPG VPTDIPKPVL  20580
ASDPLSEPDP PRKLEVTEMT KNSATLAWLP PLRDGGAKID GYIISYREED QPADRWTEYS  20640
VVKDLSLIVT GLKEGKKYKF RVAARNAVGV SMPREAEGVY EAKEQLLPPK ILMPEQITIK  20700
AGKKLRVEAH VYGKPNPICK WKKGDDEVVT SSHLAIHKAD GSSVLIIKDV TRKDSGYYSL  20760
TAENSSGSDT QKIKVTVMDA PGPPQPPFDI SEIDADACSL SWHIPLEDGG SNITNYIVEK  20820
CDVSRGDWVT ALASVTKTSC RVGKLIPGQE YIFRVRAENR FGISEPLTSP KMLAKFPFDV  20880
PSEPKNARVT KVNKDCIFVA WDRPDSDGGS PITGYLIERK ERNSLLWVKA NDTIVRSTEY  20940
PCAGLVEGLE YSFRIYALNK AGSSPPSKPT EYVTARMPVD PPGKPEVVDV TKNSASLIWA  21000
RPKHDGGSRI IGYFVEACKL PGDKWVRCNT TPHQIPQEEY TATGLEENAQ YQFRAIAKTA  21060
VNISQPSEPS DPVTILAENV PPRIELSVEM KSLLTVKAGT NVCLDATVFG KPMPTVSWKK  21120
DTTPIKQAEG IKMAMKRNLC TLELFSVNRK DSGDYTITAE NSSGSKSATI KLKVLDKPGP  21180
PASVKINKMY ADRAMLSWEP PLEDGGSEIT NYIIDKRETS RPNWAQVSAT VPITSCTVEK  21240
LIEGHEYQFR ICAENKYGVG DPILTEPAIA KNPYDPPGRC DPPVISNITK DHMTVSWKAP  21300
ADDGGSPITG YLVEKRETQA VNWTKVNRKP VIERTLKATG LQEGTEYEFR VTAINKAGPG  21360
KPSDASKAVY AQDPLYPPGP PAFPKVYDTT RSSVSLSWGK PAYDGGSPII GYLVEVKRAD  21420
SDHWVRCNLP EKLQKTRFEV TGLMENTEYQ FRVYAVNKIG YSDPSDVPDK HCPKDILIPP  21480
EGELDAELRK TLILRAGVTM RLYVPVKGRP PPKITWSKPN VNLRERIGLD IKSTDFDTFL  21540
RCENVNKYDA GKYILTLENS CGKKEYTIVV KVLDTPGPPV NVTVKEVSKD SAYVTWDPPI  21600
IDGGSPIINY VVEKRDAERK SWSTVTTECS KTSFRVSNLE EGKSYFFRVF AENEYGIGDP  21660
GETRDAVKAS ETPGPVVDLK ALAITKSSCT IGWKKPRSDG GSRITGYVVD FLTEENKWQR  21720
VMKSLSLQYS TKDLKEGKEY TFRVSAENEN GEGTPSEIVV VAKDDVVAPD LDLKDLPDLC  21780
YLAKENSNFR LKIPIKGKPA PSVSWKKGED PLATDTRVSV ESTAVNTTLV VYDCQKSDAG  21840
KYTITLKNVA GTKEGTLSIK VVGKPGIPTG PIKFDEVTAE AMTLKWGPPK DDGGSEITNY  21900
VLEKRDSVNN KWVTCASAVQ KTTFRVTRLH EGIEYTFRVS AENKYGVGEG LKSEPIVAKH  21960
PFDVPDAPPP PNIVDVRHDS VSLTWTDPKK TGGSPITGYH IEFKERNSLL WKRANKTPIR  22020
MKDFKVTGLT EGLEYEFRVM AINLAGVGKP SLPSEPVVAL DPIDPPGKPE VISVTRNSVT  22080
LIWTEPKYDG GHKLTGYIVE KRDLPSKSWM KANHVNVPDC AFTVTDLVEG GKYEFRIRAK  22140
NTAGAISAPS ESTGTIICKD EYEAPTIVLD PTIKDGLTVK AGDSIVLSAI SILGKPLPKS  22200
SWSRAGKDIR PSDIAQITST PTSSMLTVKY ATRKDAGEYT ITATNPFGTK EEHVKVSVLD  22260
VPGPPGPIEI SNVSAEKATL TWTPPLEDGG SPIKAYVLEK RETSRLLWTV VSEDIQACRH  22320
VVTKLIQGNE YLFRVSAVNH YGKGEPVQSE PVKMVDRFGP PGPPGKPEIS NVTKNTATVS  22380
WKRPTDDGGS EITGYHVERR EKKGLRWVRA TKTPVSDLRC KVTGLQEGNT YEFRVSAENR  22440
AGIGPPSDAS NPVLMKDVAY PPGPPSNAHV TDTTKKSASL AWGKPHYDGG LEITGYVVEH  22500
QKVGDDAWIK DTTGTALRIT QFVVPDLQTK EKYNFRISAI NDAGVGEPAV IPNVEIVEKE  22560
VAPDFELDAE LRRTLVVRAG LSIRIFVPIK GRPAPEVTWT KDNINLKHRA NIENTESFTL  22620
LIIPECNRYD TGKFVMTIEN PAGKKSGFVN VRVLDTPGPV LNLRPTDITK DSVTLHWDLP  22680
LIDGGSRITN YIVEKREATR KSYSTVTTKC HKCTYKVTGL TEGCEYFFRV MAENEYGVGE  22740
PTETTEPVRA SEAPLPPDSL NIMDITKNTV SLAWPKPRHD GGSKITGYVI EAQRKGSDQW  22800
THISTVKGLE CVVRNLTEGE EYTFQVMAVN SAGRSAPRES RPVIVKEQTM LPELDLRGIY  22860
QKLVIARAGD NIKVEIPVLG RPKPTVTWKK GDQILKQTQR VNVENTATST ILNINECVRS  22920
DSGPYPLTAK NTVGEVGDVI TIQVHDIPGP PTGPIKFDEV SSDFVTFSWE PPENDGGVPI  22980
SNYVVEMRQT DSTTWVELAT TVIRTTYKAT RLTTGVEYQF RVRAQNRYGV GPGITSASVV  23040
ANYPFKVPGP PGTPQVTAVT KDSMTISWHE PLSDGGSPIL GYHIERKERN GILWQTVSKA  23100
LVPGNIFKST GLTDGIAYEF RVIAENMAGK SKPSKPSEPM FALDPIDPPG KPVPLNITRH  23160
TVALKWAKPE YTGGFKITSY VVEKRDLPNG RWLKANFSNI LENEFTVSGL TEDAAYEFRV  23220
IAKNAAGAIS PPSEPSDAIT CRDDLEAPRI MVDVRFKDTI TLKAGEAFKL EADVSGRPPP  23280
TMEWTKDGKE LEGTGKLEIK IADFSTHLIN KDSSRTDSGA YILTATNPGG FAKHIFNVKV  23340
LDRPGPPEGP LAVSDVTSEK CVLSWLPPLD DGGAKIDHYI VQKRETSRLA WTNVATEVQV  23400
```

-continued

```
TKLKVTKLLK GNEYIFRVMA VNKYGVGEPL ESEPVLAVDP YGPPDPPKNP EVTTITKDSM  23460
VVCWGHPDSD GGSEIINYIV ERRDKAGQRW VKCNKKALTD LRFKVSGLTE GHEYEFRIMA  23520
ENAAGISAPS ATSPFYKACD TVFKPGPPGN PRVLDTSRSS ISIAWNKPIY DGGSEITGYM  23580
VEIALPEEDE WQVVTPPAGL KATSYTITSL IENQEYKIRI YAMNSEGLGE PALVPGTPKA  23640
EERMLPPEIE LDADLRKVVT IRACCTLRLF VPIKGRPAPE VKWAREHGES LDKASIESTS  23700
SYTLLVVGNV NRFDSGKYIL TVENSSGSKS AFVNVRVLDT PGPPQNLKIK EVTKTSVTLT  23760
WEPPLLDGGS KIKNYIVEKR ESTRKAYSTV ATNCHKTSWK VDQLQEGCSY YFRVLAENEY  23820
GIGLPAETAE SVKASERPLP PGKITLTDVT RNSVSLSWEK PEHDGGSRIL GYIVEMQSKG  23880
SDRWATCATV KVTEATITGL IQGEEYSFRV SAQNEKGISD PRQLSVPVIA KDLVIPPAFK  23940
LLFNTFTVLA GEDLKIDVPF IGRPPPAVTW HKDDIPLKQT TRVNAESTEN NSLLTIKEAC  24000
REDVGHYTVK LTNSAGEATE TLNVIVLDKP GPPTGPVKMD EVTADSVTLS WEPPKYDGGS  24060
SINNYIVEKR DTSTTAWQIV SATVARTTIK ACRLKTGCEY QFRIAAENRY GKSTYLNSEP  24120
VVAQYPFKVP GPPGTPFVTL ASKDSMEVQW HEPVSDGGSK VIGYHLERKE RNSILWVKLN  24180
KTPIPQTKFK TTGLEEGIEY EFRVSAENIV GIGKPSKPSE CYAAHDPCDP PGRPEAIIVT  24240
RNSVTLQWKK PTYDGGSKIT GYIVEKKELP DGRWMKASFT NIIDTQFEVT GLLEDHRYEF  24300
RVIARNAAGV FSEPSESTGA ITARDEVEPP RISMDPKYRD TVVVQAGESF KIDADIYGKP  24360
IPTTQWVKGD QELSSTARLE IKSTDFATSL SVKDAVRVDS GNYILKAKNV AGEKSVTINV  24420
KVLDRPGPPE GPVAISGVTA EKCTLAWKPP LQDGGSDITN YIVERRETSR LVWTLVDANV  24480
QTLSCKVLKL LEGNEYIFRI MAVNKYGVGE PLESESLIAK NPFVVPDAPK APEVTAVTKD  24540
SMIVVWERPA SDGGSEILGY VLEKRDKEGI RWTRCHKRLI GELRLRVTGL LENHNYEFRV  24600
SAENAAGLSE PSPPSAYQKA CDPIYKPGPP NNPKVMDVTR SSVFLSWTKP IYDGGCEIQG  24660
YIVEKCDVSV GEWTMCTPPT GINKTNLEVE KLLEKHEYNF RICAINKAGV GEHADVPGPV  24720
MVEEKLEAPD IDLDLELRKV INIRAGGSLR LFVPIKGRPT PEVKWGKVDG DIRDAAIIDV  24780
TSSFTSLVLD NVNRYDSGKY TLTLENSSGT KSAFVTVRVL DTPSPPVNLK VTEITKDSVS  24840
ITWEPPLLDG GSKIKNYIVE KREATRKSYA AVVTNCHKNS WKIDQLQEGC SYYFRVTAEN  24900
EYGIGLPART ADPIKVAEVP QPPGKITVDD VTRNSVSLSW TKPEHDGGSK IIQYIVEMQA  24960
KNTDKWSECA RVKSLDAVIT NLTQGEEYLF RVIAVNEKGR SDPRSLAVPI IAKDLVIEPD  25020
VRPAFSSYSV QVGQDLKIEV PISGRPKPSI SWTKDGMPLK QTTRINVTDS LDLTTLSIKE  25080
THKDDGGQYG ITVSNVVGQK TAAIEIITLD KPDPPKGPVK FDEISAESIT LSWNPPLYTG  25140
GCQITNYIVQ KRDTTTTVWD VVSATVARTT LKVTKLKTGT EYQFRIFAEN RYGQSFALES  25200
EPVVAQYPYK EPGPPGTPFV TAISKESMVV QWHEPINNGG SPVIGYHLER KERNSILWTK  25260
VNKTIIHDTQ FKALNLEEGI EYEFRVYAEN IVGVGKASKN SECYVARDPC DPPGTPEAII  25320
VKRNEITLQW TKPVYDGGSM ITGYIVEKRD LPEGRWMKAS FTNVIETQFT VSGLTEDQRY  25380
EFRVIAKNAA GAISKPSDST GPITAKDEVE LPRISMDPKF RDTIVVNAGE TFRLEADVHG  25440
KPLPTIEWLR GDKEIEESAR CEIKNTDFKA LLIVKDAIRI DGGQYILRAS NVAGSKSFPV  25500
NVKVLDRPGP PEGPVQVTGV TAEKCTLAWS PPLQDGGSDI SHYVVEKRET SRLAWTVVAS  25560
EVVTNSLKVT KLLEGNKYIF RIMAVNKYGV GEPLESAPVL MKNPFVLPGP PKSLEVTNIA  25620
KDSMTVCWNR PDSDGGSEII GYIVEKRDRS GIRWIKCNKR RITDLRLRVT GLTEDHEYEF  25680
RVSAENAAGV GEPSPATVYY KACDPVFKPG PPTNAHVVDT TKNSITLAWS KPIYDGGSEI  25740
LGYVVEICKA DEEEWQIVTP QTGLRVTRFE IAKLIEHQEY KIRVCALNKV GLGEAASVPG  25800
TVKPEDKLEA PELDLDSELR KGIVVRAGGS ARIHIPFKGR PTPEITWSKE EGEFTDKVQI  25860
EKGINFTQLS IDNCDRNDAG KYILKLENSS GSKSAFVTVK VLDTPGPPQN LAVKEVRKDS  25920
VLLVWEPPII DGGAKVKNYV IDKRESTRKA YANVSSKCNK TSFRVENLTE GAIYYFRVMA  25980
ENEFGVGVPT ETSDAVKASE PPSPPGKVTL TDVSQTSASL MWEKPEHDGG SRILGYVVEM  26040
QPKGTEKWSV VAESKVCNAV VSGLSSGQEY QFRVKAYNEK GKSDPRVLGI PVIAKDLTIQ  26100
PSFKLPFNTY SVQAGEDLKI EIPVIGRPRP KISWVKDGEP LKQTTRVNVE ETATSTILHI  26160
KESSKDDFGK YSVTATNSAG TATENLSVIV LEKPGPPVGP VKFDEVSADF VVISWEPPAY  26220
TGGCQISNYI VEKRDTTTTN WQMVSATVAR TTIKISKLKT GTEYQFRIFA ENRYGKSTPL  26280
DSKPVVVQYP FKEPGPPGTP FVTSISKDQM LVQWHEPVND GGSKVTGYHL EQKEKNSILW  26340
VKLNKIPIQD TKFKTTGLDE GLEYEFRVSA ENIVGIGKPS KVSECYVARD PCDPPGRPEA  26400
IVITRNSVTL KWKKPVYDGG SKITGYIVEK KDLPDGRWMK ASFTNVVETE FTVTGLVEDQ  26460
RYEFRVIARN AADNFSEPSE SSGAITARDE IDAPNASLDP KYRDVIIVHA GETFVLEADI  26520
RGKPIPDIIW SKDGNELEET AARMEIKSTL QKTTLIVKDC IRTDGGQYTL KLSNVGGTKT  26580
IPITVKVLDR PGPPEGPLKV TGVTAEKCYL AWNPPLQDGG ASISHYIIEK RETSRLSWTQ  26640
VSNEVQALNY KVTKLLPGNE YIFRVMAVNK YGIGEALESE PVIACNPYKR PGPPSTPEAS  26700
AITKDSMVLT WTRPVDDGGA EIEGYILEKR DKEGIRWTKC NKKTLTDLRF RVTGLTEGHS  26760
YEFRVAAENA AGVGEPSEPS VFYRACDALY PPGPPSNPKV TDTSRSSVSL AWNKPIYDGG  26820
APVRGYVIEL KKAAADEWTT CTPPSGLQGK QFTVTKLKEN TEYNFRICAF NTEGVGEPAT  26880
IPGSVVAQER MEAPEIELDA DLRKVVTLRA SATLRLFVTI KGRPEPEVKW EKAEGILTER  26940
AQIEVTSSYT MLVIDNVTRF DSGRYNLTLE NNSGSKTAFV NVRVLDSPSA PVNLTIREVK  27000
KDSVTLSWEP PLIDGGAKIT NYIVEKRETT RKAYATITNN CTKNTFKIEN LQEGCSYYFR  27060
VLASNEYGIG LPAETAEPVK VSEPPLPPGR VTLVDVTRNT ATIKWEKPES DGGSKITGYV  27120
VEMQTKGSEK WSACTQVKTL ETTISGLTAG EEYVFRVAAV NEKGRSDPRQ LGVPVIAKDI  27180
EIKPSVELPF NTFNVKANDQ LKIDIPFKGR PQATVAWKKD GQVLRETTRV NVASSKTVTT  27240
LSIKEASRED VGTYELCVSN TAGSITVPIT VIVLDRPGPP GPIRIDEVSC DNVSISWNPP  27300
EYDGGCQISN YIVEKRETTS TTWQVVSQAV ARTSIKIVRL TTGSEYQFRV CAENRYGKSS  27360
YSESSAVVAE YPFSPPGPPG TPKVVHATKS TMVVSWQVPV NDGGSQVIGY HLEYKERSSI  27420
LWSKANKVLI ADTQMKVSGL DEGLMYEYRV YAENIAGIGK CSKACEPVPA RDPCDPPGQP  27480
EVTNITRKSV SLKWSKPRYD GGAKITGYIV ERRELPDGRW LKCNFTNVQE TYFEVTELTE  27540
DQRYEFRVFA RNAADSVSEP SESTGPITVK DDVEAPRIMM DVKFRDVIIV KAGEVLKINA  27600
DIAGRPLPVI SWAKDGVEIE ERAKTEIVST DYTTTLTVKD CVRRDTGQYV LTLKNVAGTR  27660
TMAVNCKVLD KPGPPAGPLE INGLTAEKCS LSWGRPQEDG GADIDYYIVE KRETSRLAWT  27720
ICEAELRTTS CKVTKLLKGN EYIFRVTGVN KYGVGEPLES MAVKALDPFT TPSPPTSLEI  27780
TSVTKDSMTL CWSRPETDGG SDISGYIIER REKNSLRWMR VNKKPVYDLR VKSTGLREGC  27840
EYEYRVFAEN AAGLSLPSET SPLVRAEDPV FLPSPPSKPK IVDSGKTTIT IGWVKPLFDG  27900
GAPITGYTVE YKKSEETDWK VAIQSFRGTE YTMSGLTTGD EYVFRVRSLN KMGASDPSDS  27960
```

-continued

```
SDPQVAKERE EEPVFDVDSE MRKTLIVKAG SSFTMTVPFR GRPIPNVSWS KPDTDLRTRA  28020
YIDSTDSRTL LTIENANRND SGKYTLTIQN VLSAASMTFV VKVLDSPGPP ANITVREVTK  28080
ETAMLSWDVP ENDGGAPVKN YHIEKREASK KAWVSVTNNC NRLSYKVTNL QEGAIYYFRV  28140
SGENEFGVGV PAETKEGVKI TEKPSPPEKL GVTSVSKDSV SLSWLKPEHD GGSRIIHYVV  28200
EALEKGQKTW VKCAVVKTTH HVVSGLRESH EYFFRVFAEN QAGLSDPREL LLPVLIKDQL  28260
EPPEIDMKNF PSHTVYVRAG SNLKVDIPIS GKPLPKVTLS RDGVPLKATM RFNTEITAEN  28320
LTINLKESVT TDAGRYEITA ANSSGTTKTF INIIVLDRPG PPTGPVAISD ITEESVTLKW  28380
EPPKYDGGSH VTNYIVLKRE TSTAVWSEVS ATVARTMIKV MKLTTGEEYQ FRIKAENRFG  28440
ISDHIDSVCV VVKLPYTTPG PPSTPWVSNV TRESITVGWH EPVSNGGSAV TGYHLEMKDR  28500
NSILWQKANK MIIRTTHFKV TTISAGLIYE FRVYAENAAG IGKPSHPSEP VLAIDACEPP  28560
RNVRITDISK NSVNLSWQQP AFDGGSKITG YIVERRDLPD GRWTKASFTN VIETQFTVSG  28620
LTQNSQYEFR VFARNAVGSV SNPSEVVGPI TCIDSYGGPV IDLPLEYTEV VKYRAGTSVK  28680
LRAGISGKPE PTIEWYKDDK ELQTNALVCV ENSTDLASIL IKDANRLNSG SYELKLRNAM  28740
GSASATIRVQ ILDKPGPPGG PIEFKTVTAE KITLLWRPPA DDGGAKITHY IVEKRETSRV  28800
VWSMVAENLE ECIVTTTKII KGNEYVFRVR AVNKYGIGEP LESEPVVAKN AFVTPGPPSI  28860
PEVTKITKNS MTVVWDRPTV DGGSEINGYF LERRDKKSLA WLKVLKETIR DTRQKVTGLT  28920
ENSDYQYRVC AVNAAGVGPF SEPSDFYKAA DPIDPPGPPA KIRIADSTKS SITLGWSKPV  28980
YDGGSDVTGY VVEMKQGDEE EWTIVSTRGE VRTTEYVVSN LKPGVNYYFQ VSAVNCAGQG  29040
EPITMTEPAQ AKDVLEEPEI DLDVALRTSV IAKAGEDVQL LIPFKGRPPP TVTWRKDEKN  29100
LGSDTRYSIQ NTDSSSLLVI PQVTRNDTGK YILTIENGVG QPKSSTVSVK VLDTPAACQK  29160
LQVKHVSLGT VTLLWDPPLI DGGSPIINYV IEKRDATKRT WSVVSHKCSG TSFKVTDLSE  29220
KTPFFFRVLA ENEIGIGEPC ETTEPVKAAE VPAPIRDLSM KDSTKTSVVL SWTKPDFDGG  29280
SIITDYLVER KGKGEQAWSH AGISKTCEIE IGQLKEQSVL EFRVSARNEK GQSDPVTIGP  29340
LTVKELVITP EVDLSEIPGA QISVRIGHNV HLELPYKGKP KPSISWLKDG LPLKESEYVR  29400
FSKTENKITL SIKNSKKEHG GKYTVILDNA VCRNSFPITI ITLGPPSKPK GPIRFDEIKA  29460
DSAIMSWDIP EDDGGGEITC YSIEKREASQ TNWKMVCSSV ARTTFKVSNL VKDSEYQFRV  29520
RAENRYGVSE PLASNIIVAK HQFRIPGPPG KPVIYNVTSD GMSLTWDAPV YDGGSEVTGF  29580
HVEKKERNSI LWQRVNTSPI SGREYRATGL IEGLDYQFRV YAENSAGLSS PSDPSKFTLA  29640
VSPVDPPGTP DYIDVTRETI TLKWNPPLRD GGSKIVAYSI EKRQGSDRWV RCNFTDVSEC  29700
QYTVTGLSPG DRYEFRIIAR NAVGTISPPS QSSGLIMTRD ENVPPTVEFG PEYFDGLVIK  29760
SGDSLRIKAL VQGRPVPRVT WFKDGVEIER RMNMEITDVL GSTSLFVRDA TRDHRGVYTV  29820
EAKNVSGSTK AEVTVKVQDT PGKVVGPIRF TNITGEKMTL WWEAPLNDGC APVTHYIIEK  29880
RETSRLAWAL IEDNCEALSY TAIKLITGNE YQFRISAVNK FGVGRPLESD PVVAQIQYTI  29940
PDAPGVPEPS NVTGNSITLT WTRPESDGGS EIQHYILERR EKKSTRWVKV ISKRPISETR  30000
FKVTGLVEGN EYEFHVMAEN AAGIGPASGI SRLIKCREPV NPPSAPSVVK VTDTSKTTVS  30060
LEWARPVFDG GMEIIGYIIE MCKADLGDWH KVNTEPCVKT RYTVTDLQAG EEYKFRVSAV  30120
NGAGKGDSCE VTGTIKAVDR LSAPELDIDA NFKQTHIVRA GVSIRLFIAY QGRPTPTAVW  30180
SKPDSNLSIR ADIHTTDSFS TLTVENCNRN DAGKYTLTVE NNSGKKSITF TVKVLDSPGP  30240
PGPITFKDVT RGSATLMWDA PLLDGGARIH HYVIEKREAS RRSWQVVSEK CTRQILKVSE  30300
LTEGVPYYFR VSAENEYGVG EPYEMPEPIV ATEQPAPPRR LDVVDTSKSS AVLAWLKPDH  30360
DGGSRITSYL LEMRQKGSDF WVEAGHTKQL TFTVERLVEN TEYEFRVKAK NDAGYSEPRE  30420
AFSSVIIKEP QIEPTADLTG ITNQLITCKA GSTFTIDVPI SGRPAPKVTW KLEEMRLKET  30480
DRMSIATTKD RTTLTVKDSM RGDSGRYFLT LENTAGVKTF TITVVVIGRP GPVTGPIEVS  30540
SVSAESCVLS WTEPKDDGGT EITNYIVEKR ESGTTAWQLI NSSVKRTQIK VTHLTKYKEY  30600
CFRVSSENRF GVSKPLESVP IVAEHPFVPP SAPTRPEVYY VSANAMSIRW EEPYHDGGSK  30660
IVGYWVEKKE RNTILWVKEN KVPCLECNYK VTGLVEGLEY QFRTYALNAA GVSKASEASR  30720
PIMAQNPVDP PGRPEVTDVT RSTVSLIWSA PVYDGGSKVV GYIIERKPVS EVGDGRWLKC  30780
NYTIVSDNFF TVTALSEGDT YEFRVLAKNA AGVISKGSES TGPVTCRDEY APPKAELDAR  30840
LQGDLVTIRA GSDLVLDAAV GGKPEPKIIW TKGDKELDLC EKISLQYTGK RATAVIKYCD  30900
RSDSGKYTLT VKNASGTKSV SVMVKVLDSP GPCGKLTVSR VTEEKCTLAW SLPQEDGGAE  30960
ITHYIVERRE TSRLNWVIVE GECLTASYVV TRLIKNNEYT FRVRAVNKYG LGVPVESEPI  31020
VARNSFTIPS QPGIPEEVGA GKEHIIIQWT KPESDGGNEI SNYLVDKREK KSLRWTRVNK  31080
DYVVYDTRLK VTSLMEGCDY QFRVTAVNAA GNSEPSEASN FISCREPSYT PGPPSAPRVV  31140
DTTKRSISLA WTKPMYDGGT DIIGYVLEMQ EKDTDQWCRV HTNATIRNNE FTVPDLKMGQ  31200
KYSFRVAAVN AKGMSDYSET TAEIEPVERL EIPDLELADD LKKTVIVRAG ASLRLMVSVS  31260
GRPSPVITWS KKGIDLANRA IIDNTESYSL LIVDKVNRYD AGKYTIEAEN QSGKKSATVL  31320
VKVYDTPGPC PSVSVKEVSR DSVTITWEIP TIDGGAPVNN YIIEKREAAM RAFKTVTTKC  31380
SKTLYRISGL VEGTMYYFRV LPENIYGIGE PCETSDAVLV SEVPLVPTKL EVVDVTKSTV  31440
TLAWEKPLYD GGSRLTGYVL EACKAGTERW MKVVTLKPTV LEHTVISLNE GEQYLFRVRA  31500
QNEKGVSEPR EIVTPVTVQD LRVLPTIDLS TMPQKTIHVP AGRPIELVIP ITGRPPPTAS  31560
WFFAGSKLRE SERVTVETHT KVTKLTIRET TIRDTGEYTL ELKNVTGTTS ETIKVIILDK  31620
PGPPTGPIKI DEIDATSVTI SWEPPELDGG APLSGYVVEQ RDAHRPGWLP VSESVTRPTF  31680
KFTRLTEGNE YVFRVAATNR FGIGSYLQSE VIECRSSISI PGPPETLQIF DVSRDGMTLT  31740
WYPPEDDGGS QVTGYIVERK EVRADRWVRV NKVPVTMTRY RSTGLIEGLE YEHRVTAINA  31800
RGTGKPSRPS KPTVAMDPIA PPGKPQNPRV TDTTRTSVSL AWSVPEDEGG SKVTGYLIEM  31860
QKVDQREWTK CNTTPTKIRE YTLTHLPQGA EYRFRVLACN AGGPGEPAEV PGTVKVTEML  31920
EYPDYELDER YQEGVFVRQG GVIRLTIPIK GKPFPVCKWT KEGQDISKRA MIATSETHTE  31980
LVIKEADRND SGTYDLVLEN KCGKKTVYIK VKVIGSPNTP EGPLEYDDIQ ARSVRVSWRP  32040
PADDGGADIL GYILERREVP KAAWYTIDSR VRGTSLVVKG LKENVEYHFR VSAENQFGIS  32100
KPLKSEEPVI PKTPLNPPEP PSNPPEVLDV TKSSVSLSWS RPKDDGGSRV TGYYIERKET  32160
STDKWVRHNK TQITTTMYTV TGLVPDAEYQ FRIIAQNDVG LSETSPASEP VVCKDPFDKP  32220
SQPGELEILS ISKDSVTLQW EKPECDGGKE ILGYWVEYRQ SGDSAWKKSN KERIKDRQFT  32280
IGGLLEATEY EFRVFAENET GLSRPRRTAM SVKTKLTSGE APGVRKEMAD VTTKLGEAAQ  32340
LSCQIVGRPL PDIKWYRFGK ELIQSRKYKM SSDGRTHTLT VMTDEQEDEG VYTCVATNEV  32400
GEVESSSKLL LQAAPQFHPG YPLKEKYYGA VGSTLRLHVM YIGRPVPAMT WFHGQKLLQN  32460
SEKITIENTE HYTHLVMKNV QRKTHAGKYK VQLSNAFGTV DATLDVEIQD KPDKPTGPIV  32520
```

-continued

```
IEALLKNSVV ISWKPPADDG GSWITNYVVE KCEAKEGAEW QLVSSAISVT TCRIVNLTEN  32580
AGYYFRVSAQ NTFGISEPLE VASIVIIKSP FEKPGVPGKP TITAVTKDSC VVAWKPPASD  32640
GGAKIRNYYL ERREKKQNKW IAVTTEEIRE TVFSVQNLIE GLEYEFRVKC ENLGGESEWS  32700
EISEPVTPKS DVPIQAPHFK EELRNLNVRY QSNATLVCKV TGHPKPIVKW YRQGKEIIAD  32760
GLKYRIQEFK GGYHQLIIAS VTDDDATVYQ VRATNQGGSV SGTASLEVEV PAKIHLPKTL  32820
EGMGAVHALR GEVVSIKIPF SGKPDPVITW QKGQDLIDNN GHYQVIVTRS FTSLVFSNGV  32880
ERKDAGFYVV CAKNRFGIDQ KTVELDVADV PDPPRGVKVS DVSRDSVNLT WTEPASDGGS  32940
KVTNYIVEKC ATTAERWLRV GQARETRYTV INLFGKTSYQ FRVIAENKFG LSKPSEPSEP  33000
TVTKEDKTRA MNYDDEVDET REVTTTKASH SKTKELYEKY MIAEDLGRGE FGIVHRCVET  33060
SSKRTFMAKF VKVKGTDQVL VKKEISILNI ARHRNILYLH ESFESMEELV MIFEFISGLD  33120
IFERINTSAF ELNEREIVSY VRQVCEALEF LHSQNIGHFD IRPENIIYQT RKNSTIKIIE  33180
FGQARQLKPG DNFRLLFTAP EYYAPEVHQH DVVSTATDMW SLGTLVYVLL SGINPFLAET  33240
NQQMIENIMN AEYTFDEEAF KEISLEAMDF VDRLLVKERK SRMTASEALQ HPWLKQRIDR  33300
VSTKVIRTLK HRRYYHTLIK KDLNMVVSAA RISCGGAIRS QRGVSVAKVK VASIEIGPVS  33360
GQIMHAIGEE GGYVKYVCKI ENYDQSTQVT WYFGVRQLEN SEKYEITYED GVATMYVKDI  33420
TKFDDGTYRC KVVNDYGEDS SYAELFVKGV REVYDYYCRR TKKVKRRTDA MRLLERPPEF  33480
TLPLYNKTAY VGENVRFGVT ITVHPEPRVT WYKSGQKIKP GDDEKKYTFE SDKGLYQLTI  33540
NSVTTDDDAE YTVVARNKHG EDSCKAKLTV TLHPPPTETT LRPMFKRLLA NAECHEGQSV  33600
CFEIRVSGIP APTLKWEKDG QPLSLGPHIE IVHEGLDYYA LHIRDTLPED TGYYRVTATN  33660
TAGSTSCQAH LQVERLRYVK QEFQSKEERE RHVQKQIDKT LRMAEILSGT ETVPLTPVAQ  33720
EALREAAILY KPAVSTKTVK GEYRLQTEEK KEERKLRMPY EVPEPRRFKQ ATVEEDQRIK  33780
QFVPMSDMKW YKKIRDQYEM PGKLDRVVQK RPKRIRLSRW EQFYVMPLPR ITDQYRPKWR  33840
IPKLTQDDLE MVRPARRRTP SPDYDLYYYR RRRRSLGDMS DEELLLPIDD YLAMKRTEEE  33900
RLRLEEELEL GFSASPPSRS PPRFELSSLR YSSPPAHVKV EDRRRDFRYS TYHVPTKEET  33960
STSYAELRER HAQASYRQPK LRQRIMAEKE EEELLRPVTT TQRLSEYKSE LDYMSKEEKS  34020
KKKSKRQRQV TEITEIEEEY EISRRAQRES SSSVSRLLRR RRSLSPTYIE LMRPVSELIR  34080
SHPRPAEEYE DDAERRSPTP ERTRPRSPSP VSSERSLSRF ERSARFDIFS RYESMKAALK  34140
TQKTSERKYE VLSQQPFTLD HAPRITLRMR SHRVPCGQNT RFILNVQSKP TAEVKWYHNG  34200
VELQESSKIH YTNTSGVLTL EILDCQTEDG GTYRAVCTNY KGEASDYATL DVTGGAYTTY  34260
ASQRRDEEVP KSVFPELTKT EAYAVSSFKR TSELEAASSV REVKSQMTET RESLSTYEHY  34320
ASAEMKSATS EEKSLEEKAT VRKIKTTLAA RILTKPRSIT VHEGESARFS CDTDGEPVPT  34380
VTWLREGQVV STSARHQVTT TKYKSTFEIS SVQASDEGNY SVVVENSDGK QEAQFTLTVQ  34440
KARVIEKAVT SPPRVKSPEP RVKSPETVKS PKRVKSPEPV TSHPKAVSPT ETKPTEKGQH  34500
LPVSAPPKIT QSLKAEASKD IAKLTCAVES SALCAKEVAW YKDGKKLKEN GHFQFHYSAD  34560
GTYELKIHNL SESDCGEYVC EVSGEGGTSK TSFQFTGQSF KSIHEQVSSI SETTKSVQKT  34620
AESPEAKKQE PIAPESISSK PVIVTGLRDT TVSSDSVAKF TIKVTGEPQP TITWTKDGKA  34680
IAQGSKYKLS SNKEEFILEI LKTETSDGGL YACTVTNSAG SVSSSCKLTI KAVKDTEAQK  34740
VSTQKTSEVT SQKKASAQEE ISQKALTSEE IKMSEVKSHE TLAIKEEASK VLIAEEVKRS  34800
AAASLEKSIV HEEVTKTSQA SEEVKTHAEI KTLSTQMNIT KGQRATLKAN IAGATDVKWV  34860
LNGTELPNSE EYRYGVSGSD QTLTIKQASH REEGILSCIG KTSQGVVKCQ FDLTLSEELS  34920
DAPSFITQPR SQNINEGQNV LFSCEVSGEP SPEIEWFKNN LPISISSNIS VSRSRNVYTL  34980
EIRNAAVSDS GKYTIKAKNF HGQCSATASL TVLPLVEEPP REVVLKTSSD VSLHGSVSSQ  35040
SVQMSASKQE ASFSSFSSSS ASSMTEMKFA SMSAQSMSSM QESFVEMSSS SFMGKSSMTQ  35100
LESSTSRMLK AGGRGIPPKI EALPSDISID EGKVLTVACA FTGEPTPEIT WSCGGRKIQN  35160
QEQQGRFHIE NTDDLTTLII MDVQKQDGGL YTLSLGNEFG SDSATVNINI RSM         35213
```

```
SEQ ID NO: 2            moltype = AA  length = 1939
FEATURE                 Location/Qualifiers
source                  1..1939
                        mol_type = protein
                        organism = Mus musculus
SEQUENCE: 2
MSSDAEMAVF GEAAPYLRKS EKERIEAQNK PFDAKSSVFV VDAKESYVKA TVQSREGGKV  60
TAKTEGGATV TVKDDQVFSM NPPKYDKIED MAMMTHLHEP AVLYNLKERY AAWMIYTYSG  120
LFCVTVNPYK WLPVYNPEVV AAYRGKKRQE APPHIFSISD NAYQFMLTDR ENQSILITGE  180
SGAGKTVNTK RVIQYFATIA VTGDKKKEEA TSGKMQGTLE DQIISANPLL EAFGNAKTVR  240
NDNSSRFGKF IRIHFGATGK LASADIETYL LEKSRVTFQL KAERSYHIFY QIMSNKKPEL  300
IEMLLITTNP YDFAYVSQGE ITVPSIDDQE ELMATDTAVD ILGFSADEKV AIYKLTGAVM  360
HYGNMKFKQK QREEQAEPDG TEVADKAAYL TSLNSADLLK ALCYPRVKVG NEYVTKGQTV  420
QQVYNSVGAL AKSMYEKMFL WMVTRINQQL DTKQPRQYFI GVLDIAGFEI FDFNTLEQLC  480
INFTNEKLQQ FFNHHMFVLE QEEYKKEGID WEFIDFGMDL AACIELIEKP MGIFSILEEE  540
CMFPKATDTS FKNKLYEQHL GKSNNFQKPK PAKGKAEAHF SLVHYAGTVD YNIIGWLDKN  600
KDPLNETVVG LYQKSGLKTL AFLFSGGQAA EAEGGGGKKG GKKKGSSFQT VSALFRENLN  660
KLMTNLKSTH PHFVRCLIPN ETKTPGAMEH ELVLHQLRCN GVLEGIRICR KGFPSRILYA  720
DFKQRYKVLN ASAIPEGQFI DSKKASEKLL GSIDIDHTQY KFGHTKVFFK AGLLGTLEEM  780
RDEKLAQLIT RTQAVCRGYL MRVEFKKMME RRESIFCIQY NVRAFMNVKH WPWMKLYFKI  840
KPLLKSAETE KEMANMKEDF EKAKEDLAKS EAKRKELEEK MVALMQEKND LQLQVQAEAD  900
GLADAEERCD QLIKTKIQLE AKIKELTERA EDEEEINAEL TAKKRKLEDE CSELKKDIDD  960
LELTLAKVEK EKHATENKVK NLTEEMAGLD ENIAKLTKEK KALQEAHQQT LDDLQAEEDK  1020
VNTLTKAKTK LEQQVDDLEG SLEQEKKLRM DLERAKRKLE GDLKLAQEST MDIENDKQQL  1080
DEKLKKKEFE MSNLQSKIED EQALGMQLQK KIKELQARIE ELEEEIEAER ASRAKAEKQR  1140
SDLSRELEEI SERLEEAGGA TSAQIEMNKK REAEFQKMRR DLEEATLQHE ATAAALRKKH  1200
ADSVAELGEQ IDNLQRVKQK LEKEKSELKM EIDDLASNME TVSKAKGNLE KMCRTLEDQL  1260
SEVKTKEEEQ QRLINELSTQ KARLHTESGE FSRQLDEKDA MVSQLSRGKQ AFTQQIEELK  1320
RQLEEESKAK NALAHALQSA RHDCDLLREQ YEEEQEAKAE LQRAMSKANS EVAQWRTKYE  1380
TDAIQRTEEL EEAKKKLAQR LQDAEEHVEA VNSKCASLEK TKQRLQNEVE DLMIDVERSN  1440
```

-continued

```
AACAALDKKQ RNFDKVLAEW KQKYEETQAE LEASQKESRS LSTELFKVKN AYEESLDQLE  1500
TLKRENKNLQ QEISDLTEQI AEGGKHIHEL EKIKKQIDQE KSELQASLEE AEASLEHEEG  1560
KILRIQLELN QVKSEIDRKI AEKDEEIDQL KRNHLRVVES MQSTLDAEIR SRNDALRIKK  1620
KMEGDLNEME IQLNHANRQA AEAIRNLRNT QGMLKDTQLH LDDALRGQDD LKEQLAMVER  1680
RANLMQAEIE ELRASLEQTE RSRRVAEQEL LDASERVQLL HTQNTSLINT KKKLETDISQ  1740
IQGEMEDIVQ EARNAEEKAK KAITDAAMMA EELKKEQDTS AHLERMKKNM EQTVKDLQHR  1800
LDEAEQLALK GGKKQIQKLE ARVRELENEV ENEQKRNIEA VKGLRKHERR VKELTYQTEE  1860
DRKNVLRLQD LVDKLQTKVK AYKRQAEEAE EQSNVNLAKF RKIQHELEEA EERADIAESQ  1920
VNKLRVKSRE VHTKVISEE                                               1939

SEQ ID NO: 3            moltype = AA  length = 7152
FEATURE                 Location/Qualifiers
source                  1..7152
                        mol_type = protein
                        organism = Mus musculus
SEQUENCE: 3
MADEEEYEEV IEYYTEETVY EEVPGETITE VYETTTTRTS DYGQSETATP ALAQPQPAKP  60
VERKKVIRKK VDSSKFMTPY IQHSQKMQDL FSSNKYKENY EKAKGQPYAI TSDTPELRRI  120
KKVQDQLSEV KYRMDGDVAK TICHVDEKAK DIEHAKKVSQ QVSKVLYKQN WEDTKDKYLL  180
PPDAPELVQA VKNTAMFSKK LYTEDWEADK ALFYPYNDSP ELRRVAQAQK ALSDIAYKKG  240
LTEQQTQFTS LPDPPEVEFA KKVTNQVSKQ KYKEDYEKKV KGKWSETPCF EIATARMNAD  300
NLSSRKYQED FENMKDQIYF MQTETPEYKM NKQAGVAASK VKYKQDYEKN KGKADYNVLP  360
ASENPLLRQL KTAGNVLSDK LYKENYEKTK ARSINYCETP KFQLDTVLQN FSSDTKYKDS  420
YLKNILGHYV GSFEDPYHTH CMKVTAQNSD KNYKAEYEED RGKGFFPQTI TQEYEAIKKL  480
DQCKDHAYKV HPDKTKFTQV TDSPVLVQAQ VNSKQLSDLN YKAKHESEKF KCHIPPDAPA  540
FIQHRVNAYN LSDNVYKQDW EKSKAKKFDI KVDAIPLLAA KANSKNASDV MYKKDYEKSK  600
GKMIGALSIN DDPKMLHSLK TAKNQSDRLY RENYEKTKAK SMNYCETPKY QLDTLLKNFS  660
EAKYKDSYVK NMLGHYVGSF EDPYHTHCMK VSAQNSDKSY KAEYEEDKGK CYFPQTITQE  720
YESIKKLDQC KDHTYKVHPD KTRFTAVTDT PVLLQAQLNT KQLSDLNYKA KHEGEKFKCS  780
VPADAPEFIQ HRVNAYNLSD NVYKHDWEKS KAKKFDIKVD AIPLLAAKAN SKNASDVMYK  840
KDYEKSKGKM IGALSINDDP KMLHSLKTAK NQSDREYRKD YEKSKTIYTA PLDMIQVTQA  900
KKSQAIASDV DYKHVLHNYS YPPDSINVDL AKKAYALQSD VEYKADYNSW MKGCGWVPFG  960
SLEMEKAKRA SEILNEKKYR QHPDTLKFTS IEDAPITVQS KINQAQRSDI AYKAKGEEIL  1020
HKYNLPADLP QFIQAKVNAY NISENMYKAD LKDLSKKGYD LRIDAIPIKA AKAARQAASD  1080
IHYKKDYEKA KGKMVGFQSL QDDPKLVHYM NVAKIQSDRE YKKDYEKTKT KYNTPHDMFD  1140
VVAAKKAQDV ISNTNYKHSL HHYTYLPDAM DLELSKNMMH IQSDNAYKED YNTWMKGIGW  1200
IPIGSLDQEK VKKAGDALSE KKYRQHPDTL KFTSIVDSPV MVQAKQNTQQ ASDILYKAKG  1260
EDVKHKYTLN PDLPQFLQAK CNAYNISDVC YKRDWHDLIA KGNNVLADAI PITAAKSSRN  1320
IASDYKYKEA YEKAKGKHVG FRSLQDDPKL VHYMNVAKLQ SDREYKKNYE NTKTSYHTPG  1380
DMVSITAARM AQDVATNVNY KQPLHHYTYL PDALSLEHTR NANQIQSDNV YKDEYNSFFK  1440
GIGWIPIGSL EVEKAKKAGE ALNERKYRQH PDTIKFTSVP DSMGMVLAQQ NTKQLSDLNY  1500
KVEGEKMKHK YTMDPELPQF IQAKVNAINM SDAHYKADWK KTLAKGYDLR PDAIPIVAAK  1560
SSRNIASDCK YKEAYEKARG KQIGFLSLQD DPKLVHYMNV AKIQSDREYK KGYEASKTRY  1620
HTPLDMFSVT AAKKSQEVAT NSNYRQPYHN YTLLPDALNV EHSRNAMQIQ SDNLYKSDFT  1680
HWMKGIGWVP IDSLEVEKAK KAGEILSERK YRQHPEKLKF TYSMDTMEQA LNKSNKLTMD  1740
KRLYTEKWNK DKTSIHVMPD TPDILLSRVN QITMSDKLYK AGWEEEKTKG YDLKPDAIAI  1800
KAAKASRDIA SDYKYKQAYE QSKGKHIGFR SLEDDPKLVH FMQVAKIQSD REYKKGYEKS  1860
KTSFHTPVDM FSVVAAKKSQ EVATNANYRN IIHTYNMLPD AMSFELAKNM MQIQSDNQYK  1920
ADYADFMKGI GWLPLGSLEA EKNKKAMEII SEKKYRQHPD TLKYSTLMDS MNMVLAQNNA  1980
KIMNKNLYKQ AWEADKTKIH IMPDIPEIVL AKANAINMSD KLYKLSLEES RKKGYDLRTD  2040
AIPIKAAKAS RDIASDYKYK HSYEKEKGKM VGFRSLEDDP KLVHSMQVAK MQSDREYKKN  2100
YEKTKTSYHT PADMLSVTAA KDAQANITNT NYKRLIHKYI LLPDAMNIEL TRNMNHIQSD  2160
NEYKQDYREW YRGLGWSPAG SLEVEKAKKA TEYASDQKYR QHPSTFQFKK LTDSMDMVLA  2220
KQNAHTMNKH LYTIDWNKDK TKIHVMPDTP DILQAKQNQT LYSQKSYRLG WEEALKKGYD  2280
LPVDAISVQL AKTSRDIASD FKYKQGYRKQ LGHHIGFLSV QDDPKLVLSM NVAKMQSDRE  2340
YKKDFEKWKT KYTSPVDMLG VVLAKKCQAL VSDADYRNYL HQWTCLPDQN DVIQAKKVYE  2400
LQSENMYKSD LEWLRGIGWS PLGSLEAEKN KRASEIISEK KYRQPPDRNK FTSIPDAMDI  2460
VLAKTNAKNR SDRLYREAWD KDKTQIHIMP DTPDIILAKA NLINTSDKLY KMGYEELKKK  2520
GYDLPVDAIP IKAAKASREI ASEYKYKEGF RKQLGHHIGA RNIKDDPKMM WSMHVAKIQS  2580
DREYKKDFEK WKTKYSSPVD MLGVVLAKKC QTLVSDADYR NYLHQWTCLP DQNDVIHARQ  2640
AYDLQSDNVY KADLQWLKGI GWIPSGSLED EKNKRATQIL SDHVYRQHPN KFKFSSLMDS  2700
MPMVLAKNNA ITMNHHLYTE AWNKDKTTVH IMPDTPEVIL AKQNKVNYSE KLYKLGLEEA  2760
KRKGYDMRLD AIPIKTAKAS RDIASDFKYK EGYRKQLGHH IGARGIHDDP KMMWSMHVAK  2820
IQSDREYKKD FEKWKTKYSS PVDMLGVVLA KKCQTLVSDA DYRNYLHQWT CLPDQNDVIH  2880
ARQAYDLQSD NMYKSDLQWM RGIGWVPIGS LDVEKCKRAT EILSDKIYRQ PPDKFKFTSV  2940
TDSLEQVLAK NNAITMNKRL YTEAWDKDKT QIHIMPDTPE IMLARMNKVN YSESLYKLAN  3000
EEAKKKGYDL RSDAIPIVAA KASRDIASDY KYKDGYRKQL GHHIGARNIK DDPKMMWSMH  3060
VAKIQSDREY KKDFEKWKTK YSSPVDMLGV VLAKKCQILV SDIDYKHPLH EWTCLPDQND  3120
VIHARQAYDL QSDNVYKSDL QWMRGIGWVP IGSLDVVKCK RAAEILSDNI YRQPPNKFKF  3180
TSVTDSLEQV LAKSNALNMN KRLYTEAWDK DKIQIHVMPD TPEIMLARQN RINYSESLYK  3240
LANEEAKKKG YDLRVDAIPI VAAKASRDIA SDYKYKVGYR KQLGHHIGAR NIEDDPKMMW  3300
SMHVAKIQSD REYKKDFEKW KTKYSSPVDM LGVVLAKKCQ TLVSDVDYRN YLHQWTCLPD  3360
QNDVIHARQA YDLQSDNVYK SDLQWMRGIG WVPIGSLDVV KCKRAAEILS DNIYRQPPNK  3420
FKFTSVTDSL EQVLAKSNAL NMNKRLYTEA WDKDKTQIHI MPDTPEIMLA RQNKINYSES  3480
LYRQAMEEAK KEGYDLRSDA IPIVAAKASR DIASDYKYKE AYRKQLGHHI GARAIHDDPK  3540
MMWSLHIAKV QSDREYKKEF EKYKTRYSSP VDMLGIVLAK KCQTLVSDVD YKHPLHEWTC  3600
```

-continued

```
LPDQNDVIQA RKAYDLQSDN LYKADLEWMK GIGWVPIGSL EVVKAKRATE ILSDNIYRQR  3660
PDTLKFTSIT DSLEQVLAKN NAITMNKRLY TEAWDNDKKT IHVMPDTPEI MLAKLNRINY  3720
SDKLYKVSLE EARRQGYDLR VDAIPIRSAK ASREIASDYK YKEGYRKQLG HHIGARNIKD  3780
DPKMMWSMHV AKIQSDREYK KDFEKWKTKY SSPVDMLGVV LAKKCQILVS DIDYKHPLHE  3840
WTCLPDQNDV IQARKAYDLQ SDAVYKADLE WLRGIGWVPI GSVEVEKVKR AGEILSDRKY  3900
RQPADQLKFT CITDTPEIVL AKSNAITMSK HLYTEAWDAD KASIHVMPDT PEILLAKSNS  3960
ANISHKLYTK GWDESKMKDY DLREDAISIK SAKASRDIAS DYKYKEAYEK QKGHHIGAQS  4020
VEDDPRILCA MHAGKIYSDR EYKKEFQKWK TKFSSPVDML GVVLAKKCQT LVSDVDYRNY  4080
LHNWTCLPDQ NDIIQAKKAY ELQSDSVYKA DLEWLRGIGW MPEGSIEMTR VKGAQDLVNE  4140
RLYRTRPEAL KFTSIVDTPE VVLAKANSLQ MSERLYQEAW NKDKTNISLP LDTPAMLQAQ  4200
INALQISNKL YQKDWDEAKQ KGYDLRADAI EIKHAKASRE IASEYKYKEG YRKQLGHHVG  4260
FRSLQDDPKL VWSIHAAKIQ SDREYKKAYE KSKGIYNTPL DMMSIVQAKK CQVLVSDVDY  4320
RNYLHQWTCL PDQNDVIQAK KAYELQSDAI YKADLEWLRG IGWMPEGSPE VLRVKNAQEI  4380
LCDSVYRTPV VKLKYTSIVD TPEVVLAKSN AENISIPKYR EVWDKDKTSI HIMPDTPEIN  4440
LARANALNVS NKLYREGWDE VKMSCDVRLD AIPIQAAKAS REIASDYKYK LDHEKQKGHY  4500
VGTLTARDDN KIRWALIAGK IQNEREYRLQ WAKWKTKFQS PVDMLSILHS KKCQTLVSDI  4560
DYRNYLHEWT CMPDQNDVIQ AKKAYELQSD AVYKADLEWL RGIGWMPNDS VSVNHAKHAA  4620
DIFSEKKYRT KIETLNFTPV DDRVDYVTAK QSSEIRDDIK YRKDWNAIKS KYTLTETPLL  4680
HTAQEAARIL DQYLYKEGWE KQKATGYILP PDAVPFVHAH HSSDVQSELK YKAEHVKQRG  4740
HYVGVPTMRD DPKLVWFEHA GQIQNDRLYK ESYHKSKAKI NIPADMVSVV AAKEGQNLVS  4800
DIDYRQYLHQ WTCHPDQNDV IQAKIAYDLQ SDNVYKADLE WLRGIGWIPL DSVDHVRVTK  4860
NQEMVNQIKY KKDALANYPN FTSVEDPPEI VLAKINSVNQ SDVKYKETFN KAKGKYIFSP  4920
DTPYISHSKD MGKLYSTILY KGAWEGTKAY GYTLDERYIP IVGAKHADLV NSELKYKETY  4980
EKQKGHYLAG KEISEFPGVV HCLDFQQMRS ALNYRRDYED TKANVHIPND MMNHVLAKRC  5040
QYILSDLEYR HYFHQWTSLP EEPNVIRARN AQEILSDNVY KDDLNWLKGI GCYVWDTPQI  5100
LHAKKSYDLQ SQLQYTAASK ENLQNYNLVT DTPVYVTALQ SGINASEVKY KENYHQIKDK  5160
YTTVLETADY DRTKHLKSLY SSNIYKEAWD KVKATSYNLP TSTLSLTHAK NQKHLASHIK  5220
YREEYEKFKA LYTLPKSVDD DPNTARCLRV GKLNIDRLYR SVYEKNKMKI NIVPDMVEMV  5280
TAKDSQKKVS EIDYRLHLHE WICHPDLQVN SHVRKVTDQI SDIVYKDDLT WLKGIGCYVW  5340
DTPEILHAKH AYDLRNDIKY KAHVQKTRND YKLVTDTPVY VQAVKSGKQL SDAVYHYDYV  5400
HSVRGKVAPT TKTVDLDRAL HAYKLQSENL YKKGLRFLPT GYRLPVDTPH FNHTKDTRYM  5460
SSYFKYKEAY EHIKANGYTL GPKDVPFVHV RRVNNVTSER LYRELYHKLK DKIHTTPDTP  5520
EIRQVKKTQE AVSELIYKSD FFKMQGHMIS LPYTPQVLHC RYVGDITSDI KYKEDLQILK  5580
GLGCFLYDTP DMVRSRHLRK LWSNYLYTDN ARKMRDKYKV VLDTPEYRKV QELKTHLSEL  5640
VYRASGKKQK SIFTSVPDTP DLLRAKRGQK LQSQYLYVEL ATKERPHHHA GNQTTALQHA  5700
KDVKDMVSEK KYKIQYEKMK DKYTPVADTP ILIRAKRAYW NASDLRYKET FQKTKGRYHT  5760
VKDALDIVYH RKVTDDISKV KYKENYMSQL GIWSSIPDRP EHFHHRAVTD AVSDVKYKED  5820
LTWLKGIGCY AYDTPDLTLA EQNKTLYSKY KYKEVFERTK SNFKYVADCP INRHFKYATQ  5880
LMNEKKYRAD YEQRKDKYHL VVDEPRHLLA KTAGDQISQI KYRKKYEKSK DKFTSVVDTP  5940
EHLRTTKVNK QISDILYKLE YNKARPRGYT TIHDTPMLLH VRKVKDEVSD LKYKEVYQRN  6000
KSNCTIKPDA VHIKAAKDAY KVNTNLDYKK QYEANKAYWK WTPDRPDFIQ AAKSSLQQSD  6060
FEYKLDREFL KGCKLSVTDD KDMVLALRNS LIESDLKYKE KHVKERGSCH AVPDTPQILL  6120
AKTVSNLVSE NKYKDHVKKH LAQGSYTTLP ETRDTVHVKA VTKHVSDTNY KKKFVKEKGK  6180
SNYSIMQEPP EVKHAMEVAK KQSDVAYKKD AKENLHYTTV ADRPDIKKAT QAAKQASEVE  6240
YRAKHRKEGS HGLSMLGRPD IEMAKKAAKL SSQVKYRENF DKEKGKTPKY NPKDSQLYKT  6300
MKDANNLASE VKYKADLKKL HKPVTDMKES LIMHHVLNTS QLASSYQYKK NYEKSKGHYH  6360
TIPDNLEQLH HREATELQSI VKYKEKYEKE RGKPMLDFET PTYITAKESQ QMQSGKEYRK  6420
DYEESIKGRN LTGLEVTPAL LHVRHATKIA SEKEYRKDLE ESIRGRGLTE MEDTPDMLRA  6480
KNATQILNEK EYKRDLELEV KGRGLNAMAN ETPDFMRARN ATDIASQIKY KQSAEMEKAN  6540
FTSVVDTPEI IHAQQVKNLS SQKKYKEDAE KSMSYYETVL DTPEMQRVRE NQKNFSLLQY  6600
QCDLKNSKGK ITVVQDTPEI LRVKENQKNF SSVLYKEDVS PGTAIGKTPE MLRVKQTQDH  6660
ISSVKYKEAI GQGTPIPDLP EVKRVKETQK HISSVMYKEN PGTGIPTTVT PEIERVKRNQ  6720
ENFSSVLYKE SLGKGTPTAI TPEMERVKRN QENFSSVLYK ENMGKGTPLP VTPEMERVKH  6780
NQENISSVLY KENVGKATPT PVTPEMQRVK RNQENISSVL YKENMGKATP TPFTPEMERV  6840
KRNQENFSSV LYKENMRKAT PTPVTPDMER AKRNQENISS VLYSDSFRKQ IQGKAAYVLD  6900
TPEMRRVRET QRHISTVKYH EDFEKHKGCF TPVVTDPITE RVKKNTQDFS DINYRGIQRK  6960
VVEMEQKRND QDQDTITGLR VWRTNPGSVF DYDPAEDNIQ SRSLHMINVQ AQRRSRDQSR  7020
SASALSISGG EEKSEHSEAA DQRLSTYSDG GAFFSATSTA YKHVKTTQLP QQRSSSVATQ  7080
QTTLSSIPSH PSTAGKIFRA IYDYIAADAD EVSFKDGDAI VNAQAIDEGW MYGTVQRTGR  7140
TGMLPANYVE AI                                                      7152

SEQ ID NO: 4            moltype = AA  length = 377
FEATURE                 Location/Qualifiers
source                  1..377
                        mol_type = protein
                        organism = Mus musculus
SEQUENCE: 4
MCDEDETTAL VCDNGSGLVK AGFAGDDAPR AVFPSIVGRP RHQGVMVGMG QKDSYVGDEA  60
QSKRGILTLK YPIEHGIITN WDDMEKIWHH TFYNELRVAP EEHPTLLTEA PLNPKANREK  120
MTQIMFETFN VPAMYVAIQA VLSLYASGRT TGIVLDSGDG VTHNVPIYEG YALPHAIMRL  180
DLAGRDLTDY LMKILTERGY SFVTTAEREI VRDIKEKLCY VALDFENEMA TAASSSSLEK  240
SYELPDGQVI TIGNERFRCP ETLFQPSFIG MESAGIHETT YNSIMKCDID IRKDLYANNV  300
MSGGTTMYPG IADRMQKEIT ALAPSTMKIK IIAPPERKYS VWIGGSILAS LSTFQQMWIT  360
KQEYDEAGPS IVHRKCF                                                 377

SEQ ID NO: 5            moltype = AA  length = 381
```

-continued

```
FEATURE                 Location/Qualifiers
source                  1..381
                        mol_type = protein
                        organism = Mus musculus
SEQUENCE: 5
MPFGNTHNKF KLNYKPQEEY PDLSKHNNHM AKVLTPDLYN KLRDKETPSG FTLDDVIQTG  60
VDNPGHPFIM TVGCVAGDEE SYTVFKDLFD PIIQDRHGGY KPTDKHKTDL NHENLKGGDD  120
LDPNYVLSSR VRTGRSIKGY TLPPHCSRGE RRAVEKLSVE ALNSLTGEFK GKYYPLKSMT  180
EQEQQQLIDD HFLFDKPVSP LLLASGMARD WPDARGIWHN DNKSFLVWVN EEDHLRVISM  240
EKGGNMKEVF RRFCVGLQKI EEIFKKAGHP FMWNEHLGYV LTCPSNLGTG LRGGVHVKLA  300
NLSKHPKFEE ILTRLRLQKR GTGGVDTAAV GAVFDISNAD RLGSSEVEQV QLVVDGVKLM  360
VEMEKKLEKG QSIDDMIPAQ K                                            381

SEQ ID NO: 6            moltype = AA  length = 900
FEATURE                 Location/Qualifiers
source                  1..900
                        mol_type = protein
                        organism = Mus musculus
SEQUENCE: 6
MMMVMQPEGL GAGEGPFSGG GGGEYMEQEE DWDRDLLLDP AWEKQQRKTF TAWCNSHLRK  60
AGTQIENIEE DFRNGLKLML LLEVISGERL PRPDKGKMRF HKIANVNKAL DFIASKGVKL  120
VSIGAEEIVD GNLKMTLGMI WTIILRFAIQ DISVEETSAK EGLLLWCQRK TAPYRNVNVQ  180
NFHTSWKDGL ALCALIHRHR PDLIDYAKLR KDDPIGNLNT AFEVAEKYLD IPKMLDAEDI  240
VNTPKPDEKA IMTYVSCFYH AFAGAEQAET AANRICKVLA VNQENEKLME EYEKLASELL  300
EWIRRTVPWL ENRVGEPSMS AMQRKLEDFR DYRRLHKPPR VQEKCQLEIN FNTLQTKLRL  360
SHRPAFMPSE GKLVSDIANA WRGLEQVEKG YEDWLLSEIR RLQRLQHLAE KFQQKASLHE  420
AWTRGKEEML NQHDYESASL QEVRALLRRH EAFESDLAAH QDRVEHIAAL AQELNELDYH  480
EAASVNSRCQ AICDQWDNLG TLTQKRRDAL ERMEKLLETI DQLQLEFARR AAPFNNWLDG  540
AIEDLQDVWL VHSVEETQSL LTAHEQFKAT LPEADRERGA ILGIQGEIQK ICQTYGLRPK  600
SGNPYITLSS QDINNKWDTV RKLVPSRDQT LQEELARQQV NERLRRQFAA QANAIGPWIQ  660
GKVEEVGRLA AGLAGSLEEQ MAGLRQQEQN IINYKSNIDR LEGDHQLLQE SLVFDNKHTV  720
YSMEHIRVGW EQLLTSIART INEVENQVLT RDAKGLSQEQ LNEFRASFNH FDRKRNGMME  780
PDDFRACLIS MGYDLGEVEF ARIMTMVDPN AAGVVTFQAF IDFMTRETAE TDTAEQVVAS  840
FKILAGDKNY ITPEELRREL PAEQAEYCIR RMAPYKGSGA PSGALDYVAF SSALYGESDL  900

SEQ ID NO: 7            moltype = AA  length = 1942
FEATURE                 Location/Qualifiers
source                  1..1942
                        mol_type = protein
                        organism = Mus musculus
SEQUENCE: 7
MSSDAEMAVF GEAAPYLRKS EKERIEAQNK PFDAKSSVFV VDAKESFVKA TVQSREGGKV  60
TAKTEGGTTV TVKDDQVYPM NPPKYDKIED MAMMTHLHEP AVLYNLKERY AAWMIYTYSG  120
LFCVTVNPYK WLPVYNAEVV AAYRGKKRQE APPHIFSISD NAYQFMLTDR ENQSILITGE  180
SGAGKTVNTK RVIQYFATIA VTGEKKKEEA TSGKMQGTLE DQIISANPLL EAFGNAKTVR  240
NDNSSRFGKF IRIHFGTTGK LASADIETYL LEKSRVTFQL KAERSYHIFY QIMSNKKPDL  300
IEMLLITTNP YDYAFVSQGE ITVPSIDDQE ELMATDSAID ILGFTSDERV SIYKLTGAVM  360
HYGNMKFKQK QREEQAEPDG TEVADKAAYL QNLNSADLLK ALCYPRVKVG NEYVTKGQTV  420
QQVYNSVGAL AKAVYEKMFL WMVTRINQQL DTKQPRQYFI GVLDIAGFEI FDFNSLEQLC  480
INFTNEKLQQ FFNHHMFVLE QEEYKKEGIE WEFIDFGMDL AACIELIEKP MGIFSILEEE  540
CMFPKATDTS FKNKLYEQHL GKSNNFQKPK PAKGKVEAHF SLVHYAGTVD YNIAGWLDKN  600
KDPLNETVVG LYQKSSMKTL AYLFSGAAAA AEAESGGGGG KKGAKKKGSS FQTVSALFRE  660
NLNKLMTNLR STHPHFVRCI IPNETKTPGA MEHELVLHQL RCNGVLEGIR ICRKGFPSRI  720
LYADFKQRYK VLNASAIPEG QFIDSKKASE KLLGSIDIDH TQYKFGHTKV FFKAGLLGLL  780
EEMRDDKLAQ LITRTQAMCR GYLARVEYQK MVERRESIFC IQYNVRAFMN VKHWPWMKLY  840
FKIKPLLKSA ETEKEMANMK EEFEKAKENL AKAEAKRKEL EEKMVALMQE KNDLQLQVQS  900
EADSLADAEE RCDQLIKTKI QLEAKIKEVT ERAEDEEEIN AELTAKKRKL EDECSELKKD  960
IDDLELTLAK VEKEKHATEN KVKNLTEEMA GLDETIAKLT KEKKALQEAH QQTLDDLQAE  1020
EDKVNTLTKA KIKLEQQVDD LEGSLEQEKK IRMDLERAKR KLEGDLKLAQ ESTMDVENDK  1080
QQLDEKLKKK EFEMSNLQSK IEDEQALGMQ LQKKIKELQA RIEELEEEIE AERASRAKAE  1140
KQRSDLSREL EEISERLEEA GGATSAQIEM NKKREAEFQK MRRDLEEATL QHEATAATLR  1200
KKHADSVAEL GEQIDNLQRV KQKLEKEKSE MKMEIDDLAS NMEVISKSKG NLEKMCRTLE  1260
DQVSELKTKE EEQQRLINEL TAQRGRLQTE SGEYSRQLDE KDSLVSQLSR GKQAFTQQIE  1320
ELKRQLEEEI KAKSALAHAL QSSRHDCDLL REQYEEEQEA KAELQRAMSK ANSEVAQWRT  1380
KYETDAIQRT EELEEAKKKL AQRLQDAEEH VEAVNAKCAS LEKTKQRLQN EVEDLMIDVE  1440
RTNAACAALD KKQRNFDKIL AEWKQKYEET HAELEASQKE SRSLSTELFK IKNAYEESLD  1500
HLETLKRENK NLQQEISDLT EQIAEGGKRI HELEKIKKQI EQEKSELQAA LEEAEASLEH  1560
EEGKILRIQL ELNQVKSEID RKIAEKDEEI DQLKRNHIRV VESMQSTLDA EIRSRNDAIR  1620
LKKKMEGDLN EMEIQLNHSN RMAAEALRNY RNTQGILKDT QLHLDDALRG QEDLKEQLAM  1680
VERRANLLQA EIEELRATLE QTERSRKIAE QELLDASERV QLLHTQNTSL INTKKKLETD  1740
ISQIQGEMED IVQEARNAEE KAKKAITDAA MMAEELKKEQ DTSAHLERMK KNLEQTVKDL  1800
QHRLDEAEQL ALKGGKKQIQ KLEARVRELE GEVENEQKRN VEAIKGLRKH ERRVKELTYQ  1860
TEEDRKNVLR LQDLVDKLQS KVKAYKRQAE EAEEQSNVNL AKFRKIQHEL EEAEERADIA  1920
ESQVNKLRVK SREVHTKIIS EE                                           1942

SEQ ID NO: 8            moltype = AA  length = 842
```

-continued

```
FEATURE                 Location/Qualifiers
source                  1..842
                        mol_type = protein
                        organism = Mus musculus
SEQUENCE: 8
MSRPLSDQDK RKQISVRGLA GVENVSELKK NFNRHLHFTL VKDRNVATPR DYYFALAHTV  60
RDHLVGRWIR TQQHYYEKDP KRIYYLSLEF YMGRTLQNTM VNLALENACD EATYQLGLDM  120
EELEEIEEDA GLGNGGLGRL AACFLDSMAT LGLAAYGYGI RYEFGIFNQK ICGGWQMEEA  180
DDWLRYGNPW EKARPEFTLP VHFYGRVEHT SQGAKWVDTQ VVLAMPYDTP VPGYRNNVVN  240
TMRLWSAKAP NDFNLKDFNV GGYIQAVLDR NLAENISRVL YPNDNFFEGK ELRLKQEYFV  300
VAATLQDIIR RFKSSKFGSR DPVRTNFDAF PDKVAIQLND THPSLAIPEL MRILVDLERL  360
DWDKAWDVTV KTCAYTNHTV LPEALERWPV HLMETLLPRH LQIIYEINQR FLNRVAAAFP  420
GDVDRLRRMS LVEEGAVKRI NMAHLCIAGS HAVNGVARIH SEILKKTIFK DFYELEPHKF  480
QNKTNGITPR RWLVLCNPGL AEVIAERIGE DYISDLDQLR KLLSYVDDEA FIRDVAKVKQ  540
ENKLKFSAYL EREYKVHINP NSLFDVQVKR IHEYKRQLLN CLHIITLYNR IKREPNRFMV  600
PRTIMIGGKA APGYHMAKMI IKLITAIGDV VNHDPAVGDR LRVIFLENYR VSLAEKVIPA  660
ADLSEQISTA GTEASGTGNM KFMLNGALTI GTMDGANVEM AEEAGEENFF IFGMRVEDVE  720
RLDQRGYNAQ EYYDRIPELR QIIEQLSSGF FSPKQPDLFK DIVNMLMHHD RFKVFADYEE  780
YIKCQDKVSE LYKNPREWTR MVIRNIATSG KFSSDRTIAQ YAREIWGVEP SRQRLPAPDE  840
KI                                                                 842

SEQ ID NO: 9            moltype = AA  length = 364
FEATURE                 Location/Qualifiers
source                  1..364
                        mol_type = protein
                        organism = Mus musculus
SEQUENCE: 9
MPHPYPALTP EQKKELSDIA HRIVAPGKGI LAADESTGSI AKRLQSIGTE NTEENRRFYR  60
QLLLTADDRV NPCIGGVILF HETLYQKADD GRPFPQVIKS KGGVVGIKVD KGVVPLAGTN  120
GETTTQGLDG LSERCAQYKK DGADFAKWRC VLKIGEHTPS ALAIMENANV LARYASICQQ  180
NGIVPIVEPE ILPDGDHDLK RCQYVTEKVL AAVYKALSDH HVYLEGTLLK PNMVTPGHAC  240
TQKFSNEEIA MATVTALRRT VPPAVTGVTF LSGGQSEEEA SINLNAINKC PLLKPWALTF  300
SYGRALQASA LKAWGGKKEN LKAAQEEYIK RALANSLACQ GKYTPSGQSG AAASESLFIS  360
NHAY                                                               364

SEQ ID NO: 10           moltype = AA  length = 188
FEATURE                 Location/Qualifiers
source                  1..188
                        mol_type = protein
                        organism = Mus musculus
SEQUENCE: 10
MAPKKDVKKP AAAPAPAPAP APAPAKPKEE KIDLSAIKIE FSKEQQEDFK EAFLLFDRTG  60
ECKITLSQVG DVLRALGTNP TNAEVKKVLG NPSNEEMNAK KIEFEQFLPM MQAISNNKDQ  120
GGYEDFVEGL RVFDKEGNGT VMGAELRHVL ATLGEKMKEE EVEALLAGQE DSNGCINYEA  180
FVKHIMSV                                                           188

SEQ ID NO: 11           moltype = AA  length = 333
FEATURE                 Location/Qualifiers
source                  1..333
                        mol_type = protein
                        organism = Mus musculus
SEQUENCE: 11
MVKVGVNGFG RIGRLVTRAA ICSGKVEIVA INDPFIDLNY MVYMFQYDST HGKFNGTVKA  60
ENGKLVINGK PITIFQERDP TNIKWGEAGA EYVVESTGVF TTMEKAGAHL KGGAKRVIIS  120
APSADAPMFV MGVNHEKYDN SLKIVSNASC TTNCLAPLAK VIHDNFGIVE GLMTTVHAIT  180
ATQKTVDGPS GKLWRDGRGA AQNIIPASTG AAKAVGKVIP ELNGKLTGMA FRVPTPNVSV  240
VDLTCRLEKP AKYDDIKKVV KQASEGPLKG ILGYTEDQVV SCDFNSNSHS STFDAGAGIA  300
LNDNFVKLIS WYDNEYGYSN RVVDLMAYMA SKE                               333

SEQ ID NO: 12           moltype = AA  length = 284
FEATURE                 Location/Qualifiers
source                  1..284
                        mol_type = protein
                        organism = Mus musculus
SEQUENCE: 12
MDAIKKKMQM LKLDKENALD RAEQAEADKK AAEDRSKQLE DELVSLQKKL KGTEDELDKY  60
SEALKDAQEK LELAEKKATD AEADVASLNR RIQLVEEELD RAQERLATAL QKLEEAEKAA  120
DESERGMKVI ESRAQKDEEK MEIQEIQLKE AKHIAEDADR KYEEVARKLV IIESDLERAE  180
ERAELSEGKC AELEEELKTV TNNLKSLEAQ AEKYSQKEDK YEEEIKVLSD KLKEAETRAE  240
FAERSVTKLE KSIDDLEDEL YAQKLKYKAI SEELDHALND MTSI                   284

SEQ ID NO: 13           moltype = AA  length = 1136
FEATURE                 Location/Qualifiers
source                  1..1136
                        mol_type = protein
                        organism = Mus musculus
```

-continued

```
SEQUENCE: 13
MPEAKPAAKK ASKGKDAPKE APAKQTPEEP PKEAPPEDQS PTAEEPTGIF LKKPDSVSVE  60
TGKDAVILAK VNGKELPGKP TIKWFKGKWQ ELGSKSGARF IFKESHDSTS NVYTVELHIG  120
KVVLGDRGDY RLEIKAKDVC DSCSFNVDVE APRQDSSGQS LESFKRSGDG KSEDAGELDF  180
SGLLKKREVV EEEKKKKKDD DDLGIPPEIW ELLKGAKKSE YEKIAFQYGI TDLRGMLKRL  240
KKAKVEVKKS AAFTKKLDPA YQVDRGNKIK LVVEISDPDL PLKWFKNGQE IKPSSKYVFE  300
NVGKKRILTI NKCTLADDAA YEVAVQDEKC FTELFVKEPP VLIVTPLEDQ QVFVGDRVEM  360
SVEVSEEGAQ VMWMKDGVEM TREDSYKARY RFKKDGKRHI LIYSDVAQED GGRYQVITNG  420
GQCEAELIVE EKQLEVLQDI ADLTVKAAEQ AVFKCEVSDE KVTGKWYKNG VEVRPSKRIT  480
ISHVGRFHKL VIDDVRPEDE GDYTFVPDGY ALSLSAKLNF LEIKVEYVPK QEPPKIHLDC  540
SGKTSDNSIV VVAGNKLRLD VAITGEPPPT ATWLRGDEVF TATEGRTHIE QRPDCSSFVI  600
ESAERSDEGR YTIKVTNPVG EDVASIFLRV VDVPDPPEAV RVTSVGEDWA ILVWEPPKYD  660
GGQPVTGYLM ERKKKGSQRW MKINFEVFTD TTYESTKMIE GVLYEMRVFA VNAIGVSQPS  720
MNTKPFMPIA PTSAPQHLTV EDVTDTTTTL KWRPPDRIGA GGIDGYLVEY CLEGSEEWVP  780
ANKEPVERCG FTVKDLPTGA RILFRVVGVN IAGRSEPATL LQPVTIREIV EQPKIRLPRH  840
LRQTYIRKVG EALNLVIPFQ GKPRPQVVWT KGGAPLDTSR VNVRTSDFDT VFFVRQAARS  900
DSGEYELSVQ IENMKDTATI RIRVVEKAGP AENVMVKEVW GTNALVEWQP PKDDGNSEIT  960
GYFVQKADKK TMEWFNVYEH NRHTSCTVSD LIVGNEYYFR IFSENICGLS DSPGVSKNTA  1020
RILKTGITLK PLEYKEHDFR TAPKFLTPLM DRVVVAGYTA ALNCAVRGHP KPKVVWMKNK  1080
MEIHEDPKFL ITNYQGILTL NIRRPSPFDA GTYSCRAFNE LGEALAECKL DVRVPQ      1136

SEQ ID NO: 14          moltype = AA  length = 1270
FEATURE                Location/Qualifiers
source                 1..1270
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 14
MPEPGKKPVS AFNKKPRSAE VTAGSAAVFE AETERSGVMV RWQRDGSDIT ANDKYGLAAE  60
GKRHTLTVRD ASPDDQGSYA VIAGSSKVKF DLKVTEPAPP EKAESEVAPG APEEVPAPAT  120
ELEESVSSPE GSVSVTQDGS AAEHQGAPDD PIGLFLMRPQ DGEVTVGGSI VFSARVAGAS  180
LLKPPVVKWF KGKWVDLSSK VGQHLQLHDS YDRASKVYLF ELHITDAQTT SAGGYRCEVS  240
TKDKFDSCNF NLTVHEAIGS GDLDLRSAFR RTSLAGAGRR TSDSHEDAGT PDFSSLLKKR  300
DSFRRDSKLE APAEEDVWEI LRQAPPSEYE RIAFQHGVEA CHRPLKRLKG MKQDEKKSTA  360
FQKKLEPAYQ VNKGHKIRLT VELADPDAEV KWLKNGQEIQ MSGSKYIFES VGAKRTLTIS  420
QCSLADDAAY QCVVGGEKCS TELFVKEPPV LITRSLEDQL VMVGQRVEFE CEVSEEGAQV  480
KWLKDGVELT REETFKYRFK KDGRKHHLII NEATLEDAGH YAVRTSGGQS LAELIVQEKK  540
LEVYQSIADL AVGAKDQAVF KCEVSDENVR GVWLKNGKEL VPDNRIKVSH IGRVHKLTID  600
DVTPADEADY SFVPEGFACN LSAKLHFMEV KIDFVPRQEP PKIHLDCPGS TPDTIVVVTG  660
NKLRLDVPIS GDPAPTVVWQ KTVTQGKKAS AGPHPDAPED AGADEEWVFD KKLLCETEGR  720
VRVETTKDRS VFTVEGAEKE DEGVYTVTVK NPVGEDQVNL TVKVIDVPDA PAAPKISNVG  780
EDSCTVQWEP PAYDGGQPVL GYILERKKKK SYRWMRLNFD LLRELSHEAR RMIEGVAYEM  840
RVYAVNAVGM SRPSPASQPF MPIGPPGEPT HLAVEDVSDT TVSLKWRPPE RVGAGGLDGY  900
SVEYCQEGCS EWTPALQGLT ERRSMLVKDL PTGARLLFRV RAHNVAGPGG PIVTKEPVTV  960
QEILQRPRLQ LPRHLRQTIQ KKVGEPVNLL IPFQGKPRPQ VTWTKEGQPL AGEEVSIRNS  1020
PTDTILFIRA ARRTHSGTYQ VTVRIENMED KATLILQIVD KPSPPQDIRI VETWGFNVAL  1080
EWKPPQDDGN TEIWGYTVQK ADKKTMEWFT VLEHYRRTHC VVSELIIGNG YYFRVFSHNM  1140
VGSSDKAAAT KEPVFIPRPG ITYEPPKYKA LDFSEAPSFT QPLANRSIIA GYNAILCCAV  1200
RGSPKPKISW FKNGLDLGED ARFRMFCKQG VLTLEIRKPC PYDGGVYVCR ATNLQGEAQC  1260
ECRLEVRVPQ                                                         1270

SEQ ID NO: 15          moltype = AA  length = 1463
FEATURE                Location/Qualifiers
source                 1..1463
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 15
MSLVAVPFYQ KRHKHFDQSY RNIQTRYLLD QYALKKQATT QSSSQRSLTE RSSSKRASSQ  60
SSAGAMTCRL CAKRVSASEE EEVENENRYR SQAASYGEAK RQRFLSELAQ LEENVHLARS  120
QARDKLDKYF MEQTVEDNLA WERHSFEDRM SRAPEILVRL RSHTIWERMS VRLCFTVQGF  180
PTPVVQWYKD GSLICQAGEP GKYRIESRYG VHTLEINRAN FEDTATYSAV ATNSHGQVST  240
NAAVVVRRYR GEEEPFHSVG LPIGLPLSSV IPYTHFDVQF LEKFGVTFRR EGETVTLKCT  300
LLVTPDLKRV QPRAEWYRDD VLLKESKWTK MFFGEGQASL SFSHLNKDDE GLYTLRIVSR  360
GGVSDHSAFM FVRDADPLVT GAPGAPMDLQ CHDANRDYVI VTWKPPNTTT ESPVIGYFID  420
KCEVGTNNWV QCNDAPVKIC KYPVTGLFEG RSYVFRVRAV NNAGISRPSR ISDAVAALDP  480
VDLRRLQAIH LEGEKEIVIY QDDLEGDVQI PGPPTNVQAS EVSRNYVVLS WDPPSPRGKD  540
PLMYFIEKSA VGSGSWQRVN AQTAVRSPRY AVFDLAEGKS YVFRVLSANK HGLSDPSEIT  600
PPIQAQDMIV VPSAPGRVLA SRNTKTSVVV QWDRPKHEED LLGYYVDCCV AGTNMWEPCN  660
HKPIGYNRFV VHGLTTGEQY IFRVKAVNAV GTSENSQESE VIKVQAALTV PSHPYGITLL  720
NCDGHSMTLG WKVPKFSGGS AIIGYYLDKR EVHHKNWHEV NSSPVKERIL TVEGLTEGSL  780
YEFKIAATNL AGIGQPSDPS EHFKCEAWTD PEPGPAYDLT FCEVRDTSLV ILWKAPVYSG  840
SSPVSGYFVD FKEEDSGEWK TTSEAATPNR YLKVCDLQQG KTYVFRIRAV NASGPGKPSD  900
TSEPVLVEAR PGTKEISAGV DEEGNIYLGF DCQEMTDASQ FTWCKAYEEI ADEERFEVHT  960
EGDHSKLYFK NPDKEDLGTY SVSVSDTDGV SSSFVLDEEE LERLMALSNE IKNPTIPLKS  1020
ELAYEIFDKG QVRFWLQAEH LSPDASFRFI INDREVSDSD THRIKCDRST GMIEMVMDRF  1080
TIENEGTYTV QIQDGKAKNQ SSLVLIGDAF KAVLEEAEFQ RKEFLRKQGP HFAEYLHWDV  1140
TEECEVRLVC KVANTKRETV FKWLKDDVLY ETETPPDLEK GVCELLIPKL SKKDHGEYKA  1200
```

-continued

```
TLKDDRGQDV SVLEVGGKVY EDMILAMSRV CGASASPLKV LCTPEGIRLQ CFMKYFTEEM  1260
KVSWYHKEAK ISSSEHMRIG GSEEMAWLQI CEPTEKDKGK YTFEIFDGKD SHQRSLDLSG  1320
QAFDEAYAEF QQLKAAAFAE KNRGKVIGGL PDVVTIMEGK TLNLTCTVFG NPDPEVVWFK  1380
NDKDIELSEH FLVKMEQSKY VSLTIQGVTA EDSGKYSINV KNKYGGEKID VTVSVYKHGE  1440
KIPDISPPQQ AKPKLIPAST SSD                                          1463

SEQ ID NO: 16          moltype = AA  length = 608
FEATURE                Location/Qualifiers
source                 1..608
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 16
MKWVTFLLLL FVSGSAFSRG VFRREAHKSE IAHRYNDLGE QHFKGLVLIA FSQYLQKCSY  60
DEHAKLVQEV TDFAKTCVAD ESAANCDKSL HTLFGDKLCA IPNLRENYGE LADCCTKQEP  120
ERNECFLQHK DDNPSLPPFE RPEAEAMCTS FKENPTTFMG HYLHEVARRH PYFYAPELLY  180
YAEQYNEILT QCCAEADKES CLTPKLDGVK EKALVSSVRQ RMKCSSMQKF GERAFKAWAV  240
ARLSQTFPNA DFAEITKLAT DLTKVNKECC HGDLLECADD RAELAKYMCE NQATISSKLQ  300
TCCDKPLLKK AHCLSEVEHD TMPADLPAIA ADFVEDQEVC KNYAEAKDVF LGTFLYEYSR  360
RHPDYSVSLL LRLAKKYEAT LEKCCAEANP PACYGTVLAE FQPLVEEPKN LVKTNCDLYE  420
KLGEYGFQNA ILVRYTQKAP QVSTPTLVEA ARNLGRVGTK CCTLPEDQRL PCVEDYLSAI  480
LNRVCLLHEK TPVSEHVTKC CSGSLVERRP CFSALTVDET YVPKEFKAET FTFHSDICTL  540
PEKEKQIKKQ TALAELVKHK PKATAEQLKT VMDDFAQFLD TCCKAADKDT CFSTEGPNLV  600
TRCKDALA                                                           608

SEQ ID NO: 17          moltype = AA  length = 434
FEATURE                Location/Qualifiers
source                 1..434
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 17
MAMQKIFARE ILDSRGNPTV EVDLHTAKGR FRAAVPSGAS TGIYEALELR DGDKARYLGK  60
GVLKAVEHIN KTLGPALLEK KLSVVDQEKV DKFMIELDGT ENKSKFGANA ILGVSLAVCK  120
AGAAEKGVPL YRHIADLAGN PDLVLPVPAF NVINGGSHAG NKLAMQEFMI LPVGASSFKE  180
AMRIGAEVYH HLKGVIKAKY GKDATNVGDE GGFAPNILEN NEALELLKTA IQAAGYPDKV  240
VIGMDVAASE FYRNGKYDLD FKSPDDPARH ISGEKLGELY KNFIQNYPVV SIEDPFDQDD  300
WATWTSFLSG VDIQIVGDDL TVTNPKRIAQ AVEKKACNCL LLKVNQIGSV TESIQACKLA  360
QSNGWGVMVS HRSGETEDTF IADLVVGLCT GQIKTGAPCR SERLAKYNQL MRIEEALGDK  420
AVFAGRKFRN PKAK                                                    434

SEQ ID NO: 18          moltype = AA  length = 169
FEATURE                Location/Qualifiers
source                 1..169
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 18
MAPKKAKRRA GAEGSSNVFS MFDQTQIQEF KEAFTVIDQN RDGIIDKEDL RDTFAAMGRL  60
NVKNEELDAM MKEASGPINF TVFLTMFGEK LKGADPEDVI TGAFKVLDPE GKGTIKKQFL  120
EELLTTQCDR FSQEEIKNMW AAFPPDVGGN VDYKNICYVI THGDAKDQE              169

SEQ ID NO: 19          moltype = AA  length = 574
FEATURE                Location/Qualifiers
source                 1..574
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 19
MSVQENELPQ QLWPWIFKSQ KDLAKSALSG APGGPAGYLR RASVAQLTQE LGTAFFQQQQ  60
LPAAMADTFL EHLCLLDIDS EPVAARSTSI IATIGPASRS VDRLKEMIKA GMNIARLNFS  120
HGSHEYHAES IANIREAAES FATSPLSYRP VAIALDTKGP EIRTGVLQGG PESEVEIVKG  180
SQVLVTVDPK FRTRGDAKTV WVDYHNITQV VAVGGRIYID DGLISLVVRK IGPEGLVTEV  240
EHGGFLGNRK GVNLPNAEVD LPGLSEQDLL DLRFGVEHYV DIIFASFVRK ASDVVAVRDA  300
LGPEGRGIKI ISKIENHEGV KKFDEILEVS DGIMMARGDL GIEIPAEKVF LAQKMMIGRC  360
NLAGKPVVCA TQMLESMITK ARPTRAETSD VANAVLDGAD CIMLSGETAK GSFPVEAVKM  420
QHAIAREAEA AVYHRQLFEE LRRAAPLSRD PTEVTAIGAV EASFKCCAAA IIVLTKTGRS  480
AQLLSRYRPR AAVIAVTRSA QAARQVHLSR GVFPLLYREP PEAVWADDVD RRVQFGIESG  540
KLRGFLRVGD LVIVVTGWRP GSGYTNIMRV LTIS                              574

SEQ ID NO: 20          moltype = AA  length = 260
FEATURE                Location/Qualifiers
source                 1..260
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 20
MAKEWGYASH NGPDHWHELY PIAKGDNQSP IELHTKDIKH DPSLQPWSAS YDPGSAKTIL  60
NNGKTCRVVF DDTYDRSMLR GGPLSGPYRL RQFHLHWGSS DDHGSEHTVD GVKYAAELHL  120
VHWNPKYNTF GEALKQPDGI AVVGIFLKIG REKGEFQILL DALDKIKTKG KEAPFTHFDP  180
SCLFPACRDY WTYHGSFTTP PCEECIVWLL LKEPMTVSSD QMAKLRSLFS SAENEPPVPL  240
```

-continued

```
VGNWRPPQPV KGRVVRASFK                                                260

SEQ ID NO: 21          moltype = AA  length = 332
FEATURE                Location/Qualifiers
source                 1..332
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 21
MATLKDQLIV NLLKEEQAPQ NKITVVGVGA VGMACAISIL MKDLADELAL VDVMEDKLKG  60
EMMDLQHGSL FLKTPKIVSS KDYCVTANSK LVIITAGARQ QEGESRLNLV QRNVNIFKFI  120
IPNIVKYSPH CKLLIVSNPV DILTYVAWKI SGFPKNRVIG SGCNLDSARF RYLMGERLGV  180
HALSCHGWVL GEHGDSSVPV WSGVNVAGVS LKSLNPELGT DADKEQWKEV HKQVVDSAYE  240
VIKLKGYTSW AIGLSVADLA ESIMKNLRRV HPISTMIKGL YGINEDVFLS VPCILGQNGI  300
SDVVKVTLTP EEEARLKKSA DTLWGIQKEL QF                                332

SEQ ID NO: 22          moltype = AA  length = 1942
FEATURE                Location/Qualifiers
source                 1..1942
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 22
MSSDAEMAVF GEAAPYLRKS EKERIEAQNR PFDAKTSVFV AEPKESFVKG TIQSKDAGKV  60
TVKTEAGATL TVKEDQIFPM NPPKYDKIED MAMMTHLHEP AVLYNLKERY AAWMIYTYSG  120
LFCVTVNPYK WLPVYNPEVV AAYRGKKRQE APPHIFSISD NAYQFMLTDR ENQSILITGE  180
SGAGKTVNTK RVIQYFATIA VTGDKKKEEA TSGKMQGTLE DQIISANPLL EAFGNAKTVR  240
NDNSSRFGKF IRIHFGTTGK LASADIETYL LEKSRVTFQL KAERSYHIFY QITSNKKPEL  300
IEMLLITTNP YDYPFVSQGE ISVASIDDQE ELMATDSAID ILGFTNDEKV SIYKLTGAVM  360
HYGNMKFKQK QREEQAEPDG TEVADKAAYL QGLNSADLLK ALCYPRVKVG NEYVTKGQTV  420
EQVTNAVGAL AKAMYEKMFL WMVTRINQQL DTKQPRQYFI GVLDIAGFEI FDFNSLEQLC  480
INFTNEKLQQ FFNHHMFVLE QEEYKKEGIE WTFIDFGMDL AACIELIEKP MGIFSILEEE  540
CMFPKATDTS FKNKLYEQHL GKSANFQKPK VVKGKAEAHF SLIHYAGTVD YNITGWLDKN  600
KDPLNETVVG LYQKSSVKTL AYLFSGAQTA EAEASSGGAA KKGAKKKGSS FQTVSALFRE  660
NLNKLMTNLR STHPHFVRCI IPNETKTPGA MEHELVLHQL RCNGVLEGIR ICRKGFPSRI  720
LYADFKQRYK VLNASAIPEG QYIDSKKASE KLLGSIDIDH TQYKFGHTKV FFKAGLLGLL  780
EEMRDDKLAQ LITRTQAMCR GFLARVEYQK MVERRESIFC IQYNIRAFMN VKHWPWMKLF  840
FKIKPLLKSA ETEKEMATMK EEFQKTKDDL AKSEAKRKEL EEKMVSLLKE KNDLQLQVQA  900
EAEGLADAEE RCDQLIKTKI QLEAKIKEVT ERAEDEEEIN AELTAKKRKL EDECSELKKD  960
IDDLELTLAK VEKEKHATEN KVKNLTEEMA GLDETIAKLT KEKKALQEAH QQTLDDLQAE  1020
EDKVNTLTKA KIKLEQQVDD LEGSLEQEKK LRMDLERAKR KLEGDLKLAQ ESIMDIENEK  1080
QQLDERLKKK EFEMSNLQSK IEDEQAIGIQ LQKKIKELQA RIEELEEEIE AERASRAKAE  1140
KQRSDLSREL EEISERLEEA GGATSAQIEM NKKREAEFQK MRRDLEEATL QHEATAATLR  1200
KKHADSVAEL GEQIDNLQRV KQKLEKEKSE MKMEIDDLAS NVETVSKAKG NLEKMCRTLE  1260
DQVSELKSKE EEQQRLINDL TTQRGRLQTE SGEFSRQLDE KEALVSQLSR GKQAFTQQIE  1320
ELKRQLEEEV KAKNALAHAL QSSRHDCDLL REQYEEEQES KAELQRALSK ANSEVAQWRT  1380
KYETDAIQRT EELEEAKKKL AQRLQAAEEH VEAVNAKCAS LEKTKQRLQN EVEDLMLDVE  1440
RTNAACAALD KKQRNFDKIL AEWKQKYEET HAELEASQKE ARSLGTELFK MKNAYEESLD  1500
QLETLKRENK NLQQEISDLT EQIAEGGKRI HELEKIKKQV EQEKCELQAA LEEAEASLEH  1560
EEGKILRIQL ELNQVKSEID RKIAEKDEEI DQLKRNHIRV VESMQSTLDA EIRSRNDAIR  1620
IKKKMEGDLN EMEIQLNHSN RMAAEALRNY RNTQGILKDT QLHLDDALRG QEDLKEQLAM  1680
VERRANLLQA EIEELRATLE QTERSRKIAE QELLDASERV QLLHTQNTSL INTKKKLETD  1740
ISQIQGEMED IVQEARNAEE KAKKAITDAA MMAEELKKEQ DTSAHLERMK KNMEQTVKDL  1800
QLRLDEAEQL ALKGGKKQIQ KLEARVRELE GEVESEQKRN AEAVKGLRKH ERRVKELTYQ  1860
TEEDRKNILR LQDLVDKLQA KVKSYKRQAE EAEEQSNTNL SKFRKIQHEL EEAEERADIA  1920
ESQVNKLRVK SREVHTKIIS EE                                           1942

SEQ ID NO: 23          moltype = AA  length = 1124
FEATURE                Location/Qualifiers
source                 1..1124
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 23
MPEPTKKEEN EVSAPAPPPE EPNKEKEAGT APAKEWSLGE SPAGGEEQDK QNANSQLSTL  60
FVEKPQTGSV KVGANITFVA KVRAEDLLRK PTVKWFKGKW MDLASKAGKH LQLKETFERQ  120
TRVYTFEMQI IKAKENYAGN YRCEVTYKDK FDSCSFDLEV HESTGTTPNI DIRSAFKRSG  180
EGQEDAGELD FSGLLKRREV KQQEEEPEID VWELLKNANP NEYEKIAFQY GITDLRGMLK  240
RLKRMRRVEK KSAAFAKILD PAYQVDKGGK VRFVVELADP KLEVKWFKNG QEIRPSTKYI  300
FEHKGNERIM FINNCSLTDD SEYYVTAGDE KCSTELFVRE PPIMVTKQLE DMNAYCGERV  360
EMEVEVSEDD ANVKWFKNGE EIFPGPKSRY KIKVEGKKHT LIIEGATKAD SAEYSAMTTG  420
GQSSAKLSVD LRPLKITTPL TDQTVKLGKE VCLKCEISEN VPGKWTKNGL PVQEGERLKV  480
VHKGRIHKLV IANALIEDEG EYVFTPDAYN VPLSAKVHVI DPPKIILDGL DADNTVTVIA  540
GSKLRLEIPV TGEPPPKAIW SRADKAIMEG SGRIRAESYP DSSTLVIDVA ERDDSGVYNI  600
NLKNEAGEAH ASIKIKVVDI PDPPVAPNVT EVGDDWCIMN WEPPAYDGGS PILGYFIERK  660
KKQSSRWMRL NFDLCKETTF EPKKMIEGVA YEVRIFAVNA IGISKPSMPS KPFVPLAVTS  720
PPTLLAVDSV TDSSVTMKWR PPDQIGAAGL DGYVLEYCFE GTEDWIVANT DLIDKTKFTI  780
NGLPTDAKIF VRVKAINAAG ASEPKYYSQP ILVKEIIEPP KIRIPRHLKQ TYIRRVGEAV  840
NLVIPFQGKP RPELTWKKDG EEIDKNQINI RNSETDTIIF IRKAERSHSG KYDLQVKVDK  900
```

-continued

```
YVENASIDIQ IVDRPGPPQT VTIEDVWGEN VALTWTPPKD DGNAAITGYT IQKADKKSME  960
WFTVIEHYHR TNATITELVI GNEYYFRVFA ENMCGLSEDA TMTKESAVIA KDGKIYKNPV  1020
YEDFNFTEAP MFTQPLVNTY AIAGYNATLN CSVRGNPKPK ITWMKNKVAI VDDPRYRMFS  1080
NQGVCTLEIR KPSPYDGGTY CCKAVNDLGT VEIECKLEVK VVAQ                   1124

SEQ ID NO: 24          moltype = AA  length = 562
FEATURE                Location/Qualifiers
source                 1..562
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 24
MVKIVTVKTQ AYPDQKPGTS GLRKRVKVFQ SNANYAENFI QSIVSTVEPA LRQEATLVVG  60
GDGRFYMTEA IQLIVRIAAA NGIGRLVIGQ NGILSTPAVS CIIRKIKAIG GIILTASHNP  120
GGPNGDFGIK FNISNGGPAP EAITDKIFQI SKTIEEYAIC PDLKVDLGVL GKQQFDLENK  180
FKPFTVEIVD SVEAYATMLR NIFDFNALKE LLSGPNRLKI RIDAMHGVVG PYVKKILCEE  240
LGAPANSAVN CVPLEDFGGH HPDPNLTYAA DLVETMKSGE HDFGAAFDGD GDRNMILGKH  300
GFFVNPSDSV AVIAANIFSI PYFQQTGVRG FARSMPTSGA LDRVANATKI ALYETPTGWK  360
FFGNLMDASK LSLCGEESFG TGSDHIREKD GLWAVLAWLS ILATRKQSVE DILKDHWQKF  420
GRNFFTRYDY EEVEAEGANK MMKDLEALML DRSFVGKQFS ANDKVYTVEK ADNFEYSDPV  480
DGSISKNQGL RLIFADGSRI IFRLSGTGSA GATIRLYIDS YEKDVAKINQ DPQVMLAPLI  540
SIALKVSQLQ ERTGRTAPTV IT                                           562

SEQ ID NO: 25          moltype = AA  length = 284
FEATURE                Location/Qualifiers
source                 1..284
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 25
MDAIKKKMQM LKLDKENAID RAEQAEADKK QAEDRCKQLE EEQQALQKKL KGTEDEVEKY  60
SESVKDAQEK LEQAEKKATD AEADVASLNR RIQLVEEELD RAQERLATAL QKLEEAEKAA  120
DESERGMKVI ENRAMKDEEK MELQEMQLKE AKHIAEDSDR KYEEVARKLV ILEGELERSE  180
ERAEVAESKC GDLEEELKIV TNNLKSLEAQ ADKYSTKEDK YEEEIKLLEE KLKEAETRAE  240
FAERSVAKLE KTIDDLEDEV YAQKMKYKAI SEELDNALND ITSL                   284

SEQ ID NO: 26          moltype = AA  length = 249
FEATURE                Location/Qualifiers
source                 1..249
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 26
MAPTRKFFVG GNWKMNGRKK CLGELICTLN AANVPAGTEV VCAPPTAYID FARQKLDPKI  60
AVAAQNCYKV TNGAFTGEIS PGMIKDLGAT WVVLGHSERR HVFGESDELI GQKVSHALAE  120
GLGVIACIGE KLDEREAGIT EKVVFEQTKV IADNVKDWSK VVLAYEPVWA IGTGKTATPQ  180
QAQEVHEKLR GWLKSNVNDG VAQSTRIIYG GSVTGATCKE LASQPDVDGF LVGGASLKPE  240
FVDIINAKQ                                                          249

SEQ ID NO: 27          moltype = AA  length = 253
FEATURE                Location/Qualifiers
source                 1..253
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 27
MTTHRLVMVR HGESLWNQEN RFCGWFDAEL SEKGAEEAKR GATAIKDAKI EFDICYTSVL  60
KRAIRTLWTI LDVTDQMWVP VVRTWRLNER HYGGLTGLNK AETAAKHGEE QVKIWRRSFD  120
TPPPPMDEKH NYYTSISKDR RYAGLKPEEL PTCESLKDTI ARALPFWNEE IAPKIKAGQR  180
VLIAAHGNSL RGIVKHLEGM SDQAIMELNL PTGIPIVYEL DQNLKPTKPM RFLGDEETVR  240
KAMEAVAAQG KAK                                                     253

SEQ ID NO: 28          moltype = AA  length = 472
FEATURE                Location/Qualifiers
source                 1..472
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 28
MKALLLLCCF LASLLLSGQA EVEDASEEAP LRDRSHIDKT LMLNEDKPAD DYSAVLQRLR  60
KIYHTSIKPL EQSYKYNELR QHEITDGEIT SKPMVLFLGP WSVGKSTMIN YLLGLEDTRY  120
QLYTGAEPTT SEFTVLMHGP KLKTIEGIVM AADSARSFSP LEKFGQNFLE KLIGIEVPHK  180
LLERVTFVDT PGIIENRKQQ ERGYPFNDVC QWFIDRADLI FVVFDPTKLD VGLELEMLFR  240
QLKGRESQIR IILNKADNLA TQMLMRVYGA LFWSLAPLIN VTEPPRVYVS SFWPQDYKPD  300
THRELFLKEE ISLLEDLNQV IENRLENKIA FIRQHAIRVR IHALLVDRYL QTYKDKMTFF  360
SDGELVFKDI VEDPDKFYIF KTILAKTNVS KFDLPNREAY KDFFGINPIS NFKLLSQQCS  420
YMGGCFLEKI ERAITQELPS LLGSIGLGKN PGAPNCDKTG CGETPKNRYK KH          472

SEQ ID NO: 29          moltype = AA  length = 417
FEATURE                Location/Qualifiers
source                 1..417
```

-continued

```
                        mol_type = protein
                        organism = Mus musculus
SEQUENCE: 29
MSLSNKLTLD KLDVKGKRVV MRVDFNVPMK NNQITNNQRI KAAVPSIKFC LDNGAKSVVL  60
MSHLGRPDGV PMPDKYSLEP VAAELKSLLG KDVLFLKDCV GPEVENACAN PAAGTVILLE  120
NLRFHVEEEG KGKDASGNKV KAEPAKIDAF RASLSKLGDV YVNDAFGTAH RAHSSMVGVN  180
LPQKAGGFLM KKELNYFAKA LESPERPFLA ILGGAKVADK IQLINNMLDK VNEMIIGGGM  240
AFTFLKVLNN MEIGTSLYDE EGAKIVKDLM SKAEKNGVKI TLPVDFVTAD KFDENAKTGQ  300
ATVASGIPAG WMGLDCGTES SKKYAEAVGR AKQIVWNGPV GVFEWEAFAR GTKSLMDEVV  360
KATSRGCITI IGGGDTATCC AKWNTEDKVS HVSTGGGASL ELLEGKVLPG VDALSNV     417

SEQ ID NO: 30           moltype = AA  length = 4691
FEATURE                 Location/Qualifiers
source                  1..4691
                        mol_type = protein
                        organism = Mus musculus
SEQUENCE: 30
MVAGMLMPLD RLRAIYEVLF REGVMVAKKD RRPRSLHPHV PGVTNLQVMR AMASLKARGL  60
VRETFAWCHF YWYLTNEGID HLRQYLHLPP EIVPASLQRV RRPVAMVIPA RRRSPHVQTM  120
QGPLGCPPKR GPLPAEDPAR EERQVYRRKE REEGAPETPV VSATTVGTLA RPGPEPAPAT  180
DERDRVQKKT FTKWVNKHLI KHWRAEAQRH ISDLYEDLRD GHNLISLLEV LSGDSLPREK  240
GRMRFHKLQN VQIALDYLRH RQVKLVNIRN DDIADGNPKL TLGLIWTIIL HFQISDIQVS  300
GQSEDMTAKE KLLLWSQRMV EGYQGLRCDN FTTSWRDGRL FNAIIHRHKP MLIDMNKVYR  360
QTNLENLDQA FSVAERDLGV TRLLDPEDVD VPQPDEKSII TYVSSLYDAM PRVPGAQDGV  420
RANELQLRWQ EYRELVLLLL QWIRHHTAAF EERKFPSSFE EIEILWCQFL KFKETELPAK  480
EADKNRSKVI YQSLEGAVQA GQLKIPPGYH PLDVEKEWGK LHVAILEREK QLRSEFERLE  540
CLQRIVSKLQ MEAGLCEEQL NQADALLQSD IRLLASGKVA QRAGEVERDL DKADGMIRLL  600
FNDVQTLKDG RHPQGEQMYR RVYRLHERLV AIRTEYNLRL KAGVGAPVTQ VTLQSTQRRP  660
ELEDSTLRYL QDLLAWVEEN QRRIDSAEWG VDLPSVEAQL GSHRGMHQSI EEFRAKIERA  720
RNDESQLSPA TRGAYRDCLG RLDLQYAKLL NSSKARLRSL ESLHGFVAAA TKELMWLNEK  780
EEEEVGFDWS DRNTNMAAKK ESYSALMREL EMKEKKIKEI QNTGDRLLRE DHPARPTVES  840
FQAALQTQWS WMLQLCCCIE AHLKENTAYF QFFSDVREAE EQLQKLQETL RRKYSCDRTI  900
TVTRLEDLLQ DAQDEKEQLN EYKGHLSGLA KRAKAIVQLK PRNPAHPVRG HVPLIAVCDY  960
KQVEVTVHKG DQCQLVGPAQ PSHWKVLSGS SSEAAVPSVC FLVPPPNQEA QEAVARLEAQ  1020
HQALVTLWHQ LHVDMKSLLA WQSLSRDIQL IRSWSLVTFR TLKPEEQRQA LRNLELHYQA  1080
FLRDSQDAGG FGPEDRLVAE REYGSCSRHY QQLLQSLEQG EQEESRCQRC ISELKDIRLQ  1140
LEACETRTVH RLRLPLDKDP ARECAQRIAE QQKAQAEVEG LGKGVARLSA EAEKVLALPE  1200
PSPAAPTLRS ELELTLGKLE QVRSLSAIYL EKLKTISLVI RSTQGAEEVL KTHEEQLKEA  1260
QAVPATLQEL EATKASLKKL RAQAEAQQPV FNTLRDELRG AQEVGERLQQ RHGERDVEVE  1320
RWRERVTQLL ERWQAVLAQT DVRQRELEQL GRQLRYYRES ADPLSAWLQD AKRRQEQIQA  1380
VPIANCQAAR EQLRQEKALL EEIERHGEKV EECQKFAKQY INAIKDYELQ LITYKAQLEP  1440
VASPAKKPKV QSGSESVIQE YVDLRTRYSE LTTLTSQYIK FISETLRRME EEERLAEQQR  1500
AEERERLAEV EAALEKQRQL AEAHAQAKAQ AELEAQELQR RMQEEVARRE EAAVDAQQQK  1560
RSIQEELQHL RQSSEAEIQA KAQQVEAAER SRMRIEEEIR VVRLQLETTE RQRGGAEGEL  1620
QALRARAEEA EAQKRQAQEE AERLRRQVQD ESQRKRQAEA ELALRVKAEA EAAREKQRAL  1680
QALDELRLQA EEAERRLRQA EAERARQVQV ALETAQRSAE VELQSKRASF AEKTAQLERT  1740
LQEEHVTVAQ LREEAERRAQ QQAEAERARE EAERELERWQ LKANEALRLR LQAEEVAQQK  1800
SLAQADAEKQ KEEAEREARR RGKAEEQAVR QRELAEQELE KQRQLAEGTA QQRLAAEQEL  1860
IRLRAETEQG EQQRQLLEEE LARLQHEATA ATQKRQELEA ELAKVRAEME VLLASKARAE  1920
EESRSTSEKS KQRLEAEAGR FRELAEEAAR LRALAEEAKR QRQLAEEDAA RQRAEAERVL  1980
TEKLAAISEA TRLKTEAEIA LKEKEAENER LRRLAEDEAF QRRRLEEQAA LHKADIEERL  2040
AQLRKASESE LERQKGLVED TLRQRRQVEE EIMALKVSFE KAAAGKAELE LELGRIRSNA  2100
EDTMRSKEQA ELEAARQRQL AAEEEQRRRE AEERVQRSLA AEEEAARQRK VALEEVERLK  2160
AKVEEARRLR ERAEQESARQ LQLAQEAAQK RLQAEEKAHA FVVQQREEEL QQTLQQEQNM  2220
LDRLRSEAEA ARRAAEEAEE AREQAEREAA QSRKQVEEAE RLKQSAEEQA QAQAQAQAAA  2280
EKLRKEAEQE AARRAQAEQA ALKQKQAADA EMEKHKKFAE QTLRQKAQVE QELTTLRLQL  2340
EETDHQKSIL DEELQRLKAE VTEAARQRSQ VEEELFSVRV QMEELGKLKA RIEAENRALI  2400
LRDKDNTQRF LEEEAEKMKQ VAEEAARLSV AAQEAARLRQ LAEEDLAQQR ALAEKMLKEK  2460
MQAVQEATRL KAEAELLQQQ KELAQEQARR LQEDKEQMAQ QLVEETQGFQ RTLEAERQRQ  2520
LEMSAEAERL KLRMAEMSRA QARAEEDAQR FRKQAEEIGE KLHRTELATQ EKVTLVQTLE  2580
IQRQQSDHDA ERLREAIAEL EREKEKLKQE AKLLQLKSEE MQTVQQEQIL QETQALQKSF  2640
LSEKDSLLQR ERFIEQEKAK LEQLFQDEVA KAKQLREEQQ RQQQQMEQEK QELMASMEEA  2700
RRRQREAEEG VRRKQEELQH LEQQRQQQEK LLAEENQRLR ERLQRLEEEH RAALAHSEIA  2760
TTQAASTKAL PNGRDAPDGP SVEAEPEYTF EGLRQKVPAQ QLQEAGILSQ EELQRLAQGH  2820
TTVAELTQRE DVYRYLKGRS SIAGLLLKPT NEKLSVYTAL QRQLLSPGTA LILLEAQAAS  2880
GFLLDPVRNR RLTVNEAVKE GVVGPELHHK LLSAERAVTG YKDPYTGEQI SLFQAMKKDL  2940
IVRDHGVRLL EAQIATGGII DPVHSHRVPV DVAYKRGYFD EEMNRILSDP SDDTKGFFDP  3000
NTHENLTYLQ LLERCVEDPE TGLRLLPLTD KAAKGGELVY TDTEARDVFE KATVSAPFGK  3060
FQGRTVTIWE IINSEYFTAE QRRDLLQQFR TGHITVEKII KIVITVVEEH ERKGQLCFEG  3120
LRALVPAAEL LDSGVISHEL YQQLQRGERS VREVAEADSV RQALRGTNVI AGVWLEEAGQ  3180
KLSIYEALKK DLLQPEVAVA LLEAQAGTGH IIDPATSARL TVDEAVRAGL VGPELHEKLL  3240
SAEKAVTGYR DPYSGQSVSL FQALKKGLIP REQGLRLLDA QLSTGGIVDP SKSHRVPLDV  3300
AYARGYLDKE TNRALTSPRD DARVYHDPST QEPVTYSQLQ QRCRSDQLTG LSLLPLSEKA  3360
VRARQEEVYS ELQARETLEQ AKVEVPVGSF KGRAMTVWEL ISSEYFTEEQ RQELLRQFRT  3420
GKVTVEKVIK IVITIVEEVE TRRQERLSFS GLRAPVPASE LLDAKILSRA QFDQLKDGKT  3480
SVKELSEVGS VRTLLQGSGC LAGIYLEDSK EKVTIYEAMR RGLLRPSTAT LLLEAQAATG  3540
```

-continued

```
FLVDPVRNQR LYVHEAVKAG VVGPELHEKL LSAEKAVTGY KDPYSGNTIS LFQAMKKGLV  3600
LRDHAIRLLE AQVATGGIID PVHSHRLPVD VAYQRGYFDE EMNRVLADPS DDTKGFFDPN  3660
THENLTYLQL LERCVEDPET GLRLLPLKGA EKTEVVETTQ VYTEEETRRA FEETQIDIPG  3720
GGSHGGSSMS LWEVMQSNMI PEDQRARLMA DFQAGRVTKE RMIIIIIEII EKTEIIRQQN  3780
LASYDYVRRR LTAEDLYEAR IISLETYNLF REGTKNLREV LEMESAWRYL YGTGAVAGVY  3840
LPGSRQTLTI YQALKKGLLS AEVARLLLEA QAATGFLLDP VKGERLTVDE AVRKGLVGPE  3900
LHDRLLSAER AVTGYRDPYT EQTISLFQAM KKELIPAEEA LRLLDAQLAT GGIVDPRLGF  3960
HLPLEVAYQR GYLNKDTHDQ LSEPSEVRSY VDPSTDERLS YTQLLKRCRR DDPSGQMLLL  4020
LSDARKLTFR GLRKQITVEE LVRSQVMDEA TALQLQEGLT SIEEVTKNLQ KFLEGTSCIA  4080
GVFVDATKER LSVYQAMKKG IIRPGTAFEL LEAQAATGYV IDPIKGLKLT VEEAVRMGIV  4140
GPEFKDKLLS AERAVTGYKD PYSGKLISLF QAMKKGLILK DHGIRLLEAQ IATGGIIDPE  4200
ESHRLPVEVA YKRGLFDEEM NEILTDPSDD TKGFFDPNTE ENLTYLQLME RCITDPQTGL  4260
CLLPLKEKKR ERKTSSKSSV RKRRVVIVDP ETGKEMSVYE AYRKGLIDHQ TYLELSEQEC  4320
EWEEITISSS DGVVKSMIID RRSGRQYDID DAITKNLIDR SALDQYRAGT LSITEFADML  4380
SGNAGGFRSR SSSVGSSSSY PISSAGPRTQ LASWSDPTEE TGPVAGILDT ETLEKVSITE  4440
AMHRNLVDNI TGQRLLEAQA CTGGIIDPST GERFPVTEAV NKGLVDKIMV DRINLAQKAF  4500
CGFEDPRTKT KMSAAQALKK GWLYYEAGQR FLEVQYLTGG LIEPDTPGRV SLDEALQRGT  4560
VDARTAQKLR DVSAYSKYLT CPKTKLKISY KDALDRSMVE EGTGLRLLEA AAQSSKGYYS  4620
PYSVSGSGST AGSRTGSRTG SRAGSRRGSF DATGSGFSMT FSSSSYSSSG YGRRYASGPS  4680
ASLGGPESAV A                                                       4691
```

```
SEQ ID NO: 31          moltype = AA  length = 894
FEATURE                Location/Qualifiers
source                 1..894
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 31
MNQIEPGVQY NYVYDEDEYM IQEEEWDRDL LLDPAWEKQQ RKTFTAWCNS HLRKAGTQIE  60
NIEEDFRNGL KLMLLLEVIS GERLPKPDRG KMRFHKIANV NKALDYIASK GVKLVSIGAE  120
EIVDGNVKMT LGMIWTIILR FAIQDISVEE TSAKEGLLLW CQRKTAPYRN VNIQNFHTSW  180
KDGLGLCALI HRHRPDLIDY SKLNKDDPIG NINLAMEIAE KHLDIPKMLD AEDIVNTPKP  240
DERAIMTYVS CFYHAFAGAE QAETAANRIC KVLAVNQENE RLMEEYERLA SELLEWIRRT  300
IPWLENRTPE KTMQAMQKKL EDFRDYRRKH KPPKVQEKCQ LEINFNTLQT KLRISNRPAF  360
MPSEGKMVSD IAGAWQRLEQ AEKGYEEWLL NEIRRLERLE HLAEKFRQKA STHETWAYGK  420
EQILLQKDYE SASLTEVRAL LRKHEAFESD LAAHQDRVEQ IAAIAQELNE LDYHDAVNVN  480
DRCQKICDQW DRLGTLTQKR REALERTEKL LETIDQLHLE FAKRAAPFNN WMEGAMEDLQ  540
DMFIVHSIEE IQSLITAHEQ FKATLPEADG ERQSILAIQN EVEKVIQSYS IRISSSNPYS  600
TVTMDELRNK WDKVKQLVPV RDQSLQEELA RQHANERLRR QFAAQANAIG PWIQNKMEEI  660
ARSSIQITGA LEDQMNQLKQ YEHNIINYKN NIDKLEGDHQ LIQEALVFDN KHTNYTMEHI  720
RVGWELLLTT IARTINEVET QILTRDAKGI TQEQMNEFRA SFNHFDRRKN GLMDHEDFRA  780
CLISMGYDLG EAEFARIMTL VDPNGQGTVT FQSFIDFMTR ETADTDTAEQ VIASFRILAS  840
DKPYILAEEL RRELPPDQAQ YCIKRMPPYS GPGSVPGALD YTAFSSALYG ESDL        894
```

```
SEQ ID NO: 32          moltype = AA  length = 780
FEATURE                Location/Qualifiers
source                 1..780
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 32
MTHEEHHAAK TLGIGKAIAV LTSGGDAQGM NAAVRAVVRV GIFTGARVFF VHEGYQGLVD  60
GGEHIREATW ESVSMMLQLG GTVIGSARCK DFREREGRLR AAHNLVKRGI TNLCVIGGDG  120
SLTGADTFRS EWSDLLNDLQ KDGKITAEEA TKSSYLNIVG LVGSIDNDFC GTDMTIGTDS  180
ALHRIVEIVD AITTTAQSHQ RTFVLEVMGR HCGYLALVTS LSCGADWVFI PECPPDDDWE  240
EHLCRRLSET RTRGSRLNII IVAEGAIDKN GKPITSEDIK NLVVKRLGYD TRVTVLGHVQ  300
RGGTPSAFDR ILGSRMGVEA VMALLEGTPD TPACVVSLSG NQAVRLPLME CVQVTKDVTK  360
AMDEKRFDEA IKLRGRSFMN NWEVYKLLAH VRPPVSKGGL HTVAVMNVGA PAAGMNAAVR  420
STVRIGLIQG NRVLVVHDGF EGLAKGQIEE AGWSYVGGWT GQGGSKLGTK RTLPKKNLEQ  480
ISANITKFNI QGLVIIGGFE AYTGGLELME GRKQFDELCI PFVVIPATVS NNVPGSDFSI  540
GADTALNTIC TTCDRIKQSA AGTKRRVFII ETMGGYCGYL ATMAGLAAGA DAAYIFEEPF  600
TIRDLQVNVE HLVQKMKTTV KRGLVLRNEK CNENYTTDFI FNLYSEEGKG IFDSRKNVLG  660
HMQQGGSPTP FDRNFATKMG AKAMNWMSGK IKESYRNGRI FANTPDSGCV LGMRKRALVF  720
QPVTELKDQT DFEHRIPKEQ WWLKLRPILK ILAKYEIDLD TSDHAHLEHI SRKRSGEAAV  780
```

```
SEQ ID NO: 33          moltype = AA  length = 338
FEATURE                Location/Qualifiers
source                 1..338
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 33
MLSALARPAG AALRRSFSTS AQNNAKVAVL GASGGIGQPL SLLLKNSPLV SRLTLYDIAH  60
TPGVAADLSH IETRANVKGY LGPEQLPDCL KGCDVVVIPA GVPRKPGMTR DDLFNTNATI  120
VATLTAACAQ HCPEAMVCII ANPVNSTIPI TAEVFKKHGV YNPNKIFGVT TLDIVRANTF  180
VAELKGLDPA RVNVPVIGGH AGKTIIPLIS QCTPKVDFPQ DQLATLTGRI QEAGTEVVKA  240
KAGAGSATLS MAYAGARFVF SLVDAMNGKE GVVECSFVQS KETECTYFST PLLLGKKGLE  300
KNLGIGKITP FEEKMIAEAI PELKASIKKG EDFVKNMK                          338
```

-continued

```
SEQ ID NO: 34          moltype = AA  length = 154
FEATURE                Location/Qualifiers
source                 1..154
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 34
MGLSDGEWQL VLNVWGKVEA DLAGHGQEVL IGLFKTHPET LDKFDKFKNL KSEEDMKGSE  60
DLKKHGCTVL TALGTILKKK GQHAAEIQPL AQSHATKHKI PVKYLEFISE IIIEVLKKRH  120
SGDFGADAQG AMSKALELFR NDIAAKYKEL GFQG                              154

SEQ ID NO: 35          moltype = AA  length = 147
FEATURE                Location/Qualifiers
source                 1..147
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 35
MVHLTDAEKA AVSCLWGKVN SDEVGGEALG RLLVVYPWTQ RYFDSFGDLS SASAIMGNAK  60
VKAHGKKVIT AFNDGLNHLD SLKGTFASLS ELHCDKLHVD PENFRLLGNM IVIVLGHHLG  120
KDFTPAAQAA FQKVVAGVAT ALAHKYH                                      147

SEQ ID NO: 36          moltype = AA  length = 2726
FEATURE                Location/Qualifiers
source                 1..2726
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 36
MMNNSNYSDA SGLGLVDEAD EMPSTEKDLA EDAPWKKIQQ NTFTRWCNEH LKCVGKRLTD  60
LQRDLSDGLR LIALLEVLSQ KRMYRKFHPR PNFRQMKLEN VSVALEFLER EHIKLVSIDS  120
KAIVDGNLKL ILGLIWTLIL HYSISMPMWE DEDDEDARKQ TPKQRLLGWI QNKVPQLPIT  180
NFNRDWQDGK ALGALVDNCA PGLCPDWEAW DPNQPVQNAR EAMQQADDWL GVPQVIAPEE  240
IVDPNVDEHS VMTYLSQFPK AKLKPGAPVR SKQLNPKKAI AYGPGIEPQG NTVLQPAHFT  300
VQTVDAGVGE VLVYIEDPEG HTEEAKVVPN NDKDRTYAVS YVPKVAGLHK VTVLFAGQNI  360
ERSPFEVNVG MALGDANKVS ARGPGLEPVG NVANKPTYFD IYTAGAGTGD VAVVIVDPQG  420
RRDTVEVALE DKGDNTFRCT YRPVMEGPHT VHVAFAGAPI TRSPFPVHVA EACNPNACRA  480
SGRGLQPKGV RVKEVADFKV FTKGAGSGEL KVTVKGPKGT EEPVKVREAG DGVFECEYYP  540
VVPGKYVVTI TWGGYAIPRS PFEVQVSPEA GAQKVRAWGP GLETGQVGKS ADFVVEAIGT  600
EVGTLGFSIE GPSQAKIECD DKGDGSCDVR YWPTEPGEYA VHVICDDEDI RDSPFIAHIQ  660
PAPPDCFPDK VKAFGPGLEP TGCIVDRPAE FTIDARAAGK GDLKLYAQDA DGCPIDIKVI  720
PNGDGTFRCS YVPTKPIKHT IIVSWGGVNV PKSPFRVNVG EGSHPERVKV YGPGVEKTGL  780
KANEPTYFTV DCSEAGQGDV SIGIKCAPGV VGPVEADIDF DIIKNDNDTF TVKYTPPGAG  840
HYTIMVLFAN QEIPASPFHI KVDPSHDASK VKAEGPGLSR TGVEVGKPTH FTVLTKGAGK  900
AKLDVHFAGA AKGEAVRDFE IIDNHDYSYT VKYTAVQQGN MAVTVTYGGD PVPKSPFVVN  960
VAPPLDLSKV KVQGLNSKVA VGQEQAFSVN TRGAGGQGQL DVRMTSPSRR PIPCKLEPGG  1020
GAEAQAVRYM PPEEGPYKVD ITYDGHPVPG SPFAVEGVLP PDPSKVCAYG PGLKGGLVGT  1080
PAPFSIDTKG AGTGGLGLTV EGPCEAKIEC QDNGDGSCAV SYLPTEPGEY TINILFAEAH  1140
IPGSPFKATI QPVFDPSKVR ASGPGLERGK AGEAATFTVD CSEAGEAELT IEILSDAGVK  1200
AEVLIQNNAD GTYHITYSPA FPGTYTITIK YGGHPIPKFP TRVHVQPAVD TSGIKVSGPG  1260
VEPHGVLREV TTEFTVDARS LTATGGNHVT ARVLNPSGAK TDTYVTDNGD GTYRVQYTAY  1320
EEGVHLVEVL YDEVAVPKSP FRVGVTEGCD PTRVRAFGPG LEGGLVNKAN RFTVETRGAG  1380
TGGLGLAIEG PSEAKMSCKD NKDGSCTVEY IPFTPGDYDV NITFGGQPIP GSPFRVPVKD  1440
VVDPGKVKCS GPGLGTGVRA RVPQTFTVDC SQAGRAPLQV AVLGPTGVAE PVEVRDNGDG  1500
THTVHYTPAT DGPYTVAVKY ADQEVPRSPF KIKVLPSHDA SKVRASGPGL NASGIPASLP  1560
VEFTIDARDA GQGLLTVQIL DPEGKPKKAN IRDNGDGTYT VSYLPDMSGR YTITIKYGGD  1620
EIPYSPFRIH ALPTGDASKC LVTVSIGGHG LGACLGPRIQ IGEETVITVD AKAAGKGKVT  1680
CTVSTPDGAE LDVDVVENHD GTFDIYYTAP EPGKYVITIR FGGEHIPNSP FHVLACDPLP  1740
HVEEPAEMLQ MRQPYAPLRP GTCPTHWATE EPVVPVEPLE SMLRPFNLVI PFTVQKGELT  1800
GEVRMPSGKT ARPNITDNKD GTITVRYAPT EKGLHQMGIK YDGNHIPGSP LQFYVDAINS  1860
RHVSAYGPGL SHGMVNKPAT FTIVTKDAGE GGLSLAVEGP SKAEITCKDN KDGTCTVSYL  1920
PTAPGDYSII VRFDDKHIPG SPFTAKITGD DSMRTSQLNV GTSTDVSLKI TEGDLSQLTA  1980
SIRAPSGNEE PCLLKRLPNR HIGISFTPKE VGEHVVSVRK SGKHVTNSPF KILVGPSEIG  2040
DASKVRVWGK GLSEGQTFQV AEFIVDTRNA GYGGLGLSIE GPSKVDINCE DMEDGTCKVT  2100
YCPTEPGTYI INIKFADKHV PGSPFTVKVT GEGRMKESIT RRRQAPSIAT IGSTCDLNLK  2160
IPGNWFQMVS AQERLTRTFT RSSHTYTRTE RTEISKTRGG ETKREVRVEE STQVGGDPFP  2220
AVFGDFLGRE RLGSFGSITR QQEGEASSQD MTAQVTSPSG KTEAAEIVEG EDSAYSVRFV  2280
PQEMGPHTVT VKYRGQHVPG SPFQFTVGPL GEGGAHKVRA GGTGLERGVA GVPAEFSIWT  2340
REAGAGGLSI AVEGPSKAEI AFEDRKDGSC GVSYVVQEPG DYEVSIKFND EHIPDSPFVV  2400
PVASLSDDAR RLTVTSLQET GLKVNQPASF AVQLNGARGV IDARVHTPSG AVEECYVSEL  2460
DSDKHTIRFI PHENGVHSID VKFNGAHIPG SPFKIRVGEQ SQAGDPGLVS AYGPGLEGGT  2520
TGVSSEFIVN TQNAGSGALS VTIDGPSKVQ LDCRECPEGH VVTYTPMAPG NYLIAIKYGG  2580
PQHIVGSPFK AKVTGPRLSG GHSLHETSTV LVETVTKSSS SRGASYSSIP KFSSDASKVV  2640
TRGPGLSQAF VGQKNSFTVD CSKAGTNMMM VGVHGPKTPC EEVYVKHMGN RVYNVTYTVK  2700
EKGDYILIVK WGDESVPGSP FKVNVP                                       2726

SEQ ID NO: 37          moltype = AA  length = 780
FEATURE                Location/Qualifiers
source                 1..780
```

-continued

```
                      mol_type = protein
                      organism = Mus musculus
SEQUENCE: 37
MAPYSLLVTR LQKALGVRQY HVASVLCQRA KVAMSHFEPS EYIRYDLLEK NINIVRKRLN  60
RPLTLSEKIV YGHLDDPANQ EIERGKTYLR LRPDRVAMQD ATAQMAMLQF ISSGLPKVAV  120
PSTIHCDHLI EAQVGGEKDL RRAKDINQEV YNFLATAGAK YGVGFWRPGS GIIHQIILEN  180
YAYPGVLLIG TDSHTPNGGG LGGICIGVGG ADAVDVMAGI PWELKCPKVI GVKLTGSLSG  240
WTSPKDVILK VAGILTVKGG TGAIVEYHGP GVDSISCTGM ATICNMGAEI GATTSVFPYN  300
HRMKKYLSKT GRTDIANLAE EFKDHLVPDP GCQYDQVIEI NLNELKPHIN GPFTPDLAHP  360
VADVGTVAEK EGWPLDIRVG LIGSCTNSSY EDMGRSAAVA KQALAHGLKC KSQFTITPGS  420
EQIRATIERD GYAQILRDVG GIVLANACGP CIGQWDRKDI KKGEKNTIVT SYNRNFTGRN  480
DANPETHAFV TSPEIVTALA IAGTLKFNPE TDFLTGKDGK KFKLEAPDAD ELPRSDFDPG  540
QDTYQHPPKD SSGQRVDVSP TSQRLQLLEP FDKWDGKDLE DLQILIKVKG KCTTDHISAA  600
GPWLKFRGHL DNISNNLLIG AINIENGKAN SVRNAVTQEF GPVPDTARYY KKHGIRWVVI  660
GDENYGEGSS REHAALEPRH LGGRAIITKS FARIHETNLK KQGLLPLTFA DPSDYNKIHP  720
VDKLTIQGLK DFAPGKPLKC VIKHPNGTQE TILLNHTFNE TQIEWFRAGS ALNRMKELQQ  780

SEQ ID NO: 38          moltype = AA  length = 110
FEATURE                Location/Qualifiers
source                 1..110
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 38
MSMTDVLSAE DIKKAIGAFA AADSFDHKKF FQMVGLKKKN PDEVKKVFHI LDKDKSGFIE  60
EDELGSILKG FSSDARDLSA KETKTLLAAG DKDGDGKIGV EEFSTLVAES             110

SEQ ID NO: 39          moltype = AA  length = 142
FEATURE                Location/Qualifiers
source                 1..142
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 39
MVLSGEDKSN IKAAWGKIGG HGAEYGAEAL ERMFASFPTT KTYFPHFDVS HGSAQVKGHG  60
KKVADALASA AGHLDDLPGA LSALSDLHAH KLRVDPVNFK LLSHCLLVTL ASHHPADFTP  120
AVHASLDKFL ASVSTVLTSK YR                                           142

SEQ ID NO: 40          moltype = AA  length = 142
FEATURE                Location/Qualifiers
source                 1..142
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 40
MSLMKNERAI IMSMWEKMAA QAEPIGTETL ERLFCSYPQT KTYFPHFDLH HGSQQLRAHG  60
FKIMTAVGDA VKSIDNLSSA LTKLSELHAY ILRVDPVNFK LLSHCLLVTM AARFPADFTP  120
EVHEAWDKFM SILSSILTEK YR                                           142

SEQ ID NO: 41          moltype = AA  length = 646
FEATURE                Location/Qualifiers
source                 1..646
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 41
MSKGPAVGID LGTTYSCVGV FQHGKVEIIA NDQGNRTTPS YVAFTDTERL IGDAAKNQVA  60
MNPTNTVFDA KRLIGRRFDD AVVQSDMKHW PFMVVNDAGR PKVQVEYKGE TKSFYPEEVS  120
SMVLTKMKEI AEAYLGKTVT NAVVTVPAYF NDSQRQATKD AGTIAGLNVL RIINEPTAAA  180
IAYGLDKKVG AERNVLIFDL GGGTFDVSIL TIEDGIFEVK STAGDTHLGG EDFDNRMVNH  240
FIAEFKRKHK KDISENKRAV RRLRTACERA KRTLSSSTQA SIEIDSLYEG IDFYTSITRA  300
RFEELNADLF RGTLDPVEKA LRDAKLDKSQ IHDIVLVGGS TRIPKIQKLL QDFFNGKELN  360
KSINPDEAVA YGAAVQAAIL SGDKSENVQD LLLLDVTPLS LGIETAGGVM TVLIKRNTTI  420
PTKQTQTFTT YSDNQPGVLI QVYEGERAMT KDNNLLGKFE LTGIPPAPRG VPQIEVTFDI  480
DANGILNVSA VDKSTGKENK ITITNDKGRL SKEDIERMVQ EAEKYKAEDE KQRDKVSSKN  540
SLESYAFNMK ATVEDEKLQG KINDEDKQKI LDKCNEIISW LDKNQTAEKE EFEHQQKELE  600
KVCNPIITKL YQSAGGMPGG MPGGFPGGGA PPSGGASSGP TIEEVD                 646

SEQ ID NO: 42          moltype = AA  length = 182
FEATURE                Location/Qualifiers
source                 1..182
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 42
MGDEEKRNRA ITARRQHLKS VMLQIAATEL EKEESRRESE KENYLSEHCP PLHIPGSMSE  60
VQELCKQLHA KIDVAEEEKY DMEVKVQKSS KELEDMNQKL FDLRGKFKRP PLRRVRMSAD  120
AMLKALLGSK HKVCMDLRAN LKQVKKEDTE KERDLRDVGD WRKNIEEKSG MEGRKKMFES  180
ES                                                                 182

SEQ ID NO: 43          moltype = AA  length = 405
```

-continued

```
FEATURE                 Location/Qualifiers
source                  1..405
                        mol_type = protein
                        organism = Mus musculus
SEQUENCE: 43
MRATDRMGAR AVSELRLALL FVLVLGTPRL GVQGEDGLDF PEYDGVDRVI NVNAKNYKNV  60
FKKYEVLALL YHEPPEDDKA SQRQFEMEEL ILELAAQVLE DKGVGFGLVD SEKDAAVAKK  120
LGLTEEDSVY VFKGDEVIEY DGEFSADTLV EFLLDVLEDP VELIEGEREL QAFENIEDEI  180
KLIGYFKSKD SEHYKAYEDA AEEFHPYIPF FATFDSKVAK KLTLKLNEID FYEAFMEEPM  240
TIPDKPNSEE EIVSFVEEHR RSTLRKLKPE SMYETWEDDL DGIHIVAFAE EADPDGYEFL  300
ETLKAVAQDN TENPDLSIIW IDPDDFPLLV PYWEKTFDID LSAPQIGVVN VTDADSIWME  360
MDNEEDLPSA DELEDWLEDV LEGEINTEDD DDDDDDDDDD DDDDD                  405

SEQ ID NO: 44           moltype = AA  length = 8886
FEATURE                 Location/Qualifiers
source                  1..8886
                        mol_type = protein
                        organism = Mus musculus
SEQUENCE: 44
MDHSFSGAPR FLTRPKAFVV SVGKDATLSC QIVGNPTPHV SWEKDRQPVE AGARFRLAQD  60
GDVYRLTILD LALGDSGQYV CRARNAIGEA FAAVGLRVDS EGTCAEQAPH FLLRPTSIRV  120
REGADATFRC RVGGSPQPAV SWSKDGRRLG PPDAPHVRVE EHGESSALRI RSARPRDGGT  180
YEVRAENPLG SASAAAALVV DSDAEVAGPP GTSTATLLAH LQQRREAMRA EGIPPSPPGA  240
GTRTCTVTEG KHARLSCFVT GEPKPETVWK KDGQLVTEGR RHVVYEDEQE NFVLKILFCK  300
QSDRGLYTCT ASNLVGQTYS SVLVVVREPT VPFKKRLQDL EVREKESATF QCEVAQPATE  360
AAWFKEETRL WASAKYDIEE EGTERRLTVR NVSADDDAVY ICETTEGSRT VAELSVQGNL  420
TRKLPRKTAV RTGDTAIFWV ELAVPEGPVR WLRNQEEMVA GGRVAITAEG TCHTLTIFQC  480
TLEDMGEVAF VSGGCRTTTQ FCVSAPRRPP LYPPADPVVK AKTESSVTLS WSAPPHGDRP  540
VTIDGYVLEK RKLGAYAWSR CHEAGWLATT EFTITGVAEE GDFQFRVSAI NHFGQGPYLE  600
FPGTMHLVPM LAVKTPLKAV EAVEGGEVTF SVDLTVASSG EWFLDGEALK ASSIYVIRCD  660
RTRHMLTIRE VPARLHGAQL KFVANGIETS IQMVVRAALG LPSSKLPAAA AREVLAQLHE  720
EAQLLAELSD QAAAVTWLKD GRELSLGPKY EMQVSAGRRA LLVRDVAQDD AGLYECVSRG  780
SRTAYQLLVQ DITDGYRDWG PAGPQKHMCK CAGAKIARYL GSSCYRFLQY DKGVWHWLEA  840
ALDTRQGKGT SSCSLHEKPK LVFAKGQQAH SEVKAEAGNS ATLSCEVTQA QTEVTWFKDG  900
KKLSSSSKVR MEASGCSRRL VVQQAGKADA GEYSCEAGGQ KLSFRLDVAE PKLVFAKGQQ  960
AHSEVKAEAG ASATLSCEVA QAQTEVTWFK DGKKLSSSSK VRMEASGCSR RLVVQQAGKA  1020
DAGEYSCEAG GQKLSFRLDV AEPKMVFAKE QQARSEVKAE AGASATLSCE VAQAQTEVTW  1080
FKDGKKLSSS SKVRMEASGC SRRLVVQQAG KADAGEYSCE AGGQKLSFRL DVTEPKLVFA  1140
KEQQARSEVK AEVGNSATLS CEVAQAQTEV TWFKDGKKLS SSSKVRMEAS GCSRRLVVQQ  1200
AGKADAGEYS CEAGGQKLSF RLDVAEPKLV FAKEQQARSE VKAEAGNSAT LSCEVAQAQT  1260
EVTWFKDGKK LSSSSKVRME ASGCSRRLVV QQAGKADAGE YSCEAGGQKL SFHLDVTEPK  1320
LVFAKEQQAH SEVKAEAGAS ATLSCEVAQA QTEVTWFKDG KKLSSSSKVR MEASGCSRRL  1380
VVQQAGKADA GEYSCEAEGQ KLSFRLDVAE PKLVFAKEQQ ARSEVKAEAG ASATLSCEVA  1440
QAQTEVTWFK DGKKLSSSSK VRVEASGCSR RLVVQQAGKA DAGEYSCEAG GQKLSFRLDV  1500
AEPKLVFAKE QQANSEVKAE AGASATLSCE VAQAQTEVTW FKDGKKLSSS SKVRVEASGC  1560
SRRLVVQQAG KADAGEYSCE AGGQKLSFRL DVAEPKLAFA KEQQAHSEVK AEAGASATLS  1620
CEVAQAQTEV TWFKDGKKLS SSSKVRVEAS GCSRRLVVQQ AGKADAGEYS CEAGGQKLSF  1680
RLDVAEPKLA FAKEQQAHSE VKAEAGASAT LSCEVAQAQT EVTWFKDGKK LSSSSKVRVE  1740
ASGCSRRLVV QQAGKADAGE YSCEAGGQKL SFRLDVAEPK LAFAKEQQAH SEVKAEAGAS  1800
ATLSCEVAQA QTEVTWFKDG KKLSSSSKVR VEASGCSRRL VVQQAGKADA GEYSCEAGGQ  1860
KLFFRLDVAE PKLMFAKEQQ AHSEVKAEAG ASATLSCEVA QAQTEVTWFK DGKKLSSSSK  1920
VRVEASGCSR RLVVQQAGKA DAGEYSCEAG GQKLSFRLDV AEPKLVFAKE QQAHSEVKAE  1980
AGASATLSCE VAQAQTEVTW FKDGKKLSSS SKVRVEASGC SRRLVVQQVG KADAGEYSCE  2040
ARGQKLSFRL DVADTRLMFA KEQQARTEVK AEAGNSATLS CEVAQAQTEV TWFKDGKKLS  2100
SSSKVRVEAS GCSRRLVVQQ AGKADAGEYS CEAGGQKLSF RLDVAEAESQ IPERPSRREP  2160
LVVKEHETII LTATIAAPSV AAVTWLKDGV EIRRSKRHEA TSLGDTHTLT VRGAQVLDSA  2220
IYSCRVGKEG QDFPVQVEEV AAKFSKPLEP VEGELGGTVM LVCELSPEQA EVVWRCGNTQ  2280
LRPGKRFQMT SEGPRRTLTV SGLREDDAEE YVCESRDDRT SARLTVKVPR VVKFTSGLSA  2340
MVAEEGQEAT FQCVVSPSDA GVTWYKDGMQ LQPSEKFVMV ESGASRSLTI LGLTLEDAGQ  2400
VTVEAEGASS SAALRVREAP VLFKKKLEPQ TVEERTSVTL EVELTRPWPE VKWTRNAAVL  2460
TPSENVEIRA EGARHCLVLR SVGFADRGFF GCETPDDKTQ AKLNVEMRQV RLVRGLQEVE  2520
AKEQGTASMD VELSHAEVEG SWTRDGLRLQ PGPKCHLAVQ GPVHILTLSA LQPQDSGLVA  2580
FRAEGVHTSA RLIVTELPVS FTRVLQDVVA TQKEKVTLEC ELSRPVDVRW LKDGVELRAG  2640
KAIGIVAQGT CRSLVIYRCE TGDQGVYVCD ALDAQTSASL RVQGRTYTLI FRRVLAEDAG  2700
EVKFVAENAE SRAHLRVKEL PVTLLRPLRD KIAMEKHRGV LECQVSRASA QVRWFKGGVE  2760
LQSGPKYEVV SDGLYRKLVI NDVQPEDEDT YTCDAGNVKT SAQFFVEEQS ITIVRGLKDM  2820
TVMEPAPAWF ECETSIPSVR PPKWLLGKTV LQAGGNVGLE QDGTVHRLTL HKTCSTMTGP  2880
VHFTIGKSRS SAQLVVSDIP VVLTRPLEPK AGRELQSVVL SCDFRPAPKA VQWYKDDTPL  2940
SPSEKFKMAL EGQMAELRIL RLTPADAGVY RCQAGSAQSS AEVTVEAREV TVIQPLQDAE  3000
AMEEGRVCFS CELSHKDEDI EWSLNGTPLY NDSFHEISHE GCLHTLVLKS VRQADTGTVC  3060
ATSPKVSVSA RLVVKGKPVV FLKALDDVSA EERGTLTLQC EVSDPEARVV WRKDGVELGP  3120
SDKYDFLHKA GARGLTVHDL SHEDAGLYTC QVGSKETQSK VSVHDLHVGI TKRLKTVEVL  3180
EGESCSFECV LSHESPSDPA VWTVGGKTVG GSGHFHAVRQ GRKYTLTVKD AALSDAGEVV  3240
FSVLGLTSKA SLIIRERPVD ITKPLEDQRT TLGEDVMLSC ELSRAGTSVR WLKDGKAIRK  3300
SQKYDLLSEG TRAVLVVRKA SLKDSGEYTC ETEASKSTAK LCVEEKANRF TEELADLQVE  3360
EKGRAVFTCK TEHPASVVTW RKGLLELRAS GKHVPSQEGL TLKLTINALE RTDSDTYTCD  3420
```

-continued

```
IGQARTQARL LVHGQKVRVI EDLEDTAVQE GSSAKFCCRI APADYGPVHW FLDKTPLHSN  3480
ELNEITVQPG GYHVLTLRQL ALKDSGTVYF EAGDQRTSAA LRVTEKPSIF SRPLTDVTVT  3540
EGEDLTLVCE TTTVDSSVRW TKDGKTLRPS ARCQLSHEGC QAQLLITATT PQDGGRYKCE  3600
IGGASSSSIV RVHALPVRFR ESLKDVEVPE GKAATLRCVL SSVAAPVEWR HGDDVLKSSN  3660
KYSLRQEGAV LELVIRDLQP QDSGQYSCSF GDQTTSATLT VKTSSAQFVG KLRNKEATEG  3720
TTVTLRCELT KEAPVEWKKG TETLRNGDKY SLKQDGAVCE LQICSLLVAD AGEYSCVCGQ  3780
EKTSATLTVK ALLVHFVRRL RSKEATEGDT TTLQCELSKA APVEWRKGTE TLRDGDRYSL  3840
KQDGAVCELQ IRSLTIADAG EYLCTCGQEK TSATLTVRAL PAKFKDSLKN EEATEGTTAT  3900
LSCELSKAAP VTWKKGPKTL QSGDKYVLRQ DGAVCGLQIH GLTMADAGEY SCVCGQEKTS  3960
ATLTVRGLPA KFIEDLRSQE ATEGATAILR CELSKAAPVE WRKGSETLKD GDRYTLRQDG  4020
AVCELQIRGL AVVDTGTYSS LPTKFTEGLR NEEATEGTMA TLRCQMSKAA PVEWRKGSET  4080
LRDGDRYSLR QDGAMCELQI RGLTIEDSGE YTCVCGQEKT SATLSVKALP SRFIEDLRSQ  4140
EATEGTMATL RCQMSKTAPV EWKKGSETLR DGGRYSLRQD GPVCELQICD LAVEDAGEYS  4200
CVCGQEKTSA TLSIKALPPR FIEDLRSQEA TEGTMATLRC QMSKAAPVEW RKGSETLGDG  4260
GRYSLRQNGA VCELQIHDLA VEDTGEYSCV CGQEKTSATL NVKALPPRFI EDLRSQEATE  4320
GTMATLRCQM SKAAPVEWRK GSETLRDGGR YSLRQDGAVC ELQIHDLDVE DAGQYSCVCG  4380
QEKTSAVLTV DALPPKFTEG LKKEEATEGT MVTLRCQMSK EATVEWRKGA KTLSDGGRYS  4440
LRQDGAMCEL QICGLAVEDA GEYSCVCGQE KTSATLSVKA LPPRFIEDLR SQEATEGTMA  4500
TLRCQMSKAA PVEWRKGSET LRDGDRYSLR QDGAVCELQI RDLAVEDAGE YLCVCGQEKT  4560
SATLSVKALP PRFIEDLRSQ EATEGTMATL RCQMSKAAPV EWRKGSKTLR DGDRYSLRQD  4620
GAMCELQICD LAVEDTGDYS CVCGQEKTSA TLSVKALPPR FIEDLRSQEA REGTVATLRC  4680
RMSKAAPVEW RKGSETLKDG DRYSLRQEGN LCELQIRDLA VEDTEEYSCV CGQEKTSATL  4740
SVKALPAKFI EDLRSQEAPE SSTVTLRCKL SKKASVVWKK GSETLRNGAR YSLRQDGAVC  4800
ELEIRDLTVE DTGEYSCTCG QERTSATLSI MAPQVVFQQP LQNLQAEEGS MASLRCELSV  4860
PNAAMVWSKG GLELQGDTRR EARQQGCVAE LLLRDLRRED AGEYSCTCGS QTTSATLMVT  4920
AAPVRFLREL QAQDVDEGAT ARLRCELSRE AVSVEWRKGS LQLFPCAKYQ MVQEGTTAEL  4980
LVHGVEQEDA GEYTCDAGHT QSIARLSVRA PKPKFKTDLQ STEQEAGGTA RLCCQLSEAE  5040
PGTPVQWLKE GVELHVGSKY EMRRQGAVCE LLIHGLEAKD TGEYACLVGG QKTLASLRVK  5100
EPEVTIVRGL VDMEVQADED VEFTCKVSQA GATDVQWHLQ GLPLQSNEVT EVAVLADGCT  5160
HVLQLKGVTL EDAGTVSFHV GGLSSSAQLT VRVPEVTVLE PLKDVQLSEG QDAHFRCRLS  5220
RASGQEARWA LGGVPLQCNE MNDITVEQGT LYLLTLHKVT LEDAGTITLQ VGSCSSEAQL  5280
KVTAKNTVLR GLENVDALEG GEALFECQLS QPEVAAHTWL LDDEPVHTSE KVEVVYFENG  5340
LRHLLLLKNL KPQDSCRVTF LAGDVVTSAF LTVRGWRLEV LEPPHDASVK AGMQVRFTCI  5400
LSEAVPVGEA TWYINGAAIQ PDDTDWTVTT DGSHHALTLS NAQPQHAGEV TFAARDAVAS  5460
ARLSVLALPD PPEDAEVVGR SDHSVTLSWV APMSDGGGGL CGYRVEMKEA STGQWQLCHD  5520
LVPGPECVVD DLVPGKTYRF RVAAVGPAGA GEPVHLPQMV KIAPAPAPAP APAPAPAPET  5580
RQAVVGEDIC LELEVAADGG EVVWHKGTER IQPGGHFEVL SRGQRQMLVI KGFRTEDQGE  5640
YRCGPIQGLP SSGAATFNVV MTSGSGDEVP AQPSLPPEAA QEGDLHLLWE ALARKRRMSR  5700
EPTLDSISEL PEEDSRVQHL RQEAEETAPD LSEGYSTADE LARTGEADLS HTSSDDESRA  5760
GTPSLVTYLK KAGGPGISPL ASKHEAQVTT SVKPQKQQEP VVPTCPPPGD LSAADLMDPS  5820
LDKAAVKIQA AFKGYKVRKE MKQQEGPVFS RTFGDTEAQV GDVLRLECVL ATKTDMRACW  5880
LKDGIELTDG RHYHIDQLKD GTCSLLVTGL APTDSGRYTC QVSTKSGRVS HSACVVVSGT  5940
ESEAESSSGG ELDDAFRRAA RRLHRLFRTK SPAELSEEEL FLSADEGPGE PEEPADWQTY  6000
REDENFVCIR FESLAEARQA VTCFRNMFAT MGIGVEISLG EQGPRGVEMR IGKVAPTVTP  6060
AVPLAKTPGL QTSDAAPVFL TELQNQDVQD GYPMSFDCVV TGQPVPSVRW FKDGKLLEED  6120
DHYMINEDQQ GGHQLIITAV VPADMGVYRC LAENSMGVSS TKAELRVELT STDYDTAADA  6180
TETSSYFSAQ GYLSSREQEG TESDEGQLPQ VLEELKDLQV APGTRLAKFQ LKVKGYPAPK  6240
LYWFKDGQPL TTSDHIRMTD KKTLHTLEIV SVTREDSGQY AAYISNAVGA AYSSARLLVR  6300
GPSEPEEKPA SDVHERLVPP RILEKFTPKK VKRGSSITFS VKVEGHPAPS VHWLKEEAEK  6360
GVLWIGPDTP GYTMASSSKQ HSLVLLDVGR QHQGTYTCIA TNPAGQALCS ASLHISGLAK  6420
EEEQERVKEA LISSFLQGTS QAVSAQMSES AGFADLVGQS KGESLVAEEA HSHLSLAEVG  6480
TEEFLQKLTS QITEMVSAKI SQAKLQVPGG DSDEETKTPS ASPRHGRSRP SSSVQESSSE  6540
SEDGDSRGEI FDIYVVTADY LPLGAEQDAI ILREGQYVEV LDSAHPLRWL VRTKPTKSSP  6600
SRQGWVSPAY LDKRLKLSPE WGPTEAPEFP GEAVSEDEYR TRLSSVIQEL LSSEQAFVGE  6660
LQFLESHHMK HLERSPRVPA AVASQKTVIF RNVQDISHFH SSFLKELQGC GTDDDVAMCF  6720
IKNQEAFEKY LEFLVGRVQA ESVVVSTPVQ EFYKKYAEET LSAKDPTQPP PPPLQHYLEQ  6780
PVERVQKYQA LLKELIRNKA RNRQNCALLE QAYAVVSALP QRAENKLHVS LMENYPGTLE  6840
ALGEPIRQGH FIVWEGAPGA RMPWKGHNRH VFLFRNYLVI CKPRRDSRTD TFSYVFRNMM  6900
KLSSIDLNDQ VEGDDRAFEV WHEREDSVRK YLLQARTVII KNSWVKEICG IQQRLAQPVW  6960
RPPEFEEELA DCTAELGETV KLACRVTGTP KPIVSWYKDG KPVEVDPHHI LIEDPDGSCT  7020
LILDNLTGID SGQYMCFAAS AAGNASTLGK ILVQVPPRFV NKVRATPFVE GEDAQITCTV  7080
EGAPYPQIRW YKDGTLLAPG NRYRMLNEPR SGVLVLVIQA ASKEDLGHYE CELVNRLGST  7140
RGGGELYMQS PALRARDQHH REQIVAAVED TSVEGSAHSA QDGADQQAAS VLWRLLGSEA  7200
LGPSPGDLPN TRQSEPPAFE EAASQIPGAA SGTPEVSQPG THKGLEQETT SSGSQGWTVP  7260
IRVEGTAWPG AGTGQLLLDV HSQVIMETTQ RTYVCQAPDT GVTRAPSMQV TIEDVQVQVG  7320
DMAQFDAVIE GHPPPIVTWY KGSTQLTSSA RLSQRQDGTT YSLVLTDVAP HDAGVYTCVA  7380
NNAGGQVLCK AELLVHGGDK LDAENQVYRR KLHSFYDVQE EIGRGVFGFV KRVQHKGNKM  7440
FCAAKFIPLR SKTRAQAYQE RDILATLGHP LVTGLLDQFE TRKTLILILE LCSSEELLDR  7500
LFKKGVVTEA EVKVYIQQLV EGLHYLHSHG ILHLDIKPPN ILMVHPARED IKICDFGFAQ  7560
KITPSEPQYS KYGSPEFVSP EIIEQNPVSE GSDIWAMGVI SYLSLTCSSP FAGESDRATL  7620
LNVLEGRVSW SSPTAAHLSE DAKDFIKATL QRTPRARPST SQCLAHPWFL KSMPAEEAHF  7680
INTKQLKFLL ARSRWQRSLM SYKSILVMRS IPELLQGPPD SPSLGVARHL RGEASGASSS  7740
SSSSDNELAP FARAKSLPPS PVTHSPLLHP RGFLRPSASL PEETEASMPT ADAAVPASPQ  7800
SAGPPASPGC VPRHSVISSL FYQQAGEGAE RGNKTSGAKR HPARRRHLLK GGYIARALPG  7860
LREPLMEYSL LEEEAAREEQ ASLMTKTPSF ETALRLPSSS VREVPGRSHS LDNPPVTTGP  7920
SPEACKEQLL FPPSTGLTHE TTAKDRGHKE GFLQESVPFP PMSGDSRPGK QEGSSQDSCR  7980
```

-continued

```
GKPASSCHSE LGSGSQEGCG PPSSQSLGSL PPQSLKKELS TSCGPLFSEQ PQAAPFPTQV  8040
SPLLGSEKEP QDGSLSEGPV PVPSSSPGSA SQVDASLDTE GLSEAGDTCD FTPPPQRPQE  8100
QATTRKFSLE SRGGYAGVAG YGTFAFGGDA GGMLGQGPLW ARMAWAVSQS SEEQDEAATE  8160
SPQPLESLGP IAEASGVPLR TSPSLTPWEE VEQVSLVQIR DLSGDAEAAD TISLDISEVD  8220
PAYLNLSDLY DIKYLPFEFM IFRRVPKPIE QPESPGSETE AGQGLADFLE EAAWPWPGEL  8280
GLRAGLEITE EPEEPGDLEA LLGEAAVGRK RKWSPSRGLF QFPGRCLSGE EPVELGLRQR  8340
VKASMAHISR ILKGRPEGPE REGPPRKKAG LASFRLSGLK GRDQELSDEA VVLGQSVTLA  8400
CQVLAQPTAQ ATWSKDGVLL ESSGHLLISS TLKNFQLLTI LVVKEEDLGT YTCCVSNPLG  8460
TAVTTGVLRK AERPSSSPRP EVGELYKDAV LLVWKPVESC GPVTYIVQCC IEGGSWTTLA  8520
SDISDCCYLT GKLSRGGMYI FRTACVSKAG MGPYSSPSEQ VLLGGPNHLA SEEESSRGRP  8580
AQLLPSTKTF AFQMQIRRGR FSVVRQCREK ASGRALAAKI VPYQPEDKTA VLREYEALKR  8640
LHHPHLAQLH AAYLSPRHLV LILELCSGPE LLPSLAERES YSESDVKDYL WQMLSATQYL  8700
HAQHILHLDL RSENMMVTEY NLLKVIDLGN AQSLDQEKVP APENFKDYLE TMAPELLEGQ  8760
GAVPQTDIWA IGVTAFIMLS GEYPESSEGT RDLQKGLRKG LIRLSRCYAG LSGGAVAFLQ  8820
SSLCAQPWGR PCASTCLQCG WLTEEGPTGS RPTPVTFPTV RLRAFVRERE KRRALLYKKH  8880
NLAQVR                                                             8886

SEQ ID NO: 45          moltype = AA  length = 1532
FEATURE                Location/Qualifiers
source                 1..1532
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 45
MEHSKQIRIL LLNEMEKLEK TLFRLEQGFE LQFRLGPTLQ GKAVTVYTNY PLPGEAFNRE  60
KFRSLDWENP TEREDDSDKY CKLHLQQSGS FQYYFLQGNE KSGGGYIVVD PILRIGVDNH  120
VLPLDCVTLQ TFLAKCLGPF DEWESRLRVA KESGYNMIHF TPLQTLGLSR SCYSLADQLE  180
LNPDFSRPSK RYTWSDVGQL VEKLKREWNI LCITDVVYNH TATNSKWILE HPESAYNLVN  240
SPHLKPAWVL DRALWHFSCD VADGKYREKG VPALIENDQH MNCIRKIIWE DIFPRIQLWE  300
FFQVDVHKAV EQFRRLLSQE NRRVTKSEPK EHLKIIQDPE YRRRGCAVDM DTALATFIPH  360
DNGPAAIEEC CNWFRKRLEE LNSEKHHLTS CHQEQAVNCL LGNVFYERLA GHGPKLGPVT  420
RKYPLVTRYF TFPFGEMALS AEEALIHLPD KACFLMAHNG WVMGDDPLRN FAEPGSDVYL  480
RRELICWGDS VKLRYGNKPE DCPYLWAHMK KYTEITATHF QGVRLDNCHS TPLHVAEYML  540
DAARKLQPNL YVVAELFTGS EELDNIFVTR LGISSLIREA MSAYNSHEEG RLVYRYGGEP  600
VGSFVQPCLR PLMPAIAHAL FMDITHDNEC PIVHRSAYDA LPSTTVVSMA CCASGSTRGY  660
DELVPHQISV VAEERFYTKW NPGASPADTG DVNVHSGIIA ARCAINRLHQ ELGAKGFIQV  720
YVDQVDEDIV AVTRHSPSIH QSVVAVSRTA FRNPKTSFYS KEVPQMCIPG KIEEVVLEAR  780
TIERNTKPYK KDENSINGMP NMTVELREHI QLHESKIVRQ AGVATKGPNE YIQEIEFENL  840
SPGSVIIFRV SLDPHAQVAV GILRNHLTQF SSHFKSGSLA VDNADPILKI PFASIASKLT  900
LAELNQVLYR CESEEQEDGG GCYDIPNWSS LKYAGLQGLM SVLAEIRPKN DLGHPFCENL  960
RSGDWMIDYV SGRLISRSGS IAEVGKWLQA MFFYLKQIPR YLIPCYFDAI LIGAYTTLLD  1020
VAWKQMSSFV QNGSTFVKHL SLGSVQMCGV GKCPCLPLLS PSLLDVPCRL NEITKEKEQC  1080
CASLAAGLPH FSSGLFRCWG RDTFIALRGM LLVTGRYLEA RNIILAFAST LRHGLIPNLL  1140
GEGTYARYNC RDAVWWWLQC IQDYCRTVPN GLDILKCPVS RMYPTDDSAP LPAGTLDQPL  1200
FEVIQEAMQR HMQGIQFRER NAGPQIDRNM KDEGFNITAG IDEETGFVYG GNRFNCGTWM  1260
DKMGESDRAR NRGIPATPRD GSAVEIVGLC KSAVRWLLEL SKKNIFPYHE VRVKRHGKVV  1320
AVSYDEWNRK IQNNFEKLFH VSEDPSDPNE KHPNLVHKRG IYKDSYGASS PWCDYQLRPN  1380
FTIAMVVAPE LFTAEKAWKA LEIAEKKLLG PLGMKTLDPD DMVYCGVYDN ALDNDNYNLA  1440
RGFNYHQGPE WLWPIGYFLR AKLYFSKMMG PETAAKTVFL VKNVLSRHYV HLERSPWKGL  1500
PELTNENGQY CPFSCETQAW SMAVVLETLY DL                                1532

SEQ ID NO: 46          moltype = AA  length = 463
FEATURE                Location/Qualifiers
source                 1..463
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 46
MGKEKTHINI VVIGHVDSGK STTTGHLIYK CGGIDKRTIE KFEKEAAEMG KGSFKYAWVL  60
DKLKAERERG ITIDISLWKF ETTKYYITII DAPGHRDFIK NMITGTSQAD CAVLIVAAGV  120
GEFEAGISKN GQTREHALLA YTLGVKQLIV GVNKMDSTEP AYSEKRYDEI VKEVSAYIKK  180
IGYNPATVPF VPISGWHGDN MLEPSPNMPW FKGWKVERKE GNASGVSLLE ALDTILPPTR  240
PTDKPLRLPL QDVYKIGGIG TVPVGRVETG ILRPGMVVTF APVNITTEVK SVEMHHEALS  300
EALPGDNVGF NVKNVSVKDI RRGNVCGDSK ADPPQEAAQF TSQVIILNHP GQISAGYSPV  360
IDCHTAHIAC KFAELKEKID RRSGKKLEDN PKSLKSGDAA IVEMVPGKPM CVESFSQYPP  420
LGRFAVRDMR QTVAVGVIKN VEKKSGGAGK VTKSAQKAQK AGK                    463

SEQ ID NO: 47          moltype = AA  length = 413
FEATURE                Location/Qualifiers
source                 1..413
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 47
MAPPSVFAQV PQAPPVLVFK LTADFRDDPD PRKVNLGVGA YRTDESQPWV LPVVRKVEQK  60
IANDNSLNHE YLPILGLAEF RSCASRLVLG DNSLAIRENR VGGVQSLGGT GALRIGADFL  120
GRWYNGTDNK NTPIYVSSPT WENHNAVFSA AGFKDIRPYC YWDAEKRGLD LQGFLNDLEN  180
APEFSIFVLH ACAHNPTGTD PTPEQWKQIA AVMQRRFLFP FFDSAYQGFA SGDLEKDAWA  240
IRYFVSEGFE LFCAQSFSKN FGLYNERVGN LTVVGKESDS VLRVLSQMEK IVRITWSNPP  300
```

-continued

```
AQGARIVAAT LSDPELFKEW KGNVKTMADR ILTMRSELRA RLEALKTPGT WSHITEQIGM  360
FSFTGLNPKQ VEYLVNEKHI YLLPSGRINM CGLTTKNLDY VATSIHEAVT KIQ         413

SEQ ID NO: 48          moltype = AA  length = 194
FEATURE                Location/Qualifiers
source                 1..194
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 48
MEEKLKKAKI IFVVGGPGSG KGTQCEKIVQ KYGYTHLSTG DLLRAEVSSG SERGKKLSAI  60
MEKGELVPLD TVLDMLRDAM LAKVDSSNGF LIDGYPREVK QGEEFEQKIG QPTLLLYVDA  120
GAETMTQRLL KRGETSGRVD DNEETIKKRL ETYYNATEPV ISFYDKRGIV RKVNAEGTVD  180
TVFSEVCTYL DSLK                                                    194

SEQ ID NO: 49          moltype = AA  length = 558
FEATURE                Location/Qualifiers
source                 1..558
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 49
MAALTRNPQF QKLLEWHRAN SANLKLRELF EADPERFNNF SLNLNTNHGH ILVDYSKNLV  60
NKEVMQMLVE LAKSRGVEAA RDNMFSGSKI NYTENRAVLH VALRNRSNTP IKVDGKDVMP  120
EVNRVLDKMK SFCQRVRSGD WKGYTGKSIT DIINIGIGGS DLGPLMVTEA LKPYSKGGPR  180
VWFVSNIDGT HIAKTLASLS PETSLFIIAS KTFTTQETIT NAETAKEWFL EAAKDPSAVA  240
KHFVALSTNT AKVKEFGIDP QNMFEFWDWV GGRYSLWSAI GLSIALHVGF DHFEQLLSGA  300
HWMDQHFLKT PLEKNAPVLL ALLGIWYINC YGCETHALLP YDQYMHRFAA YFQQGDMESN  360
GKYITKSGAR VDHQTGPIVW GEPGTNGQHA FYQLIHQGTK MIPCDFLIPV QTQHPIRKGL  420
HHKILLANFL AQTEALMKGK LPEEARKELQ AAGKSPEDLE KLLPHKVFEG NRPTNSIVFT  480
KLTPFILGAL IAMYEHKIFV QGIMWDINSF DQWGVELGKQ LAKKIEPELE GSSAVTSHDS  540
STNGLISFIK QQRDTKLE                                                558

SEQ ID NO: 50          moltype = AA  length = 469
FEATURE                Location/Qualifiers
source                 1..469
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 50
MSQAYSSSQR VSSYRRTFGG APGFSLGSPL SSPVFPRAGF GTKGSSSSMT SRVYQVSRTS  60
GGAGGLGSLR SSRLGTTRAP SYGAGELLDF SLADAVNQEF LATRTNEKVE LQELNDRFAN  120
YIEKVRFLEQ QNAALAAEVN RLKGREPTRV AELYEEEMRE LRRQVEVLTN QRARVDVERD  180
NLIDDLQRLK AKLQEEIQLR EEAENNLAAF RADVDAATLA RIDLERRIES LNEEIAFLKK  240
VHEEEIRELQ AQLQEQQVQV EMDMSKPDLT AALRDIRAQY ETIAAKNISE AEEWYKSKVS  300
DLTQAANKNN DALRQAKQEM MEYRHQIQSY TCEIDALKGT NDSLMRQMRE LEDRFASEAN  360
GYQDNIARLE EEIRHLKDEM ARHLREYQDL LNVKMALDVE IATYRKLLEG EESRINLPIQ  420
TFSALNFRET SPEQRGSEVH TKKTVMIKTI ETRDGEVVSE ATQQQHEVL              469

SEQ ID NO: 51          moltype = AA  length = 419
FEATURE                Location/Qualifiers
source                 1..419
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 51
MASAFSKLLT GRNASLLFTT LGTSALTTGY LLNRQKVSAD AREQHKLFPP SADYPDLRKH  60
NNCMAECLTP TIYAKLRNKM TPSGYTLDQC IQTGVDNPGH PFIKTVGMVA GDEESYEVFA  120
DLFDPVIKLR HNGYDPRVMK HPTDLDASKI THGQFDERYV LSSRVRTGRS IRGLSLPPAC  180
SRAERREVEN VAITALEGLK GDLAGRYYKL SEMTEQDQQR LIDDHFLFDK PVSPLLTCAG  240
MARDWPDARG IWHNYDKTFL IWINEEDHTR VISMEKGGNM KRVFERFCRG LKEVERLIQE  300
RGWEFMWNER LGYILTCPSN LGTGLRAGVH VRIPKLSKDP RFSKILENLR LQKRGTGGVD  360
TAAVADVYDI SNIDRIGRSE VELVQIVIDG VNYLVDCEKK LERGQDIKVP PPLPQFGRK   419

SEQ ID NO: 52          moltype = AA  length = 3284
FEATURE                Location/Qualifiers
source                 1..3284
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 52
MRKHRHLPLV AVFSLLLSGI ATTHAQQHGD VKNGAAADIV FLVDSSWSAG KDRFLLVQEF  60
LSDVVESLAV GDNDFHFALV RLNGNPHTEF LLNTYHSKQE VLSHIVNMSY IGGSNQTGKG  120
LEYVIHSHLT EASGSRAADG VPQVIIVLTD GQSEDGFALP SAELKSADVN VFAVGVEGAD  180
ERALGEVASE PLSMHVFNLE NVTSLHGLVG NLVSCIHSSV NPERAGDKES LKDITAQDSA  240
DIIFLIDGSQ NTGNANFDVI RDFLVNVLER LSVGNQQVQV GVVQYSEEPI TMFSLNSYPS  300
KAAVLDAVKG LSLVGGESAN IGQALDFVVE NHFTRAGGSR VEEGVPQVLV LISAGPSSDE  360
IRDSVVALKQ ASVFSFGLGA QAASRAELQH IATDDSLVFT VPEFRSFGDL QEQILPYLVG  420
VAQRHIVLQP PAIVTQVMEV NKRDIVFLVD GSSSLGPSNF NAIRDFVTRV IQRLEIGQDL  480
VQVSVAQYAD TVKPEFYLNS YTNKRDAITA VRKMRALNGS ALYTGSSLDF VRNNLFTSSA  540
GHRAAEGVPK LLVLITGGKS LDEVSQPAQE LKRGSIMALA VGSKAADEDE LKEIAFDSSL  600
```

-continued

```
VFIPAEFRPA PLQNMLPSLM APLRTLTGTT EVHVNKRDII FLLDGSDNVG KNNFPYVRDF  660
VTNLVNSLDV GSDNIRVGLV QFSDTPVTEF SLDTYQTKSE LLAHLRRLQL KGGSGLNAGS  720
ALSYIHANHF TEAGGSRTRE HVPQLLLLLM AGPSEDAYLQ AANALVRSGV LTFCVGTNRA  780
DKAELEHIAF NPSLVYLMDD FRSLPSLPQQ LIQPLTTYVS GGVQEVPLSQ PESKRDILFL  840
FDGSVNVLGQ FPAVRDFLYR IIEELDVKPD GTRVAIAQFS DDVRLESRFS EHQTKAEILN  900
LVKKMKLKTG KALNLGYALD YALRNIFVRS AGSRIEDNVQ QFLVLLVAGR SSDAVAGPAS  960
SLKQRGVVPF IFQAKNANPS ELEQIVLSPA FILAAESLPK IGDLQSQIVS LLKAEQGSGP  1020
VSGEKDVVFL IDGSEGVRSG FPLLKDFVQR VVESLDVGPD RVRVALVQYS DRTRPEFYLN  1080
SHMDQQGVIS AIRRLTLLGG PTPNTGAALE FVLRNILTSS TGSRIAEGVP QLLIVLTAEP  1140
SGDDVRGPSV VLKQGGAVPI GIGIGNADIS EMQTISFIPD FAVAIPTFRE LGTIQQVISE  1200
RVIQLNREEL SSLKPILTPS TGAGVGSKKD VVFLIDGSRN AGPEFQYIRT LIERIVEYLD  1260
IGFDTTRVAV IQFSEDSKME FPLNAHFSKD EVQNAVRRLR PKGGSQVYIG NALEYVLKNI  1320
FQRPLGSRIE EGVPQFLVLI SSGKSDDEVD DSAVELKQFG VAPLTIARHT DQEELVKISL  1380
SPEYVYSVST FRELPRLEQK LLTPITTLTS QQIHQILAST RYPPSVVESD AADIVFLIDS  1440
SDAVKPDGIA HIRDFVSRIV RRLNIGPSKV RIGVVQFSND VFPEFYLKTH KSQSSVLEAI  1500
RRLRFKGGSP LNTGRALEFV ARNLFVKSAG SRIEDGVPQH LVLFLGGKSQ DDVARHAQVI  1560
SSSGIVSLGI GDRNIDRTDL QTITNDPRLV FTVREFRELP NIEERVMLSF GPSGATPQPP  1620
GVDLPSPSRP EKKKADIVFL LDGSINFRRD SFQEVLRFAS EIVDTVYEDG DSIRVGLVQY  1680
NSDPTDEFFL RDFSTKRQII DAINKVVYKG GRHANTRVGI EHLLRNHFVP EAGSRLDERV  1740
PQIAFVITGG KSVEDAQDVS LALTQKGVKV FAVGVRNIDS EEVGKIASNS ATAFRVGSVQ  1800
ELSELSETVL ETLHDAMHET LCPGVTDVSK ACNLEVILGF DGSRDQNVFV SQKGLESKVD  1860
IILNRISQIQ RISCSGNQLP TVRVSVMANT PSGPVEAFDF AEYQPELFEK FRNMRSQRPY  1920
VLTADTLKLY QNKFRQSSPD TVKVVIHFTD GADGDMADLY RASEELRQAG AQALILVGLE  1980
RVANLERLMH LEFGRGFMYD RPLRLNLLDL DYELAEQLDN IAEKACCGVP CKCSGERGDR  2040
GPIGSIGPKG ISGEDGYRGY PGDEGGPGER GPPGVNGTQG FQGCPGQRGV KGSRGFPGEK  2100
GELGEIGLDG LDGEEGDKGL PGSSGEKGSP GRRGDKGPKG DKGERGDVGI RGDPGDSGRD  2160
SQQRGPKGET GDIGPMGLPG RDGIPGSPGD PGKDGGSGRR GPAGAKGNRG GPGQPGFEGE  2220
QGTRGSQGPP GPIGPPGLIG EQGIPGPRGG GGTAGAPGER GRTGPLGRKG EPGEPGPKGS  2280
IGNRGPRGET GDDGRDGVGS EGRRGKKGER GFPGYPGPKG TPGEPGADGP PGPKGIRGRR  2340
GNSGPPGATG QKGDPGYPGP SGHKGNRGDS VDQCALIQSI RDKCPCCYGP LECPVFPTEL  2400
AFALDTSEGV TQDTFSRMRE VLLGIVGDLT IAESNCPRGA RVAVVTYNNE VTTEIRFADS  2460
KKKSALLDSI QNLQVALTSK QQSLETAMSF VARNTFKRVR SGFLMRKVAV FFSNKPTRAS  2520
PQLREAVLKL SDAGITPLFL TSQEDRQLIN ALQINNTAVG HALVLPARRD LTDFLKNVLT  2580
CHVCLDICNI DPSCGFGSWR PSFRDRRAAG SDVDIDLAFI LDSSEATTLF QFNEMKKYIG  2640
YVIRQLDLSP DPKASQHFAR VAVVQQSTYE SVDNASVPPV KVEFSLTDYG AKEKLLDFLS  2700
RRMTQLQGTM GLGNAIEYTI ENIFESAPNP RDLKIMVLML TGDMQRQQLE EAQRAILQAK  2760
CKGYFFVVLG IGRKVNIKEV YSFASEPNDV FFKFVDKSTE LNEEPLMRFG RLLPSFVSSE  2820
NAFYLPPDLR KQCDWFQGDQ PVKNGVKFGH KQINPPHTAN SSLTSKVVTT MKPVTTTKPT  2880
AIVNLPPAKP APARPAPAQP VLAKPDPAKP AQARPAPAKP ASAKLVPPQP VHVQPAPAQT  2940
ASVRPAPAKP APPQPAAAKP VPAKPAVPAQ PAPPQPAAAK PVPAKPAVPA QPAPPQPAAA  3000
KPVPAKPAVP AQPAAAQPMP AQPVLTKSAA VKPASANKPV AAKPVATNTA TATARPALAA  3060
KPAAAKPAAT RPLAAAIRPV ATKPEAPRQQ AKPAATKPAT TKPLARVSRE VQVSEVTENS  3120
ARLHWERPEP SSSFFYDLTV TSAHDQSLVL RQNLTVTDRV IGGLLAGQLY HVVVVSYLQS  3180
QVRAIYQGSF NTKKTQPPPL QAAHRASSST INLMVNTEPL FLTKTDICKL SRDAGTCVDF  3240
KLLWHYDLES KSCKRFWYGG CGGNENRFHS QEECEKMCSP DLLV                   3284

SEQ ID NO: 53          moltype = AA  length = 745
FEATURE                Location/Qualifiers
source                 1..745
                       mol_type = protein
                       organism = Mus musculus
SEQUENCE: 53
MPLFKLTGQG KQIDDAMRSF AEKVFASEVK DEGGRHEISP FDVDEICPIS LHEMQAHIFH  60
MENLSMDGRR KRRFQGRKTV NLSIPQSETS STKLSHIEEF ISSSPTYESV PDFQRVQITG  120
DYASGVTVED FEVVCKGLYR ALCIREKYMQ KSFQRFPKTP SKYLRNIDGE ALVGNESFYP  180
VFTPPPKKGE DPFRTEDLPA NLGYHLKMKA GVIYIYPDEA AANRDEPKPY PYPNLDDFLD  240
DMNFLLALIA QGPVKTYAHR RLKFLSSKFQ VHQMLNEMDE LKELKNNPHR DFYNCRKVDT  300
HIHAAACMNQ KHLLRFIKKS YHIDADRVVY STKEKSLTLK ELFAKLNMHP YDLTVDSLDV  360
HAGRQTFQRF DKFNDKYNPV GASELRDLYL KTDNYINGEY FATIIKEVGA DLVEAKYQHA  420
EPRLSIYGRS PDEWNKLSSW FVCNRIYCPN MTWMIQVPRI YDVFRSKNFL PHFGKMLENI  480
FLPVFEATIN PQAHPDLSVF LKHITGFDSV DDESKHSGHM FSSKSPKPEE WTMENNPSYT  540
YYAYYMYANI TVLNSLRKER GMNTFLFRPH CGEAGALTHL MTAFMIADNI SHGLNLKKSP  600
VLQYLFFLAQ IPIAMSPLSN NSLFLEYAKN PFLDFLQKGL MISLSTDDPM QFHFTKEPLM  660
EEYAIAAQVF KLSTCDMCEV ARNSVLQCGI SHEEKAKFLG NNYLEEGPVG NDIRRTNVAQ  720
IRMAYRYETW CYELNLIAEG LKATE                                        745

SEQ ID NO: 54          moltype = AA  length = 379
FEATURE                Location/Qualifiers
source                 1..379
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 54
FGNTHNKYKL NYKSEEEYPD LSKHNNHMAK VLTPDLYKKL RDKETPSGFT LDDVIQTGVD  60
NPGHPFIMTV GCVAGDEESY TVFKDLFDPI IQDRHGGFKP TDKHKTDLNH ENLKGGDDLD  120
PHYVLSSRVR TGRSIKGYTL PPHCSRGERR AVEKLSVEAL NSLTGEFKGK YYPLKSMTEQ  180
EQQQLIDDHF LFDKPVSPLL LASGMARDWP DARGIWHNDN KSFLVWVNEE DHLRVISMEK  240
```

-continued

```
GGNMKEVFRR FCVGLQKIEE IFKKAGHPFM WNEHLGYVLT CPSNLGTGLR GGVHVKLAHL  300
SKHPKFEEIL TRLRLQKRGT GGVDTAAVGS VFDISNADRL GSSEVEQVQL VVDGVKLMVE  360
MEKKLEKGQS IDDMIPAQK                                               379

SEQ ID NO: 55          moltype = AA  length = 900
FEATURE                Location/Qualifiers
source                 1..900
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 55
MMMVMQPEGL GAREGPFAGG GGGGEYMEQE EDWDRDLLLD PAWEKQQRKT FTAWCNSHLR  60
KAGTQIENIE EDFRNGLKLM LLLEVISGER LPRPDKGKMR FHKIANVNKA LDFIASKGVK  120
LVSIGAEEIV DGNLKMTLGM IWTIILRFAI QDISVEETSA KEGLLLWCQR KTAPYRNVNV  180
QNFHTSWKDG LALCALIHRH RPDLIDYAKL RKDDPIGNLN TAFEVAEKYL DIPKMLDAED  240
IVNTPKPDEK AIMTYVSCFY HAFAGAEQAE TAANRICKVL AVNQENEKLM EEYEKLASEL  300
LEWIRRTVPW LENRVGEPSM SAMQRKLEDF RDYRRLHKPP RVQEKCQLEI NFNTLQTKLR  360
LSHRPAFMPS EGKLVSDIAN AWRGLEQVEK GYEDWLLSEI RRLQRLQHLA EKFQQKASLH  420
ETWTRGKEDM LTQRDYESAS LQEVRALLRR HEAFESDLAA HQDRVEHIAA LAQELNELDY  480
HEAASVNSRC QAICDQWDNL GTLTQKRRDA LERMEKLLET IDQLQLEFAR RAAPFNNWLD  540
GAIEDLQDVW LVHSVEETQS LLTAHDQFKA TLPEADRERG AILGIQGEIQ KICQTYGLRP  600
SSTNPYINLS PQDINTKWDT VRKLVPSRDQ TLQEELARQQ MNERLRRQFA AQANAVGPWI  660
QAKVEEVGRL AAGLAGSLEE QMAGLRQQEQ NIINYKSNID RLEGDHQLLQ ESLVFDNKHT  720
VYSMEHIRVG WEQLLTSIAR TINEVENQVL TRDAKGLSQE QLNEFRASFN HFDRKRNGMM  780
EPDDFRACLI SMYDLGEVEF ARIMTMVDPN AAGVVTFQAF IDFMTRETAE TDTAEQVVAS  840
FKILAGDKNY ITPEELRREL PAEQAEYCIH RMAPYKGAGA PAGALDYVAF SSALYGESDL  900

SEQ ID NO: 56          moltype = AA  length = 842
FEATURE                Location/Qualifiers
source                 1..842
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 56
SRPLSDQEKR KQISVRGLAG VENVTELKKN FNRHLHFTLV KDRNVATPRD YYFALAHTVR  60
DHLVGRWIRT QQHYYEKDPK RIYYLSLEFY MGRTLQNTMV NLALENACDE ATYQLGLDME  120
ELEEIEEDAG LGNGGLGRLA ACFLDSMATL GLAAYGYGIR YEFGIFNQKI CGGWQMEEAD  180
DWLRYGNPWE KARPEFTLPV HFYGRVEHTS QGAKWVDTQV VLAMPYDTPV PGYRNNVVNT  240
MRLWSAKAPN DFNLKDFNVG GYIQAVLDRN LAENISRVLY PNDNFFEGKE LRLKQEYFVV  300
AATLQDIIRR FKSSKFGCRD PVRTNFDAFP DKVAIQLNDT HPSLAIPELM RVLVDLERLD  360
WDKAWEVTVK TCAYTNHTVL PEALERWPVH LLETLLPRHL QIIYEINQRF LNRVAAAFPG  420
DVDRLRRMSL VEEGAVKRIN MAHLCIAGSH AVNGVARIHS EILKKTIFKD FYELEPHKFQ  480
NKTNGITPRR WLVLCNPGLA EIIAERIGEE YISDLDQLRK LLSYVDDEAF IRDVAKVKQE  540
NKLKFAAYLE REYKVHINPN SLFDVQVKRI HEYKRQLLNC LHVITLYNRI KKEPNKFVVP  600
RTVMIGGKAA PGYHMAKMII KLITAIGDVV NHDPVVGDRL RVIFLENYRV SLAEKVIPAA  660
DLSEQISTAG TEASGTGNMK FMLNGALTIG TMDGANVEMA EEAGEENFFI FGMRVEDVDR  720
LDQRGYNAQE YYDRIPELRQ IIEQLSSGFF SPKQPDLFKD IVNMLMHHDR FKVFADYEEY  780
VKCQERVSAL YKNPREWTRM VIRNIATSGK FSSDRTIAQY AREIWGVEPS RQRLPAPDEK  840
IP                                                                 842

SEQ ID NO: 57          moltype = AA  length = 362
FEATURE                Location/Qualifiers
source                 1..362
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 57
HSHPALTPEQ KKELSDIAHR IVAPGKGILA ADESTGSIAK RLQSIGTENT EENRRFYRQL  60
LLTADDRVNP CIGGVILFHE TLYQKADDGR PFPQVIKSKG GVVGIKVDKG VVPLAGTNGE  120
TTTQGLDGLS ERCAQYKKDG ADFAKWRCVL KIGEHTPSAL AIMENANVLA RYASICQQNG  180
IVPIVEPEIL PDGDHDLKRC QYVTEKVLAA VYKALSDHHI YLEGTLLKPN MVTPGHACTQ  240
KYSHEEIAMA TVTALRRTVP PAVTGVTFLS GGQSEEEASI NLNAINKCPL LKPWALTFSY  300
GRALQASALK AWGGKKENLK AAQEEYVKRA LANSLACQGK YTPSGQAGAA ASESLFISNH  360
AY                                                                 362

SEQ ID NO: 58          moltype = AA  length = 191
FEATURE                Location/Qualifiers
source                 1..191
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 58
APKKDVKKPA AAAAPAPAPA PAPAPAPAKP KEEKIDLSAI KIEFSKEQQD EFKEAFLLYD  60
RTGDSKITLS QVGDVLRALG TNPTNAEVKK VLGNPSNEEM NAKKIEFEQF LPMLQAISNN  120
KDQGTYEDFV EGLRVFDKEG NGTVMGAELR HVLATLGEKM KEEEVEALMA GQEDSNGCIN  180
YEAFVKHIMS I                                                       191

SEQ ID NO: 59          moltype = AA  length = 333
FEATURE                Location/Qualifiers
source                 1..333
```

-continued

```
                      mol_type = protein
                      organism = Oryctolagus cuniculus
SEQUENCE: 59
MVKVGVNGFG RIGRLVTRAA FNSGKVDVVA INDPFIDLHY MVYMFQYDST HGKFHGTVKA  60
ENGKLVINGK AITIFQERDP ANIKWGDAGA EYVVESTGVF TTMEKAGAHL KGGAKRVIIS  120
APSADAPMFV MGVNHEKYDN SLKIVSNASC TTNCLAPLAK VIHDHFGIVE GLMTTVHAIT  180
ATQKTVDGPS GKLWRDGRGA AQNIIPASTG AAKAVGKVIP ELNGKLTGMA FRVPTPNVSV  240
VDLTCRLEKA AKYDDIKKVV KQASEGPLKG ILGYTEDQVV SCDFNSATHS STFDAGAGIA  300
LNDHFVKLIS WYDNEFGYSN RVVDLMVHMA SKE                               333

SEQ ID NO: 60          moltype = AA  length = 284
FEATURE                Location/Qualifiers
source                 1..284
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 60
MDAIKKKMQM LKLDKENALD RAEQAEADKK AAEDRSKQLE DELVSLQKKL KGTEDELDKY  60
SEALKDAQEK LELAEKKATD AEADVASLNR RIQLVEEELD RAQERLATAL QKLEEAEKAA  120
DESERGMKVI ESRAQKDEEK MEIQEIQLKE AKHIAEDADR KYEEVARKLV IIESDLERAE  180
ERAELSEGKC AELEEELKTV TNNLKSLEAQ AEKYSQKEDK YEEEIKVLSD KLKEAETRAE  240
FAERSVTKLE KSIDDLEDEL YAQKLKYKAI SEELDHALND MTSI                   284

SEQ ID NO: 61          moltype = AA  length = 1134
FEATURE                Location/Qualifiers
source                 1..1134
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 61
EAKPAARKAP KGKDAPKEEP PKEAAAEPPK EAPPEDQSPT AEEPTGIFLK KPDSVSVETG  60
KDAVIMAKLN GKELPAKPTI KWFKGKWLEL GSKSGARFSF KETHDAASNV YSLELHIGKV  120
VLGDRGDYRL EVKAKDVCDS CSFNIDVEAP RQDSAGQSLE SFKRAGDAKP DGAGELDFSG  180
LLKKREVVEE EKKKKKDDDD LGIPPEIWEL LKGAKKSEYE RIAFQYGITD LRGMLKRLKK  240
AKVEVKKSAA FTKKLDPAYQ VDRGNKVKLV VEISDPDLPL KWFKNGQEIK PSLKYVFENV  300
GKKRILTINK CTLADDAAYE VVVNDEKCFT ELFVKEPPVL IVTPLEDQQV FVGDRVEMSV  360
EVSEEGAQVM WMKDGVELTR EGSFKSNYRF KKDGKRHILI YSEVNLEDRG HYQVMTNGGQ  420
CEAELIVEEK QLEVLQDIAD LTVKASEQAM FKCEVSDEKV TGKWYKNGVE VRPSKRITIS  480
HVGRFHKLVI DDVRPEDEGD YTFVPDGYAL SLSAKLNFLE IKVEYVPKQE PPKIHLDCSG  540
KTSDNAIVVV AGNKLRLDVS ITGEPPPVAT WMKGDEVFTV TEGRVRIEQR PDSSSFVIES  600
AERGDEGRYT IKVTNPAGED VATVMLRVVD VPDPPEAVRV TSVGEDWAIL VWDPPKYDGG  660
QPVTGYLLER KKKGSQRWMK LNFEVFTETT YESTKMIEGI LYEMRVFAVN AIGVSQPSMN  720
TKPFMPIAPT SAPQHLTVED VTDTTTTLKW RPPDRIGAGG IDGYLVEYCV EGSDEWIPAN  780
KEPVERCGFT VKGLPTGAKI LFRVIGVNIA GRSEPATLVQ PVTIREIVEQ PKIRLPRHLR  840
QTYVRKVGEH INLVIPFQGK PRPQVVWTKG GAPVDTSRVH VRTSDFDTVF FVREAARSDS  900
GEYELSVQIE NMKDTATVRI RVVEKAGPAE NVMVKEVWGT NALVEWQPPK DNGNSEITGY  960
LVQKADKKTM EWFTVYEHNR HTSCTVSDLI VGNEYYFRVY TENVCGLSDS PGVSKNTAHI  1020
LKTGITFKPF EYKEHDFRSP PKFLTPLIDR VVVAGYAAAL NCAVRGHPKP KVVWMKNKVE  1080
IHEDPKFLMT NHQGVLTLNI RRPSPFDAGT YSCRAFNELG EALAECRLDV RVPQ        1134

SEQ ID NO: 62          moltype = AA  length = 1415
FEATURE                Location/Qualifiers
source                 1..1415
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 62
MSLVAVPFYQ KRHKHFDQSY RNIQTKYLMD EYTAKKRAST QMSTHKSAST QMATHKSALQ  60
KASSQTSSGG TTCRLCAKRL SAAEEEEQES KNRYRSLVAS YGEAKRQRFL SELAQLEESV  120
HLARSQARDV VDKYALQQMV EDKMAWERHV FEERLSRAPE ILVQLRSHTV WEKMSVKLCF  180
TVQGFPTPVV QWYKDGSLIC QAGEPGKYRI DSKYGVHTLE ISRASFDDTA TYSAVATNAH  240
GQVSTNAAVV VRRYRGDEEP FYSVGLPIGY VLVKESKWTR MFFGEGQASL SFSHLNKDDE  300
GLYTLRIVSR GGVSDHSAFL FVRDADPLVT GAPGAPMDLQ CHDANRDYVI VTWKPPNTTT  360
ESPVIGYFVD RCEVGTNNWV QCNDAPVKIC KYPVTGLYEG RSYIFRVRAV NSAGISRPSR  420
VSDAVAALDP ADLRRLQAIH LEGEKEVVIY QDDLEGDVQV PGPPTNVHAS EISRNYVVLS  480
WEPPVPRGKE PLMYFIEKSV VGSGSWQRVN AQTAVRSPRY AVFDLAEGKP YVFRVLSANK  540
HGLSDPSEIT SPIQAQDTIA VPSAPGRVLA SRNTKTSVVV QWDRPKHEED LLGYYVDCCV  600
AGTNAWEPCN HKPIGYNRFV VHGLTTGEQY IFRVKAVNAL GTSENSQESE VIKVQAALTV  660
PSHPYGVTLL NCDGHSMTLG WKVPKFSGGS PILGYYVDKR EVHHKNWHEV NSSPIQERIL  720
TVEGLTEGSF YEFKIAATNL AGIGQPSDPS EHFKCEAWTA PEPGPAYDLT FCEVRDTSLV  780
MLWKAPVYSG SSPVSGYFVD FKEEDSGEWS TVNEAATPNR YLKVSDLHQG RTYVFRVRAV  840
NASGVGKPSD VSEPVLVEAR PGTKEISAGV DEEGNIYLGF DCQEMTDASH FTWCKSYEEI  900
ADPERFHVET KGDHSKLYFK NPDKADLGTY SVSVSDTDGV SSSFVLDEEE LERLLALSNE  960
IKNPTIPLKS ELAYEIFDKG QVRFWLQTEP LSPDASFRFV INDREVSDSE THRMKCDKAS  1020
GIIEMVMDRF TIENEGTYTV QIHDGKAKNQ SSLVLIGDAF RKVLEEAEFQ RKEFLRKQGP  1080
HFAEYLHWDV TEECEVRLIC KVANTKKETV FKWLKDDVLY ETETPPDLEK GVCELLIPKL  1140
SKKDHGEYKA TLKDERGQDA SILEIAGKVY EDMILAMSRV CGMSASPLKI LCTPEGIRLQ  1200
CFMKYFTEEM RVNWYHKDAK ISSSEHMRIG GSEDMAWLQI CEPTEKDKGK YTFEIFDGKD  1260
NHQRSLDLSG QAFDEAFAEF QRLKAAAFAE KNRGKVIGGL PDVVTIMEGK TLNLTCTVFG  1320
```

-continued

```
NPDPEVVWFK NDQDIELSEH FSVKVEQAKY VSLTIKGVTS EDSGKYSINV KNKYGGEKID  1380
VTVSVYKHGE AIPEVAPPQQ AKPKLIPASA SATAP                             1415

SEQ ID NO: 63          moltype = AA  length = 433
FEATURE                Location/Qualifiers
source                 1..433
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 63
AMQKIFAREI LDSRGNPTVE VDLHTAKGRF RAAVPSGAST GIYEALELRD GDKSRYLGKG  60
VLKAVEHINK TLGPALLEKK LSVVDQEKVD KFMIELDGTE NKSKFGANAI LGVSLAVCKA  120
GAAEKGVPLY RHIADLAGNH DLVLPVPAFN VINGGSHAGN KLAMQEFMIL PVGASSFREA  180
MRIGAEVYHH LKGVIKAKYG KDATNVGDEG GFAPNILENN EALELLKTAI QAAGYPDKVV  240
IGMDVAASEF HRNGKYDLDF KSPDDPARHI TGQKLGELYK SFIKNYPVVS IEDPFDQDDW  300
ATWTSFLSGV DIQIVGDDLT VTNPKRIAQA VEKKACNCLL LKVNQIGSVT ESIQACKLAQ  360
SNGWGVMVSH RSGETEDTFI ADLVVGLCTG QIKTGAPCRS ERLAKYNQLM RIEEALGDKA  420
VFAGRKFRNP KAK                                                     433

SEQ ID NO: 64          moltype = AA  length = 169
FEATURE                Location/Qualifiers
source                 1..169
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 64
APKKAKRRAA AEGGSSNVFS MFDQTQIQEF KEAFTVIDQN RDGIIDKEDL RDTFAAMGRL  60
NVKNEELDAM MKEASGPINF TVFLTMFGEK LKGADPEDVI TGAFKVLDPE GKGTIKKQFL  120
EELLTTQCDR FSQEEIKNMW AAFPPDVGGN VDYKNICYVI THGDAKDQE              169

SEQ ID NO: 65          moltype = AA  length = 1273
FEATURE                Location/Qualifiers
source                 1..1273
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 65
IFSISDNAYQ FMLTDRENQS ILITGESGAG KTVNTKRVIQ YFATIAITGD KKKEEPTPGK  60
MQGTLEDQII SANPLLEAFG NAKTVRNDNS SRFGKFIRIH FGATGKLASA DIETYLLEKS  120
RVTFQLKAER SYHIFYQIMS NKKPELIEML LITTNPYDFA FVSQGEITVP SIDDQEELMA  180
TDSAVDILGF TSDEKVAIYK LTGAVMHYGN LKFKQKQREE QAEPDGTEVA DKAAYLTSLN  240
SADLLKALCY PRVKVGNEYV TKGQTVQQVY NAVGALAKAI YEKMFLWMVT RINQQLDTKQ  300
PRQYFIGVLD IAGFEIFDFN SLEQLCINFT NEKLQQFFNH HMFVLEQEEY KKEGIEWEFI  360
DFGMDLAACI ELIEKPMGIF SILEEECMFP KATDTSFKNK LYEQHLGKSN NFQKPKPAKG  420
KAEAHFSLVH YAGTVDYNIT GWLDKNKDPL NETVVGLYQK SGMKTLAFLF SGAQTGEEGG  480
GGGKKGGKKK GSSFQTVSAL FRENLNKLMT NLRSTHPHFV RCLIPNETKT PGAMEHELVL  540
HQLRCNGVLE GIRICRKGFP SRILYADFKQ RYKVLNASAI PEGQFIDSKK ASEKLLGSID  600
VDHTQYKFGH TKVFFKAGLL GLLEEMRDEK LAQLITRTQA ICRGYLTRVE FKKMMERRES  660
IFCIQYNIRA FMNVKHWPWM KLYFKIKPLL KSAETEKEMA NMKEEFEKTK EELAKSEAKR  720
KELEEKMVAL MQEKNDLQLQ VQAEADGLAD AEERCDQLIK TKIQLEAKIK EVTERAEDEE  780
EINAELTAKK RKLEDECSEL KKDIDDLELT LAKVEKEKHA TENKVKNLTE EMAGLDENIA  840
KLTKEKKALQ EAHQQTLDDL QAEEDKVNTL TKAKTKLEQQ VDDLEGSLEQ EKKLRMDLER  900
AKRKLEGDLK LAQESTMDIE NDKQQLDEKL KKKEFEMTNL QSKIEDEQAL AMQLQKKIKE  960
LQARIEELEE EIEAERASRA KAEKQRSDLS RELEEISERL EEAGGATSAQ IEMNKKREAE  1020
FQKMRRDLEE ATLQHEATAA ALRKKHADSV AELGEQIDNL QRVKQKLEKE KSELKMEIDD  1080
LAGNMETVSK AKGNLEKMCR TLEDQLSEVK TKEEEHQRLI NELSAQKARL HTESGEFSRQ  1140
LDEKDAMVSQ LSRGKQAFTQ QIEELKRQLE EETKAKSALA HALQSSRHDC DLLREQYEEE  1200
QEAKAELQRA MSKANSEVSQ WRTKYETDAI QRTEELEEAK KKLAQRLQDA EEHVEAVNSK  1260
CASLEKTKQR LQN                                                     1273

SEQ ID NO: 66          moltype = AA  length = 1161
FEATURE                Location/Qualifiers
source                 1..1161
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 66
IFYQIMSNKK PELIEMLLIT TNPYDFAFVS QGEITVPSID DQEELMATDS AVDILGFTSD  60
EKVAIYKLTG AVMHYGNLKF KQKQREEQAE PDGTEVADKA AYLTSLNSAD LLKALCYPRV  120
KVGNEYVTKG QTVQQVYNAV GALAKAIYEK MFLWMVTRIN QQLDTKQPRQ YFIGVLDIAG  180
FEIFDFNSLE QLCINFTNEK LQQFFNHHMF VLEQEEYKKE GIEWEFIDFG MDLAACIELI  240
EKPMGIFSIL EEECMFPKAT DTSFKNKLYE QHLGKSNNFQ KPKPAKGKAE AHFSLVHYAG  300
TVDYNITGWL DKNKDPLNET VVGLYQKSGM KTLAFLFSGA QTGEEGGGGG KKGGKKKGSS  360
FQTVSALFRE NLNKLMTNLR STHPHFVRCL IPNETKTPGA MEHELVLHQL RCNGVLEGIR  420
ICRKGFPSRI LYADFKQRYK VLNASAIPEG QFIDSKKASE KLLGSIDVDH TQYKFGHTKV  480
FFKAGLLGLL EEMRDEKLAQ LITRTQAICR GYLTRVEFKK MMERRESIFC IQYNIRAFMN  540
VKHWPWMKLY FKIKPLLKSA ETEKEMANMK EEFEKTKEEL AKSEAKRKEL EEKMVALMQE  600
KNDLQLQVQA EADGLADAEE RCDQLIKTKI QLEAKIKEVT ERAEDEEEIN AELTAKKRKL  660
EDECSELKKD IDDLELTLAK VEKEKHATEN KVKNLTEEMA GLDENIAKLT KEKKALQEAH  720
QQTLDDLQAE EDKVNTLTKA KTKLEQQVDD LEGSLEQEKK LRMDLERAKR KLEGDLKLAQ  780
```

-continued

```
ESTMDIENDK QQLDEKLKKK EFEMTNLQSK IEDEQALAMQ LQKKIKELQA RIEELEEEIE  840
AERASRAKAE KQRSDLSREL EEISERLEEA GGATSAQIEM NKKREAEFQK MRRDLEEATL  900
QHEATAAALR KKHADSVAEL GEQIDNLQRV KQKLEKEKSE LKMEIDDLAG NMETVSKAKG  960
NLEKMCRTLE DQLSEVKTKE EEHQRLINEL SAQKARLHTE SGEFSRQLDE KDAMVSQLSR  1020
GKQAFTQQIE ELKRQLEEET KAKSALAHAL QSSRHDCDLL REQYEEEQEA KAELQRAMSK  1080
ANSEVSQWRT KYETDAIQRT EELEEAKKKL AQRLQDAEEH VEAVNSKCAS LEKTKQRLQN  1140
EVEDLMIDVE RSNAACAALD K                                            1161

SEQ ID NO: 67          moltype = AA  length = 667
FEATURE                Location/Qualifiers
source                 1..667
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 67
TLDDLQAEED KVNTLTKAKT KLEQQVDDLE GSLEQEKKLR MDLERAKRKL EGDLKLAQES  60
TMDIENDKQQ LDEKLKKKEF EMTNLQSKIE DEQALAMQLQ KKIKELQARI EELEEEIEAE  120
RASRAKAEKQ RSDLSRELEE ISERLEEAGG ATSAQIEMNK KREAEFQKMR RDLEEATLQH  180
EATAAALRKK HADSVAELGE QIDNLQRVKQ KLEKEKSELK MEIDDLAGNM ETVSKAKGNL  240
EKMCRTLEDQ LSEVKTKEEE HQRLINELSA QKARLHTESG EFSRQLDEKD AMVSQLSRGK  300
QAFTQQIEEL KRQLEEETKA KSALAHALQS SRHDCDLLRE QYEEEQEAKA ELQRAMSKAN  360
SEVSQWRTKY ETDAIQRTEE LEEAKKKLAQ RLQDAEEHVE AVNSKCASLE KTKQRLQNEV  420
EDLMIDVERS NAACAALDKK QRNFDKVLAE WKHKYEETQA ELEASQKESR SLSTEVFKVK  480
NAYEESLDHL ETLKRENKNL QQEISDLTEQ IAESGKHIHE LEKVKKQIDQ EKSELQAALE  540
EAEGSLEHEE GKILRIQLEL NQVKSEIDRK IAEKDEEIDQ LKRNHLRVVE SMQSTLDAEI  600
RSRNDALRIK KKMEGDLNEM EIQLNHANRQ AAEAIKNLRN TQGILKDTQL HLDDAVRGQD  660
DLKEQLA                                                            667

SEQ ID NO: 68          moltype = AA  length = 830
FEATURE                Location/Qualifiers
source                 1..830
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 68
LAQESTMDIE NDKQQLDEKL KKKEFEMTNL QSKIEDEQAL AMQLQKKIKE LQARIEELEE  60
EIEAERASRA KAEKQRSDLS RELEEISERL EEAGGATSAQ IEMNKKREAE FQKMRRDLEE  120
ATLQHEATAA ALRKKHADSV AELGEQIDNL QRVKQKLEKE KSELKMEIDD LAGNMETVSK  180
AKGNLEKMCR TLEDQLSEVK TKEEEHQRLI NELSAQKARL HTESGEFSRQ LDEKDAMVSQ  240
LSRGKQAFTQ QIEELKRQLE EETKAKSALA HALQSSRHDC DLLREQYEEE QEAKAELQRA  300
MSKANSEVSQ WRTKYETDAI QRTEELEEAK KKLAQRLQDA EEHVEAVNSK CASLEKTKQR  360
LQNEVEDLMI DVERSNAACA ALDKKQRNFD KVLAEWKHKY EETQAELEAS QKESRSLSTE  420
VFKVKNAYEE SLDHLETLKR ENKNLQQEIS DLTEQIAESG KHIHELEKVK KQIDQEKSEL  480
QAALEEAEGS LEHEEGKILR IQLELNQVKS EIDRKIAEKD EEIDQLKRNH LRVVESMQST  540
LDAEIRSRND ALRIKKKMEG DLNEMEIQLN HANRQAAEAI KNLRNTQGIL KDTQLHLDDA  600
VRGQDDLKEQ LAMVERRANL MQAEIEELRA SLEQTERSRR VAEQELLDAS ERVQLLHTQN  660
TSLINTKKKL ETDISQIQGE MEDIVQEARN AEEKAKKAIT DAAMMAEELK KEQDTSAHLE  720
RMKKNMEQTV KDLQHRLDEA EQLALKGGKK QIQKLEARVR ELENEVESEQ KRNVEAVKGL  780
RKHERRVKEL TYQTEEDRKN VLRLQDLVDK LQSKVKAYKR QAEEAEEQSN             830

SEQ ID NO: 69          moltype = AA  length = 1218
FEATURE                Location/Qualifiers
source                 1..1218
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 69
VFVADPKESF VKATVQSREG GKVTAKTEAG ATVTVKEDQV FPMNPPKYDK IEDMAMMTHL  60
HEPAVLYNLK ERYAAWMIYT YSGLFCVTVN PYKWLPVYNA EVVTAYRGKK RQEAPPHIFS  120
ISDNAYQFML TDRENQSILI TGESGAGKTV NTKRVIQYFA TIAVTGDKKK EEATSGKMQG  180
TLEDQIISAN PLLEAFGNAK TVRNDNSSRF GKFIRIHFGT TGKLASADIE TYLLEKSRVT  240
FQLKAERSYH IFYQIMSNKK PDLIEMLLIT TNPYDYAFVS QGEITVPSID DQEELMATDS  300
AIDILGFTSD ERVSIYKLTG AVMHYGNMKF KQKQREEQAE PDGTEVADKA AYLQSLNSAD  360
LLKALCYPRV KVGNEYVTKG QTVQQVYNAV GALAKAVYEK MFLWMVTRIN QQLDTKQPRQ  420
YFIGVLDIAG FEIFDFNSLE QLCINFTNEK LQQFFNHHMF VLEQEEYKKE GIEWEFIDFG  480
MDLAACIELI EKPMGIFSIL EEECMFPKAT DTSFKNKLYE QHLGKSNNFQ KPKPAKGKVE  540
AHFSLVHYAG TVDYNITGWL DKNKDPLNET VVGLYQKSAM KTLAFLFTGT AAAEAEGGGK  600
KGGKKKGSSF QTVSALFREN LNKLMTNLRS THPHFVRCII PNETKTPGAM EHELVLHQLR  660
CNGVLEGIRI CRKGFPSRIL YADFKQRYKV LNASAIPEGQ FIDSKKASEK LLGSIDVDHT  720
QYKFGHTKVF FKAGLLGLLE EMRDDKLAQL ITRTQAMCRG FLARVEYKKM VERRESIFCI  780
QYNIRAFMNV KHWPWMKLYF KIKPLLKSAE TEKEMANMKE EFEKTKESLA KAEAKRKELE  840
EKMVALMQEK NDLQLQVQAE ADSLADAEER CDQLIKTKIQ LEAKIKEVTE RAEDEEEINA  900
ELTAKKRKLE DECSELKKDI DDLELTLAKV EKEKHATENK VKNLTEEMAG LDETIAKLTK  960
EKKALQEAHQ QTLDDLQAEE DKVNTLTKAK TKLEQQVDDL EGSLEQEKKI RMDLERAKRK  1020
LEGDLKLAQE STMDIENDKQ QLDEKLKKKE FEMSNLQSKI EDEQALAMQL QKKIKELQAR  1080
IEELEEEIEA ERASRAKAEK QRSDLSRELE EISERLEEAG GATSAQIEMN KKREAEFQKM  1140
RRDLEEATLQ HEATAATLRK KHADSVAELG EQIDNLQRVK QKLEKEKSEL KMEIDDLASN  1200
METVSKAKGN LEKMCRTL                                                1218
```

-continued

```
SEQ ID NO: 70          moltype = AA  length = 1574
FEATURE                Location/Qualifiers
source                 1..1574
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 70
TLVSDVDYRN YLHQWTCMPD QNDVIQAKKA YELQSDAVYK ADLEWLRGIG WMPSDSVSVN  60
HAKHAADIFS EKKYRTKIET LNFTPVDDRV DYVTAKQSGE ILDDIKYRKD WNSIKSKYTL  120
TETPLLHTVQ EAARILDQYL YKESWEKQKA TGYILPPDAV PFVHAHHSSD VQSELKYKAE  180
HVKQRGHYVG VPTMRDDPKL VWFEHAGQIQ NDRLYKEDYH KTKAKINIPA DMVSVVAAKE  240
GQNLVSDIDY RNYLHQWTCH PDQNDVIQAR KAYDLQSDNV YKADLEWLRG IGWIPLDSVD  300
HVRVTKNQEM VNQIKYKKDA LDNYPNFTSV VDPPAIVLAK INAVNQSDVK YKETFNKAKG  360
KYTFSPDTPH ISHSKDMGKL YSTILYKGAW EGTKAYGYTL DERYIPIVGA KHADFVNSEL  420
KYKETYEKQK GHYLAGKVIS EFPGVVHCLD FQKMRSVLNY RRDYENTKAN VHIPNDMMNH  480
VLAKKCQYIL SDLEYRHYFH QWTSLPDEPK VLLARNAQEI LSDNVYKDDL NWLKGIGCYV  540
WDTPQILHAK KSYDLQSQLQ YTAAGKENLQ NYNLVTDTPL YVTALQSGIN ASEVKYKENY  600
HQIKDKYTTV LETVDYDRTR NLKDLYSSNL YKEAWNKVKA TSYILPTSTL SLTHAKNQKH  660
LASHIKYREE YEKFKALYTL PKSVEDDPNT ARCLRVGKLN IDRLYRSVYE KNKMKIHIVP  720
DMVEMVTAKD TQKKVSEIDY RLHLHEWICH PDLQVNSHVR KVTDQISDIV YKDDLNWLKG  780
IGCYVWDTPE ILHAKHAYDL RNDIKYKAHM MKTRNDYKLV TDTPVYVQAV KSGKQLSDAV  840
YHHDYVHSVR GKVAPTTKTV DLERALHAYK LQSENLYKTS LRTLPTGYRL PGDTPHFKHN  900
KDIRYMSSYF KYKEVYEHIK ANGYTLGPND VPFVNVRRVN KVTSERLYRE LYHKLKDKIH  960
TTPDTPEIRQ VKKTQEAVSE LIYKSDFFKM QGHMISLPYT PHVLHCRYVG DITSDIKYKE  1020
DLQILKGLGC FLYDTPDMVR SRHLRKLWSN YLYTDNARKM RDKYKVVLDT PEYRKVQELK  1080
THLSELVYRA SGRKQKSIFT SVPDTPDLLR AKRGQKLQSQ YLYVELATKE RPHHHAGNQT  1140
TALKHARDVK DMVSEKKYKV QYEKMKDKYT PIPDTPILIR AKNAYWNASD LRYKETFQKT  1200
KGKYHTVKDA LDIVYHRKVT DDISKVKYKE NYMRHLGIWR SIPDRPEHFH HRAVTDAVSD  1260
IKYKEDLTWL KGIGCYAYDT PDFTLAEKNK TLYSKYKYKE VFERTKSDFK YVADCPINRH  1320
FKYATQLMNE KKYRADYEQR KDKYQLVVDE PRHLLAKTAG DQISQIKYRK QYEKSKDKFT  1380
SIVDTPEHLR TTKVNKQISD ILYKLEYNKA KPRGYTTIHD TPMLLHVRKV KDEVSDLKYK  1440
EVYQRNKSNC TIKPDAVHIK AAKDAYKVNT NLDYKKQYEA NKAHWRWTPD RPDFIQAAKS  1500
SLQQSDFEYK LDREFLKGCK LSVTDDKDMV LALKNSLIES DLKYKEKHVK ERGSCHAVPD  1560
TPQILLAKTV SNLV                                                    1574

SEQ ID NO: 71          moltype = AA  length = 1500
FEATURE                Location/Qualifiers
source                 1..1500
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 71
EHTRNVNQIQ SDNVYKDEYN NFFKGIGWIP IGSLEVEKVK KAGDALNEKK YRQHPDTIKF  60
TSVPDSMGMV LAQHNTKQLS DLNYKVEGEK VKHKYTMDPE VPQFIQARVN AFNLSDAHYK  120
ADWKKTLAKG YDLRPDAIPI VAAKSSRNIA SDCKYKEAYE KAKGKQIGFL SLQDDPKLVH  180
YMNVAKIQSD REYRKGYEAS KTRYHTPLDM LSVTAAKKSQ EVATNANYKQ PFHNYTLLPD  240
ALPVEHSRNA MQIQSDNLYK SDFTGWMKGI GWIPIQSLEV EKAKKAGEIL SEKRYRQHPE  300
KLKFTYAMDT MEQALNKANK LNMDKRLYTK KWDKDKTTIH VMPDTPDILL SRVNQITMSD  360
KLYRAGWEEE KKKGYDLRPD AISIKAAKAS RDIASDYKYK QAYEQAKGKH IGFRSLEDDP  420
KLVHFMQVAK MQSDREYKKA YEKSKTSFHT PVDMLSVVAA KKSQEVATNA NYRNVIHTYN  480
MLPDAMGFEL AKNMMQIQSD NQYKADYADF MKGIGWLPLG SLEAEKNKKA MEIISEKKYR  540
QHPDTLKYST LMDSMNMVLA QNNAKIMNEH LYKQAWEADK TKVHIMPDIP QIVLAKANAI  600
NMSDKLYKLS MEEAKKKGYD LRTDAIPIKA AKASRDIASD YKYKYNYEKQ KGKMVGFRSL  660
EDDPKLVHSM QVAKMQSDRE YKKNYEKTKT SYHTPADMLS VTAAKDAQAN ITNTNYKHLI  720
HKYILLPDAM SLELSRNMNR IQSDNEYKQD YNEWYKGLGW SPAGSLDVEK AKKATEYASD  780
QKYRQHPSNF QFKKLTDSMD MVLAKQNAHT MNKHLYTIDW NKDKTKIHVM PDTPDILQAK  840
QNQTLFSQKF YKLGWEEALK KGYDLPVDAI SVQLAKASRD IASDFKYKQG YRKQLGHHVG  900
FRSLQDDPKL VLSMNVAKMQ SEREYKKDFE KWKTKFSSPV DMLGVVLAKK CQALVSDVDY  960
KNYLHQWTCL PDQNDVIHAK KAYELQSENM YKSDLEWLKG IGWSPLGSLE AEKNKRASEI  1020
LSEKKYRQPP DTNKFTSIPD AMDIVLAKTN AKNRSDRLYR EAWDKDKTQV HIMPDTPDII  1080
LAKANLINTS DKLYRMGYEE LKKKGYDLPV DAIPIKAAKA SREIASEYKY KEGFRKQLGH  1140
HIGARNIKDD PKMMWSMHVA KIQSDREYKK DFEKWKTKFS SPVDMLGVVL AKKCQTLVSD  1200
IDYKNYLHQW TCLPDQNDVT HARHAYDLQS DNLYKSDLQW LKGIGWIPTG SIEDEKNKRA  1260
TQILSDHVYR QHPDKFKFSI LMDSIPMVLA KNNATNMNHR LYTEAWDKDK TTVHIMPDTP  1320
EVLLAKQNQV NCSEKLYKLG LEETKRKGYD MRVDAIPIKA AKASRDIASE FKYKEGYRKQ  1380
LGHHIGARAI HDDPKMMWSM HVAKIQSDRE YKKDFEKWKT KFSSPVDMLG VVLAKKCQAL  1440
VSDVDYKNYL HQWTCLPDQN DVMHARQAYD LQSDNLYKSD LQWLRGIGWV PIGSLDVEKC  1500

SEQ ID NO: 72          moltype = AA  length = 1400
FEATURE                Location/Qualifiers
source                 1..1400
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 72
DVTHARHAYD LQSDNLYKSD LQWLKGIGWI PTGSIEDEKN KRATQILSDH VYRQHPDKFK  60
FSILMDSIPM VLAKNNATNM NHRLYTEAWD KDKTTVHIMP DTPEVLLAKQ NQVNCSEKLY  120
KLGLEETKRK GYDMRVDAIP IKAAKASRDI ASEFKYKEGY RKQLGHHIGA RAIHDDPKMM  180
WSMHVAKIQS DREYKKDFEK WKTKFSSPVD MLGVVLAKKC QALVSDVDYK NYLHQWTCLP  240
```

-continued

```
DQNDVMHARQ AYDLQSDNLY KSDLQWLRGI GWVPIGSLDV EKCKKATEIL SDKIYRQPPD  300
RFKFTSVTDS LEQVLAKNNA INMNKRLYTE AWDKDKTQIH IMPDTPEIML ARMNKINYSE  360
SLYKLANEEA KKKGYDLRSD AIPIVAAKAS RDIASDYKYK DGYRKQLGHH IGARDIKDDP  420
KMMWSMHVAK IQSDREYKKD FEKWKTKFSS PVDMLGVVLA KKCQTLVSDI DYKNYLHQWT  480
CLPDQNDVMH ARQAYDLQSD NIYKSDLQWL RGIGWVPIGS LEVAKCKRAA EILSDNIYRQ  540
PPDKLKFTSV PDSLEQVLAK NNAINMNKRL YTEAWDKDKT QIHIMPDTPE IMLARQNKIN  600
YSESLYKLAN EEAKKKGYDL RSDAIPIVAA KASRDIASDY KYKDGYRRQL GHHIGARDIK  660
DDPKMMWSMH VAKIQSDREY KKDFEKWKTK FSSPVDMLGV VLAKKCQTLV SDIDYKNYLH  720
QWTCLPDQND VMHARQAYDL QSDNIYKSDL QWLRGIGWVP IGSVDLVKCK RAAEILSDNI  780
YRQSPDKLKF TSVTDSLEQV LAKNNAINMN KRLYTEAWDK DKTQIHMMPD TPEIMLARQN  840
KMNYSENFYR QAMEEAKKEG YDLRSDAIPI VAAKASRDIA SDYKYKEAYR KQLGHHIGAR  900
AIHDDPKMMW SLHIAKVQSD REYKKEFEKY KTRYSSPVDM LGIVLAKKCQ TLVSDVDYKN  960
LLHEWTCLPD QNDVIQARKA YDLQSDNLYK SDLEWLRGIG WVPIGSLEVV KAKRATEILS  1020
DNLYRTRPDT LKFTSVTDSL EQVLAKNNAI NMNKRLYTEA WENDKKTIHV MPDTPEIMLA  1080
KLNKINYSDK LYKLALEESK KEGYDLRVDA IPIQAAKASR DIASDYKYKE GYRKQLGHHI  1140
GARNIKDDPK MMWSMHVAKI QSDREYKKDF EKWKTKFSSP VDMLGVVLAK KCQILVSDID  1200
YKHPLHEWIC LPDQNDVIQA RKAYDLQSDA VYKADLEWLR GIGWVPIGSV EVEKVKRAGE  1260
ILSDRKYRQP ADKLKFTSIT DTPEIVLAKS NALTMSKHLY TEAWDADKTS IHVMPDTPEI  1320
LLAKSNSANI SQKLYTKGWD ESKMKDYDLK ADAIPIKSAK ASRDIASDYK YKEAYEKQKG  1380
HHIGAQSIED DPKIMCAIHA                                              1400
```

```
SEQ ID NO: 73          moltype = AA  length = 1500
FEATURE                Location/Qualifiers
source                 1..1500
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 73
DKRETSRPNW AQVSANVPIT SCSVEKLIEG HEYQFRICAE NKYGVGDPVF TEPAIAKNPY  60
DPPGRCDPPV ISNVTKDHMT VSWKPPADDG GSPITGYLLE KRETHAVNWT KVNRKPVIER  120
TIKATGLQEG TEYEFRVTAI NKAGPGKPSD ASKAVYAQDP LYPPGPPAFP KVYDTTRSSV  180
SLSWGKPAYD GGSPIIGYLV EVKRADTDNW VRCNLPQKLQ KTRFEVTGLM ENTEYQFRVY  240
AVNKVGYSDP SDVPDKHCPK DILIPPEGEL DADLRKTLIL RAGVTMRLYV PVKGRPPPKI  300
TWSKPNVNLR ERIGLDIKST DFDTFLRCEN VNKYDAGKYI LTLENSCGKK EYTIVVKVLD  360
TPGPPVNVTV KEISRDSAYI TWDPPIVDGG SPIINYVVEK RDAERKSWST VTTECSKTSF  420
RVSNLEEGKS YFFRVFAENE YGIGDPGETR DAVKASETPG PVVDLKVLTV TKSSCNIGWK  480
KPRSDGGSRI TGYVVDFLTE ENKWQRVMKS LSLQYSTKDL NEGKQYTFRV SAENENGEGT  540
PSEITVVAKD DVVAPDLDLK DLPDLCYLAK ENSNFRLKIP IKGKPAPSVT WKKGEDPLAT  600
DTRVSVESSA VNTTLVVYDC QKSDAGKYTI TLKNVAGTKE GTLSIKVVGK PGIPTGPIKF  660
DEVTAEAITL KWGPPKDDGG SEITNYILEK RDSVNNKWVT CASAVQKTTF RVTRLHEGME  720
YTFRVSAENK YGVGEGLKSE PIVAKHPFDV PDAPPPPNIV DVRHDSVSLT WTDPRKTGGS  780
PITGYHIEFK ERNSLLWKRA NKTPIRMKDF KVTGLTEGLE YEFRVMAINL AGVGKPSLPS  840
EPVVALDPID PPGKPEVINV TRNSVTLIWT EPKYDGGHKL TGYIVEKRDL PSKTWMKANH  900
INVPDCAFTV TDLVEGGKYE FRIRAKNTAG AISAPSESTG TIICKDEYEA PTIVLDPTIK  960
DGLTIKAGDT IVLNAISILG KPLPKSSWSK AGKDIRPSDI TQITSTPTSS MLTVKYASRK  1020
DAGEYTITAT NPFGTKEEHV RVTVLDVPGP PGPIEISNVS AEKATLTWTP PLEDGGSPIK  1080
SYVLEKRETS RLLWTVVAED IQSCRHVVTK LIQGNEYLFR VSAVNHYGKG EPVQSEPVKM  1140
VDRFGPPGPP GKPEVSNVTK NTATVSWKRP TDDGGSEITG YYVERREKKG LRWVRATKTP  1200
VSDLRCKVTG LQEGNTYEFR VSAENRAGIG PPSDASNYVL MKDVAYPPGP PSNARVTDTT  1260
KKSASLAWGK PHYDGGLEIT GYVVEHQKVG DETWVKDTTG PALRITEFVV PDLHTKEKYN  1320
FRISAINDAG VGEPAVIPDV EIVEREMAPD FELDAELRRT LVVRAGLSIR IFVPIKGRPA  1380
PEVTWTKDDI NLKTRANIEN TESFTLLIIP ECNRYDTGKF VMTIENPAGK KSGFVNVRVL  1440
DTPGPVLNLR PTDITKDSVT LHWDLPLIDG GSRITNYIVE KREATRKSYS TVTTKCHKCT  1500
```

```
SEQ ID NO: 74          moltype = AA  length = 338
FEATURE                Location/Qualifiers
source                 1..338
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 74
DKETPSGFTL DDVIQTGVDN PGHPFIMTVG CVAGDEESYT VFKDLFDPII QDRHGGFKPT  60
DKHKTDLNHE NLKGGDDLDP HYVLSSRVRT GRSIKGYTLP PHCSRGERRA VEKLSVEALN  120
SLTGEFKGKY YPLKSMTEQE QQQLIDDHFL FDKPVSPLLL ASGMARDWPD ARGIWHNDNK  180
SFLVWVNEED HLRVISMEKG GNMKEVFRRF CVGLQKIEEI FKKAGHPFMW NEHLGYVLTC  240
PSNLGTGLRG GVHVKLAHLS KHPKFEEILT RLRLQKRGTG GVDTAAVGSV FDISNADRLG  300
SSEVEQVQLV VDGVKLMVEM EKKLEKGQSI DDMIPAQK                          338
```

```
SEQ ID NO: 75          moltype = AA  length = 808
FEATURE                Location/Qualifiers
source                 1..808
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 75
GGGGGGEYME QEEDWDRDLL LDPAWEKQQR KTFTAWCNSH LRKAGTQIEN IEEDFRNGLK  60
LMLLLEVISG ERLPRPDKGK MRFHKIANVN KALDFIASKG VKLVSIGAEE IVDGNLKMTL  120
GMIWTIILRF AIQDISVEET SAKEGLLLWC QRKTAPYRNV NVQNFHTSWK DGLALCALIH  180
RHRPDLIDYA KLRKDDPIGN LNTAFEVAEK YLDIPKMLDA EDIVNTPKPD EKAIMTYVSC  240
```

-continued

```
FYHAFAGAEQ AETAANRICK VLAVNQENEK LMEEYEKLAS ELLEWIRRTV PWLENRVGEP  300
SMSAMQRKLE DFRDYRRLHK PPRVQEKCQL EINFNTLQTK LRLSHRPAFM PSEGKLVSDI  360
ANAWRGLEQV EKGYEDWLLS EIRRLQRLQH LAEKFQQKAS LHETWTRGKE DMLTQRDYES  420
ASLQEVRALL RRHEAFESDL AAHQDRVEHI AALAQELNEL DYHEAASVNS RCQAICDQWD  480
NLGTLTQKRR DALERMEKLL ETIDQLQLEF ARRAAPFNNW LDGAIEDLQD VWLVHSVEET  540
QSLLTAHDQF KATLPEADRE RGAILGIQGE IQKICQTYGL RPSSTNPYIN LSPQDINTKW  600
DTVRKLVPSR DQTLQEELAR QQMNERLRRQ FAAQANAVGP WIQAKVEEVG RLAAGLAGSL  660
EEQMAGLRQQ EQNIINYKSN IDRLEGDHQL LQESLVFDNK HTVYSMEHIR VGWEQLLTSI  720
ARTINEVENQ VLTRDAKGLS QEQLNEFRAS FNHFDRKRNG MMEPDDFRAC LISMYDLGEV  780
EFARIMTMVD PNAAGVVTFQ AFIDFMTR                                     808

SEQ ID NO: 76          moltype = AA  length = 761
FEATURE                Location/Qualifiers
source                 1..761
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 76
IYYLSLEFYM GRTLQNTMVN LALENACDEA TYQLGLDMEE LEEIEEDAGL GNGGLGRLAA  60
CFLDSMATLG LAAYGYGIRY EFGIFNQKIC GGWQMEEADD WLRYGNPWEK ARPEFTLPVH  120
FYGRVEHTSQ GAKWVDTQVV LAMPYDTPVP GYRNNVVNTM RLWSAKAPND FNLKDFNVGG  180
YIQAVLDRNL AENISRVLYP NDNFFEGKEL RLKQEYFVVA ATLQDIIRRF KSSKFGCRDP  240
VRTNFDAFPD KVAIQLNDTH PSLAIPELMR VLVDLERLDW DKAWEVTVKT CAYTNHTVLP  300
EALERWPVHL LETLLPRHLQ IIYEINQRFL NRVAAAFPGD VDRLRRMSLV EEGAVKRINM  360
AHLCIAGSHA VNGVARIHSE ILKKTIFKDF YELEPHKFQN KTNGITPRRW LVLCNPGLAE  420
IIAERIGEEY ISDLDQLRKL LSYVDDEAFI RDVAKVKQEN KLKFAAYLER EYKVHINPNS  480
LFDVQVKRIH EYKRQLLNCL HVITLYNRIK KEPNKFVVPR TVMIGGKAAP GYHMAKMIIK  540
LITAIGDVVN HDPVVGDRLR VIFLENYRVS LAEKVIPAAD LSEQISTAGT EASGTGNMKF  600
MLNGALTIGT MDGANVEMAE EAGEENFFIF GMRVEDVDRL DQRGYNAQEY YDRIPELRQI  660
IEQLSSGFFS PKQPDLFKDI VNMLMHHDRF KVFADYEEYV KCQERVSALY KNPREWTRMV  720
IRNIATSGKF SSDRTIAQYA REIWGVEPSR QRLPAPDEKI P                      761

SEQ ID NO: 77          moltype = AA  length = 351
FEATURE                Location/Qualifiers
source                 1..351
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 77
SHPALTPEQK KELSDIAHRI VAPGKGILAA DESTGSIAKR LQSIGTENTE ENRRFYRQLL  60
LTADDRVNPC IGGVILFHET LYQKADDGRP FPQVIKSKGG VVGIKVDKGV VPLAGTNGET  120
TTQGLDGLSE RCAQYKKDGA DFAKWRCVLK IGEHTPSALA IMENANVLAR YASICQQNGI  180
VPIVEPEILP DGDHDLKRCQ YVTEKVLAAV YKALSDHHIY LEGTLLKPNM VTPGHACTQK  240
YSHEEIAMAT VTALRRTVPP AVTGVTFLSG GQSEEEASIN LNAINKCPLL KPWALTFSYG  300
RALQASALKA WGGKKENLKA AQEEYVKRAL ANSLACQGKY TPSGQAGAAA S           351

SEQ ID NO: 78          moltype = AA  length = 584
FEATURE                Location/Qualifiers
source                 1..584
                       mol_type = protein
                       organism = Oryctolagus cuniculus
SEQUENCE: 78
EAHKSEIAHR FNDVGEEHFI GLVLITFSQY LQKCPYEEHA KLVKEVTDLA KACVADESAA  60
NCDKSLHDIF GDKICALPSL RDTYGDVADC CEKKEPERNE CFLHHKDDKP DLPPFARPEA  120
DVLCKAFHDD EKAFFGHYLY EVARRHPYFY APELLYYAQK YKAILTECCE AADKGACLTP  180
KLDALEGKSL ISAAQERLRC ASIQKFGDRA YKAWALVRLS QRFPKADFTD ISKIVTDLTK  240
VHKECCHGDL LECADDRADL AKYMCEHQET ISSHLKECCD KPILEKAHCI YGLHNDETPA  300
GLPAVAEEFV EDKDVCKNYE EAKDLFLGKF LYEYSRRHPD YSVVLLLRLG KAYEATLKKC  360
CATDDPHACY AKVLDEFQPL VDEPKNLVKQ NCELYEQLGD YNFQNALLVR YTKKVPQVST  420
PTLVEISRSL GKVGSKCCKH PEAERLPCVE DYLSVVLNRL CVLHEKTPVS EKVTKCCSES  480
LVDRRPCFSA LGPDETYVPK EFNAETFTFH ADICTLPETE RKIKKQTALV ELVKHKPHAT  540
NDQLKTVVGE FTALLDKCCS AEDKEACFAV EGPKLVESSK ATLG                   584

SEQ ID NO: 79          moltype = AA  length = 381
FEATURE                Location/Qualifiers
source                 1..381
                       mol_type = protein
                       organism = Bos taurus
SEQUENCE: 79
MPFGNTHNKH KLNFKAEEEY PDLSKHNNHM AKALTLEIYK KLRDKETPSG FTLDDVIQTG  60
VDNPGHPFIM TVGCVAGDEE SYTVFKDLFD PIIQDRHGGF KPTDKHKTDL NHENLKGGDD  120
LDPNYVLSSR VRTGRSIKGY ALPPHCSRGE RRAVEKLSVE ALNSLTGEFK GKYYPLKSMT  180
EQEQQQLIDD HFLFDKPVSP LLLASGMARD WPDARGIWHN DNKSFLVWVN EEDHLRVISM  240
EKGGNMKEVF RRFCVGLQKI EEIFKKAGHP FMWNEHLGYV LTCPSNLGTG LRGGVHVKLA  300
HLSKHPKFEE ILTRLRLQKR GTGGVDTAAV GSVFDVSNAD RLGSSEVEQV QLVVDGVKLM  360
VEMEKKLEKG QSIDDMIPAQ K                                            381

SEQ ID NO: 80          moltype = AA  length = 901
```

-continued

```
FEATURE                 Location/Qualifiers
source                  1..901
                        mol_type = protein
                        organism = Bos taurus
SEQUENCE: 80
MMMVLQPEGL GTGEGPPAGG RGGGEYMEQE EDWDRDLLLD PAWEKQQRKT FTAWCNSHLR  60
KAGTQIENIE EDFRNGLKLM LLLEVISGER LPRPDKGKMR FHKIANVNKA LDFIASKGVK  120
LVSIGAEEIV DGNLKMTLGM IWTIILRFAI QDISVEETSA KEGLLLWCQR KTAPYRNVNV  180
QNFHTSWKDG LALCALIHRH RPDLIDYAKL RKDDPIGNLN TAFEVAEKYL DIPKMLDAED  240
IVNTPKPDEK AIMTYVSCFY HAFAGAEQAE TAANRICKVL AVNQENEKLM EEYEKLASEL  300
LEWIRRTVPW LENRVGEPSM SAMQRKLEDF RDYRRLHKPP RVQEKCQLEI NFNTLQTKLR  360
LSHRPAFMPS EGKLVSDIAN AWRGLEQAEK GYEDWLLSEI RRLQRLQHLA EKFQQKASLH  420
EAWTRGKEDM LSQRDYETAS LQEVRALLRR HEAFESDLAA HQDRVEHIAA LAQELNELDY  480
HEAASVNSRC QAICDQWDNL GTLTQKRRDA LERMEKLLET IDQLQLEFAR RAAPFNNWLD  540
GAVEDLQDVW LVHSVEETQS LVTAHDQFKA TLPEADRERG AILGIQGEIQ KICQTYGLRP  600
SSTNPYITLT PQDINTKWDT VRKLVPSRDQ MLQEELTRQQ VNERLRRQFA AQANAIGPWI  660
QGKVEEVGRL AAGMAGSLEE QMAGLRQQEQ NIINYKSNID RLEGDHQLLQ ESLVFDNKHT  720
VYSMEHIRVG WEQLLTSIAR TINEVENQVL TRDAKGLSQE QLNEFRASFN HFDRKRNGMM  780
EPDDFRACLI SMGYDLGEVE FARIMTMVDP NAAGVVTFQA FIDFMTRETA ETDTAEQVVA  840
SFKILAGDKN YITAEELRRE LPAEQAEYCI RRMAPYKGAG APAGALDYVA FSSALYGESD  900
L                                                                  901

SEQ ID NO: 81           moltype = AA  length = 842
FEATURE                 Location/Qualifiers
source                  1..842
                        mol_type = protein
                        organism = Bos taurus
SEQUENCE: 81
MSRPLTDQEK RKQISVRGLA GVENVTELKK NFNRHLHFTL VKDRNVATPR DYYFALAYTV  60
RDHLVGRWIR TQQHYYEKDP KRIYYLSLEF YIGRTLQNTM VNLALENACD EATYQLGLDM  120
EELEEIEEDA GLGNGGLGRL AACFLDSMAT LGLAAYGYGI RYEFGIFNQK ISGGWQMEEA  180
DDWLRYGNPW EKARPEFTLP VHFYGRVEHT SQGAKWVDTQ VVLAMPYDTP VPGYRNNVVN  240
TMRLWSAKAP NDFNLKDFNV GGYIQAVLDR NLAENISRVL YPNDNFFEGK ELRLKQEYFV  300
VAATLQDIIR RFKSSKFGCL DPVRTNFDAF PDKVAIQLND THPSLAIPEL MRILVDQERL  360
EWEKAWEVTV KTCAYTNHTV LPEALERWPV HLIETLLPRH LQIIYEINQR FLNRVAAAFP  420
GDVDRLRRMS LVEEGAVKRI NMAHLCIAGS HAVNGVARIH SEILKKTIFK DFYELEPHKF  480
QNKTNGITPR RWLVMCNPGL AEIIAERIGE EYIADLDQLR KLLSYVDDES FIRDVAKVKQ  540
ENKLKFSAYL EKEYKVHINP NSLFDIQVKR IHEYKRQLLN CLHVITLYNR IKKEPNKFFV  600
PRTVMIGGKA APGYHMAKMI IKLITAIGDV VNHDPVVGDR LRVIFLENYR VSLAEKVIPA  660
ADLSEQISTA GTEASGTGNM KFMLNGALTI GTMDGANVEM AEEAGEENFF IFGMRVEDVE  720
RLDQKGYNAQ EYYDRIPELR HVIDQLSSGF FSPKQPDLFK DIVNMLMHHD RFKVFADYEE  780
YIKCQERVSA LYKNPREWTR MVIRNIATSG KFSSDRTIAQ YAREIWGVEP TRQRMPAPDE  840
KI                                                                 842

SEQ ID NO: 82           moltype = AA  length = 364
FEATURE                 Location/Qualifiers
source                  1..364
                        mol_type = protein
                        organism = Bos taurus
SEQUENCE: 82
MPHQYPALTP EQKKELCDIA HRIVAPGKGI LAADESTGSI AKRLQSIGTE NTEENRRFYR  60
QLLLTADDRV NPCIGGVILF HETLYQKADD GRPFPQVIKA KGGVVGIKVD KGVVPLAGTN  120
GETTTQGLDG LSERCAQYKK DGADFAKWRC VLKIGEHTPS SLAIMENANV LARYASICQQ  180
NGIVPIVEPE ILPDGDHDLK RCQYVTEKVL AAVYKALSDH HIYLEGTLLK PNMVTPGHAC  240
TQKYSHEEIA MATVTALRRT VPPAVPGITF LSGGQSEEEA SINLNAINKC PLLKPWALTF  300
SYGRALQASA LKAWGGKKEN LKAAQEEYVK RALANSLACQ GKYTPSGKAG AAASESLFIS  360
NHAY                                                               364

SEQ ID NO: 83           moltype = AA  length = 192
FEATURE                 Location/Qualifiers
source                  1..192
                        mol_type = protein
                        organism = Bos taurus
SEQUENCE: 83
MAPKKDVKKP AAAAAPAPAP APAPAPAPAP PKEEKIDLSA IKIEFSKQQQ DEFKEAFLLF  60
DRTGECKITL SQVGDVLRAL GTNPTNAEVK KVLGNPSNEE MNAKKIEFEQ FLPMLQAISN  120
NKDQGTYEDF VEGLRVFDKE GNGTVMGAEL RHVLATLGEK MKEEEVEALM AGQEDSNGCI  180
NYEAFVKHIM SN                                                      192

SEQ ID NO: 84           moltype = AA  length = 333
FEATURE                 Location/Qualifiers
source                  1..333
                        mol_type = protein
                        organism = Bos taurus
SEQUENCE: 84
MVKVGVNGFG RIGRLVTRAA FNSGKVDIVA INDPFIDLHY MVYMFQYDST HGKFNGTVKA  60
```

-continued

```
ENGKLVINGK AITIFQERDP ANIKWGDAGA EYVVESTGVF TTMEKAGAHL KGGAKRVIIS  120
APSADAPMFV MGVNHEKYNN TLKIVSNASC TTNCLAPLAK VIHDHFGIVE GLMTTVHAIT  180
ATQKTVDGPS GKLWRDGRGA AQNIIPASTG AAKAVGKVIP ELNGKLTGMA FRVPTPNVSV  240
VDLTCRLEKP AKYDEIKKVV KQASEGPLKG ILGYTEDQVV SCDFNSDTHS STFDAGAGIA  300
LNDHFVKLIS WYDNEFGYSN RVVDLMVHMA SKE                               333

SEQ ID NO: 85          moltype = AA  length = 284
FEATURE                Location/Qualifiers
source                 1..284
                       mol_type = protein
                       organism = Bos taurus
SEQUENCE: 85
MDAIKKKMQM LKLDKENALD RAEQAEADKK AAEDRSKQLE DELVSLQKKL KATEDELDKY  60
SEALKDAQEK LELAEKKATD AEADVASLNR RIQLVEEELD RAQERLATAL QKLEEAEKAA  120
DESERGMKVI ESRAQKDEEK MEIQEIQLKE AKHIAEDADR KYEEVARKLV IIESDLERAE  180
ERAELSEGKC AELEEELKTV TNNLKSLEAQ AEKYSQKEDK YEEEIKVLSD KLKEAETRAE  240
FAERSVTKLE KSIDDLEDEL YAQKLKYKAI SEELDHALND MTSI                   284

SEQ ID NO: 86          moltype = AA  length = 1141
FEATURE                Location/Qualifiers
source                 1..1141
                       mol_type = protein
                       organism = Bos taurus
SEQUENCE: 86
MPEAKPAAKK APAGKDAAAK PAPKEATPQE APAAPPTEAP PEDQSPTAEE PTGVFLKKPD  60
SVSVETGKDT VIVAKLNGKE LPAKPAVKWF KGKWLELGSK SGARFSFKES HDAASNVYTI  120
ELHITKVVLG DRGDYRIEVK AKDFCDSCAF NIDVEAPRQD SAGQSLESFK RSGEAKSDTA  180
GELDFSGLLK KRQVVEEEKK KKKDDDDLGI PPEIWELLKG AKKSEYERIA FQYGITDLRG  240
MLKRLKKAKV EVKKSAAFTK KLDPAYQVDR GNKIKLVVEI SDPDLPLKWF KNGQEIKPSS  300
KYVFENVGKK RILTINKCTL ADDAAYEVVV KDEKCFTELF VKEPPVLIVT PLEDQQVFVG  360
DRVEMAVEVS EEGAQVMWLK DGVELTREDS FKARYRLKKD GKRHILIYSE VTMEDKGHYQ  420
VMTNGGQCEA ELIVEEKQLE VLQDIADLTV KASEQAVFKC EVSDEKVTGK WYKNGVEVRP  480
SKRITISHTG RFHKLVIDDV RPEDEGDYTF VPDGYALSLS AKLKFLEIKV EYVPKQEPPK  540
IHLDCSGQTS ENAIVVVAGN KLRMDVSITG EPRPVATWMK EDEVFTGTEG RVRIEQRGDI  600
SSFVIESAER GDEGRYTIKV TNPVGEDVAS ILLKVVDVPD PPEAVRVTSV GEDWAILVWE  660
PPKYDGGRPV TGYLLERKKK GSQRWMKLNF EVFTETTYES TNMIEGILYE MRVFAVNVIG  720
VSQPSMNTKP FMPIAPTSEP QHLTVEDVTD TTTTLKWRPP DKIGAGGIDG YLIEYCVEGS  780
DEWIPANTEP VERCGFTVKN LPTGAKITFR VVGVNIAGRS QPATLAQPVT IREIVEQPKI  840
RLPRHLRQTY VRKVGEHLNL VIPFQGKPRP QVVWTKGGAP IDTSHVHVRT SDFDTVFFVR  900
QAARSDSGEY ELTVQIENMK DTATVCIRVV EKAGPPQNVM VKEVWGTNAL VEWQPPKDDG  960
NSEITGYFVQ KADKKTMEWF TVYERNRHTS CTVSDLIMGN EYYFRVYSEN VCGLSDLPGV  1020
SKNTARIVKT GMTLKLPEYK EHDFRTPPKF LTPLPDRVVV AGYAAALNCA VRGHPKPKVV  1080
WMKNKMEIRE DPKFLMTNQQ GVLTLNIRRP SPFDSGTYSC RAVNELGEAL AECKLEVRVP  1140
Q                                                                  1141

SEQ ID NO: 87          moltype = AA  length = 1495
FEATURE                Location/Qualifiers
source                 1..1495
                       mol_type = protein
                       organism = Bos taurus
SEQUENCE: 87
MSLVAVPFYQ KRHKHFDQSY RNIQTRYLLD EYSAKKKRAS AQSSLQRSVT RKSSSQRESS  60
RAALGESTLG ETTLGETSLG DTTLGRTTCR LCAKRMSSAL EEEAQERKRR YQSMVAAYGE  120
AKRERYLSEL AQLEEDVHVA RTHARDQLDR LALQHAGRHG LHHAVDDRLA WERHSFEERI  180
SRAPEILVRL RSHTVWERMS VKLCFTVQGF PTPVVQWYKD GSLICQAGEP GKYRIDSKYG  240
VHTLEINRAD FDDTATYSAV ATNVHGQVST NAAVVVRRFR GDEEPFHSVG LPIGLPLSSV  300
IPYTHFDVQF IEKFGVTFRR EGETLTLKCT LLVTPDLKRV QPRAEWYRDD VLLKESKWTK  360
MFFGEGQASL SFSHLNKDDE GLYTLRIVSR GGITDHSAFL FVRDADPLVT GAPGAPMDVH  420
CHDANRDYVI VTWKPPNATK ESPVIGYFID RCEVGTDNWI QCNDAPVKIC KYPVTGLFEG  480
RSYIFRVRAV NSAGISRPSR VSEAVAALDP VDLRRLQAVH LEGEKDIVVY QEELEGEVQI  540
PGPPTNVHAS ETSRTYVVLS WDPPEPRGKE PLMYFIEKSM VGSGSWQRVN AQTAVRSPRY  600
AVFDLAEGKP YVFRVLSANK HGLSDPSEIT APIQAQDTIV VPSAPGRVLA SRNTRTSVVV  660
QWDRPKHEED LLGYYVDCSV AGSNVWEPCN HKPIGYNRFV VHGLTTGEQY IFRVKAVNAV  720
GTSENSQESD VIKVQAALTV PSHPYGITLL NCDGHSMILG WKVPKFSGGS PILGYYVDKR  780
EAHHKNWHEV NSSPLKERIL TVEGLTEGSL YEFKIAAANM AGIGQPSDPS ELFKCEAWTM  840
PEPGPAYDLT FCEVRDTSLV LLWKPPVYSG SSPVSGYFVD YKEEDSGEWL TVHETATPHR  900
YLKVCDLHQG KTYVFRVRAV NASGVGRPSD TSEPVLVEAR PGTKEVSAGV DEEGNVYLSF  960
DCPEVTDASH FTWCKSYEDI ADDDRFKVET AGDHSKLYFK NLDKEDLGTY SVSVSDTDGV  1020
SSSFVLDEEE LERLKALSNE IKNPTIPLKS ELAYEIFDKG QVRFWLQAEH LSPDSSYRFI  1080
INDREVADSE THRIKCDKST GIIEMVMDRF TIDNEGTYTV QIQDGKAKSQ SSLVLIGDAF  1140
KAVLKEAEFQ RKEFLRKQGP HFAEYLHWDV TEECEVRLVC KVANTKKETV FKWLKDDVLY  1200
ETEKMPDLEK GVCELLIPKL SKKDHGEYKA TLKDDRGQDV STLEIAGKVY EDMILAMSRV  1260
CGASASPLKI LCTPEGIRLQ CFMKYFTEEM KVNWYHKDAK ITSSEHMRIG GNEQMAWLQI  1320
CEPTEKDKGK YTFEILDGKK NHQRSLDLSG QAFDEAFAEF QQLKAAAFAE KNRGKVIGGL  1380
PDVVTIMEGK TLNLTCTVFG NPDPEVVWFK NDKDIELSDH FSVKVEQGKY VSMTIKGVTS  1440
EDSGKYSIHV KNKYGGEKID VTVSVYKHGE EIPDVVLPQQ AKPKLLPAAA PSPAH       1495
```

-continued

```
SEQ ID NO: 88          moltype = AA  length = 434
FEATURE                Location/Qualifiers
source                 1..434
                       mol_type = protein
                       organism = Bos taurus
SEQUENCE: 88
MAMQKIFARE ILDSRGNPTV EVDLHTAKGR FRAAVPSGAS TGIYEALELR DGDKSRYLGK  60
GVLKAVEHIN KTLGPALLEK KLSVVDQEKV DKFMIELDGT ENKSKFGANA ILGVSLAVCK  120
AGAAEKGVPL YRHIADLAGN PELILPVPAF NVINGGSHAG NKLAMQEFMI LPVGASSFRE  180
AMRIGAEVYH HLKGVIKAKY GKDATNVGDE GGFAPNILEN NEALELLKTA IQAAGYPDKV  240
VIGMDVAASE FYRNGKYDLD FKSPDDPARH ISGEKLGELY KNFIKNYPVV SIEDPFDQDD  300
WATWTSFLSG VNIQIVGDDL TVTNPKRIAQ AVEKKACNCL LLKVNQIGSV TESIQACKLA  360
QSNGWGVMVS HRSGETEDTF IADLVVGLCT GQIKTGAPCR SERLAKYNQL MRIEEALGDK  420
AVFAGRKFRN PKAK                                                    434

SEQ ID NO: 89          moltype = AA  length = 170
FEATURE                Location/Qualifiers
source                 1..170
                       mol_type = protein
                       organism = Bos taurus
SEQUENCE: 89
MAPKKAKRRA AAEGGSSSVF SMFDQTQIQE FKEAFTVIDQ NRDGIIDKED LRDTFAAMGR  60
LNVKNEELDA MMKEASGPIN FTVFLNMFGE KLKGADPEDV ITGAFKVLDP EGKGTIKKKF  120
LEELLTTQCD RFSQEEIKNM WAAFPPDVGG NVDYKNICYV ITHGDAKDQE             170

SEQ ID NO: 90          moltype = AA  length = 1938
FEATURE                Location/Qualifiers
source                 1..1938
                       mol_type = protein
                       organism = Bos taurus
SEQUENCE: 90
MSSDQEMAVF GEAAPYLRKS EKERIEAQNK PFDAKTSVFV ADPKESFVKA TVQSREGGKV  60
TAKTEAGATV TVKEDQVFPM NPPKFDKIED MAMMTHLHEP AVLYNLKERY AAWMIYTYSG  120
LFCVTVNPYK WLPVYNAEVV TAYRGKKRQE APPHIFSISD NAYQFMLTDR ENQSILITGE  180
SGAGKTVNTK RVIQYFATIA VTGEKKKEEP TSGKMQGTLE DQIISANPLL EAFGNAKTVR  240
NDNSSRFGKF IRIHFGTTGK LASADIETYL LEKSRVTFQL KAERSYHIFY QIMSNKKPEL  300
IEMLLITTNP YDYAYVSQGE ITVPSIDDQE ELMATDSAIE ILGFTSDERV SIYKLTGAVM  360
HYGNLKFKQK QREEQAEPDG TEVADKAAYL QGLNSADLLK ALCYPRVKVG NEFVTKGQTV  420
EQVYNAVGAL AKAVYDKMFL WMVARINQQL DTKQPRQYFI GVLDIAGFEI FDFNSLEQLC  480
INFTNEKLQQ FFNHHMFVLE QEEYKKEGIE WEFIDFGMDL AACIELIEKP MGIFSILEEE  540
CMFPKATDMS FKNKLYEQHL GKSNNFQKPK PAKGKAEAHF SLIHYAGTVD YNITGWLDKN  600
KDPLNETVVG LYQKSSVKTL ALLFSGPASG EAEGGPKKGG KKKGSSFQTV SALFRENLNK  660
LMTNLRSTHP HFVRCIIPNE TKTPGAMEHE LVLHQLRCNG VLEGIRICRK GFPSRILYAD  720
FKQRYKVLNA SAIPEGQFID SKKASEKLLA SIDVDHTQYK FGHTKVFFKA GLLGLLEEMR  780
DEKLAQLITR TQARCRGFLA RVEYQKMVER RESIFCIQYN VRAFMNVKHW PWMKLYFKIK  840
PLLKSAETEK EMANMKEEFE KTKEELAKSE AKRKELEEKM VTLTQEKNDL QLQVQSEADA  900
LADAEERCDQ LIKTKIQLEA KIKEVTERAE DEEEINAELT AKKRKLEDEC SELKKDIDDL  960
ELTLAKVEKE KHATENKVKN LTEEMAGLDE TIAKLTKEKK ALQEAHQQTL DDLQAEEDKV  1020
NTLTKAKTKL EQQVDDLEGS LEQEKKLRMD LERAKRKLEG DLKLAQESTM DIENDKQQLD  1080
EKLKKKEFEM SNLQSKIEDE QALAMQLQKK IKELQARIEE LEEEIEAERA SRAKAEKQRS  1140
DLSRELEEIS ERLEEAGGAT SAQIEMNKKR EAEFQKMRRD LEEATLQHEA TAAALRKKHA  1200
DSVAELGEQI DNLQRVKQKL EKEKSEMKME IDDLASNMET VSKAKGNLEK MCRALEDQLS  1260
ELKTKEDEQQ RLINDLTTQR ARLQTESGEF SRQLDEKDAL VSQLSRGKQA FTQQIEELKR  1320
QLEEEIKAKS ALAHALQSAR HDCDLLREQY EEEQEGKAEL QRAMSKANSE VAQWRTKYET  1380
DAIQRTEELE EAKKKLAQRL QDAEEHVEAV NAKCASLEKT KQRLQNEVED LMIDVERTNA  1440
ACAALDKKQR NFDKILSEWK QKYEETHAEL EASQKESRSL STELFKIKNA YEESLDQLET  1500
LKRENKNLQQ EISDLTEQIA EGGKRIHELE KVKKQVEQEK SEIQAALEEA EASLEHEEGK  1560
ILRIQLELNQ VKSEIDRKIA EKDEEIDQLK RNHIRIVESM QSTLDAEIRS RNDAIRLKKK  1620
MEGDLNEMEI QLNHANRMAA EALKNYRSTQ AILKDTQIHL DDALRGQEDL KEQLAMVERR  1680
ANLLQAEIEE LRATLEQTER SRKIAEQELL DASERVQLLH TQNTSLINTK KKLETDITQI  1740
QGEMEDIIQE ARNAEEKAKK AITDAAMMAE ELKKEQDTSA HLERMKKNLE QTVKDLQHRL  1800
DEAEQLALKG GKKQIQKLEA RVRELEGEVE SEQKRNVEAV KGLRKHERRV KELTYQTEED  1860
RKNILRLQDL VDKLQAKVKS YKRQAEEAEE QSNVNLSKFR KLQHELEEAE ERADIAESQV  1920
NKLRVKSREV HTKIISEE                                                1938

SEQ ID NO: 91          moltype = AA  length = 7100
FEATURE                Location/Qualifiers
source                 1..7100
                       mol_type = protein
                       organism = Bos taurus
SEQUENCE: 91
MPNFQTRTRF YETTTTRTSD YEQSETSRPA LAQPVPEKPV ERKRVIRKKV DPSKFMTPYI  60
AHSQKMQNLF STNKYKENYE KAKGKPYAIT TDTPELRRIK KVQDQLSEVK YRVDGDVAKT  120
ICHVDEKAKD IEHAKKVSQQ VSKVLYKQNW EDTKDKYLLP PDAPELVQAI KNTAMFSKKL  180
YTEDWEADKT MFYPYNDSPE LRRVAQAQKA LSDIAYKKGL AEQQTQFTSL PDPPDIEFAK  240
```

-continued

```
KVTNQVSKQK YKEDYENKVK GKWSETPCFE IATARMNSNN ISTKKYQEDF EHMKDQIYFM  300
QTETPEYKVN KQAGVAASKV KYKQDYEKNK GKADYNVLPA SENPLLRQLK AAGDALSDKL  360
YKENYEKTKA KSINYCETPK FKLDTVLHNF SSDTKYKDSY LKDILGHYVG SYEDPYHTHC  420
MRVSAQNSDK NYKAEYEEDR GKGFFPQTIT QEYEAIKKLD QCKDHTYKVH PDKTKFTQVT  480
DSPVLMQAQV NSKQLSDLNY KAKHENEKFK CHIPPDTPAF IQHKLNAYNL SDNIYKHDWE  540
KTKAKKFDIK VDAIPLLAAK ANTKIASDVM YKKDYEKSKG KMIGALSIND DPKMLHSLKT  600
AKNQSDRLYK ENYEKTKAKS MNYCETPKYQ LDTLLKNFSE AKYKDSYVQN VLGHYIGSFE  660
DPYQVHCLKI SAQNSDKNYK AEYEEDKGKC YFPQTITQEY EAIKKLDQCK DHTYKVHPDK  720
TKFTAVTDTP VLLQAQLNTK QLSDLNYKAK HEGEKFKCHI PADAPQFIQH RVNAYNLSDN  780
IYKHDWEKTK AKKFDIKVDA IPLLAAKANT KIASDVMYKK DYEKSKGKMI GALSINDDPK  840
MLHSLKTAKN QSDHEYRKDY EKSKTIYTAP LDMLPLTHAK KSQAIASDVD YKHLLHNYSY  900
PPDSVNVDLA KKAYALQSDV EYKADYNSWM KGCGWMPFGS LEMEKAKRAS DILNEKKYRQ  960
HPDTLKFTSI EDAPIIVQSK INQAQRSDVA YKAKGEEVIH KYSLPADLPQ FIQAKVNAYN 1020
ISENLYKADL KDLSKKGYDL RIDAIPIKAA KAARQAASDV QYKKDYEKAK GKMVGFQSLQ 1080
DDPKLVHYMN VAKIQSDREY KKAYEKTKTR HNTPHDMVNI VAAKKAQDVA SNVNYKHSLH 1140
HYTYLPDAMD LELSKNMMHI QSDNVYKEDY NNWMKGIGWI PIGSLEVEKV KKAGDALNEK 1200
KYRQHPDTLK FTSIVDSPVM VQAKQNTQQV SDILYKAKGE DVKHKYTMSP DLPQFLQAKC 1260
NAYNLSDVCY KRDWHDLIAK GTNVLGDAIP ITAAKASRNI ASDYKYKEAY EKAKGKQVGF 1320
RSLQDDPKLV HYMNVAKLQS DREYKKNYEN TKTSYHTPGD MVSITAAKMA QDVATNVNYK 1380
QPIHHYTYLP DALSLEHIRN VNQIQSDNVY KDEYNHFFKG MGWIPIGSLE VEKVKKAGDA 1440
LNEKKYRQHP DTIKFTSVPD SMGMVLAQHN TKQLSDLNYK VEGEKVKHKY TMDPDVPQFI 1500
QAKVNAYNMS DSHYKADWKK TLAKGYDLRP DAIPIVAAKS SRNIASDFKY KEAYEKTKGK 1560
QIGFRSLQDD PKLVHFMNVA KMQSDREYKK GYEASKTKYH TPLDMLSVTA AKKSQEVATN 1620
TNYKQPFHHY TLLPDALNVE HSRNAMQIQS DNLYKSDFTN WMKGIGWLPL ESLEVEKAKK 1680
AGEILSEKKY RQHPEKLKFT YAMDTMEQAL NKSNKLIMDK RLYTEKWNKD KTTIHVMPDT 1740
PDILLSRVNQ ITMSDKLYKA GWEEEKKKGY DLRPDAISIK AARASRDIAS DYKYKQAYEQ 1800
AKGKQIGFRS LEDDPKLVHF MQVAKMQSDR EYKKAYEKSK TSFQTPVDML SVVAAKKSQE 1860
VATNANYRNV IHTYNMLPDA MSLELAKNMM QIQSDNQYKA DYADFMKGIG WLPLGSLEAE 1920
KNKKAMEIIS EKKYRQHPDT LKYSTLMDSM NMVLAKNNAK IMNEHLYKQA WEADKTKVHI 1980
MPDIPQIILA KANAINISDK LYKLSLEEAK KKGYDLRTDA IPIKAAKASR DIASDYKYKH 2040
SYEKERGKMV GFRSLEDDPK LVHSMQVAKM QSDREYKKNY EKTKTSYHTP ADMLSVTAAK 2100
DAQANITNTN YKHLIHKYIL LPDAMNIQLS KNMNRIQSDN EYKQDYNEWY KGLGWSPAGS 2160
LEVEKAKKAT EYASDQKYRQ HPSNFQFTKL NDSMDMVLAK QNAHTMNKYL YTVDWNKDKT 2220
KIHVMPDTPD ILQAKQNQTM YSQKLYKLGW EEALKKGYDL PVDAISVQAA KASRDIVSDF 2280
KYKQGYRKQI GHHIGFRSLQ DDPKLVLSMN VAKMQSEREY KKDFEKWKTK FSSPVDMLGV 2340
VLAKKCQALV SDVDYKNYLH GWTCLPDQND VIHAKKAYDL QSENLYKSDL EWLKGIGWSP 2400
LGSLEAEKNK RASEIISEKK YRQPPDRNKF TSIPDAMDIV LAKTNAKNRS DILYREAWDK 2460
DKTQIHIMPD TPDIILAKAN LINTSDKFYR MGYEELRKKG YDLPVDAIPI KAAKASREIA 2520
SEYKYKEGFR MQLGHHIGAR NIKDDPKMMW SMHVAKIQSD REYKKDFEKW KTKFSSPVDM 2580
LGVVLAKKCQ TLVSDIDYKN YLHQWTCLPD QSDVIHARRA YDLQSDNLYK SDLQWLRGIG 2640
WLPSGSLEDE KNKRASQILS DHVYRQHPDQ FKFSSLMDSI PMVLAKNNAI TMNHRLYTEA 2700
WDKDKTTVHI MPDTPEVLLA KQNQINYSEK LYKLGLDEAK RKGYDMRIDA IPIRAAKASR 2760
DIASEFKYKE GYRRQLGHHI GARAIHDDPK MMWSMHVAKI QSDREYKKDF EKWKTKFSSP 2820
VDMLGVVLAK KCQTLVSDID YKNYLHQWTC LPDQSDVIHA RQAYDLQSDN VYKSDLQWLR 2880
GIGWVPIGSL DVEKSKRASE ILSDKLYRQP PDKFKFTSVT DSLEQVLAKN NAINMNKRLY 2940
TEAWDKDKTQ VHIMPDTPEI TLARTNKVNY SENLYKLAHE EAKKKGYDLR SDAIPIVAAK 3000
ASRDIISDYK YKDGYRKQLG HHIGARDIKD DPKMMWSMHV AKIQSDREYK KDFEKWKTKF 3060
SSPVDMLGVV LAKKCQTLVS DIDYKNYLHQ WTCLPDQSDV IHARQAYDLQ SDNVYKSDLQ 3120
WLRGIGWVPI GSVDVVKCKR AAEILSDNLY RQPPDTLKFT SVPDSLEQVL AKNNAINMNK 3180
RLYTEAWDKD KTQIHIMPDT PEIMLARQNK LNYSETLYKL ANEEAKKKGY DLRSDAIPIV 3240
AAKASRDIIS DYKYKDGYRK QLGHHIGARD IKDDPKMMWS MHVAKIQSDR EYKKDFEKWK 3300
TKFSSPVDML GVVLAKKCQT LVSDIDYKNY LHQWTCLPDQ SDVIHARRAY DLQSDNIYKS 3360
DLQWLRGIGW VPIGSVDVVK CKRAAEILSD NLYRQRPDTL KFTSVPDSLE QVLAKNNAIN 3420
MNKRLYTEAW DKDKTQIHMM PDTPEITLAR QNKINYSENL YRQAMEEAKK EGYDLRSDAI 3480
PIVAAKASRE IASDYKYKEA YRKQLGHHIG ARAIHDDPKM MWSVHVAKMQ SDREYKKDFE 3540
KYKTRFSSPV DMLGIVLAKK CQTLVSDVDY KHPLHEWTCL PDQNDIIHAR KAYDLQSDNL 3600
YKSDLEWLKG IGWVPIGSVE VLKAKRAGEI LSDNIYRQRP DTLKFTSVTD SPEQVLAKNN 3660
AINMNKRLYT EAWDNDKKTI HVMPDTPEIM LAKLNRINYS DKLYKLALEE SRKEGYDLRL 3720
DAIPIQAAKA SREIASDYKY KEGYRKQLGH HIGARNIKDD PKMMWSIHAG KLQSDLEYKK 3780
DFEKWKTKFS SPVDMMGLVQ AKKCQILVSD IDYKNLLHEW TCLPDQNDII QARKAYDLQS 3840
DAIYKADLEW LRGIGWVPID SVGVEHAKRA GEILSERKYR QPAVQLKFTS ITDTPEIVLA 3900
KNNALNVSKH LYTEAWDADK TSIHVMPDTP EILLAKSNSA NISHKLYTKG WDESKMKDYD 3960
LRADAISIKS AKASRDIASD YKYKEAYEKH KGHHIGARSV EDDPRIMCAM NAGRIQSERE 4020
YKKEFQKWKT KFSSPVDMLG ILLAKKCQTL VSDIDYRNYL HHWTCLPDQN DIIQARKAYD 4080
LQSDNVYKSD LEWLRGIGWM PEGSVEMNRV KNAQDLINER RYRTRPEALT FTSIVDTPEV 4140
VLAKANSLQM SEKLYQEAWN KDKTNISIPS DTPLILQAQI NAMQISNKLY QKDWDEAKQK 4200
GYDLRADAIE IKHAKASREI ASEYKYKEGY RKQLGHHVGF RTLQDDPKLV WSIHAAKIQS 4260
DREYRKAYEK SKGIHNTPLD MMSIVQAKKC QVLVSDIDYR NYLHQWICLP DQNDVIQAKK 4320
AYELQSDAIY KADLEWLRGI GWMPEGSPEV LRVKNAQHIF RDSVYRTPVV KLKYTSIVDT 4380
PEVVLAKSNA ENISIPKYRE VWDKDKTSIH IMPDTPEINL ARTNALNVSN KLYREGWDEM 4440
KMSCDVRLDA IPIQAAKASR EIASDYKYKL DHEKQKGHYV GTRTARDDNK IRWALIAGKI 4500
QNEREYRLHW AKWKSKFQSP PDMLSIEHSK NSQTLVSDID YRHYLHQWTC MPDQNDVIQA 4560
RKAYDLQSDA IYKADLEWLR GIGWMPADSV SVNHAKHMAD IFNEKKYRTK IETLSFTPVD 4620
DRVDYVTAKH SGEILNDIKY RKDWNDTKSK YTLTETPQLH TAQEAARILD QYLYKESWEK 4680
QKATGYILPP DAVPFVHAHH SGDVQSELKY KAEHVKQKGH YVGVPTMRDD PKLVWFEHAG 4740
QIQNDRLYKE NYHKTKAKIH IPPDMVSVLA AKEGQALASD IDYRNYLHQW ICHPDQNDVI 4800
```

-continued

```
QARKAYDLQS DNIYKADLEW LRGIGWIPLD SVDHVRVTRN QEMMNQIKYK KDALANYPNF  4860
TSVVDPPEIV LAKINSVNQS DVKYKETFNK LIKGKYIFSP DTPYITHSKD MEKLYSTILY  4920
KRAWDGTKAY GYTLDERYIP IVGAKHADFV NSELKYKETY EKLKGHYLAG KEISEFPNVV  4980
HCLDFQKMRS LLNYRRHYED TKANVHIPQD MMNHVLAKRC QYILSDLEYR HYFHQWTSLP  5040
EEPNVVRVRH AQEILSDNVY KDDLNWLKGI GCYVWDTPQI LHAKKSYDLQ SQILYTAAGK  5100
ENLKNYNLVT DTPLYVTALQ SGINASEVKY KENYHQTKDK YTTVLETVDY DRIKNLKDLF  5160
SSNLYKEAWD KVKATSYILP PNTVSLTHAK NQKYMASHIK YREEYEKFKA LYTLPRSVED  5220
DPNTARCLRV GKFNIDRLYR SVYEKNKMKI HIVPDMVEMV TAKDSQKKVS EIDYRLHLHE  5280
WICHPDLQVN SHVRKVTDQI SDIVYKDDLT WLKGIGCYVW DTPEILHAKH AYDLRNDIKY  5340
KAHVLKTRNN YKLVTDTPVY VQAVKSGKQL SDAVYHYDYV HSIRGRVAPT TKTVDLDRAL  5400
HAYKLQSENL YRKAGLHALP TGYRLPVDTP HFKHTKDTRY MSSYFKYKEV YEHMKAYGYT  5460
LGPNDVPFVN VRRVNNITSE RLYRQLYHKL KDKIHTTPDT PEIRQVKKTQ EAVSELIYKS  5520
DFFKMQGHMI SLPYTPQVLH CRYVGDITSD IKYKEDLQVL RGMGCFLYDT PDMVRSRHLR  5580
KLWSHYLYTD KARKMRDKYK VVLDTPEYRK VQELKTHLSE LVYRAAGRKQ KSIFTSVPDT  5640
PDLTRAKRGQ KLQSQYLYVE LATKERPHHH AGNQTTALKH ARHVKDMVSE NKYKIQYEKM  5700
KDKYTPVPDT PILIRAKRAY WNASDLRYKE TFQKTKGKYH TVKDALDIVY HRTVTDHISK  5760
IKYKENYMSQ LGIWRSIPDR PEHFHHRAVT DAVSDVKYKQ DLTWLKGIGC YAYDTPDFTL  5820
AEKNKTLYSK YKYKEVFERT KSNFKYVADC PINRHFKFAT QLMNEKKYRA DYEQRKDKYH  5880
LVVDEPRHLL AKIAGDQISQ IKYRKNYEET KDKFTSIVDT PEHLRTTKVN KQISDILYKL  5940
EYNKAKPRGY TTIHDTPMLL HVRKVKDEVS DLKYKEVYQR TKSNCTIEPD AVHIKAAKDA  6000
YKVNTNLDYK KKYEATKAYW KWTPDRPDFI QAAKSTLQQS DFEYKLDREY LKGCKLSVTD  6060
DKDMVLALKN SIIESDLKYK EKHVKERGSC HAVPDTPQIL LAKTVSSLVS ENKYKSYVKK  6120
HLAQGSYTTL PETRDTIHVK EVTKNVSDTN YKKKFVKEKG KSNYSIMLEP PDVKHAMDVA  6180
KKQSNVAYKK DAKENLHYTT VADRPDIKKA TQAAKQASEV EYRAKHRKEG SHGLSMLGRP  6240
DIEMAKKAAK LSSQVQYREN FNKEKGKTPK YNPKDSQLYK VMKDANTLAS EVKYKADLKK  6300
LHKPVTDMKE SLIMNHVLNT SHLASSYQYK KNYEKSKGHY HTIPDNLEQL HLKEATELQS  6360
IVKYKEKYEK ERGKPMLDFE TPTYITAKES QQMQSGKEYR KDYEESIKGR NLTGLEVTPA  6420
LLHVKYATKI ASEKEYRKDL EESIRGKGLS EMEDTPDMLR AKNATQILNE KEYKRDLELE  6480
VKGRGLNAMA NETPDFLRAR NATDIASQIK YKQSAEMEKA NFTSVVDTPE IIHAQQVKNL  6540
SSQKKYKEDA EKCMSYYETV LDTPEMQRVR ENQKNFSLLQ YQYDLKNSKG KITVVQDTPE  6600
ILRVKENQKN FSSVLYKEDV SPGTAIGKTP EMMRVKQTQD HISSVKYKEV IGQGTPIPDL  6660
PEVKRVKQTQ KHISSVMYKE NLGTGIPTPV TPEIERVKRN QENFSSNLSL KHEVTSKSII  6720
LTKYNSLKLP PHTPLPPVLY KENLGKGTPT PITPEMERVK RNQENFSSVL YKENMGKGTP  6780
LPVTPEMERI KHNQENISSV LYKENMRKAT PTPVTPEMER AKRNQENISS VLYSDSFRKQ  6840
IQGKAAYVLD TPEMRRVRET QRHISTVKYH EDFEKHKGCF TPVVTDPITE RVKKNTQDFS  6900
DICYRGIQRK VVEMEQKRND QDQETITGLR VWRTNPGSVF DYDPAEDNIQ SRSLHMINAQ  6960
AQRRSREQSR SASGLSISGG EEKSEHSEAA HLSTYSDGGV FFSAASTGYK HARTTELPQQ  7020
RSSSVATQQT TVSSIPSHPS TAGKIFRAMY DYMAADADEV SFKDGDAIVN VQAIDEGWMY  7080
GTVQRTGRTG MLPANYVEAI                                              7100

SEQ ID NO: 92          moltype = AA  length = 1935
FEATURE                Location/Qualifiers
source                 1..1935
                       mol_type = protein
                       organism = Bos taurus
SEQUENCE: 92
MSSDQEMAIF GEAAPYLRKS EKERIEAQNK PFDAKTSVFV VDPKESFVKA TVQSREGGKV  60
TAKTEAGATV TVKEDQVFPM NPPKYDKIED MAMMTHLHEP AVLYNLKERY AAWMIYTYSG  120
LFCVTVNPYK WLPVYNAEVV TAYRGKKRQE APPHIFSISD NAYQFMLTGE LDVSIGESGA  180
GKTVNTKRVI QYFATIAVTG EKKKEEPTSG KMQGTLEDQI ISANPLLEAF GNAKTVRNDN  240
SSRFGKFIRI HFGTTGKLAS ADIETYLLEK SRVTFQLKAE RSYHIFYQIM SNKKPELIEM  300
LLITTNPYDY AYVSQGEITV PSIDDQEELI ATDSAIDILG FTSDERVSIY KLTGAVMHYG  360
NLKFKQKQRE EQAEPDGTEV ADKAAYLQSL NSADLLKALC YPRVKVGNEF VTKGQTVQQV  420
YNAVGALAKA VYEKMFLWMV TRINQQLDTK QPRQYFIGVL DIAGFEIFDF NSLEQLCINF  480
TNEKLQQFFN HHMFVLEQEE YKKEGIEWEF IDFGMDLAAC IELIEKPMGI FSILEEECMF  540
PKATDTSFKN KLYEQHLGKS NNFQKPKPAK GKAEAHFSLV HYAGTVDYNI AGWLDKNKDP  600
LNETVVGLYQ KSAMKTLAFL FAGGPSAEEG GGGKKGGKKK GSSFQTVSAL FRENLNKLMT  660
NLRSTHPHFV RCIIPNETKT PGAMEHELVL HQLRCNGVLE GIRICRKGFP SRILYADFKQ  720
RYKVLNASAI PEGQFIDSKK ASEKLLGSID IDHTQYKFGH TKVFFKAGLL GILEEMRDEK  780
LAQLITRTQA ICRGFLMRVE FRKMMERRES IFCIQYNVRA FMNVKHWPWM KLYFKIKPLL  840
KSAETEKEMA NMKEEFEKTK EELAKSEAKR KELEEKMVTL MQEKNDLQLQ VQSEADGLAD  900
AEERCDQLIK TKIQLEAKIK ELTERAEDEE EINAELTAKK RKLEDECSEL KKDIDDLELT  960
LAKVEKEKHA TENKVKNLTE EMAGLDENIA KLTKEKKALQ EAHQQTLDDL QAEEDKVNTL  1020
TKSKTKLEQQ VEDLEGSLEQ EKKLRMDLER AKRKLEGDLK LAQESTMDIE NDKQQLDEKL  1080
KKKEFEMSNL QSKIEDEQAL AMQLQKKIKE LQARIEELEE EIEAERASRA KAEKLRSDLS  1140
RELEEISERL EEAGGATSAQ IEMNKKREAE FQKMRRDLEE ATLQHEATAA ALRKKHADSV  1200
AELGEQIDNL QRVKQKLEKE KSEMKMEIDD LASNMETVSK AKGNLEKMCR TLEDQLSEVK  1260
SKEEEHQRLI NELSVQKAHL HTESGEFSRQ LEEKDALVSQ LSRGKQAFTQ QIEELKRQLE  1320
EETKAKSALA HALQSARHDC DLLREQYEEE QEGKAELQRA LSKANSEVAQ WRTKYETDAI  1380
QRTEELEEAK KKLAQRLQDA EEHVEVVNAK CASLEKTKQR LQNEVEDLML DVERSNAACA  1440
ALDKKQRNFD KVLSEWKQKY EETQADLEAS QKESRSLSTE LFKVKNAYEE SLDQLETLKR  1500
ENKNLQQEIS DLTEQIAEGG KHIHELEKVK KQIEQEKGEL QAALEEAEAS LEHEEGKILR  1560
IQLELNQVKS EIDRKIAEKD EEIDQLKRNH LRVVESMQST LDAEIRSRND ALRIKKKMEG  1620
DLNEMEIQLN HANRQATEAI KNLRNTQGVL KDTQLHLDDA IRGQDDLKEQ LAMVERRANL  1680
MQAEIEELRA SLEQTERSRR VAEQELLDAS ERVQLLHTQN TSLINTKKKL ETDITQIQGE  1740
MEDIVQEARN AEEKAKKAIT DAAMMAEELK KEQDTSAHLE RMKKNMEQTV KDLQHRLDEA  1800
```

-continued

```
EQLALKGGKK QIQKLEARVR ELENEVENEQ KRNAEAVKGL RKHERRVKEL TYQTEEDRKN  1860
VLRLQDLVDK LQSKVKAYKR QAEEAEEQSN VNLAKFRKLQ HELEEAEERA DIAESQVNKL  1920
RVKSREVHTK VISEE                                                   1935

SEQ ID NO: 93          moltype = AA  length = 607
FEATURE                Location/Qualifiers
source                 1..607
                       mol_type = protein
                       organism = Bos taurus
SEQUENCE: 93
MKWVTFISLL LLFSSAYSRG VFRRDTHKSE IAHRFKDLGE EHFKGLVLIA FSQYLQQCPF  60
DEHVKLVNEL TEFAKTCVAD ESHAGCEKSL HTLFGDELCK VASLRETYGD MADCCEKQEP  120
ERNECFLSHK DDSPDLPKLK PDPNTLCDEF KADEKKFWGK YLYEIARRHP YFYAPELLYY  180
ANKYNGVFQE CCQAEDKGAC LLPKIETMRE KVLASSARQR LRCASIQKFG ERALKAWSVA  240
RLSQKFPKAE FVEVTKLVTD LTKVHKECCH GDLLECADDR ADLAKYICDN QDTISSKLKE  300
CCDKPLLEKS HCIAEVEKDA IPENLPPLTA DFAEDKDVCK NYQEAKDAFL GSFLYEYSRR  360
HPEYAVSVLL RLAKEYEATL EECCAKDDPH ACYSTVFDKL KHLVDEPQNL IKQNCDQFEK  420
LGEYGFQNAL IVRYTRKVPQ VSTPTLVEVS RSLGKVGTRC CTKPESERMP CTEDYLSLIL  480
NRLCVLHEKT PVSEKVTKCC TESLVNRRPC FSALTPDETY VPKAFDEKLF TFHADICTLP  540
DTEKQIKKQT ALVELLKHKP KATEEQLKTV MENFVAFVDK CCAADDKEAC FAVEGPKLVV  600
STQTALA                                                            607

SEQ ID NO: 94          moltype = AA  length = 381
FEATURE                Location/Qualifiers
source                 1..381
                       mol_type = protein
                       organism = Ovis aries
SEQUENCE: 94
MPFGNTHNKH KLNFKAEEEY PDLSKHNNHM AKALTLEIYK KLRDKETPSG FTLDDVIQTG  60
VDNPGHPFIM TVGCVAGDEE SYTVFKDLFD PIIQDRHGGF KPTDKHKTDL NHENLKGGDD  120
LDPHYVLSSR VRTGRSIKGY ALPPHCSRGE RRAVEKLSVE ALNSLTGEFK GKYYPLKSMT  180
EQEQQQLIDD HFLFDKPVSP LLLASGMARD WPDARGIWHN DNKSFLVWVN EEDHLRVISM  240
EKGGNMKEVF RRFCVGLQKI EEIFKKAGHP FMWNEHLGYV LTCPSNLGTG LRGGVHVKLA  300
HLSKHPKFEE ILTRLRLQKR GTGGVDTAAV GSVFDVSNAD RLGSSEVEQV QLVVDGVKLM  360
VEMEKKLEKG QSIDDMIPAQ K                                            381

SEQ ID NO: 95          moltype = AA  length = 897
FEATURE                Location/Qualifiers
source                 1..897
                       mol_type = protein
                       organism = Ovis aries
SEQUENCE: 95
MMMVLQPEGL GTGEGPFAGG GRGGGEYMEQ EEDWDRDLLL DPAWEKQQRK TFTAWCNSHL  60
RKAGTQIENI EEDFRNGLKL MLLLEVISGE RLPRPDKGKM RFHKIANVNK ALDFIASKGV  120
KLVSIGAEEI VDGNLKMTLG MIWTIILRFA IQDISVEETS AKEGLLLWCQ RKTAPYRNVN  180
VQNFHTSWKD GLALCALIHR HRPDLIDYAK LRKDDPIGNL NTAFEVAEKY LDIPKMLDAE  240
DIVNTPKPDE KAIMTYVSCF YHAFAGAEQA ETAANRICKV LAVNQENEKL MEEYEKLASE  300
LLEWIRRTVP WLENRVGEPS MSAMQRKLED FRDYRRLHKP PRVQEKCQLE INFNTLQTKL  360
RLSHRPAFMP SEGKLVSDIA NAWRGLEQAE KGYEDWLLSE IRRLQRLQHL AEKFQQKASL  420
HEAWTRGKEE MLSQRDYETA SLQEVRALLR RHEAFESDLA AHQDRVEHIA ALAQELNELD  480
YHEAASVNSR CQAICDQWDN LGTLTQKRRD ALERMEKLLE TIDQLQLEFA RRAAPFNNWL  540
DGAVEDLQDV WLVHSVEETQ SLVTAHDQFK ATLPEADRER GAILGIQGEI QKICQTYGLR  600
PSSTNPYIAL TPQDINTKWD TVRKLVPSRD QMLQEELTRQ QMNERLRRQF AAQANAIGPW  660
IQGKVEEGGR LAAGMAGSLE EQMAGLRQQE QNIINYKSNI DRLEGDHQLL QESLVFDNKH  720
TVYSMEHIRV GWEQLLTSIA RTINEVENQV LTRDAKGLSQ EQLNEFRASF NHFDRKRNGM  780
MEPDDFRACL ISMGYDLGEV EFARIMTMVD PNAAGVVTFQ AFIDFMTRET AETDTAEQVV  840
ASFKILAGDK NYITAEELRR ELPAEQARRA PDRGAGAPAG ALDYVAFSSA LYGESDL     897

SEQ ID NO: 96          moltype = AA  length = 842
FEATURE                Location/Qualifiers
source                 1..842
                       mol_type = protein
                       organism = Ovis aries
SEQUENCE: 96
MSRPLTDQEK RKQISVRGLA GVENVTELKK NFNRHLHFTL VKDRNVATPR DYYFALAYTV  60
RDHLVGRWIR TQQHYYEKDP KRIYYLSLEF YIGRTLQNTM VNLALENACD EATYQLGLDM  120
EELEEIEEDA GLGNGGLGRL AACFLDSMAT LGLAAYGYGI RYEFGIFNQK ISGGWQMEEA  180
DDWLRYGNPW EKARPEFTLP VHFYGRVEHT SQGAKWVDTQ VVLAMPYDTP VPGYRNNVVN  240
TMRLWSAKAP NDFNLKDFNV GGYIQAVLDR NLAENISRVL YPNDNFFEGK ELRLKQEYFV  300
VAATLQDIIR RFKSSKFGCL DPVRTNFDAF PDKVAIQLND THPSLAIPEL MRILVDQERL  360
EWEKAWEVTV KTCAYTNHTV LPEALERWPV HLIETLLPRH LQIIYEINQR FLNRVAAAFP  420
GDVDRLRRMS LVEEGAVKRI NMAHLCIAGS HAVNGVARIH SEILKKTIFK DFYELEPHKF  480
QNKTNGITPR RWLVMCNPGL AEVIAERIGE EYIADLDQLR KLLSYVDDES FIRDVAKVKQ  540
ENKLKFSAYL EKEYKVHINP NSLFDIQVKR IHEYKRQLLN CLHVITLYNR IKKEPNKFFV  600
PRTVMIGGKA APGYHMAKMI IRLITAIGDV VNHDPVVGDR LRVIFLENYR VSLAEKVIPA  660
ADLSEQISTA GTEASGTGNM KFMLNGALTI GTMDGANVEM AEEAGEENFF IFGMRVEDVE  720
```

-continued

```
RLDQKGYNAQ EYYDRIPELR HIIDQLSSGF FSPKQPDLFK DIVNMLMHHD RFKVFADYEE  780
YVKCQERVSA LYKNPREWTR MVIRNIATSG KFSSDRTIAQ YAREIWGVEP TRQRMPAPDE  840
KI                                                                 842

SEQ ID NO: 97          moltype = AA  length = 364
FEATURE                Location/Qualifiers
source                 1..364
                       mol_type = protein
                       organism = Ovis aries
SEQUENCE: 97
MPHQYPALTP EQKKELCDIA HRIVAPGKGI LAADESTGSI AKRLQSIGTE NTEENRRFYR  60
QLLLTADDRV NPCIGGVILF HETLYQKADD GRPFPQVIKA KGGVVGIKVD KGVVPLAGTN  120
GETTTQGLDG LSERCAQYKK DGADFAKWRC VLKIGEHTPS SLAIMENANV LARYASICQQ  180
NGIVPIVEPE ILPDGDHDLK RCQYVTEKVL AAVYKALSDH HIYLEGTLLK PNMVTPGHAC  240
TQKYSHEEIA MATVTALRRT VPPAVPGITF LSGGQSEEEA SINLNAINKC PLLKPWALTF  300
SYGRALQASA LKAWGGKKEN LKAAQEEYVK RALANSLACQ GKYTPSGKAG AAASESLFIS  360
NHAY                                                               364

SEQ ID NO: 98          moltype = AA  length = 197
FEATURE                Location/Qualifiers
source                 1..197
                       mol_type = protein
                       organism = Ovis aries
SEQUENCE: 98
ILLNKMAPKK DVKKPAAAAA PAPAPAPAPA PAPAPPKEEK IDLSAIKIEF SKQQQDEFKE  60
AFLLYDRTGD GKITLSQVGD VLRALGTNPT NAEVKKVMGN PSNEEMNAKK IEFEQFLPML  120
QAISNNKDQG TYEDFVEGLR VFDKEGNGTV MGAELRHVLA TLGEKMKEEE VEALMAGQED  180
SNGCINYEAF VKHIMSN                                                 197

SEQ ID NO: 99          moltype = AA  length = 322
FEATURE                Location/Qualifiers
source                 1..322
                       mol_type = protein
                       organism = Ovis aries
SEQUENCE: 99
IGRLVTRAAF NTGKVDIVAI NDPFIDLHYM VYMFQYDSTH GKFHGTVKAE NGKLVINGKA  60
ITIFQERDPA NIKWGDAGAE YVVESTGVFT TMEKAGAHLK GGAKKVIISA PSADAPMFVM  120
GVNHEKYNNT LKIVSNASCT TNCLAPLAKV IHDHFGIVEG LMTTVHAITA TQKTVDGPSG  180
KLWRDGRGAA QNIIPASTGA AKAVGKVIPE LNGKLTGMAF RVPTPNVSVV DLTCRLEKPA  240
KYDEIKKVVK QASEGPLKGI LGYTEDQVVS CDFNSDTHSS TFDAGAGIAL NDHFVKLISW  300
YDNEFGYSNR VVDLMVHMAS KE                                           322

SEQ ID NO: 100         moltype = AA  length = 294
FEATURE                Location/Qualifiers
source                 1..294
                       mol_type = protein
                       organism = Ovis aries
SEQUENCE: 100
MDAIKKKMQM LKLDKENALD RAEQAEADKK AAEDRSKQLE DELVSLQKKL KATEDELDKY  60
SEALKDAQEK LELAEKKATD AEADVASLNR RIQLVEEELD RAQERLATAL QKLEEAEKAA  120
DESERGMKVI ESRAQKDEEK MEIQEIQLKE AKHIAEDADR KYEEVARKLV IIESDLERAE  180
ERAELSEGQV RQLEEQLRIM DQTLKALMAA EDKYSQKEDK YEEEIKVLSD KLKEAETRAE  240
FAERSVTKLE KSIDDLEDEL YAQKLKYKAI SEELDHALND MTSIESGSCQ RRKP        294

SEQ ID NO: 101         moltype = AA  length = 1141
FEATURE                Location/Qualifiers
source                 1..1141
                       mol_type = protein
                       organism = Ovis aries
SEQUENCE: 101
MPEAKPAAKK APAGKDAAAK PAPKEAPPQE APAAPSTEAP PEDQSPTAEE PTGVFLKKPD  60
SVSVENGKDT VIVAKLNGKE LPAKPAVKWF KGKWLELGSK SGARFSFKES HDAASNVYTV  120
ELHITKVVLG DRGDYRIEVK AKDFCDSCTF NIDVESPRQD SAGQSLESFK RSSEAKSDTA  180
GELDFSGLLK KRQVVEEEKK KKKDDDDLGI PPEIWELLKG AKKSEYERIA FQYGITDLRG  240
MLKRLKKAKV EVKKSAAFTK KLDPAYQVDK GNKIKLVVEI SDPDLPLKWF KNGQEIKPSS  300
KYVFENVGKK RILTINKCTL ADDAAYEVVV KDEKCFTELF VKVEPPVLIV TPLEDQQVFV  360
GNRVEMAVEV SEEGAQVMWL KDGVELTRED SFKSRYRLKK DGKRHILIYA EVALEDKGHY  420
KVMTNGDQCE AELIVEEKQL EVLQDIADLT VKASEQAVFK CEVSDEKVTG KWYKNGVEVR  480
PSKRITISHT GRIHKLVIDD VRPEDEGDYT FVPDGYALSL SAKLNFLGAG LSPLPPEPPK  540
IHLDCSGQTS ENSVVVVAGN KLRLDVSITG EPRPVATWMK EDEVFTGTEG RVRIEQRGDI  600
SSFVIESAER GDEGRYTIKV TNPAGEDVAS ILLKVVDVPD PPEAVRVTSV GEDWAVLVWE  660
PPKYDGGRPV TGYLLERKKK GSQRWMKLNF EVFTETTYES TNMIEGILYE MRVFAVNAIG  720
VSQPSMNTKP FMPIAPTSEP QHLTVEDVTD TTTTLKWRPP EKIGAGGIDG YLIEYCVEGS  780
DEWIPANTEP VERCGFTVKN LPTGAKITFR VVGVNIAGRS QPATLAQPVT IREIVEQPKI  840
RLPRHLRQTY VRKVGEHLNL VIPFQGKPRP QVVWTKGGAP IDTNHVHVRT SDFDTVFFVR  900
QAARSDSGEY ELTVQIENMK DTATIRIRVV EKAGPPENVM VKEVWGTNAL VEWQPPRDDG  960
```

-continued

```
NSEITGYLVQ KADKKTMEWF TVYERNRHTS CTVSDLIMGN EYYFRVYSEN ICGLSDLPGV  1020
SKNTARIVKT GMTLKLPEYK EHDFRTPPKF LTPLFDRVVV AGYAAALNCA VRGHPKPKVV  1080
WMKNKMEIRE DPKFLMTNHQ GVLTLNIRRP SPFDSGTYSC RAVNELGEAL AECKLEVRGE  1140
P                                                                  1141

SEQ ID NO: 102          moltype = AA  length = 1483
FEATURE                 Location/Qualifiers
source                  1..1483
                        mol_type = protein
                        organism = Ovis aries
SEQUENCE: 102
MSLVAVPFYQ KRHKHFDQSY RNIQTRYLLD EYAAKKKRAS AQSSLQRSVT RKSSSQRESS  60
RAALGETTLG ETTLGRTTCP LCAKRMSSAL EEEAQERKRR YQSMVAAYGE AKRERYLSEL  120
AQLEEDVHVA RTHAREELDR LALQHAGRHG LHHAVDDRLA WERHWFEDRI SRAPEILVRL  180
RSHTVWERMS VKLCFTVQGF PTPVVQWYKD GSLICQAGEP GKYRIDSKYG VHTLEINRAD  240
FDDTATYSAV ATNVHGQVST NAAVVVRRFR GDEEPFHSVG LPIGLPLSSV IPYTHFDVQF  300
IEKFGVTFRR EGETLTLKCT LLVTPDLKRV QPRAEWYRDD VLLKESKWTK MFFGEGQASL  360
SFSHLNKDDE GLYTLRIVSR GGVTDHSAFL FVRDADPLVT GAPGAPMDVH CHDANRDYVI  420
VTWKPPNTTK ESPVIGYFID RCEVGTDNWV QCNDAPVKIC KYPVTGLFEG RSYIFRVRAV  480
NSAGISRPSR VSEAVAALDP VDLRRLQAVH LEGEKDIVIY QEDLEGEVQI PGPPTNVHAS  540
ETSRTYVVLS WDPPEPRGKE PLMYFIEKSM VGSGSWQRVN AQTAVRSPRY AVFDLAEGKP  600
YVFRVLSANK HGLSEPSEIT APIQAQDTIV VPSAPGRVLA SRNTRTSVVV QWDRPKHEED  660
LLGYYVDCSV AGSNVWEPCN HKPIGYNRFV VHGLTTGEQY IFRVKAVNAV GTSENSQESD  720
VIKVQAALTV PSHPYGITLL NCDGHSMILG WKVPKFSGGS PILGYYIDKR EAHHKNWHEV  780
NSSPLKERIL KVEGLTEGSL YEFKIAAANM AGMGSQQTEL FKCEAWTMPE PGPAYDLTFC  840
EVRDTSLVLL WKPPVYSGSS PVSGYFVDYK EEDSGEWLTV NEVTTPHRYL KVCDLHQGKT  900
YVFRVRAVNA SGVGRPSDTS EPVLVEARPG TKEVSAGVDE EGNIYLSFDC PEVTDASHFT  960
WCKSYEDIAD DDRFKVETAG DHSKLYFKNL DKEDLGTYSV SVSDTDGVSS SFVLDEEELE  1020
RLKALSNEIK NPTIPLKSEL AYEIFDKGQV RFWLQAEHLS PDSSYRFIIN DREVADSETH  1080
RIKCDKSTGV IEMVMDRFTV DNEGTYTVQI QDGKAKSQSS LVLIGDAFKA VLEEAEFQRK  1140
EFLRKQGPHF AEYLHWDVTE ECEVRLVCKV ANTKKETVFK WLKDDVLYET EKMPDLEKGV  1200
CELLIPKLSK KDHGEYKATL KDDRGQDVST LEIAGKVYED MILAMSRVCG ASASPLKILC  1260
TPEGIRLQCF MKYFTEEMKV NWYHKDAKIA SSEHMRIGGS EEMAWLQICE PTEKDKGKYT  1320
FEILDGKKNH QRSLDLSGQA FDEAFAEFQQ LKAAAFAEKN RGKVIGGLPD VVTIMEGKTL  1380
NLTCTVFGNP DPEVVWFKND KDIELSDHFS VKVEQAKYVS MTIKGVTSED SGKYSMHVKN  1440
KYGGEKIDVT VSVYKHGEEI PDVVLPQQAK PKLLPAAAPS PAH                    1483

SEQ ID NO: 103          moltype = AA  length = 434
FEATURE                 Location/Qualifiers
source                  1..434
                        mol_type = protein
                        organism = Ovis aries
SEQUENCE: 103
MAMQKIFARE ILDSRGNPTV EVDLHTAKGR FRAAVPSGAS TGIYEALELR DGDKSRYLGK  60
GVLKAVEHIN KTLGPALLEK KLSVVDQEKV DKFMIELDGT ENKSKFGANA ILGVSLAVCK  120
AGAAEKGVPL YRHIADLAGN PELILPVPAF NVINGGSHAG NKLAMQEFMI LPVGASSFRE  180
AMRIGAEVYH HLKGVIKAKY GKDATNVGDE GGFAPNILEN NEALELLKTA IQAAGYPDKV  240
VIGMDVAASE FYRNGKYDLD FKSPDDPARH ISGEKLGELY KSFIKNYPVV SIEDPFDQDD  300
WATWTSFLSG VNIQIVGDDL TVTNPKRIAQ AVEKKACNCL LLKVNQIGSV TESIQACKLA  360
QSNGWGVMVS HRSGETEDTF IADLVVGLCT GQIKTGAPCR SERLAKYNQL MRIEEALGDK  420
AVFAGRKFRN PKAK                                                    434

SEQ ID NO: 104          moltype = AA  length = 170
FEATURE                 Location/Qualifiers
source                  1..170
                        mol_type = protein
                        organism = Ovis aries
SEQUENCE: 104
QAPKKAKRRA AAEGGSSNVF SMFDQTQIQE FKEAFTVIDQ NRDGIIDKED LRDTFAAMGR  60
LNVKNEELDA MMKEASGPIN FTVFLTMFGE KLKGADPEDV ITGAFKVLDP EGKGTIKKQF  120
LEELLTTQCD RFSQEEIRNM WAAFPPDVGG NVDYKNICYV ITHGDAKDQE             170

SEQ ID NO: 105          moltype = AA  length = 1937
FEATURE                 Location/Qualifiers
source                  1..1937
                        mol_type = protein
                        organism = Ovis aries
SEQUENCE: 105
MSSDQEMAIF GEAAPYLRKS EKERIEAQNK PFDAKTSVFV ADPKESFVKA TVQSREGGKV  60
TAKTEAGATV TVKEDQVFPM NPPKFDKIED MAMMTHLHEP AVLYNLKERY AAWMIYTYSG  120
LFCVTVNPYK WLPVYNAEVV AAYRGKKRQE APPHIFSISD NAYQFMLTDR ENQSILITGE  180
SGAGKTVNTK RVIQYFATIA VTGEKKKEEP TSGKMQGTLE DQIISANPLL EAFGNAKTVR  240
NDNSSRFGKF IRIHFGTTGK LASADIETYL LEKSRVTFQL KAERSYHIFY QIMSNKKPEL  300
IEMLLITTNP YDYAFVSQGE ITVPSIDDQE ELMATDSAIE ILGFTSDERV SIYKLTGAVM  360
HYGNLKFKQK QREEQAEPDG TEVADKAAYL QGLNSADLLK ALCYPRVKVG NEFVTKGQTV  420
EQVYNAVGAL AKAIYDKMFL WMVTRINQQL DTKQPRQYFI GVLDIAGFEI FDFNSLEQLC  480
```

-continued

```
INFTNEKLQQ FFNHHMFVLE QEEYKKEGIE WEFIDFGMDL AACIELIEKP MGIFSILEEE  540
CMFPKATDTS FKNKLYEQHL GKSNNFQKPK PAKGKAEAHF SLIHYAGTVD YNITGWLDKN  600
KDPLNETVVG LYQKSSVKTL AFLFSGAASA EAEGGGAKKG GKKKGSSFQT VSALFRENLN  660
KLMTNLRSTH PHFVRCIIPN ETKTPGAMEH ELVLHQLRCN GVLEGIRICR KGFPSRILYA  720
DFKQRYKVLN ASAIPEGQFI DSKKASEKLL GSIDVDHTQY KFGHTKVFFK AGLLGLLEEM  780
RDEKLAQLIT RTQARCRGFL ARVEYQRMVE RRESIFCIQY NVRAFMNVKH WPWMKLYFKI  840
KPLLKSAETE KEMANMKEEF EKTKEELAKS EAKRKELEEK MVTLTQEKND LQLQVQSEAD  900
ALADAEERCD QLIKTKIQLE AKIKEVTERA EDEEEINAEL TAKKRKLEDE CSELKKDIDD  960
LELTLAKVEK EKHATENKVK NLTEEMAGLD ETIAKLTKEK KALQEAHQQT LDDLQAEEDK 1020
VNTLTKAKTK LEQQVDDLEG SLEQEKKLRM DLERAKRKLE GDLKLAQEST MDIENDKQQL 1080
DEKLKKKEFE MSNLQSKIED EQALGMQLQK KIKELQARIE ELEEEIEAER ASRAKAEKQR 1140
SDLSRELEEI SERLEEAGGA TSAQIEMNKK REAEFQKMRR DLEEATLQHE ATAAALRKKQ 1200
GGVGEQLDNL QRVKQKLEKE KSEMKMEIDD LASNMETVSK AKGNLEKMCR TLEDQLSELK 1260
TKEDEQQRLI NDLTTQRARL QTESGEFSRQ LDEKDALVSQ LSRGKQAFTQ QIEELKRQLE 1320
EEIKAKSALA HALQSARHDC DLLREQYEEE QEAKAELQRA LSKANSEVAQ WRTKYETDAI 1380
QRTEELEEAK KKLAQRLQDA EEHVEAVNAK CASLEKTKQR LQNEVEDLMI DVERTNAACA 1440
ALDKKQRNFD KILSEWKQKY EETHAELEAS QKESRSLSTE LFKIKNAYEE SLDQLETLKR 1500
ENKNLQPTEE ISDLTEQIAE GGKRIHELEK VKKQVEQEKS EIQAALEEAE ASLEHEEGKI 1560
LRIQLELNQV KSEIDRKIAE KDEEIDQLKR NHIRVVESMQ STLDAEIRSR NDAIRLKKKM 1620
EGDLNEMEIQ LNHANRMAAE ALKNYRNTQA ILKDTQLHLD DALRGQEDLK EQLAMVERRA 1680
NLLQAEIEEL RATLEQTERS RKIAEQELLD ASERVQLLHT QNTSLINTKK KLETDITQIQ 1740
GEMEDIIQEA RNAEEKAKKA ITDAAMMAEE LKKEQDTSAH LERMKKNLEQ TVKDLQHRLD 1800
EAEQLALKGG KKQIQKLEAR VRELEAEVES EQKRNVETVK GLRKHERRVK ELTYQTEEDR 1860
KNILRLQDLV DKLQAKVKSY KRQAEEAEEQ SNVNLSKFRK LQHELEEAEE RADIAESQVN 1920
KLRVKSREVH TKIISEE                                                1937

SEQ ID NO: 106          moltype = AA  length = 6911
FEATURE                 Location/Qualifiers
source                  1..6911
                        mol_type = protein
                        organism = Ovis aries
SEQUENCE: 106
MADDEEYEEV VEYYTEETVY EEVPGETITE FYETTTTRTS DYEQSDTSRP ALAKPVPEKP   60
VERKRVIRKK VDPSKFMTPY IAHSQKMQDL FSTNKYKENY EKAKGQPYAI TTDTPELRRI  120
KKVQDQLSEV KYRTDGDVAK TICHVDEKAK DIEHAKKVSQ QVSKVLYKQN WEDTKDKYLL  180
PPDAPELVQA IKNTAMFSKK LYTEDWEADK TLFYPYNDSP ELRRVAQAQK ALSDITYKKG  240
LAEQQTQFTS LPDPPDIEFA KKVTNQVSKQ KYKEDYENKV KGKWSETPCF EIATARMNSQ  300
NISTKKYQED FENTKDQIYF MQTETPEYKV NKQAGVAASK VKYKQDYEKN KGKADYNVLP  360
ASENPLLRQL KAAGDALSDK LYKENYEKTK AKSINYCETP KFKLDTVLHN FSSDTKYKDS  420
YLKNILGHYV GSYEDPYHTH CMRVSAQNSD KNYKAEYEED RGKGFFPQTI TQEYEAIKKL  480
DQCKDHTYKV HPDKTKFTQV TDSPVLVQAQ VNSKQLSDLN YKAKHENEKF KCHIPPDTPA  540
FIQHKVNAYN LSDNIYKHDW EKTKAKKFDI KVDAIPLLAA KANTKIASDV MYKKDYEKSK  600
GKMIGALSIN DDPKMLHSLK TAKNQSDRLY KENYEKTKAK SMNYCETPKY QLDTLLKNFS  660
EAKYKDSYVQ NVLGHYIGSF EDPYREHCMK VAAQNSDKNY KAEYEEDKGK CYFPQTITQE  720
YEAIKKLDQC KDHTYKVHPD KTKFTAVTDT PVLLQAQLNT KQLSDLNYKA KHEGEKFKCH  780
IPADAPQFIQ HRVNAYNLSD NIYKHDWEKT KAKKFDIKVD AIPLLAAKAN TKIASDVMYK  840
KDYEKSKGKM IGALSINDDP KMLHSLKTAK NQSDREYRKD YEKSKTIYTA PLDMLPLTHA  900
KKSQEIASDV DYKHLLHNYS YPPDSVNVDL AKKAYALQSD VEYKADYNSW MKGCGWVPFG  960
SLEMEKAKRA SDILNEKKYR QHPDTLKFTS IEDAPIIVQS KINQAQRSDV AYKAKGEEVI 1020
HKYSLPADLP QFIQAKVNAY NISENMYKAD LKDLSKKGYD LRIDAIPIKA AKAARQAASD 1080
VQYKKDYEKA KGKMVGFQSL QDDPKLVHYM NVAKIQSDRE YKKDYEKTKT RHNTPHDMVN 1140
IVAAKKAQDV ASNVNYKHSL HHYTYLPDAM DLELSKNMMH IQSDNVYKED YNNWMKGIGW 1200
IPIGSLEVEK VKKAGDALNE KKYRQHPDTL KFTSIVDSPV MVQAKQNTQQ VSDILYKAKG 1260
EDVKHKYTMS PDLPQFLQAK CNAYNLSDVC YKRDWHDLIA KGTNVLGDAI PITAAKASRN 1320
IASDYKYKEA YEKAKGKQVG FRSLQDDPKL VHYMNVAKLQ SDREYKKNYE NTKTSYHTPG 1380
DMVSITAAKM AQDVATNVNY KQPIHHYTYL PDALSLEHIR NVNQIQSDKV YKDEYNHFFK 1440
GIGWIPIGSL EVEKVKKAGD ALNEKKYRQH PDTIKFTSVP DSMGMVLAQH NTKQLSDLNY 1500
KVEGEKVKHK YTLDPDVPQF IQAKVNAFNM SDSHYKADWK KTLAKGYDLR PDAIPIVAAK 1560
SSRNIASDFK YKEGYEKAKG KQIGFRSLQD DPKLVHFMNV AKMQSDREYK KGYEASKTKY 1620
HTPLDMLSVT AAKKSQEVAT NTNYKQPFHH YTLLPDALNV EHSRNAMQIQ SDNLYKSDFT 1680
SWMKGIGWMP LESLEVEKAK KAGEILSEKK YRQHPEKLKF TYAMDTMEQA LNKSNKLIMD 1740
KRRYTEKWNK DKTTIHVMPD TPDILLSRVN QITMSDKLYK AGWEEEKKKG YDLRPDAIPI 1800
KAARASRDIA SDYKYKQAYE QAKGKQIGFR SLEDDPKLVH FMQVAKMQSD REYKKAYEKS 1860
KTSFHTPVDM LSVVAAKKSQ EVATNANYRN VIHNYNMLPD AMSLELAKNM MQIQSDNQYK 1920
ADYADFMKGI GWLPLGSLEA EKNKKAMEII SEKKYRQHPD TLKYSTLMDS MNMVLAKNNA 1980
KIMNEHLYKQ AWEADKTKVH IMPDIPQIIL AKANAINMSD KLYKLSLEEA KKKGYDLRTD 2040
AIPIKAAKAS RDIASDYKYK HNYEKERGKM VGFRSLEDDP KLVHSMQVAK MQSDREYKKN 2100
YEKTKTSYHT PADMLSVTAA KDAQANITNT NYKHLIHKYI LLPDAMNIQL SKNMNRLQSD 2160
NEYKQDYNEW YKGLGWSPAG SLEVEKAKRA TEYASDQKYR QHPSNFQFSK LNDSMDMVLA 2220
KQNAHTMNKY LYTVDWNKDK TKIHVMPDTP DILQAKQNQT IYSQKLYKLG WEEALKKGYD 2280
LPVDAISVQA AKASRDIVSD FKYKQGYRKQ IGHHIGFRSL QDDPKLVLSM NVAKMQSERE 2340
YKKDFEKWKT KFSSPVDMLG VVLAKKCQAL VSDVDYKNYL HGWTCLPDQN DVIHAKKAYD 2400
LQSENLYKSD LEWLKGIGWS PLGSLEAEKN KRASEIISEK KYRQPPDRNK FTSIPDAMDI 2460
VLAKTNAKNR SDILYREAWD KDKTQVHIMP DTPDIILAKA NLINASDKFY RMGYEELRKK 2520
GYDLPVDAIP IKAAKASREI ASEYKYKEGF RKQLGHHIGA RNIKDDPKMM WSMHVAKIQS 2580
DREYKKDFEK WKTKFSSPVD MLGVVLAKKC QTLVSDIDYK NYLHQWTCLP DQSDVIHARR 2640
```

-continued

```
AYDLQSDNLY KSDLQWLKGI GWLPSGSLED EKNKRATQIL SDHVYRQHPD KFKFSSLMDS  2700
IPMVLAKNNA ITMNHRLYTE AWDKDKTTVH IMPDTPGVLL AKQNQINYSE KLYKLGLDEA  2760
KRKGYDMRID AIPIRAAKAS RDIASEFKYK EGYRKQLGHH IGARAIHDDP KMMWSMHVAK  2820
IQSDREYKKD FEKWKTKFSS PVDMLGVVLA KKCQTLVSDI DYKNYLHQWT CLPDQSDVIH  2880
ARQAYDLQSD NVYKSDLQWL KGIGWVPIGS LDVEKCKRAS EILSDKLYRQ PPDKFKFTSV  2940
TDSLEQVLAK NNAINMNKRL YTEAWDKDKT QVHIMPDTPE ITLARMNKVN YSESLYKLAH  3000
EEAKKKGYDL RSDAIPIVAA KASRDIISDY KYKDGYRKQL GHHIGARDIQ DDPKMMWSMH  3060
VAKIQSDREY KKDFEKWKTK FSSPVDMLGV VLAKKCQTLV SDIDYKNYLH QWTCLPDQSD  3120
VIHARRAYDL QSDNVYKSDL QWLRGIGWVP IGSVDVVKCK RAAEILSDNL YRQPPDTLKF  3180
TSVPDSLEQV LAKNNAINMN KRLYTEAWDK DKTQVHIMPD TPEIMLARQN KLNYSETLYK  3240
LANEEAKKKG YDLRSDAIPI VAAKASRDII SDYKYKDGYR KQLGHHIGAR DIKDDPKMMW  3300
SMHVAKIQSD REYKKDFEKW KTKFSSPVDM LGVVLAKKCQ TLVSDIDYKN YLHQWTCLPD  3360
QSDVIHARQA YDLQSDNIYK SDLQWLRGIG WVPIGSVDVV KCKRAAEILS DNLYRQRPDT  3420
LKFTSVPDSL EQVLAKNNAI NMNKRLYTEA WDKDKTQIHM MPDTPEITLA RQNKINYSEN  3480
LYRQAMEEAK KEGYDLRSDA IPIVAAKASR EIASDYKYKE AYRKQLGHHI GARAIHDDPK  3540
MMWSVHVAKM QSDREYKKDF EKYKTRFSSP VDMLGIVLAK KCQTLVSDVD YKHPLHEWTC  3600
LPDQNDIIHA RKAYDLQSDN LYKSDLEWLK GIGWVPIGSF DVVKAKRAGE ILSDNIYRQR  3660
PDTLKFTSVT DSPEQVLAKN NAINMSKRLY TEAWDNDKKT IHVMPDTPEI MLAKLNRINY  3720
SDKLYKLALE ESKKEGYDLR LDAIPIQAAK ASREIASDYK YKEGYRKQLG HHIGARNIKD  3780
DPKMMWSIHA TKLQSDREYK KDFEKWKTKF SSPVDMMGLV QAKKCQILVS DIDYKNLLHE  3840
WICLPDQNDV IQARKAYDLQ SDAIYKADLE WLRGIGWVPI GSVGVENVKR AGEILSERKY  3900
RQPAVQLKFT SITDTPEIVL AKNNALNMSK HLYTEAWDAD KTSIHVMPDT PEILLAKSNS  3960
ANISHKLYTK GWDESKMKDY DLRADAISIK SAKASRDIAS DYKYKEAYEK QKGHHIGARS  4020
VEDDPKIMCA MNAGKIQSER EYKKEFQKWK TKFSSPVDML GILQAKKCQI LVSDIDYRNY  4080
LHHWTCLPDQ NDIIQARKAY DLQSDAIYKA DLEWLRGIGW MPEGSPEVLR VKNAQHIFRD  4140
SVYRTPVVKL KYTSIVDTPE VVLAKSNAEN ISIPKYREVW DKDKTSIHIM PDTPEINLAR  4200
TNALNVSNKL YREGWDEMKM SCDVRLDAIP IQAAKASREI ASDYKYKLDH EKQKGHYVGT  4260
RTARDDNKIR WALIAGKIQN EREYRMHWAK WKSKFQSPPD MLSIEHSKNS QTLVSDIDYR  4320
HYLHQWTCMP DQNDVIQARK AYDLQSDAIY KADLEWLRGI GWMPADSVSV NHAKHMGDIF  4380
NEKKYRTKIE TLNFTPVDDR VDYVTAKHSG EILNDIKYRK DWNDTKSKYT LTETPQLHAA  4440
QEAARILDQY LYKESWEKQK ATGYILPPDA VPFVHAHHSG DVQSELKYKA EHVKQKGHYV  4500
GVPTMRDDPK LVWFEHAGQI QNDRLYKESY HKTKAKIHIP PDMVSVLAAK EGQALASDID  4560
YRTYLHQWIC HPDQNDVIQA RKAYDLQSDN IYKADLEWLR GIGWMPLDSV DHVRVTRNQE  4620
MVNQIKYKKD ALANYPNFTS VVDPPEIVLA KINAVNQSDV KYKETFNKLI KGKYIFSPDT  4680
PYITHSKDME KLYSTILYKR AWEGTKAYGY TLDERYIPIV GAKHADFVNS ELKYKETYEK  4740
LKGHYLAGKE ISEFPNVVHC LDFQKMRSLL NYRRHYEDTK ANVHIPQDMM NHVLAKRCQY  4800
ILSDLEYRHY FHQWTSLPEE PNVVRVRHAQ EILSDNVYKD DLNWLKGIGC YVWDTPQILH  4860
AKKSYDLQSQ ILYTAAGKEN LKNYNLVTDT PLYVTALQSG INASEVKYKE NYHQIKDKYT  4920
TVLETVDYDR TRNLKDLFSS NLYKEAWDKV KATSYILPPN TVSLTHAKNQ KHMASYVKYR  4980
EEYDKFKALY TLPRSVEDDP NTARCLRVGK FNIDRLYRSV YEKNKMNIHI VPDMVEMVTA  5040
KDSQKKVSEI DYRLHLHEWI CHPDLQVNSH VRKVTDQISD IVYKDDLTWL KGIGCYVWDT  5100
PEILHAKHAY DLRNDIKYKA HVLKTRNNYK LVTDTPVYVQ AVKSGKQLSD AVYHYDYVHS  5160
IRGRVAPTTK TVDLDRALHA YKLQSENLYR KAGLHALPTG YRLPVDTPHF KHTKDTRYMS  5220
SYFKYKEVYE HMKAYGYTLG PNDVPFVNVR RVNNITSERL YRQLYHKLKD KIHTTPDTPE  5280
IRQVKKTQEA VSELIYKSDF FKMQGHMISM PYTPHVLHCR YVGDITSDIK YKEDLQILRG  5340
LGCFLYDTPD MVRSRHLRKL WSHYLYTDKA RKMRDKYQVV LDTPEYRTVQ ELKTHLSELV  5400
YRASGRKQKS IFTSVPDTPD LIRAKRGQKL QSQYLYVELA TKERPHHHAG NQTTALKHAR  5460
HVKDMVSENK YKIQYEKMKD KYTPVPDTPI LIRAKRAYWN ASDLRYKETF QKTKGRYHTV  5520
KDALDIVYHR TVTDHISKIK YKENYMSQLG IWRSIPDRPE HFHHRAVTDA VSDVKYKADL  5580
TWLKGIGCYA YDTPDFTLAE KNKTLYSKYK YKEVFERTKS NFKYVADCPI NRHFKFATQL  5640
MNEKKYRADY EQRKDKYHLV VDEPRHLLAK IAGDQISQIR YRKNYEKTKD KFTSIVDTPE  5700
HLRTTKVNKQ ISDILYKLEY NKAKPRGYTT IHDTPMLLHV RKVKDEVSDL KYKEVYQRTK  5760
SNCTIEPDAV HIKAAKDAYK VNTNLDYKKQ YEATKAYWKW TPDRPDFIQA AKSTLQQSDF  5820
EYKLDREYLK GCKLSVTDDK DMVLALKNSI IESDLKYKEK HVKERGSCHA VPDTPQILLA  5880
KTVSSLVSEN KYKQYVKKHL AQGSYTTLPE TRDTIHVKEV TKNVSDTNYK KKFVKEKGKS  5940
NYSIMLEPPD VKHAMDVAKK QSNMKVAYKK DAKENLHYTT VADRPDIKKA TQAAKQASEV  6000
EYRAKHRKEG SHGLSMLGRP DIEMAKKAAK LSSQVKYREN FNKEKGKTPK YNPKDSQLYK  6060
VMKDANTLAS EVKYKADLKK LHKPVTDMKE SLIMNHVLNT SHLASSYQYK KNYEKSKGHY  6120
HTIPDNLEQL HLKEATELQS IVKYKEKYEK ERGKPMLDFE TPTYITAKES QQMQSGKEYR  6180
RDYEESIKGR NLTGLEVTPA LLHVKYATKI ASEKEYRKDL EESIRGKGLS EMEDTPDMLR  6240
AKNATQILNE KEYKRDLELE VKGRGLNAMA NETPDFLRAR NATDIASQIK YKQSAEMEKA  6300
NFTSVVDTPE IIHAQQVKNL SSQKKYKEDA EKCMSYYETV LDTPEMQRVR ENQKNFSLLQ  6360
YQYDLKNSKG KITVVQDTPE ILRVKENQKN FSSVLYKEDV SPGTAIGKTP EMMRVKQTQD  6420
HISSVKYKEV IGQGTPIPDL PEVKRVKQTQ KHISSVMYKE NLGTGIPTPV TPEIERVKRN  6480
QENFSSVLYK ENLGKGTPTP ITPEMERVKR NQENFSSILY KENLSKGTPL PVTPEMERVK  6540
RNQENFSSVL YKENVGKGTP TPVTPEMQRV KRNQENISSV LYKENLGKAT PTPFTPEMER  6600
VKRNQENFSS VLYKENMRKA TPTPVTPEME RAKRNQENIS SVLYSDSFRK QIQGKAAYVL  6660
DTPEMRRVRE TQRHISTVKY HEDFEKHKGC FTPVVTDPIT ERVKKNMQDF SDISYRGIQR  6720
KVVEMEQKRN DQDQETVTGL RVWRTNPGSV FDYDPAEDNI QSRSLHMINA QAQRRSREQS  6780
RSASGLSVSG GEEKSEHSEA AHLSTYSDGG VFFSATSTGY KHARTTELPQ QRSSSVATQQ  6840
TTVSSIPSHP STAGKIFRAM YDYMAADADE VSFKDGDAIV NVQAIDEGWM YGTVQRTGRT  6900
GMLPANYVEA I                                                       6911
```

```
SEQ ID NO: 107          moltype = AA  length = 1948
FEATURE                 Location/Qualifiers
source                  1..1948
```

-continued

```
                        mol_type = protein
                        organism = Ovis aries
SEQUENCE: 107
MSSDQEMAIF GEAAPYLRKS EKERIEAQNK PFDAKTSVFV ADPKESFVKA TVQSREGGKV  60
TAKTEAGATV TVKEDQVFPM NPPKFDKIED MAMMTHLHEP GVLYNLKERY AAWMIYTYSG  120
LFCVTVNPYK WLPVYNAEVV TAYRGKKRQE APPHIFSISD NAYQFMLTDR ENQSILITGE  180
SGAGKTVNTK RVIQYFATIA VTGEKKKEEP TSGKMQGTLE DQIISANPLL EAFGNAKTVR  240
NDNSSRFGKF IRIHFGTTGK LASADIETYL LEKSRVTFQL KAERSYHIFY QIMSNKKPEL  300
IEMLLITTNP YDYAFVSQGE ITVPSIDDQE ELMATDSAIE ILGFTSDERV SIYKLTGAVM  360
HYGNLKFKQK QREEQAEPDG TEVADKAAYL QSLNSADLLK ALCYPRVKVG NEFVTKGQTV  420
QQVYNAVGAL AKAVYEKMFL WMVTRINQQL DTKQPRQYFI GVLDIAGFEI FDFNSLEQLC  480
INFTNEKLQQ FFNHHMFVLE QEEYKKEGIE WEFIDFGMDL AACIELIEKP MGIFSILEEE  540
CMFPKATDTS FKNKLYEQHL GKSNNFQKPK PAKGKAEAHF SLVHYAGTVD YNIAGWLDKN  600
KDPLNETVVG LYQKSAMKTL AFLFAGGPSA EVHVSEGGGG KKGGKKKGSS FQTVSALFRE  660
NLNKLMTNLR STHPHFVRCI IPNETKTPGA MEHELVLHQL RCNGVLEGIR ICRKGFPSRI  720
LYADFKQRYK VLNASAIPEG QFIDSKKASE KLLGSIDIDH TQYKFGHTKV FFKAGLLGIL  780
EEMRDEKLAQ LITRTQAICR GFLTRVEFRK MMERSRESIF CIQYNVRAFM NVKHWPWMKL  840
YFKIKPLLKS AETEKEMANM KEEFEKTKEE LAKSEAKRKE LEEKMVTLMQ EKNDLQLQVQ  900
SEADGLADAE ERCDQLIKTK IQLEAKIKEL TERAEDEEEI NAELTAKKRK LEDECSELKK  960
DIDDLELTLA KVEKEKHATE NKVKNLTEEM AGLDENIAKL TKEKKALQEA HQQTLDDLQA  1020
EEDKVNTLTK SKTKLEQQVD DLEGSLEQEK KLRMDLERAK RKLEGDLKLA QESTMDIEND  1080
KQQLDEKLKK KEFEMSNLQS KIEDEQALGM QLQKKIKELQ ARIEELEEEI EAERASRAKA  1140
EKQRSDLSRE LEEISERLEE AGGATSAQIE MNKKREAEFQ KMRRDLEEAT LQHEATAAAL  1200
RKKHADSVAE LSEQIDNLQR VKQKLEEEGS NLTVQALDII SNFKIINRYY YIEGNLEKMC  1260
RTLEDQLSEV KSKEEEHQRL INELSAQKAR LHTESGEFSR QLDEKDALVS QLSRGKQAFT  1320
QQIEELKRQL EEETKAKSAL AHALQSARHD CDLLREQYEE EQEAKAELQR SMSKANSEVA  1380
QWRTKYETDA IQRTEELEEA KKKLAQRLQE AEEHVEAVNA KCASLEKTKQ RLQNEVEDLM  1440
LDVERSNAAC AALDKKQRNF DKVLSEWKQK YEETQAELEA SQKESRSLST ELFKVKNAYE  1500
ESLDQLETLK RENKNLQQEI SDLTEQIAEG GKHIHELEKV KKQIEQEKGE LQAALEEAEA  1560
SLEHEEGKIL RIQLELNQVK SEIDRKIAEK DEEIDQLKRN HLRVVESMQS TLDAEIRSRN  1620
DALRIKKKME GDLNEMEIQL NHANRQATEA TKNLRNTQGV LKDTQLHLDD AIRGQDDLKE  1680
QLAMVERRAN LMQAEIEELR ASLEQTERSR RVAEQELLDA SERVQLLHTQ NTSLINTKKK  1740
LETDITRVHG SVVGSGIISK HLTFVYISVG CAFQAAMMAE ELKKEQDTSA HLERMKKNLE  1800
QTVKDLQHRL DEAEQLALKG GKKQIQKLEA RVRELENEVE NEQKRNAEAV KGLRKHERRV  1860
KELTYQTEED RKNVLRLQDL VDKLQSKVKA YKRQAEEAEE QANVNLSKFR KLQHELEEAE  1920
ERADIAESQV NKLRVKSREV HTKVISEE                                     1948

SEQ ID NO: 108          moltype = AA  length = 607
FEATURE                 Location/Qualifiers
source                  1..607
                        mol_type = protein
                        organism = Ovis aries
SEQUENCE: 108
MKWVTFISLL LLFSSAYSRG VFRRDTHKSE IAHRFNDLGE ENFQGLVLIA FSQYLQQCPF  60
DEHVKLVKEL TEFAKTCVAD ESHAGCDKSL HTLFGDELCK VATLRETYGD MADCCEKQEP  120
ERNECFLNHK DDSPDLPKLK PEPDTLCAEF KADEKKFWGK YLYEVARRHP YFYAPELLYY  180
ANKYNGVFQE CCQAEDKGAC LLPKIDAMRE KVLASSARQR LRCASIQKFG ERALKAWSVA  240
RLSQKFPKAD FTDVTKIVTD LTKVHKECCH GDLLECADDR ADLAKYICDH QDALSSKLKE  300
CCDKPVLEKS HCIAEVDKDA VPENLPPLTA DFAEDKEVCK NYQEAKDVFL GSFLYEYSRR  360
HPEYAVSVLL RLAKEYEATL EDCCAKEDPH ACYATVFDKL KHLVDEPQNL IKKNCELFEK  420
HGEYGFQNAL IVRYTRKAPQ VSTPTLVEIS RSLGKVGTKC CAKPESERMP CTEDYLSLIL  480
NRLCVLHEKT PVSEKVTKCC TESLVNRRPC FSDLTLDETY VPKPFDEKFF TFHADICTLP  540
DTEKQIKKQT ALVELLKHKP KATDEQLKTV MENFVAFVDK CCAADDKEGC FVLEGPKLVA  600
STQAALA                                                            607

SEQ ID NO: 109          moltype = AA  length = 381
FEATURE                 Location/Qualifiers
source                  1..381
                        mol_type = protein
                        organism = Sus scrofa
SEQUENCE: 109
MPFGNTHNKY KLNFKAEEEY PDLSKHNNHM AKALTLEIYK KLRDKETPSG FTLDDVIQTG  60
VDNPGHPFIM TVGCVAGDEE SYVVFKDLFD PIIQDRHGGY KPTDKHKTDL NHENLKGGDD  120
LDPNYVLSSR VRTGRSIKGY TLPPHCSRGE RRAVEKLSVE ALNSLTGEFK GKYYPLKSMT  180
EQEQQQLIDD HFLFDKPVSP LLLASGMARD WPDARGIWHN DNKSFLVWVN EEDHLRVISM  240
EKGGNMKEVF RRFCVGLQKI EEIFKKAGHP FMWNEHLGYV LTCPSNLGTG LRGGVHVKLA  300
HLSKHPKFEE ILTRLRLQKR GTGGVDTAAV GSVFDVSNAD RLGSSEVEQV QLVVDGVKLM  360
VEMEKKLEKG QSIDDMIPAQ K                                            381

SEQ ID NO: 110          moltype = AA  length = 902
FEATURE                 Location/Qualifiers
source                  1..902
                        mol_type = protein
                        organism = Sus scrofa
SEQUENCE: 110
MMMVMQPEGL GTREGPFAGG SGGGGEYMEQ EEDWDRDLLL DPAWEKQQRK TFTAWCNSHL  60
```

-continued

```
RKAGTQIENI EEDFRNGLKL MLLLEVISGE RLPRPDKGKM RFHKIANVNK ALDFIASKGV  120
KLVSIGAEEI VDGNLKMTLG MIWTIILRFA IQDISVEETS AKEGLLLWCQ RKTAPYRNVN  180
VQNFHTSWKD GLALCALIHR HRPDLIDYAK LRKDDPIGNL NTAFEVAEKY LDIPKMLDAE  240
DIVNTPKPDE KAIMTYVSCF YHAFAGAEQA ETAANRICKV LAVNQENEKL MEEYEKLASE  300
LLEWIRRTVP WLENRVGEPS MSAMQRKLED FRDYRRLHKP PRVQEKCQLE INFNTLQTKL  360
RLSHRPAPMP SEGKLVSDIA NAWRGLEQVE KGYEDWLLSE IRRLQRLQHL AEKFQQKASL  420
HEAWTRGKED MLSQRDYESA SLQEVRALLR RHEAFESDLA AHQDRVEHIA ALAQELNELD  480
YHEAASVNSR CQAICDQWDN LGTLTQKRRD ALERMEKLLE TIDQLQLEFA RRAAPFNNWL  540
DGAVEDLQDV WLVHSVEETQ SLVTAHEQFK ATLPEADRER GAILGIQGEI QKICQTYGLR  600
PSSTNPYIAL SPQDINTKWD TVRKLVPSRD QMLQEELTRQ QVNERLRRQF AAQANAIGPW  660
IQGKVEEVGR LAAGMAGSLE EQMAGLRQQE QNIINYKSNI DRLEGDHQLL QESLVFDNKH  720
TVYSMEHIRV GWEQLLTSIA RTINEVENQV LTRDAKGLSQ EQLNEFRASF NHFDRKRNGM  780
MEPDDFRACL ISMGYDLGEV EFARIMTMVD PNAAGVVTFQ AFIDFMTRET AETDTAEQVV  840
ASFKILAGDK NYITAEELRR ELPAEQAEYC IRRMAPYKGA GAPAGALDYV AFSSALYGES  900
DL                                                                 902

SEQ ID NO: 111          moltype = AA  length = 844
FEATURE                 Location/Qualifiers
source                  1..844
                        mol_type = protein
                        organism = Sus scrofa
SEQUENCE: 111
MAMSRPLTDQ EKRKQISVRG LAGVENVAEL KKNFNRHLHF TLVKDRNVAN ARDYYFALAY  60
TVRDHLVGRW IRTQQHYYEK DPKRIYYLSL EFYMGRTLQN TMVNLALENA CDEATYQLGL  120
DMEELEEMEE DAGLGNGGLG RLAACFLDSM ATLGLAAYGY GIRYEFGIFN QKITGGWQME  180
EADDWLRYGN AWEKARPEFT LPVHFYGRVE HTNQGAKWVD TQVVLAMPYD TPVPGYRNNV  240
VNTMRLWSAK APNDFNLKDF NVGGYIQAVL DRNLAENISR VLYPNDNFFE GKELRLKQEY  300
FVVAATLQDI IRRFKSSKFG CRDPVRTSFD AFPDKVAIQL NDTHPSLAIP ELMRILVDEE  360
RLEWEKAWDV TVRTCAYTNH TVLPEALERW PVHLMETLLP RHLQIIYEIN QRFLNRVAAA  420
YPGDVDRLRR MSLVEEGAVK RINMAHLCIA GSHAVNGVAR IHSEILKKTI FKDFYELEPH  480
KFQNKTNGIT PRRWLVMCNP GLAEVIAERI GEEYIADLDQ LRKLLSYVDD EAFIRDVAKV  540
KQENKLKFSA YLEREYKVHI NPNSLFDIQV KRIHEYKRQL LNCLHVITLY NRIKKEPNKF  600
VVPRTVMIGG KAAPGYHMAK MIIKLITAIG DVVNHDPVVG DRLRVIFLEN YRVSLAEKVI  660
PAADLSEQIS TAGTEASGTG NMKFMLNGAL TIGTMDGANV EMAEEAGEEN FFIFGMRVED  720
VERLDQKGYN AQEYYDRIPE LRHVIDQLSS GFFSPKQPDL FKDIVNMLMH HDRFKVFADY  780
EDYIKCQERV SALYKNPREW TRMVIRNIAT SGKFSSDRTI AQYAREIWGV EPSRQRLPAP  840
DEKI                                                               844

SEQ ID NO: 112          moltype = AA  length = 368
FEATURE                 Location/Qualifiers
source                  1..368
                        mol_type = protein
                        organism = Sus scrofa
SEQUENCE: 112
MPHQYPALTP EQKKELSDIA HRIVAPGKGI LAADESTGSI AKRLQSIGTE NTEENRRFYR  60
QLLLTADDRV NPCIGGVILF HETLYQKADD GRPFPQVIKA KGGVVGIKVD KGVVPLAGTN  120
GETTTQGLDG LSERCAQYKK DGADFAKWRC VLKIGEHTPS SLAIMENANV LARYASICQQ  180
VGLQNGIVPI VEPEILPDGD HDLKRCQYVT EKVLAAVYKA LSDHHIYLEG TLLKPNMVTP  240
GHACTQKYSH EEIAMATVTA LRRTVPPAVP GITFLSGGQS EEEASINLNA INKCPLLKPW  300
ALTFSYGRAL QASALKAWGG KKENLKAAQE EYVKRALANS LACQGKYTPS GKAGAAASES  360
LFISNHAY                                                           368

SEQ ID NO: 113          moltype = AA  length = 191
FEATURE                 Location/Qualifiers
source                  1..191
                        mol_type = protein
                        organism = Sus scrofa
SEQUENCE: 113
MAPKKDVKKP AAAAAPAPAP APAPAPAPAP PKEEKIDLSA IKSFSADQIA EFKEAFLLFD  60
RTGECKITLS QVGDVLRALG TNPTNAEVKK VLGNPSNEEM NAKKIEFEQF LPMLQAISNN  120
KDQGSYEDFV EGLRVFDKEG NGTVMGAELR HVLATLGEKM KEEEVEALMA GQEDSNGCIN  180
YEAFVKHIMS I                                                       191

SEQ ID NO: 114          moltype = AA  length = 333
FEATURE                 Location/Qualifiers
source                  1..333
                        mol_type = protein
                        organism = Sus scrofa
SEQUENCE: 114
MVKVGVNGFG RIGRLVTRAA FNSGKVDIVA INDPFIDLHY MVYMFQYDST HGKFHGTVKA  60
ENGKLVINGK AITIFQERDP ANIKWGDAGA TYVVESTGVF TTMEKAGAHL KGGAKRVIIS  120
APSADAPMFV MGVNHEKYDN SLKIVSNASC TTNCLAPLAK VIHDHFGIVE GLMTTVHAIT  180
ATQKTVDGPS GKLWRDGRGA AQNIIPASTG AAKAVGKVIP ELNGKLTGMA FRVPTPNVSV  240
VDLTCRLEKP AKYDDIKKVV KQASEGPLKG ILGYTEDQVV SCDFNSDTHS STFDAGAGIA  300
LNDHFVKLIS WYDNEFGYSN RVVDLMVHMA SKE                               333
```

-continued

```
SEQ ID NO: 115          moltype = AA  length = 284
FEATURE                 Location/Qualifiers
source                  1..284
                        mol_type = protein
                        organism = Sus scrofa
SEQUENCE: 115
MDAIKKKMQM LKLDKENALD RAEQAEADKK AAEDRSKRLE DELVSLQKKL KATEDELDKY  60
SEAPKDAQEK LELAEKKATD AEADVASLNR RIQLVEEELD RAQERLATAL QKLEEAEKAA  120
DESERGMKVI ESRAQKDEEK MEIQEIQLKE AKHIAEDADR KYEEVARKLV IIESDLERAE  180
ERAELSEGKC AELEEELKTV TNNLKSLEAQ AEKYSQKEDK YEEEIKVLSD KLKEAETRAE  240
FAERSVTKLE KSIDDLEDEL YAQKLKYKAI SEELDHALND MTSI                  284

SEQ ID NO: 116          moltype = AA  length = 1220
FEATURE                 Location/Qualifiers
source                  1..1220
                        mol_type = protein
                        organism = Sus scrofa
SEQUENCE: 116
MPEAKPATKK APAGKDAAKP APKEAPPKGA PEEPPKEAPP EDQSPTAEEP TGIFLKKPDS  60
VSVENGKDAL IVAKVNGKEL PGKPEIKWFK GKWLELGSKS GARFSFKESH DSASNVYTLE  120
LRIGKVVLGD RGDYRLEVKA KDFCDSCAFN IDVEAPRQDS AGQSLESFKR SGEGKPEGAG  180
ELDFSGLLKK RQVVEEEKKK KKDDDDLGIP PEIWELLKGA KKSEYEKIAF QYGITDLRGM  240
LKRLKKAKVE VKKSAAFTKK LDPAYQVDQG NKIKLVVEIS DPDLPLKWYK NGQEIKPSSK  300
YVFENVGKKR ILCINKCTLA DDAAYEVAVK DEKCFTELFV REPPVLIVTP LEDQQVYVGD  360
RVEMTVEVSE EGAQVMWLKD GVELTREDSF KSRYRFKKDG KRHILIYSDV SLEDRGRYQV  420
MTNGGQCEAE LSVEEKQLEV LQDIADLTVK ASEQAVFKCE VSDEKVTGKW YKNGVEVRPS  480
KRITISHVGR FHKLVIDDVR PEDEGDYTFV PDGYALSLSA KLNFLEIKVE YVPKQEPPKI  540
HLDCSGKTSE NAVVVVAGNK LRLDVSITGE PPPVATWMKE DEVFSGTEGR VRLEQRADIS  600
SFVIESAERG DEGRYTIKVT NPAGEDVASI FLRVVDVPDP PEAVRVTTVG EDWAILVWEP  660
PKYDGGQPVT GYLLERKKKG SQRWMKLNFE VFTETTYEST NMIEGILYEM RVFAVNAIGV  720
SQPSVNTKPF MPIAPTSEPQ HLTVEDVTDT TTTLKWRPPD RIGAGGIDGY LVEYCVEGSD  780
EWIPANTEPV ERCGFTVKNL PTGARITFRV VGVNIAGRSQ PATLAQPVTI REIVGEPPRQ  840
SSVLIRASPD SLVATTLSVP LRRVPQPYAE PAPGIVSPEQ PKIRLPRHLR QTYIRKVGEH  900
LNLVIPFQGK PRPQVVWTKG GAPIDTSRVH VRTSNFDTVF FVRQAARSDS GEYELSVQIE  960
NMKDTATVHI RVVEKAGPPQ NVMVKEVWGT NALVEWQPPK DDGNTEITGY FVQKADKKTM  1020
EWFNVYEHNR HTSCTVSDLI MGNEYYFRVY SENICGLSDS PGVSKNTARI LKTGITFKLP  1080
EFKEHDFRTP PKFLTPLPDR VVVAGYTAAL NCAVRGHPKP KVVWMKNQME IGEDPKFLMT  1140
NHQGVLTLNI RRPSPFDSGT YSCRASTSWG RPWPSASWKS ECRSEAPAST QGSWLLSPDP  1200
DLPEPLSLSR VITRFLSRTQ                                              1220

SEQ ID NO: 117          moltype = AA  length = 869
FEATURE                 Location/Qualifiers
source                  1..869
                        mol_type = protein
                        organism = Sus scrofa
SEQUENCE: 117
TMQHVYTKMS LVAVPFYQKR HRHFDQSYRN IQTRYLLDEY AAKKRASTQS SSQKSVTQES  60
SSQRESSSER ASSQTLLGGT TCRLCAKRVS TALEEEEQER KRRYHSLLAA YGEAKRERYL  120
SELAQLEEDV HLARTHARDK LDRFALQHVD THALRHAVGD RLAWERHSFE DRISRAPEIL  180
VRLRSHTVWE RMSVKLCFTV QGFPTPVVQW YKDGSLICQA GEPGKYRIES KYGVHTLEIN  240
RADFDDTATY SAVATNVHGQ VSTNAAVVVR RFHGDEEPFH SVGLPIRLPL SSVIPYTHFD  300
VQFLEKFGVT FRREGETLTL KCTLLVTPDL KRVQPRAEWY RDDVLLKESK WTKMFFGEGQ  360
ASLSFSHLNK DDEGLYTLRI VSRGGISDHS AFLFVRDADP LVMGAPGAPM DVHCHDANRD  420
YVIVTWKPPN TTKESPVLGY FIDRCEVGTD NWVQCNDAPV KICKYPVSGL VEGRSYTFRV  480
RAVNSAGVSR PSRASEAVAA LDPLDLKRLQ GIHWEGDKDI VIYQDDLEGE VQIPRPPTNV  540
HASEISRTYV VLSWDPPTPR GKEPLMYFIE KSVVGSGTWQ RENAQTAVRS PRYAVFDLAE  600
GKSYVFRVLS ANKHGLSDPS EITSPIQAQD TIVVPSAPGR VLASRNTKTS VVVQWDQPKH  660
GEDLLGYYVD CCVAGSNVWE PCNHKPIGYN RFVVHGLTTG EQYIFRVKAV NAVGTSENSQ  720
ESDVIKVQAA LTVPSHPYGI TLLNCDGHSM TLGWKVPKFS GGSPIIGYYV DKREGHHKNW  780
HEVNSSPLKE RILTVEGLTE GSTYEFKIAA PTWWGSGCRR IPVSSSSVRP GQRLSQVPTS  840
GWVGTLGFFL VLLSLYRRYL LGSGEHGDS                                    869

SEQ ID NO: 118          moltype = AA  length = 434
FEATURE                 Location/Qualifiers
source                  1..434
                        mol_type = protein
                        organism = Sus scrofa
SEQUENCE: 118
MAMQKIFARE ILDSRGNPTV EVDLHTAKGR FRAAVPSGAS TGIYEALELR DGDKSRYLGK  60
GVLKAVEHIN KTLGPALLEK KLSVVDQEKV DKFMIELDGT ENKSKFGANA ILGVSLAVCK  120
AGAAEKGVPL YRHIADLAGN PDLVLPVPAF NVINGGSHAG NKLAMQEFMI LPVGASSFRE  180
AMRIGAEVYH HLKGVIKGKY GKDATNVGDE GGFAPNILEN NEALELLKTA IQAAGYPDKV  240
VIGMDVAASE FYRNGKYDLD FKSPDDPSRH ITGEKLGELY KSFIKNYPVV SIEDPFDQDD  300
WKTWTSFLSG VNIQIVGDDL TVTNLKRIAQ AVEKKACNCL LLKVNQIGSV TESIQACKLA  360
QSNGWGVMVS HRSGETEDTF IADLVVGLCT GQIKTGAPCR SERLAKYNQL MRIEEALGDK  420
AVFAGRKFRN PKAK                                                    434
```

-continued

```
SEQ ID NO: 119          moltype = AA  length = 169
FEATURE                 Location/Qualifiers
source                  1..169
                        mol_type = protein
                        organism = Sus scrofa
SEQUENCE: 119
MAPKKAKRRA AAEGSSNVFS MFDQTQIQEF KEAFTVIDQN RDGIIDKEDL RDTFAAMGRL  60
NVKNEELDAM MKEASGPINF TVFLTMFGEK LKGADPEDVI TGAFKVLDPE GKGTIKKHFL  120
EELLTTQCDR FSQEEIKNMW AAFPPDVGGN VDYKNICYVI THGDAKDQE              169

SEQ ID NO: 120          moltype = AA  length = 1939
FEATURE                 Location/Qualifiers
source                  1..1939
                        mol_type = protein
                        organism = Sus scrofa
SEQUENCE: 120
MSSDQEMAIF GEAAPYLRKS EKERIEAQNK PFDAKTSVFV AEPKESFVKG TVQSREGGKV  60
TVKTEAGATL TVKEDQVFPM NPPKFDKIED MAMMTHLHEP AVLYNLKERY AAWMIYTYSG  120
LFCVTVNPYK WLPVYNAEVV TAYRGKKRQE APPHIFSISD NAYQFMLTDR ENQSILITGE  180
SGAGKTVNTK RVIQYFATIA VTGEKKKEEP TSGKMQGTLE DQIISANPLL EAFGNAKTVR  240
NDNSSRFGKF IRIHFGTTGK LASADIETYL LEKSRVTFQL KAERSYHIFY QIMSNKKPEL  300
IEMLLITTNP YDYAFVSQGE ITVPSIDDQE ELMATDSAIE ILGFTSDERV SIYKLTGAVM  360
HYGNLKFKQK QREEQAEPDG TEVADKAAYL QGLNSADLLK ALCYPRVKVG NEFVTKGQTV  420
QQVYNAVGAL AKAVYDKMFL WMVTRINQQL DTKQPRQYFI GVLDIAGFEI FDFNSLEQLC  480
INFTNEKLQQ FFNHHMFVLE QEEYKKEGIE WEFIDFGMDL AACIELIEKP MGIFSILEEE  540
CMFPKATDTS FKNKLYEQHL GKSNNFQKPK PAKGKVEAHF SLIHYAGTVD YNITGWLDKN  600
KDPLNETVVG LYQKSSVKTL AFLFTGAAGA DAEAGGGKKG GKKKGSSFQT VSALFRENLN  660
KLMTNLRSTH PHFVRCIIPN ETKTPGAMEH ELVLHQLRCN GVLEGIRICR KGFPSRILYA  720
DFKQRYKVLN ASAIPEGQFI DSKKASEKLL GSIDIDHTQY KFGHTKVFFK AGLLGLLEEM  780
RDEKLAQLIT RTQARCRGFL ARVEYQKMVE RRESIFCIQY NIRAFMNVKH WPWMKLYFKI  840
KPLLKSAETE KEMANMKEEF EKTKESLAKA EAKRKELEEK MVALMQEKND LQLQVQAEAD  900
SLADAEERCD QLIKTKIQLE AKIKEVTERA EDEEEINAEL TAKKRKLEDE CSELKKDIDD  960
LELTLAKVEK EKHATENKVK NLTEEMAGLD ETIAKLTKEK KALQEAHQQT LDDLQAEEDK  1020
VNTLTKAKTK LEQQVDDLEG SLEQEKKLRM DLERAKRKLE GDLKLAQEST MDIENDKQQL  1080
DEKLKKKEFE MSNLQSKIED EQALAMQLQK KIKELQARIE ELEEEIEAER ASRAKAEKQR  1140
SDLSRELEEI SERLEEAGGA TSAQIEMNKK REAEFQKMRR DLEEATLQHE ATAATLRKKH  1200
ADSVAELGEQ IDNLQRVKQK LEKEKSEMKM EIDDLASNME TVSKAKGNLE KMCRTLEDQL  1260
SELKTKEEEQ QRLINDLTAQ RARLQTESGE YSRQLDEKDT LVSQLSRGKQ AFTQQIEELK  1320
RQLEEEIKAK SALAHAVQSS RHDCDLLREQ YEEEQEAKAE LQRAMSKANS EVAQWRTKYE  1380
TDAIQRTEEL EEAKKKLAQR LQDAEEHVEA VNAKCASLEK TKQRLQNEVE DLMIDVERSN  1440
AACAALDKKQ RNFDKILAEW KQKYEETHAE LEASQKESRS LSTELFKVKN AYEESLDQLE  1500
TLKRENKNLQ QEISDLTEQI AEGGKRIHEL EKIKKQVEQE KSEIQAALEE AEASLEHEEG  1560
KILRIQLELN QVKSEVDRKI AEKDEEIDQL KRNHVRVVES MQSMLDAEIR SRNDAIRLKK  1620
KMEGDLNEME IQLNHANRMA AEALRNYRNT QGILKDTQIH LDDALRSQED LKEQLAMVER  1680
RANLLQAEIE ELRATLEQTE RSRKVAEQEL LDASERVQLL HTQNTSLINT KKKLETDISQ  1740
IQGEMEDIIQ EARNAEEKAK KAITDAAMMA EELKKEQDTS AHLERMKKNL EQTVKDLQHR  1800
LDEAEQLALK GGKKQIQKLE ARVRELEGEV ESEQKRNVET VKGLRKHERR VKELTYQTEE  1860
DRKNILRLQD LVDKLQAKVK SYKRQAEEAE EQSNVNLSKF RKLQHELEEA EERADIAESQ  1920
VNKLRVKSRE VHTKIISEE                                               1939

SEQ ID NO: 121          moltype = AA  length = 6665
FEATURE                 Location/Qualifiers
source                  1..6665
                        mol_type = protein
                        organism = Sus scrofa
SEQUENCE: 121
MADDEEYEEV VEYYTEETVY EEVPGETITE VYETTTTRTS GYEQSDTSQP AVAQPMPSKP  60
VERKKVIRKK VDPSKFMTPY IAHSQKMQDL FSTNKYKKNY EKAKGKPYAI TTDTPELRRI  120
KKVQDQLSEV KYRIDGDVAK TICHVDEKAK DIEHAKKVSQ QVSKVLYKQN WEDTKDKYLL  180
PPDAPELVQA IKNTAMFSKK LYTEDWEADK SLFYPYNDSP ELRRVAQAQK ALSDIAYKKG  240
HAEQQTQFTS LPDPPDIEFA KKVTNQVSKQ KYKEDYENKI KGKWSETPCF EIATARMNAK  300
NISTRKYQED FENMKDQIYF MQTETPEYKV NKQAGTAASK VKYKEDYEKN KGKADYNVLP  360
ASENPLLRHL KAAGDALSDK RYKENYEKTK AKSINYCETP KFKLDTVLHN FSSDTKYKDS  420
YLKNILGHYV GSFEDPYHTH CMKVAAENSD KNYKAEYEED RGKGFFPQTI TQEYEAIKKL  480
DQCKDHTYKV HPDKTKFTQV TDSPVLVQAQ VNSKQLSDLN YKAKHENEKF KCHIPPDTPA  540
FIQHKVNAYN LSDNVYKHDW EKTKAKKFDI KVDAIPLLAA KANTKNTSDV MYKKDYEKSK  600
GKMIGALSIN DDPKMLHSLK TAKNQSDRLY KENYEKTKAK SMNYCETPKY QLDTLLKNFS  660
EAKYKDSYVQ NVLGHYIGSF EDPYQLHCMK VSAQNSDKNY KAEYEEDKGK CYFPQTITQE  720
YEAIKKLDQC KDHTYKVHPD KTKFTAVTDT PVLLQAQLNT KQLSDLNYKA KHEGEKFKCN  780
IPADAPQFIQ HRVNAYNLSD NVYKHDWEKT KAKKFDIKVD AIPLLAAKAN TKNTSDVMYK  840
KDYEKSKGKM IGALSINDDP KMLHSLKTAK NQSDREYRKD YEKSKTIYTA PLDMLQVTQA  900
KKSQAIASDV DYKHLLHNYS YPPDSINVDL AKKAYALQSD VEYKADYNSW MKGCGWIPFG  960
SLEMEKAKRA SDILNEKKYR QHPETLKFTS IEDAPIIVQS KINQAQRSDI AYKAKGEEII  1020
HKYSLPADLP FIQAKVNAYN ISESMYKADL KDLSKKGYDL RTDAIPIKAA KAARQAASDV  1080
QYKKDYEKAK GKMVGFQSLQ DDPKLVHYMN VAKIQSDREY KKDYEKTKTK YNTPHDMFNV  1140
```

-continued

```
VAAKKAQDVA SNVNYRHSLH HYTYLPDAMD LELSKNMMHI QSDNAYKEDY NTWMKGIGWI  1200
PLGSLEVEKV KKAGDALNEK KYRQHPDTLK FTSIVDSPVM VQAKQNTRQV SDILYKAKGE  1260
DVKHKYTMGP DLPQFLQAKC NAYNISDVCY KRDWHDLIAK GNNVLGDAIP ITAAKASRNI  1320
ASDYKYKEAY EKAKGKHVGF RSLQDDPKLV HYMNVAKLQS DREYKKNYEN TKTSYHTPGD  1380
MVSITAAKMA QDVATNVNYK QPIHHYTYLP DALSLEHTRN VNQIQSDNVY KDEYNHFFKG  1440
IGWIPLGSLE VEKVKKAGDA LNERKYRQHP DTIKFTSVPD SMGMVLAQHN TKQLSDLNYK  1500
VEGEKLKHKY TMDPDVPQFI QAKVNAINMS DAHYKADWKK TLAKGYDLTP DAIPIVAAKS  1560
SRNIASDFKY KEAYEKAKGK QVGFRSLQDD PKLVHYMNVA KIQSDREYKK GYEASKTRYH  1620
TPLDMLSVTA AKKSQEVATN TNYKQPFHHY TLLPDALNVE HSRNAMQIQS DNLYKSDFTN  1680
WMKGIGWIPM ESLEVEKAKK AGEILSEKKY RQHPEKLKFT YAMDTMEQAL NKSNKLIMDK  1740
RLYTEKWNKD KTTIHVMPDT PDILLSRVNQ ITLSDKLYKA GWEEEKKKGY DLRPDAISIK  1800
AARASRDIAS DYKYKQAYEQ AKGKHIGFRS LEDDPKLVHF MQVAKMQSDR EYKKAYEKSK  1860
TSFHTPVDML SVVAAKKSQE VATNANYRNV IHTYNMLPDA MSFELAKNMM QLQSDNQYKA  1920
DYADFMKGIG WLPLGSLEAE KNKKAMEIIS EKKYRQHPDT LKYSTLMDSM NMVLAKNNAK  1980
IMNEHLYKQA WEADKTKVHI MPDIPQIILA KANAVNMSDK LYKLSLEESK KKGYDLRTDA  2040
ISIQAAKASR DIASDYKYKH SYEKERGKMV GFRSLEDDPK LVHSMQVAKM QSDREYKKSY  2100
EKTKTSYHTP ADMLSVTAAR DAQANITNTN YKHLIHKYIL LPDAMNIQLS KNMNHIQSDN  2160
EYKQDYNECY KGLGWSPAGS LEVEKAKKAT EYASDQKYRQ HPSNFQFQKL NDSMDMVLAK  2220
QNAHTMNKYL YTADWNKDKT KIHVMPDTPD ILQAKQNQTL FSQKLYKLGW EEALKKGYDL  2280
PVDAISVQLA KVSRDNASDF KYKQGYRKQL GHHIGFRNLQ DDPKLVLSMN VAKMQSEREY  2340
KKDFEKWKTK FTSPVDMLGV VLAKKCQALV SDVDYKNYLH QWTCLPDQND VIQAKRSYEL  2400
QSENLYKSDL EWLRGIGWSP LGSLEAEKNK RASEIISEKK YRQPPDRNKF TSIPDAMDIV  2460
LAKTNAKNRS DRLYREAWDK DKTQVHIMPD TPDIILAKAN LINSSDKLYR MGYEELKKKG  2520
YDLPVDAIPI KAAKASREIA SEYKYKEGFR KQLGHHVGAR NIEDDPKMMW SMHVAKIQSD  2580
REYKKDFEKW KTKFSSPVDM LGVVLAKKCQ TLVSDIDYKN YLHQWTCLPD QSDVIHARRA  2640
YDLQSDNLYK SDLQWLKGIG WLPSGSLEDE KNKRATQILS DHVYRQHPDK FKFSSLMDSI  2700
PMVLAKNNAI TMNHRLYTEA WDKDKTTVHI MPDTPEVLLA KQNKINYSEK LYKLGLEEAK  2760
RKGYDMRLDA IPIRAAKASR DIASEFKYKE GYRKQLGHHI GARAIHDDPK MMWSMHVAKV  2820
QSDREYKKDF EKWKTKFSSP VDMLGVVLAK KCQTLVSDID YKNYLHQWTC LPDQSDVIHA  2880
RRAYDLQSDN MYKSDLQWLR GIGWVPIGSL DVEKCKRATE ILSDKIYRQP PDTFKFTSVP  2940
DSLEQVLAKN NAINMNKRLY TEAWDKDKTQ VHIMPDTPEI MLARQNRLNY SETLYKLANE  3000
EAKKKGYDLR SDAIPIVAAK ASRDIVSDYK YKDGYRKQLG HHIGARNIED DPKMMWSMHV  3060
AKIQSDREYK KDFEKWKTKF SSPVDMLGLV QAKKCQILVS DVDYKNYLHQ WTCLPDQNDV  3120
IHARQAYDLQ SDNIYKSDLQ WLKGIGWVPI GSVDVVKCKR AAEILSDHIY RQPPDKLKFT  3180
SVPDSLEQVL AKNNAINMNK RLYTEAWDKD KTQIHIMPDT PEIMLARQNK INYSESLYRQ  3240
AMEEAKKEGY DLRSDAIPIV AAKASRDIAS DYKYKEAYRK QLGHHIGARA VHDDPKMMWS  3300
LHVAKMQSDR EYKKDFEKYK TRFSSPVDML GIVLAKKCQT LVSDVDYKHP LHEWTCLPDQ  3360
NDIIHARKAY DLQSDNLYKS DLEWMKGIGW IPLGSVEVVK AKRAGEILSD NIYRQRPDTL  3420
KFTSVTDSPE LVLAKNNALN MNKRLYTEAW DNDKKTIHVM PDTPEIMLAK LNRINYSDKL  3480
YRLALEESKK EGCDLRVDAI PIQAAKASRE IASDYKYKEA YRKQLGHHIG ARNIEDDPKM  3540
MWSIHAAKIQ SDREYKKDFE KWKTRFSSPV DMLGLVQAKK CQILVSDIDY KHPLHEWICL  3600
PDQNDIIQAR KAYDLQSDAI YKADLEWLRG IGWVPLGSVE VEKVKRAGEI LSERKYRQPA  3660
DQLKFTCITD TPEIVLAKNN ALTMSKHLYT EAWDTDKTSI HVMPDTPEIL LAKSNSANIS  3720
QKLYTKGWDE SKMKDYDLKA DAISIKSAKA SRDIASDYKY KEAYEKQKGH HIGAQSVEDD  3780
PKIMCAIHAG KIQSEREYKK DFQKWKTKFT SPVDMLGILL AKKCQTLVSD IDYRNYLHHW  3840
TCLPDQNDII QARKAYELQS DALYKADLEW LRGIGWMPQG SPEVLRVKNA QEILRDSVYR  3900
TPVVKLKYTS IVDTPEVVLA KSNAENISIP KYREVWDKDK TSVHIMPDTP EINLARANAL  3960
NVSNKIYREG WDEMKMSCDV RLDAIPIQAA KASREIASDY KYKLDHEKQK GHYVGTLTAR  4020
DDNKIRWALI AGKIQNEREY RLHWAKWKSK FQSPADMLSI THSKNCQALV SDIDYRNYLH  4080
QWTCLPDQHD VIQARKAYDL QSDAIYKADL EWLRGIGWMP SDSVSVNHAK HAADIFSEKK  4140
YRTKIETLNF TPVDDRVDYV TAKQSGEILN DIKYRKDWND TKSKYTLTET PQLHAAQEAA  4200
RILDQYLYKE SWEKQKATGY ILPPDAVPFV HAHHSSDVQS ELKYKADHVK QKGHYVGVPT  4260
MRDDPKLVWF EHAGQIQNDR LYKEHYHKTK AKINIPPDMV SVVAAKEGQI LVSDIDYRNY  4320
LHQWTCHPDQ NDVIQARKAY DLQSDNVYKA DLEWLRGIGW IPLDSVDHVR VTRNQEMVNQ  4380
IKYKKDALAN YPNFTSVVDP PAVVLAKINS VNQSDVKYKE TFNKLLKGKY IFSPDTPYIT  4440
HSKDMGKLYS TILYKGAWEG TKAYGYTLDE RYIPIVGAKH ADFVNSELKY KETYEKQKGH  4500
YLAGKEISEF PNVVHCLDFQ KMRSVLNYRR HYEDTKANVH IPNDMMNHVL AKRCQYILSD  4560
LEYRHYFHQW TSLPEEPNVI RARNAQEILS DNVYKDDLNW LKGIGCYVWD TPQILHAKKS  4620
YDLQSQILYT ADGKENLKNY NMVTDTPLYV TALQSGINAS EVKYKENYHQ IKDKYTTVLE  4680
TVDYDRTRNL KDLYSSNVYK EAWDKVKATS YILPSNTVSL THAKNQKHLA SHIKYREEYE  4740
KFKALYTLPK SVDDDPNTAR CLRVGKFNID RLYRSVYEKN KMKIHMVPDM VEMVTAKDSQ  4800
KKVSEIDYRL HLHEWICHPD LQVNSHVRKV TDQISDIVYK DDLNWLKGIG CYVWDTPEIL  4860
HAKHAYDLRN DIKYQAHVKK TRNNYKLVTD TPLYVQAVKS GKQLSDAVYH YDYVHSVRGK  4920
VAPTTKTVDL DRALHAYKLQ SENLYRKDGL NSLPTGYRLP GDTPHFKHNK DIRYMSSYFK  4980
YKEVYEHPKA NGYTLGPKDV PFVHVRRVNN VTSERLYRQL YHKLKDKIHT TPDTPEIRQV  5040
KKTQEAVSEL IYKSDFFKMQ GHMISMPYTP HVLHCRYVGD ITSDIKYKED LQILRGLGCF  5100
LYDTPDMVRS RHLRKLWSNY LYTDNARKMR DKYKVVLDTP EYRKVQELKT HLSELVYRAA  5160
GKKQKSIFTS VPDTPDLLRA KRGQKLQSQY LYVELATKER PHHHAGNQTT ALKHARDVKD  5220
MVSEKKYKVQ YEKMKDKYTP VPDTPILIRA KKAYWNASDL RYKETFQKTK GKYHTVKDAL  5280
DIVYHRKVTD HISKIKYKEN YMSQLGIWRS IPDRPEHFHH RAVTDAVSDV KYKEDLTWLK  5340
GIGCYAYDTP DFTLAEKNKT LYSKYKYKEV FERTKSDFKY EADCPINRHF KFATQLMNEK  5400
KYRADYEQRK DKYQLVVDEP RHLLAKIAGD QISQIKYRKN YEKTKDKFTS VVDTPEHLRT  5460
TKVNKQISDI LYKLEYNKAK PRGYTTIHDT PMLLHVRKVK DEVSDLKYKE VYQRTKSNCT  5520
IEPDAVHIKA AKDAYKVNTN LDYKKQYEAN KAHWRWTPDR PDFIQAAKSS LQQSDFEYKL  5580
DREFLKGCKL SVTDDKDMVL ALRNSVIESD LKYKEKHVKE RGTCHAVPDT PQILLAKTVS  5640
NLVSENKYKR HVKKHLAQGS YTTLPETRDT VHVKEVTKHV SDTNYKKKFV KERGKSNYSI  5700
```

-continued

```
MLEPPEVKHA MEVAKKQSDV AYKKDAKENL HYTTVADRPD IKKATQAAKQ ASEVEYRAKH  5760
RKEGSHGLSM LGRPDIEMAK KAATLSSQVK YRENFNKEKG KTPKYNPKES QLYKIMKDAN  5820
NLASEVKYKA DLKKLHKPVT DMKESLIMNH VLNTSHLASS YQYKKNYEKS KGHYHTIPDN  5880
LEQLHLKEAT ELQSIVKYKE KYEKERGKPM LDFETPTYIT AKESQQMQSG KEYRKDYEES  5940
IKGRNLTGLE VTPALLHVKY ATKIASEKEY RKDLEESIRG KGLSEMEDTP DMLRAKNATQ  6000
ILNEKEYKRD LELEVKGRGL NAMANETPDF MRARNATDIA SQIKYKQLAE MEKANFTSVV  6060
DTPEIIHAQQ VKNLSSQKKY KEDAEKCMSY YETVLDTPEM QRVRENQKNF SLLQYQCDLK  6120
NSKGKITVVQ DTPEILRVKE NQKNFSSVLY KEDVSPGTAI GKTPEMMRVK QTQDHISSVK  6180
YKETIGQGTP IPDLPEVKRV KETQKHISSV MYKENLGTGI PTPVTPEIER VKRNQENFSS  6240
VLYKENLGKG TPTPITPEME RVKRNQENFS SVLYKENMGK GTPLPVTPEM ERVKHNQENI  6300
SSILYKENVG KGTPTPVTPE MQRVKRNQEN ISSVIYKENL GKATPTPFTP EMERVKRNQE  6360
NFSSVLYKEN MRKATPTPVT PEMERAKRNQ ENISSVLYSD SFRKQIQGKA AYVLDTPEMR  6420
RVRETQRHIS TVKYHEDFEK HKGCFTPVVT DPITERVKKN TQDFSDINYR GIQRKVVEME  6480
QKRNDQDQET ITGLRVWRTN PGSVFDYDPA EDNIQSRSLH MLNAQAQRRS REQSRSASAL  6540
SISGGEEKSE HSEAARLSTY SDGGVFFSAT SAAYKHAKTT ELPQQRSSSV ATQQTTVSSI  6600
PSHPSTAGKI FRAMYDYMAA DADEVSFRDG DAIVNVQAID EGWMYGTVQR TGRTGMLPAN  6660
YVEAI                                                              6665

SEQ ID NO: 122          moltype = AA  length = 1937
FEATURE                 Location/Qualifiers
source                  1..1937
                        mol_type = protein
                        organism = Sus scrofa
SEQUENCE: 122
MSSDQEMAIF GEAAPYLRKS EKERIEAQNK PFDAKTSVFV AEPKESFVKG TVQSREGGKV  60
TVKTEAGATL TVKEDQVFPM NPPKFDKIED MAMMTHLHEP AVLYNLKERY AAWMIYTYSG  120
LFCVTVNPYK WLPVYNAEVV TAYRGKKRQE APPHIFSISD NAYQFMLTDR ENQSILITGE  180
SGAGKTVNTK RVIQYFATIA VTGEKKKEEP TPGKMQGTLE DQIISANPLL EAFGNAKTVR  240
NDNSSRFGKF IRIHFGTTGK LASADIETYL LEKSRVTFQL KAERSYHIFY QIMSNKKPEL  300
IEMLLITTNP YDYAFVSQGE ITVPSIDDQE ELMATDSAIE ILGFTSDERV SIYKLTGAVM  360
HYGNLKFKQK QREEQAEPDG TEVADKAAYL QGLNSADLLK ALCYPRVKVG NEFVTKGQTV  420
QQVYNAVGAL AKAVYDKMFL WMVTRINQQL DTKQPRQYFI GVLDIAGFEI FDFNSLEQLC  480
INFTNEKLQQ FFNHHMFVLE QEEYKKEGIE WEFIDFGMDL AACIELIEKP MGIFSILEEE  540
CMFPKATDTS FKNKLYEQHL GKSNNFQKPK PAKGKAEAHF SLIHYAGTVD YNITGWLDKN  600
KDPLNETVVG LYQKSSVKTL AFLFAERQSS EEGGTKKGGK KKGSSFQTVS ALFRENLNKL  660
MTNLRSTHPH FVRCIIPNET KTPGAMEHEL VLHQLRCNGV LEGIRICRKG FPSRILYADF  720
KQRYKVLNAS AIPEGQFIDS KKASEKLLGS IDIDHTQYKF GHTKVFFKAG LLGTLEEMRD  780
EKLAQLITRT QAMCRGFLMR VEFRKMMERR ESIFCIQYNI RAFMNVKHWP WMKLYFKIKP  840
LLKSAETEKE MANMKEEFEK TKEDLAKSEA KRKELEEKMV ALMQEKNDLQ LQVQAEADGL  900
ADAEERCDQL IKTKIQLEAK IKEVTERAED EEEINAELTA KKRKLEDECS ELKKDIDDLE  960
LTLAKVEKEK HATENKVKNL TEEMAGLDEN IAKLTKEKKA LQEAHQQTLD DLQAEEDKVN  1020
TLTKAKTKLE QQVDDLEGSL EQEKKLRMDL ERAKRKLEGD LKLAQESTMD IENDKQQLDE  1080
KLKKKEFEMS NLQSKIEDEQ ALAMQLQKKI KELQARTEEL EEEIEAERAS RAKAEKQRSD  1140
LSRELEEISE RLEEAGGATS AQIEMNKKRE AEFQKMRRDL EEATLQHEAT AAALRKKHAD  1200
SVAELGEQID NLQRVKQKLE KEKSELKMEI DDLASNMETV SKAKGNLEKM CRTLEDQLSE  1260
VKTKEEEHQR LINELSAQKA RLQTESGEFS RQLDEKEALV SQLSRGKQAF TQQIEELKRQ  1320
LEEETKAKSA LAHAVQSSRH DCDLLREQYE EEQEAKAELQ RAMSKANSEV AQWRTKYETD  1380
AIQRTEELEE AKKKLAQRLQ DAEEHVEAVN AKCASLEKTK QRLQNEVEDL MLDVERSNAA  1440
CAALDKKQRN FDKILAEWKH KYEETQAELE ASQKESRSLS TELFKVKNAY EESLDQLETL  1500
KRENKNLQQE ISDLTEQIAE GGKHIHELEK VKKQIEQEKS ELQAALEEAE ASLEHEEGKI  1560
LRIQLELNQV KSEIDRKIAE KDEEIDQMKR NHIRVVESMQ STLDAEIRSR NDALRIKKKM  1620
EGDLNEMEIQ LNHANRQATE AIRNLRNTQG VLKDTQLHLD DAIRGQDDLK EQLAMVERRA  1680
NLMQAEIEEL RASLEQTERS RRVAEQELLD ASERVQLLHT QNTSLINTKK KLETDISQIQ  1740
GEMEDIVQEA RNAEEKAKKA ITDAAMMAEE LKKEQDTSAH LERMKKNMEQ TVKDLQHRLD  1800
EAEQLALKGG KKQIQKLEAR VRELENEVEN EQKRNVEAVK GLRKHERRVK ELTYQTEEDR  1860
KNVLRLQDLV DKLQSKVKAY KRQAEEAEEQ SNVNLSKFRK LQHELEEAEE RADIAESQVN  1920
KLRVKSREVH TKVISEE                                                 1937

SEQ ID NO: 123          moltype = AA  length = 607
FEATURE                 Location/Qualifiers
source                  1..607
                        mol_type = protein
                        organism = Sus scrofa
SEQUENCE: 123
MKWVTFISLL FLFSSAYSRG VFRRDTYKSE IAHRFKDLGE QYFKGLVLIA FSQHLQQCPY  60
EEHVKLVREV TEFAKTCVAD ESAENCDKSI HTLFGDKLCA IPSLREHYGD LADCCEKEEP  120
ERNECFLQHK NDNPDIPKLK PDPVALCADF QEDEQKFWGK YLYEIARRHP YFYAPELLYY  180
AIIYKDVFSE CCQAADKAAC LLPKIEHLRE KVLTSAAKQR LKCASIQKFG ERAFKAWSLA  240
RLSQRFPKAD FTEISKIVTD LAKVHKECCH GDLLECADDR ADLAKYICEN QDTISTKLKE  300
CCDKPLLEKS HCIAEAKRDE LPADLNPLEH DFVEDKEVCK NYKEAKHVFL GTFLYEYSRR  360
HPDYSVSLLL RIAKIYEATL EDCCAKEDPP ACYATVFDKF QPLVDEPKNL IKQNCELFEK  420
LGEYGFQNAL IVRYTKKVPQ VSTPTLVEVA RKLGLVGSRC CKRPEEERLS CAEDYLSLVL  480
NRLCVLHEKT PVSEKVTKCC TESLVNRRPC FSALTPDETY KPKEFVEGTF TFHADLCTLP  540
EDEKQIKKQT ALVELLKHKP HATEEQLRTV LGNFAAFVQK CCAAPDHEAC FAVEGPKFVI  600
EIRGILA                                                            607
```

-continued

```
SEQ ID NO: 124          moltype = AA  length = 381
FEATURE                 Location/Qualifiers
source                  1..381
                        mol_type = protein
                        organism = Bison bison
SEQUENCE: 124
MPFGNTHNKH KLNFKAEEEY PDLSKHNNHM AKALTLEIYK KLRDKETPSG FTLDDVIQTG  60
VDNPGHPFIM TVGCVAGDEE SYTVFKDLFD PIIQDRHGGF KPTDKHKTDL NHENLKGGDD  120
LDPNYVLSSR VRTGRSIKGY ALPPHCSRGE RRAVEKLSVE ALNSLTGEFK GKYYPLKSMT  180
EQEQQQLIDD HFLFDKPVSP LLLASGMARD WPDARGIWHN DNKSFLVWVN EEDHLRVISM  240
EKGGNMKEVF RRFCVGLQKI EEIFKKAGHP FMWNEHLGYV LTCPSNLGTG LRGGVHVKLA  300
HLSKHPKFEE ILTRLRLQKR GTGGVDTAAV GSVFDVSNAD RLGSSEVEQV QLVVDGVKLM  360
VEMEKKLEKG QSIDDMIPAQ K                                            381

SEQ ID NO: 125          moltype = AA  length = 894
FEATURE                 Location/Qualifiers
source                  1..894
                        mol_type = protein
                        organism = Bison bison
SEQUENCE: 125
MNQIEPGVQY NYVYEDDEYM IQEEEWDRDL LLDPAWEKQQ RKTFTAWCNS HLRKAGTQIE  60
NIEEDFRNGL KLMLLLEVIS GERLPKPDRG KMRFHKIANV NKALDYIASK GVKLVSIGAE  120
EIVDGNVKMT LGMIWTIILR FAIQDISVEE TSAKEGLLLW CQRKTAPYRN VNIQNFHTSW  180
KDGLGLCALI HRHRPDLIDY SKLNKDDPIG NINLAMEIAE KHLDIPKMLD AEDIVNTPKP  240
DERAIMTYVS CFYHAFAGAE QAETAANRIC KVLAVNQENE RLMEEYERLA SELLEWIRRT  300
IPWLENRTPE KTMQAMQKKL EDFRDYRRKH KPPKVQEKCQ LEINFNTLQT KLRISNRPAF  360
MPSEGKMVSD IAGAWQRLEQ AEKGYEEWLL NEIRRLERVE HLAEKFRQKA STHETWAYGK  420
EQILLQKDYE SSTLTEVRAL LRKHEAFESD LAAHQDRVEQ IAAIAQELNE LDYHDAVNVN  480
DRCQKICDQW DRLGTLTQKR REALERTEKL LETIDQLHLE FAKRAAPFNN WMEGAMEDLQ  540
DMFIVHSIEE IQSLITAHEQ FKATLPEADG ERQSILAIQN EVEKVIQSYS IRISSSNPYS  600
TVTVDEIRSK WDKVKQLVPI RDQSLQEELA RQHANERLRR QFAAQANAIG PWIQNKMEEI  660
ARSSIQITGA LEDQMNQLKQ YEHNIINYKN NIDKLEGDHQ LIQEALVFDN KHTNYTMEHI  720
RVGWELLLTT IARTINEVET QILTRDAKGI TQEQMNEFRA SFNHFDRRKN GLMDHEDFRA  780
CLISMGYDLG EAEFARIMTL VDPNGQGTVT FQSFIDFMTR ETADTDTAEQ VIASFRILAS  840
DKPYILAEEL RRELPPDQAQ YCIKRMPAYS GPGSVPGALD YTAFSSALYG ESDL        894

SEQ ID NO: 126          moltype = AA  length = 842
FEATURE                 Location/Qualifiers
source                  1..842
                        mol_type = protein
                        organism = Bison bison
SEQUENCE: 126
MSRPLTDQEK RKQISVRGLA GVENVTELKK NFNRHLHFTL VKDRNVATPR DYYFALAYTV  60
RDHLVGRWIR TQQHYYEKDP KRIYYLSLEF YIGRTLQNTM VNLALENACD EATYQLGLDM  120
EELEEIEEDA GLGNGGLGRL AACFLDSMAT LGLAAYGYGI RYEFGIFNQK ISGGWQMEEA  180
DDWLRYGNPW EKARPEFTLP VHFYGRVEHT SQGAKWVDTQ VVLAMPYDTP VPGYRNNVVN  240
TMRLWSAKAP NDFNLKDFNV GGYIQAVLDR NLAENISRVL YPNDNFFEGK ELRLKQEYFV  300
VAATLQDIIR RFKSSKFGCL DPVRTNFDAF PDKVAIQLND THPSLAIPEL MRILVDQERL  360
EWEKAWEVTV KTCAYTNHTV LPEALERWPV HLIETLLPRH LQIIYEINQR FLNRVAAAFP  420
GDVDRLRRMS LVEEGAVKRI NMAHLCIAGS HAVNGVARIH SEILKKTIFK DFYELEPHKF  480
QNKTNGITPR RWLVMCNPGL AEIIAERIGE EYIADLDQLR KLLSYVDDES FIRDVAKVKQ  540
ENKLKFSAYL EKEYKVHINP NSLFDIQVKR IHEYKRQLLN CLHVITLYNR IKKEPNKFFV  600
PRTVMIGGKA APGYHMAKMI IKLITAIGDV VNHDPVVGDR LRVIFLENYR VSLAEKVIPA  660
ADLSEQISTA GTEASGTGNM KFMLNGALTI GTMDGANVEM AEEAGEENFF IFGMRVEDVE  720
RLDQKGYNAQ EYYDRIPELR HIIDQLSSGF FSPKQPDLFK DIVNMLMHHD RFKVFADYEE  780
YVKCQERVSA LYKNPREWTR MVIRNIATSG KFSSDRTIAQ YAREIWGVEP TRQRMPAPDE  840
KI                                                                 842

SEQ ID NO: 127          moltype = AA  length = 364
FEATURE                 Location/Qualifiers
source                  1..364
                        mol_type = protein
                        organism = Bison bison
SEQUENCE: 127
MPHQYPALTP EQKKELCDIA HRIVAPGKGI LAADESTGSI AKRLQSIGTE NTEENRRFYR  60
QLLLTADDRV NPCIGGVILF HETLYQKADD GRPFPQVIKA KGGVVGIKVD KGVVPLAGTN  120
GETTTQGLDG LSERCAQYKK DGADFAKWRC VLKIGEHTPS SLAIMENANV LARYASICQQ  180
NGIVPIVEPE ILPDGDHDLK RCQYVTEKVL AAVYKALSDH HIYLEGTLLK PNMVTPGHAC  240
TQKYSHEEIA MATVTALRRT VPPAVPGITF LSGGQSEEEA SINLNAINKC PLLKPWALTF  300
SYGRALQASA LKAWGGKKEN LKAAQEEYVK RALANSLACQ GKYTPSGKAG AAASESLFIS  360
NHAY                                                               364

SEQ ID NO: 128          moltype = AA  length = 192
FEATURE                 Location/Qualifiers
source                  1..192
                        mol_type = protein
```

-continued

```
                        organism = Bison bison
SEQUENCE: 128
MAPKKDVKNP AAAAAPAPAP APAPAPAPAP PKEEKIDLSA IKIEFSKQQQ DEFKEAFLLF  60
DRTGECKITL SQVGDVLRAL GTNPTNAEVK KVLGNPSNEE MNAKKIEFEQ FLPMLQAISN  120
NKDQGTYEDF VEGLRVFDKE GNGTVMGAEL RHVLATLGEK MKEEEVEALM AGQEDSNGCI  180
NYEAFVKHIM SN                                                      192

SEQ ID NO: 129          moltype = AA  length = 333
FEATURE                 Location/Qualifiers
source                  1..333
                        mol_type = protein
                        organism = Bison bison
VARIANT                 158
                        note = Any amino acid
VARIANT                 290
                        note = Any amino acid
SEQUENCE: 129
MVKVGVNGFG HIGRLVTRAA FNSGKVDVIA INDPFIDLHY IVYMFQYDTT HGKFHGTVKA  60
ENGKLIINGK AVTIFQERDP ASIKWGDAGA EYVVESTGVF TTMEKAGAHL KGGAKSVIIS  120
APSADAPMFV LGMNHEKYNN NLKIVSNASC TTNCLAPXAK VIHDHFGIVE GLMTTFHAIT  180
VTQKTVDGPS GKLWHDGGGA AQNIIPASTG AAKAVGKVIP ELNRKLTGMA FCVPTPNVSV  240
MDLTYHLEKP AKYDEIKKVV KQTSEGPLEG ILGYTQDQVV SCDFNSDSYX STFDAGAGIA  300
LNDHFVKLIS WYDNEFGYSN RVVDLMVHMA SKE                               333

SEQ ID NO: 130          moltype = AA  length = 284
FEATURE                 Location/Qualifiers
source                  1..284
                        mol_type = protein
                        organism = Bison bison
VARIANT                 30
                        note = Any amino acid
SEQUENCE: 130
MDAIKKKMQM LKLDKENALD RAEQAEADKX PVPHGHSQLE DELVSLQKKL KATEDELDKY  60
SEALKDAQEK LELAEKKATD AEADVASLNR RIQLVEEELD RAQERLATAL QKLEEAEKAA  120
DESERGMKVI ESRAQKDEEK MEIQEIQLKE AKHIAEDADR KYEEVARKLV IIESDLERAE  180
ERAELSEGKC AELEEELKTV TNNLKSLEAQ AEKYSQKEDK YEEEIKVLSD KLKEAETRAE  240
FAERSVTKLE KSIDDLEDEL YAQKLKYKAI SEELDHALND MTSM                   284

SEQ ID NO: 131          moltype = AA  length = 1183
FEATURE                 Location/Qualifiers
source                  1..1183
                        mol_type = protein
                        organism = Bison bison
SEQUENCE: 131
MPEAKPAAKK APAGKDAAAK PAPKEATPQE APAAPPTGEI DGAPAGRAEP GGSSAASRVV  60
PPRPLLRSVF PTAPGLTALP YSQPCRAPQR FQGGGLTVPL PPHSPLLPGK DTVIVAKLNG  120
KELPAKPAVK WFKGKWLELG SKSGARFSFK ESHDAASNVY TIELHITKVV LGDRGDYRIE  180
VKAKDFCDSC AFNIDVEAPR QDSAGQSLES FKRSGEAKSD TAGELDFSGL LKKRQVVEEE  240
KKKKKDDDDL GIPPEIWELL KGAKKSEYER IAFQYGITDL RGMLKRLKKA KVEVKKSAAF  300
TKKLDPAYQV DRGNKIKLVV EISDPDLPLK WFKNGQEIKP SSKYVFENVG KKRILTINKC  360
TLADDAAYEV VVKDEKCFTE LFVKEPPVLI VTPLEDQQVF VGDRVEMAVE VSEEGAQVMW  420
LKDGVELTRE DSFKARYRLK KDGKRHILIY SEVTMEDKGH YQVMTNGGQC EAELIVEEKQ  480
LEVLQDIADL TVKASEQAVF KCEVSDEKVT GKWYKNGVEV RPSKRITISH TGRFHKLVID  540
DVRPEDEGDY TFVPDGYALS LSAKLKFLEI KVEYVPKQEP PKIHLDCSGQ TSENAIVVVA  600
GNKLRMDVSI TGEPRPVATW MKEDEVFTGT EGRVHIEQRG DISSFVIESA ERGDEGRYTI  660
KVTNPVGEDV ASILLKVVDV PDPPEAVRVT SVGEDWAVLV WEPPKYDGGR PVTGYLLERK  720
KKGSQRWMKL NFEVFTETTY ESTNMIEGIL YEMRVFAVNV IGVSQPSMNT KPFMPIAPTS  780
EPRHLTVEDV TDTTTTLKWR PPDKIGAGGI DGYLIEYCVE GSDEWIPANT EPVERCGFTV  840
KNLPTGAKIT FRVVGVNIAG RSQPATLAQP VTIREIVEQP KIRLPRHLRQ TYVRKVGEHL  900
NLVIPFQGKP RPQVVWTKGG APIDTSHVHV RTSDFDTVFF VRQAARSDSG EYELTVQIEN  960
MKDTATVCIR VVEKAGPPQN VMVKEVWGTN ALVEWQPPKD DGNSEITGYF VQKADKKTME  1020
WFTVYERNRH TSCTVSDLIM GNEYYFRVYS ENVCGLSDLP GVSKNTARIV KTGMTLKLPE  1080
YKEHDFRTPP KFLTPLPDRV VVAGYAAALN CAVRGHPKPK VVWMKNKMEI REDPKFLMTN  1140
QQGVLTLNIR RPSPFDSGTY SCRAVNELGE ALAECKLEVR VPQ                    1183

SEQ ID NO: 132          moltype = AA  length = 1430
FEATURE                 Location/Qualifiers
source                  1..1430
                        mol_type = protein
                        organism = Bison bison
SEQUENCE: 132
MSLVAVPFYQ KRHKHFDQSY RNIQTRYLLD EYSAKKKRAS AQSSLQRSVT RKSSSQRESS  60
RAALGESTLG ETTLGETSLG DTTLGRTTCR LCAKRMSSAL EEEAQERKRR YQSMVAAYGE  120
ILVRLRSHTV WERMSVKLCF TVQGFPTPVV QWYKDGSLIC QAGEPGKYRI DSKYGVHTLE  180
INRADFDDTA TYSAVATNVH GQVSTNAAVV VRRFRGDEEP FHSVGLPIGL PLSSVIPYTH  240
FDVQFIEKFG VTFRREGETL TLKCTLLVTP DLKRVQPRAE WYRDDVLLKE SKWTKMFFGE  300
```

-continued

```
GQASLSFSHL NKDDEGLYTL RIVSRGGITD HSAFLFVRDA DPLVTGAPGA PMDVHCHDAN  360
RDYVIVTWKP PNTTKESPVI GYFIDRCEVG TDNWIQCNDA PVKICKYPVT GLFEGRSYIF  420
RMRAVNSAGI SRPSRVSEAV AALDPVDLRR LQAVHLEGEK DIVIYQEELE GEVQIPGPPT  480
NVHASETSRT YVVLSWDPPE PRGKEPLMYF IEKSMVGSGS WQRVNAQTAV RSPRYAVFDL  540
AEGKPYVFRV LSANKHGLSD PSEITAPIQA QDTIVVPSAP GRVLASRNTR TSVVVQWDRP  600
KHEEDLLGYY VDCSVAGSNV GEPCNHKPIG YNRFVVHGLT TGEQYIFRVK AVNAVGTSEN  660
SQESDVIKVQ AALTVPSHPY GITLLNCDGH SMILGWKVPK FSGGSPILGY YVDKREAHHK  720
NWHEVNSSPL KERILTVEGL TEGSLYEFKI AAANMAGIGQ PSDPSELFKC EAWTMPEPGP  780
AYDLTFCEVR DTSLVLLWKP PVYSGSSPVS GYFVDYKEED SGEWLTVHET ATPHRYLKVC  840
DLHQGKTYVF RVRAVNASGV GRPSDTSEPV LVEARPGTKE VSAGVDEEGN VYLSFDCPEV  900
TDASHFTWCK SYEDIADDDR FKVETAGDHS KLYFKNLDKE DLGTYSVSVS DTDGVSSSFV  960
LDEEELERLK ALSNEIKNPT IPLKSELAYE IFDKGQVRFW LQAEHLSPDS SYRFIINDRE  1020
VADSETHRIK CDKSTGIIEM VMDRFTIDNE GTYTVQIQDG KAKSQSSLVL IGDAFKAVLK  1080
EAEFQRKEFL RKQGPHFAEY LHWDVTEECE VRLVCKVANT KKETVFKWLK DDVLYETEKM  1140
PDLEKGVCEL LIPKLSKKDH GEYKATLKDD RGQDVSTLEI AGKVYEDMIL AMSRVCGASA  1200
SPLKILCTPE GIRLQCFMKY FTEEMKVNWY HKDAKITSSE HMRIGGSEQM AWLQICEPTE  1260
KDKGKYTFEI LDDKKNHQRS LDLSGQAFDE AFAEFQQLKA AAFAEKNRGK VIGGLPDVVT  1320
IMEGKTLNLT CTVFGNPDPE VVWFKNDKDI ELSDHFSVKV EQAKYVSMTI KGVTSEDSGK  1380
YSIHVKNKYG GEKIDVTVSV YKHGEEIPDV VLPQQAKPKL LPAAAPSPAH             1430

SEQ ID NO: 133          moltype = AA  length = 440
FEATURE                 Location/Qualifiers
source                  1..440
                        mol_type = protein
                        organism = Bison bison
SEQUENCE: 133
MRTLRAMAMQ KIFAREILDS RGNPTVEVDL HTAKGRFRAA VPSGASTGIY EALELRDGDK  60
SRYLGKGVLK AVEHINKTLG PALLEKKLSV VDQEKVDKFM IELDGTENKS KFGANAILGV  120
SLAVCKAGAA EKGVPLYRHI ADLAGNPELI LPVPAFNVIN GGSHAGNKLA MQEFMILPVG  180
ASSFREAMRI GAEVYHHLKG VIKAKYGKDA TNVGDEGGFA PNILENNEAL ELLKTAIQAA  240
GYPDKVVIGM DVAASEFYRN GKYDLDFKSP DDPARHISGE KLGELYKNFI KNYPVVSIED  300
PFDQDDWATW TSFLSGVNIQ IVGDDLTVTN PKRIAQAVEK KACNCLLLKV NQIGSVTESI  360
QACKLAQSNG WGVMVSHRSG ETEDTFIADL VVGLCTGQIK TGAPCRSERL AKYNQLMRIE  420
EALGDKAVFA GRKFRNPKAK                                              440

SEQ ID NO: 134          moltype = AA  length = 170
FEATURE                 Location/Qualifiers
source                  1..170
                        mol_type = protein
                        organism = Bison bison
SEQUENCE: 134
MAPKKAKRRA AAEGGSSSVF SMFDQTQIQE FKEAFTVIDQ NRDGIIDKED LRDTFAAMGR  60
LNVKNEELDA MMKEASGPIN FTVFLNMFGE KLKGADPEDV ITGAFKVLDP EGKGTIKKQF  120
LEELLTTQCD RFSHEEIKNM WAAFPPDVGG NVDYKNICYV ITHGDAKDQE             170

SEQ ID NO: 135          moltype = AA  length = 1938
FEATURE                 Location/Qualifiers
source                  1..1938
                        mol_type = protein
                        organism = Bison bison
SEQUENCE: 135
MSSDQEMAIF GEAAPYLRKS EKERIEAQNK PFDAKTSVFV ADPKESFVKA TVQSREGGKV  60
TAKTEAGATV TVKEDQVFPM NPPKFDKIED MAMMTHLHEP AVLYNLKERY AAWMIYTYSG  120
LFCVTVNPYK WLPVYNAEVV TAYRGKKRQE APPHIFSISD NAYQFMLTDR ENQSILITGE  180
SGAGKTVNTK RVIQYFATIA VTGEKKKEEP TSGKMQGTLE DQIISANPLL EAFGNAKTVR  240
NDNSSRFGKF IRIHFGTTGK LASADIETYL LEKSRVTFQL KAERSYHIFY QIMSNKKPEL  300
IEMLLITTNP YDYAYVSQGE ITVPSIDDQE ELMATDSAIE ILGFTSDERV SIYKLTGAVM  360
HYGNLKFKQK QREEQAEPDG TEVADKAAYL QGLNSADLLK ALCYPRVKVG NEFVTKGQTV  420
EQVYNAVGAL AKAVYDKMFL WMVARINQQL DTKQPRQYFI GVLDIAGFEI FDFNSLEQLC  480
INFTNEKLQQ FFNHHMFVLE QEEYKKEGIE WEFIDFGMDL AACIELIEKP MGIFSILEEE  540
CMFPKATDMS FKNKLYEQHL GKSNNFQKPK PAKGKAEAHF SLIHYAGTVD YNITGWLDKN  600
KDPLNETVVG LYQKSSVKTL ALLFSGPASG EAEGGPKKGG KKKGSSFQTV SALFRENLNK  660
LMTNLRSTHP HFVRCIIPNE TKTPGAMEHE LVLHQLRCNG VLEGIRICRK GFPSRILYAD  720
FKQRYKVLNA SAIPEGQFID SKKASEKLLA SIDVDHTQYK FGHTKVFFKA GLLGLLEEMR  780
DEKLAQLITR TQARCRGFLA RVEYQKMVER RESIFCIQYN VRAFMNVKHW PWMKLYFKIK  840
PLLKSAETEK EMANMKEEFE KTKEELAKSE AKRKELEEKM VTLTQEKNDL QLQVQSEADA  900
LADAEERCDQ LIKTKIQLEA KIKEVTERAE DEEEINAELT AKKRKLEDEC SELKKDIDDL  960
ELTLAKVEKE KHATENKVKN LTEEMAGLDE TIAKLTKEKK ALQEAHQQTL DDLQAEEDKV  1020
NTLTKAKTKL EQQVDDLEGS LEQEKKLRMD LERAKRKLEG DLKLAQESTM DIENDKQQLD  1080
EKLKKKEFEM SNLQSKIEDE QALAMQLQKK IKELQARIEE LEEEIEAERA SRAKAEKQRS  1140
DLSRELEEIS ERLEEAGGAT SAQIEMNKKR EAEFQKMRRD LEEATLQHEA TAAALRKKHA  1200
DSVAELGEQI DNLQRVKQKL EKEKSEMKME IDDLASNMET VSKAKGNLEK MCRALEDQLS  1260
ELKTKEDEQQ RLINDLTTQR ARLQTESGEF SRQLDEKDAL VSQLSRGKQA FTQQIEELKR  1320
QLEEEIKAKS ALAHALQSAR HDCDLLREQY EEEQEGKAEL QRAMSKANSE VAQWRTKYET  1380
DAIQRTEELE EAKKKLAQRL QDAEEHVEAV NAKCASLEKT KQRLQNEVED LMIDVERTNA  1440
ACAALDKKQR NFDKILSEWK QKYEETHAEL EASQKESRSL STELFKIKNA YEESLDQLET  1500
```

-continued

```
LKRENKNLQQ EISDLTEQIA EGGKRIHELE KVKKQVEQEK SEIQAALEEA EASLEHEEGK  1560
ILRIQLELNQ VKSEIDRKIA EKDEEIDQLK RNHIRVVESM QSTLDAEIRS RNDAIRLKKK  1620
MEGDLNEMEI QLNHANRMAA EALKNYRNTQ AILKDTQIHL DDALRGQEDL KEQLAMVERR  1680
ANLLQAEIEE LRATLEQTER SRKIAEQELL DASERVQLLH TQNTSLINTK KKLETDITQI  1740
QGEMEDIIQE ARNAEEKAKK AITDAAMMAE ELKKEQDTSA HLERMKKNLE QTVKDLQHRL  1800
DEAEQLALKG GKKQIQKLEA RVRELEGEVE SEQKRNVEAV KGLRKHERRV KELTYQTEED  1860
RKNILRLQDL VDKLQAKVKS YKRQAEEAEE QSNVNLSKFR KLQHELEEAE ERADIAESQV  1920
NKLRVKSREV HTKIISEE                                                1938

SEQ ID NO: 136          moltype = AA  length = 6909
FEATURE                 Location/Qualifiers
source                  1..6909
                        mol_type = protein
                        organism = Bison bison
SEQUENCE: 136
MADDEEYEEV VEYYTEETVY KEVPGETITG FYETTTTRTS DYEQSETSRP ALAQPVPEKP  60
VERKRVIRKK VDPSKFMTPY IAHSQKMQNL FSTNKYKENY EKAKGKPYAI TTDTPELRRI  120
KKVQDQLSEV KYRVDGDVAK TICHVDEKAK DIEHAKKVSQ QVSKVLYKQN WEDTKDKYLL  180
PPDAPELVQA IKNTAMFSKK LYTEDWEADK TMFYPYNDSP ELRRVAQAQK ALSDIAYKKG  240
LAEQQTQFTS LPDPPDIEFA KKVTNQVSKQ KYKEDYENKV KGKWSETPCF EIATARMNSN  300
NISTKKYQED FEHMKDQIYF MQTETPEYKV NKQAGVAASK VKYKQDYEKN KGKADYNVLP  360
ASENPLLRQL KAAGDALSDK LYKENYEKTK AKSINYCETP KFKLDTVLHN FSSDTKYKDS  420
YLKDILGHYV GSYEDPYHTH CMRVSAQNSD KNYKAEYEED RGKGFFPQTI TQEYEAIKKL  480
DQCKDHTYKV HPDKTKFTQV TDSPVLMQAQ VNSKQLSDLN YKAKHENEKF KCHIPPDTPA  540
FIQHKLNAYN LSDNIYKHDW EKTKAKKFDI KVDAIPLLAA KANTKIASDV MYKKDYEKSK  600
GKMIGALSIN DDPKMLHSLK TAKNQSDRLY KENYEKTKAK SMNYCETPKY QLDTLLKNFS  660
EAKYKDSYVQ NVLGHYIGSF EDPYQVHCLK ISAQNSDKNY KAEYEEDKGK CYFPQTITQE  720
YEAIKKLDQC KDHTYKVHPD KTKFTAVTDT PVLLQAQLNT KQLSDLNYKA KHEGEKFKCH  780
IPADAPQFIQ HRVNAYNLSD NIYKHDWEKT KAKKFDIKVD AIPLLAAKAN TKIASDVMYK  840
KDYEKSKGKM IGALSINDDP KMLHSLKTAK NQSDHEYRKD YEKSKTIYTA PLDMLPLTHA  900
KKSQAIASDV DYKHLLHNYS YPPDSVNVDL AKKAYALQSD VEYKADYNSW MKGCGWMPFG  960
SLEMEKAKRA SDILNEKKYR QHPDTLKFTS IEDAPIIVQS KINQAQRSDV AYKAKGEEVI  1020
HKYSLPADLP QFIQAKVNAY NISENLYKAD LKDLSKKGYD LRIDAIPIKA AKAARQAASD  1080
VQYKKDYEKA KGKMVGFQSL QDDPKLVHYM NVAKIQSDRE YKKAYEKTKT RHNTPHDMVN  1140
IVAAKKAQDV ASNVNYKHSL HHYTYLPDAM DLELSKNMMH IRSDNVYKED YNNWMKGIGW  1200
IPIGSLEVEK VKKAGDALNE KKYRQHPDTL KFTSIVDSPV MVQAKQNTQQ VSDILYKAKG  1260
EDVKHKYTMS PDLPQFLQAK CNAYNLSDVC YKRDWHDLIA KGTNVLGDAI PITAAKASRN  1320
IASDYKYKEA YEKAKGKQVG FRSLQDDPKL VHYMNVAKLQ SDREYKKNYE NTKTSYHTPG  1380
DMVSITAAKM AQDVATNVNY KQPIHHYTYL PDALSLEHIR NVNQIQSDNV YKDEYNHFFK  1440
GMGWIPIGSL EVEKVKKAGD ALNEKKYRQH PDTIKFTSVP DSMGMVLAQH NTKQLSDLNY  1500
KVEGEKVKHK YTMDPDVPQF IQAKVNAYNM SDSHYKADWK KTLAKGYDLR PDAIPIVAAK  1560
SSRNIASDFK YKEAYEKTKG KQIGFRSLQD DPKLVHFMNV AKMQSDREYK KGYEARKTKY  1620
HTPLDMLSVT AAKKSQEVAT NTNYKQPFHH YTLLPDALNV EHSRNAMQIQ SDNLYKSDFT  1680
NWMKGIGWLP LESLEVEKAK KAGEILSEKK YRQHPEKLKF TYAMDTMEQA LNKSNKLIMD  1740
KRLYTEKWNK DKTTIHVMPD TPDILLSRVN QITMSDKLYK AGWEEEKKKG YDLRPDAISI  1800
KAARASRDIA SDYKYKQAYE QAKGRQIGFR SLEDDPKLVH FMQVAKMQSD REYKKAYEKS  1860
KTSFHTPVDM LSVVAAKKSQ EVATNANYRN VIHTYNMLPD AMSLELAKNM MQIQSDNQYK  1920
ADYADFMKGI GWLPLGSLEA EKNKKAMEII SEKKYRQHPD TLKYSTLMDS MNMVLAKNNA  1980
KIMNEHLYKQ AWEADKTKVH IMPDIPQIIL AKANAINISD KLYKLSLEEA KKKGYDLRTD  2040
AIPIKAAKAS RDIASDYKYK HSYEKERGKM VGFRSLEDDP KLVHSMQVAK MQSDREYKKN  2100
YEKTKTSYHT PADMLSVTAA KDAQANITNT NYKHLIHKYI LLPDAMNIQL SKNMNRIQSD  2160
NEYKQDYNEW YKGLGWSPAG SLEVEKAKKA TEYASDQKYR QHPSNFQFTK LNDSMDMVLA  2220
KQNAHTMNKY LYTIDWNKDK TKIHVMPDTP DILQAKQNQT MYSQKLYKLG WEEALKKGYD  2280
LPVDAISVQA AKASRDIVSD FKYKQGYRKQ IGHHIGFRSL QDDPKLVLSM NVAKMQSERE  2340
YKKDFEKWKT KFSSPVDMLG VVLAKKCQAL VSDVDYKNYL HGWTCLPDQN DVIHAKKAYD  2400
LQSENLYKSD LEWLKGIGWS PLGSLEAEKN KRASEIISEK KYRQPPDRNK FTSIPDAMDI  2460
VLAKTNAKNR SDILYREAWD KDKTQIHIMP DTPDIILAKA NLINTSDKFY RMGYEELRKK  2520
GYDLPVDAIP IKAAKASREI ASEYKYKEGF RMQLGHHIGA RNIKDDPKMM WSMHVAKIQS  2580
DREYKKDFEK WKTKFSSPVD MLGVVLAKKC QTLVSDIDYK NYLHQWTCLP DQSDVIHARR  2640
AYDLQSDNLY KSDLQWLRGI GWLPSGSLED EKNKRASQIL SDHVYRQHPD QFKFSSLMDS  2700
IPMVLAKNNA ITMNHRLYTE AWDKDKTTVH IMPDTPEVLL AKQNQINYSE KLYKLGLDEA  2760
KRKGYDMRID AIPIRAAKAS RDIASEFKYK EGYRRQLGHH IGARAIHDDP KMMWSMHVAK  2820
IQSDREYKKD FEKWKTKFSS PVDMLGVVLA KKCQTLVSDI DYKNYLHQWT CLPDQSDVIH  2880
ARRAYDLQSD NVYKSDLQWL RGIGWVPIGS LDVEKSKRAS EILSDKLYRQ PPDKFKFTSV  2940
TDSLEQVLAK NNAINMNKRL YTEAWDKDKT QVHIMPDTPE ITLARTNKVN YSENLYKLAH  3000
EEAKKKGYDL RSDAIPIVAA KASRDIISDY KYKDGYRKQL GHHIGARDIK DDPKMMWSMH  3060
VAKIQSDREY KKDFEKWKTK FSSPVDMLGV VLAKKCQTLV SDIDYKNYLH QWTCLPDQSD  3120
VIHARRAYDL QSDNVYKSDL QWLRGIGWVP IESVDVVKCK RAAEILSDNL YRQPPDTLKF  3180
TSVPDSLEQV LAKNNAINMN KRLYTEAWDK DKTQIHIMPD TPEIMLARQN KLNYSETLYK  3240
LANEEAKKKG YDLRSDAIPI VAAKASRDII SDYKYKDGYR KQLGHHIGAR DIKDDPKMMW  3300
SMHVAKIQSD REYKKDFEKW KTKFSSPVDM LGVVLAKKCQ TLVSDIDYRN YLHQWTCLPD  3360
QSDVIHARQA YDLQSDNIYK SDLQWLRGIG WVPIGSVDVV KCKRAAEILS DNLYRQRPDT  3420
LKFTSVPDSL EQVLAKNNAI NMNKRLYTEA WDKDKTQIHM MPDTPEITLA RQNKINYSEN  3480
LYRQAMEEAK KEGYDLRSDA IPIVAAKASR EIASDYKYKE AYRKQLGHHI GARAIHDDPK  3540
MMWSVHVAKM QSDREYKKDF EKYKTRFSSP VDMLGIVLAK KCQTLVSDVD YKHPLHEWTC  3600
LPDQNDIIHA RKAYDLQSDN LYKSDLEWLK GIGWVPIGSV EVLKAKRAGE ILSDNIYRQR  3660
```

-continued

```
PDTLKFTSVT DSPEQVLAKN NAINMNKRLY TEAWDNDKKT IHVMPDTPEI MLAKLNRINY  3720
SDKLYKLALE ESRKEGYDLR LDAIPIQAAK ASREIASDYK YKEGYRKQLG HHIGARNIKD  3780
DPKMMWSIHA GKIQSDLEYK KDFEKWKTKF SSPVDMMGLV QAKKCQILVS DIDYKNLLHE  3840
WTCLPDQNDI IQARKAYDLQ SDAIYKADLE WLRGIGWVPI DSVGVEHAKR AGEILSERKY  3900
RQPAVQLKFT SITDTPEIVL AKNNALNVSK HLYTEAWDAD KTSIHVMPDT PEILLAKSNS  3960
ANISHKLYTK GWDESKMKDY DLRADAISIK SAKASRDIAS DYKYKEAYEK HKGHHIGARS  4020
VEDDPRIMCA MNAGRIQSER EYKKEFQKWK TKFSSPVDML GILLAKKCQT LVSDIDYRNY  4080
LHHWTCLPDQ NDIIQARKAY DLQSDAIYKA DLEWLRGIGW MPEGSPEVLR VKNAQHIFRD  4140
SVYRTPVVKL KYTSIVDTPE VVLAKSNAEN ISIPKYREVW DKDKTSIHIM PDTPEINLAR  4200
TNALNVSNKL YREGWDEMKM SCDVRLDAIP IQAAKASREI ASDYKYKLDH EKQKGHYVGT  4260
RTARDDNKIR WALIAGKIQN EREYRLHWAK WKSKFQSPPD MLSIEHSKNS QTLVSDIDYR  4320
HYLHQWTCMP DQNDVIQARK AYDLQSDAIY KADLEWLRGI GWMPADSVSV NHAKHMADIF  4380
NEKKYRTKIE TLSFTPVDDR VDYVTAKHSG EILNDIKYRK DWNDTKSKYT LTETPQLHTA  4440
QEAARILDQY LYKESWEKQK ATGYILPPDA VPFVHAHHSG DVQSELKYKA EHVKQKGHYV  4500
GVPTMRDDPK LVWFEHAGQI QNDRLYKENY HKTKAKIHIP PDMVSVLAAK EGQALASDID  4560
YRNYLHQWIC HPDQNDVIQA RKAYDLQSDN IYKADLEWLR GIGWIPLDSV DHVRVTRNQE  4620
MMNQIKYKKD ALANYPNFTS VVDPPEIVLA KINSVNQSDV KYKETFNKLI KGKYIFSPDT  4680
PYITHSKDME KLYSTILYKR AWDGTKAYGY TLDERYIPIV GAKHADFVNS ELKYKETYEK  4740
LKGHYLAGKE ISEFPNVVHC LDFQKMRSLL NYRRHYEDTK ANVHIPQDMM NHVLAKRCQY  4800
ILSDLEYRHY FHQWTSLPEE PNVVRVRHAQ EILSDNVYKD DLNWLKGIGC YVWDTPQILH  4860
AKKSYDLQSQ ILYTAAGKEN LKNYNLVTDT PLYVTALQSG INASEVKYKE NYHQIKDKYT  4920
TVLETVDYDR IKNLKDLFSS NLYKEAWDKV KATSYILPPN TVSLTHAKNQ KYMASHIKYR  4980
EEYEKFKALY TLPRSVEDDP NTARCLRVGK FNIDRLYRSV YEKNKMKIHI VPDMVEMVTA  5040
KDSQKKVSEI DYRLHLHEWI CHPDLQVNSH VRKVTDQISD IVYKDDLTWL KGIGCYVWDT  5100
PEILHAKHAY DLRNDIKYKA HVLKTRNNYK LVTDTPVYVQ AVKSGKQLSD AVYHYDYVHS  5160
IRGRVAPTTK TVDLDRALHA YKLQSENLYR KAGLHALPTG YRLPVDTPHF KHTKDTRYMS  5220
SYFKYKEVYE HMKAYGYTLG PNDVPFVNVR RVNNITSERL YRQLYHKLKD KIHTTPDTPE  5280
IRQVKKTQEA VSELIYKSDF FKMQGHMISL PYTPQVLHCR YVGDITSDIK YKEDLQVLRG  5340
MGCFLYDTPD MVRSRHLRKL WSHYLYTDKA RKMRDKYKVV LDTPEYRKVQ ELKTHLSELV  5400
YRAAGRKQKS IFTSVPDTPD LTRAKRGQKL QSQYLYVELA TKERPHHHAG NQTTALKHAR  5460
HVKDMVSENK YKIQYEKMKD KYTPVPDTPI LIRAKRAYWN ASDLRYKETF QKTKGKYHTV  5520
KDALDIVYHR TVTDHVSKIK YKENYMSQLG IWRSIPDRPE HFHHRAVTDA VSDVKYKQDL  5580
TWLKGIGCYA YDTPDFTLAE KNKTLYSKYK YKEVFERTKS NFKYVADCPI NRHFKFATQL  5640
MNEKKYRADY EQRKDKYHLV VDEPRHLLAK IAGDQISQIK YRKNYEETKD KFTSIVDTPE  5700
HLRTTKVNKQ ISDILYKLEY NKAKPRGYTT IHDTPMLLHV RKVKDEVSDL KYKEVYQRTK  5760
SNCTIEPDAV HIKAAKDAYK VNTNLDYKKK YEATKAYWKW TPDRPDFIQA AKSTLQQSDF  5820
EYKLDREYLK GCKLSVTDDK DMVLALKNSI IESDLKYKEK HVKERGSCHA VPDTPQILLA  5880
KTVSSLVSEN KYKSYVKKHL AQGSYTTLPE TRDTIHVKEV TKNVSDTNYK KKFVKEKGKS  5940
NYSIMLEPPD VKHAMDVAKK QSNVAYKKDA KENLHYTTVA DRPDIKKATQ AAKQASEVEY  6000
RAKHRKEGSH GLSMLGRPDI EMAKKAAKLS SQVQYRENFN KEKGKTPKYN PKDSQLYKVM  6060
KDANTLASEV KYKADLKKLH KPVTDMKESL IMNHVLNTSH LASSYQYKKN YEKSKGHYHT  6120
IPDNLEQLHL KEATELQSIV KYKEKYEKER GKPMLDFETP TYITAKESQQ MQSGKEYRKD  6180
YEESIKGRNL TGLEVTPALL HVKYATKIAS EKEYRKDLEE SIRGKGLSEM EDTPDMLRAK  6240
NATQILNEKE YKRDLELEVK GRGLNAMANE TPDFLRARNA TDIASQIKYK QSAEMEKANF  6300
TSVVDTPEII HAQQVKNLSS QKKYKEDAEK CMSYYETVLD TPEMQRVREN QKNFSLLQYQ  6360
YDLKNSKGKI TVVQDTPEIL RVKENQKNFS SVLYKEDVSP GTAIGKTPEM MRVKQTQDHI  6420
SSVKYKEVIG QGTPIPDLPE VKRVKQTQKH ISSVMYKENL GTGIPTPVTP EIERVKRNQE  6480
NFSSVLYKEN LGKGTPTPIT PEMERVKRNQ ENFSSVLYKE NMGKGTPLPV TPEMERIKHN  6540
QENISSVLYK ENVGKGTPTP VTPEMQRVKR NQENISSVLY KENLGKATPT PFTPEMERVK  6600
RNQENFSSVL YKENMRKATP TPVTPEMERA KRNQENISSV LYSDSFRKQI QGKAAYVLDT  6660
PEMRRVRETQ RHISTVKYHE DFEKHKGCFT PVVTDPITER VKKNTQDFSD ICYRGIQRKV  6720
VEMEQKRNDQ DQETITGLRV WRTNPGSVFD YDPAEDNIQS RSLHMINAQA QRRSREQSRS  6780
ASGLSISGGE EKSEHSEAAH LSTYSDGGVF FSAASTGYKH ARTTELPQQR SSSVATQQTT  6840
VSSIPSHPST AGKIFRAMYD YMAADADEVS FKDGDAIVNV QAIDEGWMYG TVQRTGRTGM  6900
LPANYVEAI                                                          6909
```

```
SEQ ID NO: 137          moltype = AA  length = 1938
FEATURE                 Location/Qualifiers
source                  1..1938
                        mol_type = protein
                        organism = Bison bison
SEQUENCE: 137
MSSDQEMAIF GEAAPYLRKS EKERIEAQNK PFDAKTSVFV VDPKESFVKA TVQSREGGKV  60
TAKTEAGATV TVKEDQIFPM NPPKYDKIED MAMMTHLHEP AVLYNLKERY AAWMIYTYSG  120
LFCVTVNPYK WLPVYNAEVV TAYRGKKRQE APPHIFSISD NAYQFMLTDR ENQSILITGE  180
SGAGKTVNTK RVIQYFATIA VTGEKKKEEP TSGKMQGTLE DQIISANPLL EAFGNAKTVR  240
NDNSSRFGKF IRIHFGTTGK LASADIETYL LEKSRVTFQL KAERSYHIFY QIMSNKKPEL  300
IEMLLITTNP YDYAFVSQGE ITVPSIDDQE ELIATDSAID ILGFTSDERV SIYKLTGAVM  360
HYGNLKFKQK QREEQAEPDG TEVADKAAYL QSLNSADLLK ALCYPRVKVG NEFVTKGQTV  420
QQVYNAVGAL AKAVYEKMFL WMVTRINQQL DTKQPRQYFI GVLDIAGFEI FDFNSLEQLC  480
INFTNEKLQQ FFNHHMFVLE QEEYKKEGIE WEFIDFGMDL AACIELIEKP MGIFSILEEE  540
CMFPKATDTS FKNKLYEQHL GKSNNFQKPK PAKGKAEAHF SLVHYAGTVD YNIAGWLDKN  600
KDPLNETVVG LYQKSAMKTL AFLFAGGPSA EEGGGGKKGG KKKGSSFQTV SALFRENLNK  660
LMTNLRSTHP HFVRCIIPNE TKTPGAMEHE LVLHQLRCNG VLEGIRICRK GFPSRILYAD  720
FKQRYKVLNA SAIPEGQFID SKKASEKLLG SIDIDHTQYK FGHTKVFFKA GLLGILEEMR  780
DEKLAQLITR TQAICRGFLM RVEFRKMMER RESIFCIQYN VRAFMNVKHW PWMKLYFKIK  840
```

-continued

```
PLLKSAETEK EMANMKEEFE KTKEELAKSE AKRKELEEKM VTLMQEKNDL QLQVQSEADG  900
LADAEERCDQ LIKTKIQLEA KIKELTERAE DEEEINAELT AKKRKLEDEC SELKKDIDDL  960
ELTLAKVEKE KHATENKVKN LTEEMAGLDE NIAKLTKEKK ALQEAHQQTL DDLQAEEDKV 1020
NTLTKSKTKL EQQVDDLEGS LEQEKKLRMD LERAKRKLEG DLKLAQESTM DIENDKQQLD 1080
EKLKKKEFEM SNLQSKIEDE QALAMQLQKK IKELQARIEE LEEEIEAERA SRAKAEKQRS 1140
DLSRELEEIS ERLEEAGGAT SAQIEMNKKR EAEFQKMRRD LEEATLQHEA TAAALRKKHA 1200
DSVAELGEQI DNLQRVKQKL EKEKSEMKME IDDLASNMET VSKAKGNLEK MCRTLEDQLS 1260
EVKSKEEEHQ RLINELSAQK ARLHAESGEF SRQLEEKDAL VSQLSRGKQA FTQQIEELKR 1320
QLEEETKAKS ALAHALQSAR HDCDLLREQY EEEQEGKAEL QRALSKANSE VAQWRTKYET 1380
DAIQRTEELE EAKKKLAQRL QDAEEHVEAV NAKCASLEKT KQRLQNEVED LMLDVERSNA 1440
ACAALDKKQR NFDKVLSEWK QKYEETQAEL EASQKESRSL STELFKVKNA CEESLDQLET 1500
LKRENKNLQQ EISDLTEQIA EGGKHIHELE KVKKQIEQEK GELQAALEEA EASLEHEEGK 1560
ILRIQLELNQ VKSEIDRKIA EKDEEIDQLK RNHLRVMEST QSTLDAEIRS RNDALRIKKK 1620
MEGDLNEMEI QLNHANRQAT EAIKNLRNTQ GVLKDTQLHL DDAIRGQDDL KEQLAMVERR 1680
ANLMQAEIEE LRASLEQTER SRRVAEQELL DASERVQLLH TQNTSLINTK KKLETDITQI 1740
QGEMEDILQE ARNAEEKAKK AITDAAMMAE ELKKEQDTSA HLERMKKNME QTVKDLQHRL 1800
DEAEQLALKG GKKQIQKLEA RVRELENEVE NEQKRNAEAV KGLRKHERRV KELTYQTEEY 1860
RKNVLRLQDL VDKLQSKVKA YKRQAEEAEE QSNVNLAKFR KLQHELEEAE ERADIAESQV 1920
NKLRVKSREV HTKVISEE                                               1938

SEQ ID NO: 138          moltype = AA  length = 607
FEATURE                 Location/Qualifiers
source                  1..607
                        mol_type = protein
                        organism = Bison bison
SEQUENCE: 138
MKWVTFISLL LLFSSAYSRG VFRRDTHKSE IAHRFKDLGE EHFKGLVLIA FSQYLQQCPF  60
DEHVKLVNEL TEFAKTCVAD ESHAGCEKSL HTLFGDELCK VASLRETYGN MADCCEKQEP 120
ERNECFLSHK DDSPDLPKLK PDPNTLCDEF KADEKKFWGK YLYEIARRHP YFYAPELLYY 180
ANKYNGVFQE CCQAEDKGAC LLPKIETMRE KVLTSSARQR LRCASIQKFG ERALKAWSVA 240
RLSQKFPKAE FVEVTKLVTD LTKVHKECCH GDLLECADDR ADLAKYICDN QDTISSKLKE 300
CCDKPLLEKS HCIAEVEKDA VPENLPPLTA DFAEDKDVCK NYQEAKDAFL GSFLYEYSRR 360
HPEYAVSVLL RLAKEYEATL EECCAKDDPH ACYSTVFDKL KHLVDEPQNL IKQNCDQFEK 420
LGEYGFQNAL IVRYTRKVPQ VSTPTLVEVS RSLGKVGTRC CTKPESERMP CTEDYLSLIL 480
NRLCVLHEKT PVSEKVTKCC TESLVNRRPC FSALTPDETY VPKAFDEKLF TFHADICTLP 540
DTEKQIKKQT ALVELLKHKP KATEEQLKTV MENFVAFVDK CCAADDKEAC FAVEGPKLVV 600
STQTALA                                                           607

SEQ ID NO: 139          moltype = AA  length = 667
FEATURE                 Location/Qualifiers
source                  1..667
                        mol_type = protein
                        organism = Cervus elaphus
SEQUENCE: 139
MLDAEDIVNT PKPDEKAIMT YVSCFYHAFA GAEQAETAAN RICKVLAVNQ ENEKLMEEYE  60
KLASELLEWI RRTVPWLENR VGEPSMSAMQ RKLEDFRDYR RLHKPPRVQE KCQLEINFNT 120
LQTKLRLSHR PAFMPSEGKL VSDIAKAWRG LEQVEKGYED WLLSEIRRLQ RLQHLAEKFQ 180
QKASLHEAWT RGKEEMLSQR DYETASLQEV RALLRRHEAF ESDLAAHQDR VEHIAALAQE 240
LNELDYHEAA SVNSRCQAIC DQWDNLGTLT QKRRDALERM EKLLETIDQL QLEFARRAAP 300
FNNWLDGAVE DLQDVWLVHS VEETQSLVTA HDQFKATLPE ADRERGAILG IQGEIQKICQ 360
TYGLRPSSTN PYIALTPQDI NTKWDTVRKL VPSRDQMLQE ELTRQQVNER LRRQFAAQAN 420
AIGPWIQGKV EEVGRLAAGM AGSLEEQMAG LRQQEQNIIN YKSNIDRLEG DHQLLQESLV 480
FDNKHTVYSM EHIRVGWEQL LTSIARTINE VENQVLTRDA KGLSQEQLNE FRASFNHFDR 540
KRNGMMEPDD FRACLISMGY DLGEVEFARI MTMVDPNAAG VVTFQAFIDF MTRETAETDT 600
AEQVVASFKI LAGDKNYITA EELRRELPAE QAEYCIRRMA PYKGAGAPAG ALDYVAFSSA 660
LYGESDL                                                           667

SEQ ID NO: 140          moltype = AA  length = 792
FEATURE                 Location/Qualifiers
source                  1..792
                        mol_type = protein
                        organism = Cervus elaphus
SEQUENCE: 140
MSRPLTDQEK RKQISVRGLA GVENVTELKK NFNRHLHFTL VKDRNVATPR DYYFALAYTV  60
RDHLVGRWIR TQQHYYEKDP KRIYYLSLEF YMGRTLQNTM VNLALENACD EATYQLGLDM 120
EELEEIEEDA GLGNGGLGRL AACFLDSMAT LGLAAYGYGI RYEFGIFNQK ISGGWQMEEA 180
DDWLRYGNPW EKARPEFTLP VHFYGRVEHT SQGAKWVDTQ VVLAMPYDTP VPGYRNNVVN 240
TMRLWSAKAP NDFNLKDFNV GGYIQAVLDR NLAENISRVL YPNDNFFEGK ELRLKQEYFV 300
VAATLQDIIR RFKSSKFGCL DPVAIQLNDT HPSLAIPELM RILVDQERLE WEKAWEVTVK 360
TCAYTNHTVL PEALERWPVH LIETLLPRHL QIIYEINQRF LNRVAAAFPG DVDRLRRMSL 420
VEEGAVKRIN MAHLCIAGSH AVNGVARIHS EILKKTMTSK SWSLINSRIR PMGSPLGAGW 480
LCRIGEEYIA DLDQLRKLLS YVDDESFIRD VAKVKQENKL KFSAYLEKEY KVHINPNSLF 540
DIQVKRIHEY KRQLLNCLHA APGYHMAKMI IKLITAIGDV VNHDPVVGDR LRVIFLENYR 600
VSLAEKVIPA ADLSEQISTA GTEASGTGNM KFMLNGALTI GTMDGANVEM AEEAGEENFF 660
IFGMRVEDVE RLDQRGYNAQ EYYDRIPELR HVIDQLSSGF FSPKQPDLFK DIVNMLMHHD 720
RFKVFADYEE YVKCQERVSA LYKNPREWTR MVIRNIATSG KFSSDRTIAQ YAREIWGVEP 780
```

-continued

```
NRQRMPAPDE KI                                                      792

SEQ ID NO: 141          moltype = AA  length = 412
FEATURE                 Location/Qualifiers
source                  1..412
                        mol_type = protein
                        organism = Cervus elaphus
SEQUENCE: 141
MAQRKPEGSG FHMTCLSMAL AYSFRPDANI QPHPQLGNTQ KQTELGKELT GTMPHQYPAL  60
TPEQKKELCD IAHRIVAPGK GILAADESTG SIAKRLQSIG TENTEENRRF YRQLLLTADD  120
RVNPCIGGVI LFHETLYQKA DDGRPFPQVI KAKGGVVDKG VVPLAGTNGE TTTQGLDGLS  180
ERCAQYKKDG ADFAKWRCVL KIGEHTPSSL AIMENANVLA RYASICQQNG IVPIVEPEIL  240
PDGDHDLKRC QYVTEKVLAA VYKALSDHHI YLEGTLLKPN MVTPGHACTQ KYSHEEIAMA  300
TVTALRRTVP PAVPGITFLS GGQSEEEASI NLNAINKCPL LKPWALTFSY GRALQASALK  360
AWGGKKENLK AAQEEYVKRA LANSLACQGK YTPSGKAGAA ASESLFISNH AY          412

SEQ ID NO: 142          moltype = AA  length = 167
FEATURE                 Location/Qualifiers
source                  1..167
                        mol_type = protein
                        organism = Cervus elaphus
SEQUENCE: 142
PAPAPPKEEK IDLSAIKIEF SKQQQDEFKE AFLLFDRTGE CKITLNQVGD VLRALGTNPT  60
NAEVKKVLGN PSNEEMNAKK IEFEQFLPML QAISNNKDQG TYEDFVEGLR VFDKEGNGTV  120
MGAELRHVLA TLGEKMKEEE VEALMAGQED SNGCINYEAF VKHIMSN                167

SEQ ID NO: 143          moltype = AA  length = 206
FEATURE                 Location/Qualifiers
source                  1..206
                        mol_type = protein
                        organism = Cervus elaphus
SEQUENCE: 143
MFVMGVNHEK YNNTLKIVSN ASCTTNCLAP LAKVIHDHFG IVEGLMTTVH AITATQKTVD  60
GPSGKLWRDG RGAAQNIIPA STGAAKAVGK VIPELNGKLT GMAFRVPTPN VSVVDLTCRL  120
EKPAKYDEIK KVVKQASEGP LKGILGYTED QVVSCDFNSD THSSTFDAGA GIALNDHFVK  180
LISWYDNEFG YSNRVVDLMV HMASKE                                       206

SEQ ID NO: 144          moltype = AA  length = 428
FEATURE                 Location/Qualifiers
source                  1..428
                        mol_type = protein
                        organism = Cervus elaphus
SEQUENCE: 144
MRTLRAMAMQ KIFAREILDS RGNPTVEVDL HTAKGRFRAA VPSGASTGIY EALELRDGDK  60
SRYLGKGVLK AVEHINKTLG PALLEKASLG RLYLLPPAKF GANAILGVSL AVCKAGAAEK  120
GVPLYRHIAD LAGNPELILP VPAFNVINGG SHAGNKLAMQ EFMILPVGAS SFKEAMRIGA  180
EVYHHLKGVI KAKYGKDATN VGDEGGFAPN ILENNEALEL LKTAIQAAGY PDKVVIGMDV  240
AASEFYRNGK YDLDFKSPDD PARHISGEKL GELYKNFIKN YPVVSIEDPF DQDDWATWTS  300
FLSGVNIQIV GDDLTVTNPK RIAQAVEKKA CNCLLLKVNQ IGSVTESIQA CKLAQSNGWG  360
VMVSHRSGET EDTFIADLVV GLCTGQIKTG APCRSERLAK YNQLMRIEEA LGDKAVFAGR  420
KFRNPKAK                                                           428

SEQ ID NO: 145          moltype = AA  length = 181
FEATURE                 Location/Qualifiers
source                  1..181
                        mol_type = protein
                        organism = Cervus elaphus
SEQUENCE: 145
MFDQTQIQEF KEAPKKAKRR AAAEGGSSSV FSMFDQTQIQ EFKEAFTVID QNRDGIIDKE  60
DLRDTFAAMG RLNVKNEELD AMMKEASGPI NFTVFLNMFG EKLKGADPED VITGAFKVLD  120
PEGKGTIKKQ FLEELLTTQC DRFSQEEMRN MWAAFPPDVG GNVDYKNICY VITHGDAKDQ  180
E                                                                  181

SEQ ID NO: 146          moltype = AA  length = 340
FEATURE                 Location/Qualifiers
source                  1..340
                        mol_type = protein
                        organism = Cervus elaphus
SEQUENCE: 146
MCEEEDSTAL VCDNGSGLCK AGFAGDDAPR AVFPSIVGRP RHQGVMVGMG QKDSYVGDEA  60
QSKRGILTLK YPIEHGIITN WDDMEKIMFE TFNVPAMYVA IQAVLSLYAS GRTTGIVLDS  120
GDGVTHNVPI YEGYALPHAI MRLDLAGRDL TDYLMKILTE RGYSFVTTAE REIVRDIKEK  180
LCYVALDFEN EMATAASSSS LEKSYELPDG QVITIGNERF RCPETLFQPS FIGMESAGIH  240
ETTYNSIMKC DIDIRKDLYA NNVLSGGTTM YPGIADRMQK EITALAPSTM KIKIIAPPER  300
KYSVWIGGSI LASLSTFQQM WISKQEYDEA GPSIVHRKCF                        340
```

-continued

```
SEQ ID NO: 147          moltype = AA  length = 605
FEATURE                 Location/Qualifiers
source                  1..605
                        mol_type = protein
                        organism = Cervus elaphus
SEQUENCE: 147
MKWVTFISLL LLFSSAYSRG VFRRDTHKSE IAHRFKDLGE DNFQGLVLIA FSQYLQQCPF  60
DEHVKLVKEL TDFAKECVAD ESHAGCEKSL HTLFGDELCK VASLRETYGD MADCCEKQEP  120
ERNECFLKHK DDNPDLPKLK PDPDTLCAEF KADEKKFWGK YLYEVARRHP YFYAPELLYY  180
ANKYNGVFQE CCQADDKAAC LLPKIDTMRE KVLTSSARQR LKCASIQKFG ERALKAWTVA  240
LLSQKFPKAD FAEVTKIVTD LTKVHKECCH GDLLECADDR ADLAKYICDN QDTLSSKLKE  300
CCDKPVLEKS HCIAEIEKDA VPENLPPLTA DFAEDKEVCK NYQEAKDVFL GTFLYEYARR  360
HPEYAVSVLL RVAKEYEATL EDCCAKDDPH ACYATVFDKL QHLVDEPQNL IKQNCDLFEK  420
VGEYGFQNAL IVRYTKKAPQ VSTPTLVEVS RNLGRVGTKC CTKPESERMS CTEDYLSLVL  480
NRLCVLHEKT PVSEKVTKCC TESLVNRRPC FSALTLDETY VPKAFDEKLF TFHADICTVP  540
DAEKQIKKQT ALVELLKHKP KATDEQLKTV MENFVAFVDK CCAAVDKEAC FAVEPTLRVR  600
GKKEK                                                              605

SEQ ID NO: 148          moltype = AA  length = 381
FEATURE                 Location/Qualifiers
source                  1..381
                        mol_type = protein
                        organism = Gallus gallus
SEQUENCE: 148
MPFSSTHNKH KLKFSAEEEF PDLSKHNNHM AKVLTPELYK RLRDKETPSG FTLDDVIQTG  60
VDNPGHPFIM TVGCVAGDEE SYEVFKDLFD PVIQDRHGGY KPTDKHRTDL NHENLKGGDD  120
LDPKYVLSSR VRTGRSIKGY SLPPHCSRGE RRAVEKLSVE ALNSLEGEFK GRYYPLKAMT  180
EQEQQQLIDD HFLFDKPVSP LLLASGMARD WPDARGIWHN DNKTFLVWVN EEDHLRVISM  240
EKGGNMKEVF RRFCVGLKKI EEIFKKAGHP FMWTEHLGYI LTCPSNLGTG LRGGVHVKLP  300
KLSQHPKFEE ILHRLRLQKR GTGGVDTAAV GAVFDISNAD RLGFSEVEQV QMVVDGVKLM  360
VEMEKKLEQN QPIDDMIPAQ K                                            381

SEQ ID NO: 149          moltype = AA  length = 897
FEATURE                 Location/Qualifiers
source                  1..897
                        mol_type = protein
                        organism = Gallus gallus
SEQUENCE: 149
MNSMNQIETN MQYTYNYEED EYMTQEEEWD RDLLLDPAWE KQQRKTFTAW CNSHLRKAGT  60
QIENIEEDFR NGLKLMLLLE VISGERLPKP DRGKMRFHKI ANVNKALDYI ASKGVKLVSI  120
GAEEIVDGNV KMTLGMIWTI ILRFAIQDIS VEETSAKEGL LLWCQRKTAP YRNVNIQNFH  180
LSWKDGLGLC ALIHRHRPDL IDYSKLNKDD PIGNINLAME IAEKHLDIPK MLDAEDIVNT  240
PKPDERAIMT YVSCFYHAFA GAEQAETAAN RICKVLAVNQ ENERLMEEYE RLASELLEWI  300
RRTIPWLENR TPEKTMQAMQ KKLEDFRDYR RKHKPPKVQE KCQLEINFNT LQTKLRISNR  360
PAFMPSEGKM VSDIAGAWQR LEQAEKGYEE WLLNEIRRLE RLEHLAEKFR QKASTHEQWA  420
YGKEQILLQK DYESASLTEV RAMLRKHEAF ESDLAAHQDR VEQIAAIAQE LNELDYHDAA  480
SVNDRCQKIC DQWDSLGTLT QKRREALERT EKLLETIDQL HLEFAKRAAP FNNWMEGAME  540
DLQDMFIVHS IEEIQSLISA HDQFKATLPE ADGERQAILS IQNEVEKVIQ SYSMRISASN  600
PYSTVTVEEI RTKWEKVKQL VPQRDQSLQE ELARQHANER LRRQFAAQAN VIGPWIQTKM  660
EEIARSSIEM TGPLEDQMNQ LKQYEQNIIN YKHNIDKLEG DHQLIQEALV FDNKHTNYTM  720
EHIRVGWELL LTTIARTINE VETQILTRDA KGITQEQMND FRASFNHFDR RKNGLMDHDD  780
FRACLISMGY DLGEAEFARI MSLVDPNGQG TVTFQSFIDF MTRETADTDT AEQVIASFRI  840
LASDKPYILA DELRRELPPE QAQYCIKRMP QYTGPGSVPG ALDYTSFSSA LYGESDL     897

SEQ ID NO: 150          moltype = AA  length = 843
FEATURE                 Location/Qualifiers
source                  1..843
                        mol_type = protein
                        organism = Gallus gallus
SEQUENCE: 150
MAAPLSDGER RKQISVRGIA GLGDVAEVRK SFNRHLHFTL VKDRNVATPR DYFFALAHTV  60
RDHLVGRWIR TQQHYYERDP KRIYYLSLEF YMGRTLQNTM VNLGLQNACD EAIYQLGLDL  120
EELEEIEEDA GLGNGGLGRL AACFLDSMAT LGLAAYGYGI RYEFGIFNQK IIDGWQVEEA  180
DDWLRYGNPW EKARPEYMLP VHFYGRVEHS PEGVKWVDTQ VVLAMPYDTP VPGYKNNTVN  240
TMRLWSAKAP NDFNLQEFNV GDYIEAVLDR NLAENISRVL YPNDNFFEGK ELRLKQEYFV  300
VAATLQDIIR RFKSSKFGCQ DPVRTSFETF PDKVAIQLND THPALSIPEL MRILVDVEKV  360
DWDKAWEITK QTCAYTNHTV LPEALERWPV SMFEKLLPRH LEIIYALNQM HLDRVAALYP  420
GDVDRLRRMS VIEEGDCKRI NMAHLCVIGS HAVNGVARIH SDIVKNTVFK DFYELEPEKF  480
QNKTNGITPR RWLLLCNPGL ADVIAEKIGE DFVTDLSQLK KLLDFINNET FIRDVAKVKQ  540
ENKLKFSAYL EERYKVKINP SSMFDIQVKR IHEYKRQLLN CLHAITLYNR IRSDPSKSFV  600
PRTIMIGGKA APGYHMAKMI IKLITSIGEV VNNDPYVGDK LKVIFLENYR VSLAEKVIPA  660
ADLSQQISTA GTEASGTGNM KFMLNGALTI GTMDGANVEM AEEAGEENLF IFGMRVEEVE  720
ALDRRGYNAR EYYDRIPELR QAIDQISSGF FSPRDPGCFR DVVNMLMHHD RFKVFADYEA  780
YIKCQGQVDQ LFMDPREWTK KVIRNIACSG KFSSDRTITE YAREIWGVEP SATKIPPPNL  840
PRD                                                                843
```

-continued

```
SEQ ID NO: 151          moltype = AA  length = 364
FEATURE                 Location/Qualifiers
source                  1..364
                        mol_type = protein
                        organism = Gallus gallus
SEQUENCE: 151
MTHQYPALTA EQKKELSDIA LRIVAPGKGI LAADESVGSM AKRLNQIGVE NTEENRRLYR  60
QILFSADSRV KKCIGGVIFF HETMYQKADD GTPFVQMIKD KGIVVGIKVD KGVVPLAGTD  120
GETTTQGLDG LSERCAQYKK DGADFAKWRC VLKISDNTPS ALAIMENANV LARYASICQQ  180
NGIVPIVEPE ILPDGDHDLK RCQYVTEKVL AAVYKALSDH HVYLEGTLLK PNMVTPGHSC  240
PTKYSPEEIA MATVTALRRT VPPAVPGVTF LSGGQSEEEA SINLNAINTC PLVRPWALTF  300
SYGRALQASA LSAWRGQRDN ANAATEEFVK RAEVNGLAAL GKYEGSGDDS GAAGQSLYVA  360
NHAY                                                               364

SEQ ID NO: 152          moltype = AA  length = 192
FEATURE                 Location/Qualifiers
source                  1..192
                        mol_type = protein
                        organism = Gallus gallus
SEQUENCE: 152
MAPKKDVKKP AAAAAPAPAP APAPAPAPAK PKEPAIDLKS IKIEFSKEQQ DDFKEAFLLF  60
DRTGDAKITL SQVGDIVRAL GQNPTNAEIN KILGNPSKEE MNAKKITFEE FLPMLQAAAN  120
NKDQGTFEDF VEGLRVFDKE GNGTVMGAEL RHVLATLGEK MTEEEVEELM KGQEDSNGCI  180
NYEAFVKHIM SV                                                      192

SEQ ID NO: 153          moltype = AA  length = 333
FEATURE                 Location/Qualifiers
source                  1..333
                        mol_type = protein
                        organism = Gallus gallus
SEQUENCE: 153
MVKVGVNGFG RIGRLVTRAA VLSGKVQVVA INDPFIDLNY MVYMFKYDST HGHFKGTVKA  60
ENGKLVINGH AITIFQERDP SNIKWADAGA EYVVESTGVF TTMEKAGAHL KGGAKRVIIS  120
APSADAPMFV MGVNHEKYDK SLKIVSNASC TTNCLAPLAK VIHDNFGIVE GLMTTVHAIT  180
ATQKTVDGPS GKLWRDGRGA AQNIIPASTG AAKAVGKVIP ELNGKLTGMA FRVPTPNVSV  240
VDLTCRLEKP AKYDDIKRVV KAAADGPLKG ILGYTEDQVV SCDFNGDSHS STFDAGAGIA  300
LNDHFVKLVS WYDNEFGYSN RVVDLMVHMA SKE                               333

SEQ ID NO: 154          moltype = AA  length = 284
FEATURE                 Location/Qualifiers
source                  1..284
                        mol_type = protein
                        organism = Gallus gallus
SEQUENCE: 154
MDAIKKKMQM LKLDKENALD RAEQAEADKK AAEERSKQLE DELVALQKKL KGTEDELDKY  60
SESLKDAQEK LELADKKATD AESEVASLNR RIQLVEEELD RAQERLATAL QKLEEAEKAA  120
DESERGMKVI ENRAQKDEEK MEIQEIQLKE AKHIAEEADR KYEEVARKLV IIEGDLERAE  180
ERAELSESKC AELEEELKTV TNNLKSLEAQ AEKYSQKEDK YEEEIKVLTD KLKEAETRAE  240
FAERSVTKLE KSIDDLEDEL YAQKLKYKAI SEELDHALND MTSI                   284

SEQ ID NO: 155          moltype = AA  length = 1132
FEATURE                 Location/Qualifiers
source                  1..1132
                        mol_type = protein
                        organism = Gallus gallus
SEQUENCE: 155
MPEPSKAAPK KEAKKKEEKK EEKKEAPPPQ EHKDEAPDDV HPPETPDPEG LFLSKPQNVM  60
VESGRDVTVS ARVAGAALPC APAVKWFKGK WAELGDKSAR CRLRHSVDDD KVHTFELTIT  120
KVAMGDRGDY RCEVTAKEQK DSCSFSIDVE APRSSEGNVL QAFKRTGEGK DDTAGELDFS  180
GLLKKREVQV EEKKKKKDED DQFPPEIWEL LKGVTKKSEY ERIAFQYGIT DLRGMLKRLK  240
KVHVEPKKSE AFIRKLDPAY QVDKGNKIKL VVELSDPDLP LKWYKNGQLL KPSTKYVFEN  300
VGLKRILTIH KCSLADDAAY ECRVNDEKCF TEVFVKEPPV TVVRGLEDQQ VVVGDRVVLE  360
AEVSEEGAQV MWLKDGVDVT RDDAFKYRFK KDGKKHFLII NEAELSDSAH YKIMTNGGES  420
EAELSVEEKQ LEVLQDMADL TVKASEQAVF KCEVSDEKVT GRWFRNGVEV KPSKRIHISH  480
NGRFHKLVID DVRPEDEGDY TFIPDGYALS LSAKLNFLEI KVEYVPKQEP PKIHLDCSGK  540
AAENTIVVVA GNKVRLDVPI SGEPAPTVTW KRGDQLFTAT EGRVHIDSQA DLSSFVIESA  600
ERSDEGRYCI TVTNPVGEDS ATLHVRVVDV PDPPQSVRVT SVGEDWAVLS WEAPPFDGGM  660
PITGYLMERK KKGSMRWMKL NFEVFPDTTY ESTKMIEGVF YEMRVFAVNA IGVSQPSLNT  720
QPFMPIAPTS EPTHVVLEDV TDTTATIKWR PPERIGAGGV DGYLVEWCRE GSNEWVAANT  780
ELVERCGLTA RGLPTGERLL FRVISVNMAG KSPPATMAQP VTIREIVERP KIRLPRHLRQ  840
TYIRRVGEQV NLVIPFQGKP RPQVTWSREG GALPAEVQTR TSDVDSVFFI RSAARPLSGN  900
YEMRVRIDNM EDCATLRLRV VERPGPPQAV RVMEVWGSNA LLQWEPPKDD GNAEISGYTV  960
QKADTRTMEW FTVLEHSRPT RCTVSELVMG NEYRFRVYSE NVCGTSQEPA TSHNTARIAK  1020
EGLTLKMVPY KERDLRAAPQ FLTPLVDRSV VAGYTVTLNC AVRGHPKPKV TWLKNSVEIG  1080
ADPKFLSRHG LGVLSLLIRR PGPFDGGTYG CRAVNEMGEA TTECRLDVRV PQ          1132
```

-continued

```
SEQ ID NO: 156          moltype = AA  length = 1450
FEATURE                 Location/Qualifiers
source                  1..1450
                        mol_type = protein
                        organism = Gallus gallus
SEQUENCE: 156
MSSVAVPFYQ RRHKHFDQSY RNIQTRYVLE EYAARKAASR QAAHYESTGL GKTTCRLCAR  60
RARSLAHEAM QESRKRTHEQ KSHASDEKRI KFASELSSLE REIHMARHHA REQLDRLAIQ  120
RMVEENMALE RHVVEEKISR APEILVRLRS HTVWEKMSVR LCFTVQGFPS PVVQWYKNEE  180
LITPASDPAK YSVENKYGVH VLHINRADFD DSATYSAVAT NIHGQASTNC AVVVRRFRES  240
EEPHPAGIMP FHLPLSYDVC FTHFDVQFLE KFGVTFATEG ETLTLKCSVL VTPELKRLRP  300
RAEWYRDDVL IKDSKWTKLY FGEGQAALSF THLNKDDEGL YTLRMVTKGG VNECSAFLFV  360
RDADALIAGA PGAPMDVKCH DANRDYVIVT WKPPNTTSQN PVIGYFVDKC EVGLENWVQC  420
NDAPVKICKY PVTGLYEGRS YIFRVRAVNS AGISRPSRVS EPVAALDPVD LERTQTVHVD  480
EGRKIVISKD DLEGDIQIPG PPTNVHASEI SKTYVVLSWD PPVPRGREPL TYFIEKSMVG  540
SGSWQRVNAQ VAVKSPRYAV FDLAEGKPYV FRVLSANKHG ISDPSEITEP IQPQDIVVVP  600
SAPGRVVATR NTKTSVVVQW DKPKHEENLY GYYIDYSVVG SNQWEPANHK PINYNRFVVH  660
GLETGEQYIF RVKAVNAVGF SENSQESEAI KVQAALTCPS YPHGITLLNC DGHSMTLGWK  720
APKYSGGSPI LGYYIDKREA NHKNWHEVNS SVISRTIYTV EDLTEDAFYE FKIAAANVVG  780
IGHPSDPSEH FKCKAWTMPE PGPAYDLTVC EVRNTSLVLL WKAPVYEGKS PITGYLVDYK  840
EVDTEDWITA NEKPTSHRYF KVTDLHQGHT YVFKVRAVND AGVGKSSEIS EPVFVEASPG  900
TKEIFSGVDE EGNIYLGFEC KEATDASHFL WGKSYEEIED SDKFKIETKG DHSKLYFKHP  960
DKSDLGTYCI SVSDTDGVSS SFVLDEEELE RLMTLSNEIK NPTIPLKSEL AYEVLDKGEV  1020
RFWIQAESLS PNSTYRFVIN DKEVENGDRH KISCDHSNGI IEMVMDKFTI DNEGTYTVQI  1080
QDGKAKNQSS LVLIGDAFKA ILAESELQRK EFLRKQGPHF SEFLYWEVTE ECEVLLACKI  1140
ANTKKETVFK WYRNGSGIDV DEAPDLQKGE CHLTVPKLSR KDEGVYKATL SDDRGHDVST  1200
LELSGKVYND IILALSRVSG KTASPLKILC TEEGIRLQCF LKYYNEEMKV TWSHRESKIS  1260
SGEKMKIGGG EDVAWLQITE PTEKDKGNYT FEIFSDKESF KRTLDLSGQA FDDALTEFQR  1320
LKAAAFAEKN RGKVIGGLPD VVTIMDGKTL NLTCTVFGNP DPEVVWFKND KALELNEHYL  1380
VSLEQGKYAS LTIKGVTSED SGKYSIYVKN KYGGETVDVT VSVYRHGEKI PEVNQGQLAK  1440
PRLIPPSSST                                                         1450

SEQ ID NO: 157          moltype = AA  length = 434
FEATURE                 Location/Qualifiers
source                  1..434
                        mol_type = protein
                        organism = Gallus gallus
SEQUENCE: 157
MSIQKIHARE ILDSRGEPTV EVDLHTAKGH FRAAVPSGAS TGIHEALEPR DGDKKRFLGK  60
GVLKAVEHIN KTIGPALIEK KISVVEQEKI DKVMIEMDGT ENKSKFGANA ILGVSLAVCK  120
AGAAEKGVPL YRHIADLAGN TELILPVPAF NVINGGSHAG NKLAMQEFMV LPVGAASFHD  180
AMRVGAEVYH SLKGVIKAKY GKDATNVGGE GGFAPNILDN HEALELLKAA IAQAGYTDKV  240
VIGMDVAASE FCRDGRYHLD FKSPPHTKRY ITGEQLGEIY RGFIKDYPVV SIEDPFDQDD  300
WEAWKRFVFH VDIQVVGDDL TVTNPKRIAH GAEQHACNCL LLKVNQIGSV TESIQACKLA  360
QSHGWGVMVS HRSGETEDTF IADLVVGLCT GQIKTGAPCR SERLAKYNQL MRIEEALGDK  420
AKFAGRKFRN PKAK                                                    434

SEQ ID NO: 158          moltype = AA  length = 168
FEATURE                 Location/Qualifiers
source                  1..168
                        mol_type = protein
                        organism = Gallus gallus
SEQUENCE: 158
MAPKKAKRRA AEGSSNVFSM FDQTQIQEFK EAFTVIDQNR DGIIDKDDLR ETFAAMGRLN  60
VKNEELDAMI KEASGPINFT VFLTMFGEKL KGADPEDVIM GAFKVLDPDG KGSIKKSFLE  120
ELLTTQCDRF TPEEIKNMWA AFPPDVAGNV DYKNICYVIT HGEDKEGE               168

SEQ ID NO: 159          moltype = AA  length = 1940
FEATURE                 Location/Qualifiers
source                  1..1940
                        mol_type = protein
                        organism = Gallus gallus
SEQUENCE: 159
MATDADMAIF GEAAPYLRKS EKERIEAQNK PFDAKSSVFV VHAKESYVKS TIQSKESGKV  60
TVKTEGGETL TVKEDQIFSM NPPKYDKIED MAMMTHLHEP AVLYNLKERY AAWMIYTYSG  120
LFCVTVNPYK WLPVYNPEVV LAYRGKKRQE APPHIFSISD NAYQFMLTDR ENQSILITGE  180
SGAGKTVNTK RVIQYFATIA ASGDKKKEEQ PAGKMQGTLE DQIISANPLL EAFGNAKTVR  240
NDNSSRFGKF IRIHFGATGK LASADIETYL LEKSRVTFQL KAERSYHIFY QIMSNKKPEL  300
IEMLLITTNP YDYQYVSQGE ITVPSINDQE ELMATDSAID ILGFTPDEKT AIYKLTGAVM  360
HYGNLKFKQK QREEQAEPGG TEVADKAAYL MGLNSADLLK ALCYPRVKVG NEYVTKGQTV  420
QQVYNSVGAL AKSVFEKMFL WMVVRINQQL DTKQPRQYFI GVLDIAGFEI FDFNSLEQLC  480
INFTNEKLQQ FFNHHMFVLE QEEYKKEGIE WEFIDFGMDL AACIELIEKP MGIFSILEEE  540
CMFPKATDTS FKNKLYDQHL GKSNNFQKPK PGKGKAEAHF SLVHYAGTVD YNITGWLEKN  600
KDPLNETVVG LYQKSSLKTL ALLFASVGGA EAESGAGGKK GGKKKGSSFQ TVSALFRENL  660
NKLMSNLRST HPHFVRCLIP NETKTPGAME HELVLHQLRC NGVLEGIRIC RKGFPIRILY  720
ADFKQRYKVL NASAIPEGQF IDSKKASEKL LGSIDVDHTQ YKFGHTKVFF KAGLLGLLEE  780
```

-continued

```
MRDEKLAQLI TRTQARCRGF LMRVEFKKMM ERRESIFCIQ YNVRAFMNVK HWPWMKLFFK  840
IKPLLKSAES EKEMANMKEE FEKTKEELAK SEAKRKELEE KMVSLLQEKN DLQLQVQAEA  900
DGLADAEERC DQLIKTKIQL EAKIKELTER AEDEEEMNAE LTAKKRKLED ECSELKKDID  960
DLELTLAKVE KEKHATENKV KNLTEEMAAL DETIAKLTKE KKALQEAHQQ TLDDLQAEED  1020
KVNTLTKAKT KLEQQVDDLE GSLEQEKKLR MDLERAKRKL EGDLKMTQES TMDLENDKQQ  1080
LDEKLKKKDF EISQIQSKIE DEQALGMQLQ KKIKELQARI EELEEEIEAE RTSRAKAEKH  1140
RADLSRELEE ISERLEEAGG ATAAQIDMNK KREAEFQKMR RDLEEATLQH EATAAALRKK  1200
HADSTADVGE QIDNLQRVKQ KLEKEKSELK MEIDDLASNM ESVSKAKANL EKMCRSLEDQ  1260
LSEIKTKEEE QQRTINDISA QKARLQTESG EYSRQVEEKD ALISQLSRGK QAFTQQIEEL  1320
KRHLEEEIKA KKCPAHALQS ARHDCDLLRE QYEEEQEAKG ELQRALSKAN SEVAQWRTKY  1380
ETDAIQRTEE LEEAKKKLAQ RLQDAEEHVE AVNSKCASLE KTKQRLQNEV EDLMIDVERS  1440
NAACAALDKK QKNFDKILSE WKQKYEETQA ELEASQKESR SLSTELFKMK NAYEESLDHL  1500
ETLKRENKNL QQEISDLTEQ IAEGGKAIHE LEKVKKQIEQ EKSELQTALE EAEASLEHEE  1560
GKILRVQLEL NQVKSDIDRK IAEKDEEIDQ LKRNHLRVVD SMQSTLDAEI RSRNEALRLK  1620
KKMEGDLNEI EIQLSHANRQ AAEAQKNLRN TQGVLKDTQI HLDDALRSQE DLKEQVAMVE  1680
RRANLLQAEI EELRAALEQT ERSRKVAEQE LLDASERVQL LHTQNTSLIN TKKKLESDIS  1740
QIQSEMEDTI QEARNAEEKA KKAITDAAMM AEELKKEQDT SAHLERMKKN LDQTVKDLQH  1800
RLDEAEQLAL KGGKKQIQKL EARVRELEGE VDAEQKRSAE AVKGVRKYER RVKELTYQSE  1860
EDRKNVLRLQ DLVDKLQMKV KSYKRQAEEA EELSNVNLSK FRKIQHELEE AEERADIAES  1920
QVNKLRAKSR EIGKKAESEE                                              1940

SEQ ID NO: 160          moltype = AA  length = 1940
FEATURE                 Location/Qualifiers
source                  1..1940
                        mol_type = protein
                        organism = Gallus gallus
SEQUENCE: 160
MATDADMAIF GEAAPYLRKS EKERIEAQNK PFDAKSSVFV VHAKESYVKS TIQSKESGKV  60
TVKTEGGETL TVKEDQIFSM NPPKYDKIED MAMMTHLHEP AVLYNLKERY AAWMIYTYSG  120
LFCVTVNPYK WLPVYNPEVV LAYRGKKRQE APPHIFSISD NAYQFMLTDR ENQSILITGE  180
SGAGKTVNTK RVIQYFATIA ASGDKKKEEQ PAGKMQGTLE DQIISANPLL EAFGNAKTVR  240
NDNSSRFGKF IRIHFGATGK LASADIETYL LEKSRVTFQL KAERSYHIFY QIMSNKKPEL  300
IEMLLITTNP YDYQYVSQGE ITVPSINDQE ELMATDSAID ILGFTPDEKT AIYKLTGAVM  360
HYGNLKFKQK QREEQAEPDG TEVADKAAYL MGLNSADLLK ALCYPRVKVG NEYVTKGQTV  420
QQVYNSVGAL AKSVFEKMFL WMVVRINQQL DTKQPRQYFI GVLDIAGFEI FDFNSLEQLC  480
INFTNEKLQQ FFNHHMFVLE QEEYKKEGIE WEFIDFGMDL AACIELIEKP MGIFSILEEE  540
CMFPKATDTS FKNKLYDQHL GKSNNFQKPK PGKGKAEAHF SLVHYAGTVD YNITGWLEKN  600
KDPLNETVVG LYQKSSLKTL ALLFASVGGA EAESGAGGKK GGKKKGSSFQ TVSALFRENL  660
NKLMSNLRST HPHFVRCLIP NETKTPGAME HELVLHQLRC NGVLEGIRIC RKGFPSRILY  720
ADFKQRYKVL NASAIPEGQF IDSKKASEKL LGSIDVDHTQ YKFGHTKVFF KAGLLGLLEE  780
MRDEKLAQLI TRTQARCRGF LMRVEFKKMM ERRESIFCIQ YNVRAFMNVK HWPWMKLFFK  840
IKPLLKSAES EKEMANMKEE FEKTKEELAK SEAKRKELEE KMVSLLQEKN DLQLQVQAEA  900
DGLADAEERC DQLIKTKIQL EAKIKELTER AEDEEEMNAE LTAKKRKLED ECSELKKDID  960
DLELTLAKVE KEKHATENKV KNLTEEMAAL DETIAKLTKE KKALQEAHQQ TLDDLQAEED  1020
KVNTLTKAKT KLEQQVDDLE GSLEQEKKLR MDLERAKRKL EGDLKMTQES TMDLENDKQQ  1080
LDEKLKKKDF EISQIQSKIE DEQALGMQLQ KKIKELQARI EELEEEIEAE RTSRAKAEKH  1140
RADLSRELEE ISERLEEAGG ATAAQIDMNK KREAEFQKMR RDLEEATLQH EATAAALRKK  1200
HADSTAELGE QIDNLQRVKQ KLEKEKSELK MEIDDLASNM ESVSKAKANL EKMCRSLEDQ  1260
LSEIKTKEEE QQRTINDISA QKARLQTESG EYSRQVEEKD ALISQLSRGK QAFTQQIEEL  1320
KRHLEEEIKA KNALAHALQS ARHDCDLLRE QYEEEQEAKG ELQRALSKAN SEVAQWRTKY  1380
ETDAIQRTEE LEEAKKKLAQ RLQDAEEHVE AVNSKCASLE KTKQRLQNEV EDLMIDVERS  1440
NAACAALDKK QKNFDKILSE WKQKYEETQA ELEASQKESR SLSTELFKMK NAYEESLDHL  1500
ETLKRENKNL QQEISDLTEQ IAEGGKAIHE LEKVKKQIEQ EKSELQAALE EAEASLEHEE  1560
GKILRVQLEL NQVKSDIDRK IAEKDEEIDQ LKRNHLRVVD SMQSTLDAEI RSRNEALRLK  1620
KKMEGDLNEI EIQLSHANRQ AAEAQKNLRN TQGVLKDTQI HLDDALRSQE DLKEQVAMVE  1680
RRANLLQAEI EELRAALEQT ERSRKVAEQE LLDASERVQL LHTQNTSLIN TKKKLESDIS  1740
QIQSEMEDTI QEARNAEEKA KKAITDAAMM AEELKKEQDT SAHLERMKKN LDQTVKDLQH  1800
RLDEAEQLAL KGGKKQIQKL EARVRELEGE VDAEQKRSAE AVKGVRKYER RVKELTYQSE  1860
EDRKNVLRLQ DLVDKLQMKV KSYKRQAEEA EELSNVNLSK FRKIQHELEE AEERADIAES  1920
QVNKLRAKSR EIGKKAESEE                                              1940

SEQ ID NO: 161          moltype = AA  length = 615
FEATURE                 Location/Qualifiers
source                  1..615
                        mol_type = protein
                        organism = Gallus gallus
SEQUENCE: 161
MKWVTLISFI FLFSSATSRN LQRFARDAEH KSEIAHRYND LKEETFKAVA MITFAQYLQR  60
CSYEGLSKLV KDVVDLAQKC VANEDAPECS KPLPSIILDE ICQVEKLRDS YGAMADCCSK  120
ADPERNECFL SFKVSQPDFV QPYQRPASDV ICQEYQDNRV SFLGHFIYSV ARRHPFLYAP  180
AILSFAVDFE HALQSCCKES DVGACLDTKE IVMREKAKGV SVKQQYFCGI LKQFGDRVFQ  240
ARQLIYLSQK YPKAPFSEVS KFVHDSIGVH KECCEGDMVE CMDDMARMMS NLCSQQDVFS  300
GKIKDCCEKP IVERSQCIME AEFDEKPADL PSLVEKYIED KEVCKSFEAG HDAFMAEFVY  360
EYSRRHPEFS IQLIMRIAKG YESLLEKCCK TDNPAECYAN AQEQLNQHIK ETQDVVKTNC  420
DLLHDHGEAD FLKSILIRYT KKMPQVPTDL LLETGKKMTT IGTKCCQLGE DRRMACSEGY  480
LSIVIHDTCR KQETTPINDN VSQCCSQLYA NRRPCFTAMG VDTKYVPPPF NPDMFSFDEK  540
```

-continued

```
LCSAPAEERE VGQMKLLINL IKRKPQMTEE QIKTIADGFT AMVDKCCKQS DINTCFGEEG  600
ANLIVQSRAT LGIGA                                                   615

SEQ ID NO: 162          moltype = AA  length = 380
FEATURE                 Location/Qualifiers
source                  1..380
                        mol_type = protein
                        organism = Meleagris gallopavo
SEQUENCE: 162
MPFSNSHNLL KMKHSADDEF PDLSAHNNHM AKVLTLDLYK KLRDRQTSSG FTLDDVIQTG  60
VDNPGHPFIM TVGCVAGDEE SYEVFKELFD PVIEDRHGGY KPTDEHKTDL NADNLQGGDD  120
LDPNYVLSSR VRTGRSIRGF CFPHCSRGER RAIEKLSVEA LGSLGGDLKG KYYALRNMTD  180
AEQQQLIDDH FLFDKPVSPL LLASGMARDW PDARGIWHND NKTFLVWINE EDHLRVISMQ  240
KGGNMKEVFT RFCTGLTQIE TLFKSKNYEF MWNPHLGYIL TCPSNLGTGL RAGVHIKLPN  300
LGKHEKFGEV LKRLRLQKRG TGGVDTAAVG GVFDVSNADR LGFSEVELVQ MVVDGVKLLI  360
EMEKRLEKGQ SIDDLMPAQK                                              380

SEQ ID NO: 163          moltype = AA  length = 897
FEATURE                 Location/Qualifiers
source                  1..897
                        mol_type = protein
                        organism = Meleagris gallopavo
SEQUENCE: 163
MNSMNQIETN MQYTYNYEED EYMTQEEEWD RDLLLDPAWE KQQRKTFTAW CNSHLRKAGT  60
QIENIEEDFR NGLKLMLLLE VISGERLPKP DRGKMRFHKI ANVNKALDYI ASKGVKLVSI  120
GAEEIVDGNV KMTLGMIWTI ILRFAIQDIS VEETSAKEGL LLWCQRKTAP YRNVNIQNFH  180
LSWKDGLGLC ALIHRHRPDL IDYSKLNKDD PIGNINLAME IAEKHLDIPK MLDAEDIVNT  240
PKPDERAIMT YVSCFYHAFA GAEQAETAAN RICKVLAVNQ ENERLMEEYE RLASELLEWI  300
RRTIPWLENR TPEKTMQAMQ KKLEDFRDYR RKHKPPKVQE KCQLEINFNT LQTKLRISNR  360
PAFMPSEGKM VSDIAGAWQR LEQAEKGYEE WLLNEIRRLE RLEHLAEKFR QKASTHEQWA  420
YGKEQILLQK DYESASLTEV RAMLRKHEAF ESDLAAHQDR VEQIAAIAQE LNELDYHDAA  480
SVNDRCQKIC DQWDSLGTLT QKRREALERT EKLLETIDQL HLEFAKRAAP FNNWMEGAME  540
DLQDMFIVHS IEEIQSLISA HDQFKATLPE ADGERQAILS IQNEVEKVIQ SYSMRISASN  600
PYSTVTVEEI RTKWEKVKQL VPQRDQSLQE ELARQHANER LRRQFAAQAN VIGPWIQTKM  660
EEIARSSIEM TGPLEDQMNQ LKQYEQNIIN YKHNIDKLEG DHQLIQEALV FDNKHTNYTM  720
EHIRVGWELL LTTIARTINE VETQILTRDA KGITQEQMND FRASFNHFDR RKNGLMDHDD  780
FRACLISMGY DLGEAEFARI MSLVDPNGQG TVTFQSFIDF MTRETADTDT AEQVIASFRI  840
LASDKPYILA DELRRELPPE QAQYCIKRMP QYTGPGSVPG ALDYTSFSSA LYGESDL     897

SEQ ID NO: 164          moltype = AA  length = 839
FEATURE                 Location/Qualifiers
source                  1..839
                        mol_type = protein
                        organism = Meleagris gallopavo
SEQUENCE: 164
MAAPLSDGER RKQISVRGIA GLGDVAEVRK SFNRHLHFTL VKDRNVATPR DYFFALAHTV  60
RDHLVGRWIR TQQHYYERDP KRIYYLSLEF YMGRTLQNTM VNLGLQNACD EAIYQLGLDL  120
EELEEIEEDA GLGNGGLGRL AACFLDSMAT LGLAAYGYGI RYEFGIFNQK IVDGWQVEEA  180
DDWLRYGNPW EKARPEYMLP VHFYGRVEHS PEGVKWVDTQ VVLAMPYDTP VPGYKNNTVN  240
TMRLWSAKAP NDFNLQEFNV GDYIEAVLDR NLAENISRVL YPNDNFFEGK ELRLKQEYFV  300
VAATLQDIIR RFKSSKFGCQ DPVRTSFETF PDKVAIQLND THPALSIPEL MRILVDVEKV  360
DWDKAWEITK QTCAYTNHTV LPEALERWPV SMFEKLLPRH LEIIYALNQM HLDRVAALYP  420
GDIDRLRRMS VIEEGDCKRI NMAHLCVIGS HAVNGVARIH SDIVKNTVFK DFYELDPEKF  480
QNKTNGITPR RWLLLCNPGL ADVIAEKIGE DFVTDLSQLK KLLDFIDDET FIRDVAKVKQ  540
ENKLKFSAYL EEHYKVKINP SSMFDIQVKR IHEYKRQLLN CLHAITLYNR KSDPSKSFVP  600
RTIMIGGKAA PGYHMAKMII KLITSIGEVV NNDPYVGDKL KVIFLENYRV SLAEKVIPAA  660
DLSQQISTAG TEASGTGNMK FMLNGALTIG TMDGANVEMA EEAGEENLFI FGMRVEEVEA  720
LDRQGYNARE YYDRIPELRQ AIDQISSGFF SPRDPGCFRD VVNMLMHHDR FKVFADYEAY  780
IKCQGQVDQL FMDPREWTKK VIRNIACSGK FSSDRTITEY AREIWGVEPS GDVFPAWEA   839

SEQ ID NO: 165          moltype = AA  length = 364
FEATURE                 Location/Qualifiers
source                  1..364
                        mol_type = protein
                        organism = Meleagris gallopavo
SEQUENCE: 165
MTHQYPALTA EQKKELSDIA LRIVAPGKGI LAADESVGSM AKRLNQIGVE NTEENRRLYR  60
QILFSADSRV KKCIGGVIFF HETMYQKADD GTPFVQMIKD KGIVVGIKVD KGVVPLAGTD  120
GETTTQGLDG LSERCAQYKK DGADFAKWRC VLKISDNTPS ALAIMENANV LARYASICQQ  180
NGIVPIVEPE ILPDGDHDLK RCQYVTEKVL AAVYKALSDH HVYLEGTLLK PNMVTPGHSC  240
PTKYSPEEIA MATVTALRRT VPPAVPGVTF LSGGQSEEEA SINLNAINTC PLVRPWALTF  300
SYGRALQASA LSAWRGQRDN ANAATEEFVK RAEVNGLAAL GKYEGSGDDS GAAGQSLYVA  360
NHAY                                                               364

SEQ ID NO: 166          moltype = AA  length = 192
FEATURE                 Location/Qualifiers
```

-continued

```
source                  1..192
                        mol_type = protein
                        organism = Meleagris gallopavo
SEQUENCE: 166
MAPKKDVKKP AAAAAPAPAP APAPAPAPAK PKEPAIDLKS IKIEFSKEQQ DDFKEAFLLF  60
DKTGDAKITL SQVGDIVRAL GQNPTNAEMN KILGNPSKEE MNAKKITFEE FLPMLQAAAN  120
NKDQGTYEDF VEGLRVFDKE GNGTVMGAEL RHVLATLGEK MTEEEVEELM KGQEDSNGCI  180
NYEAFVKHIM SV                                                      192

SEQ ID NO: 167          moltype = AA  length = 333
FEATURE                 Location/Qualifiers
source                  1..333
                        mol_type = protein
                        organism = Meleagris gallopavo
SEQUENCE: 167
MVKVGVNGFG RIGRLVTRAA VLSGKVQVVA INDPFIDLNY MVYMFKYDST HGHFKGTVKA  60
ENGKLVINGH AITIFQERDP SNIKWADAGA EYVVESTGVF TTMEKAGAHL KGGAKRVIIS  120
APSADAPMFV MGVNHEKYDK SLKIVSNASC TTNCLAPLAK VIHDNFGIVE GLMTTVHAIT  180
ATQKTVDGPS GKLWRDGRGA AQNIIPASTG AAKAVGKVIP ELNGKLTGMA FRVPTPNVSV  240
VDLTCRLEKP AKYDDIKRVV KAAADGPLKG ILGYTEDQVV SCDFNGDSHS STFDAGAGIA  300
LNDHFVKLVS WYDNEFGYSN RVVDLMVHMA SKE                               333

SEQ ID NO: 168          moltype = AA  length = 284
FEATURE                 Location/Qualifiers
source                  1..284
                        mol_type = protein
                        organism = Meleagris gallopavo
SEQUENCE: 168
MEAIKKKMQM LKLDKENALD RAEQAEAEQK QAEERSKQLE DELAAMQKKL KGTEDELDKY  60
SEALKDAQEK LELAEKKAAD AEAEVASLNR RIQLVEEELD RAQERLATAL QKLEEAEKAA  120
DESERGMKVI ENRALKDEEK MELQEIQLKE AKHIAEEADR KYEEVARKLV IIEGDLERTE  180
ERAELAESKC SELEEELKNV TNNLKSLEAQ AEKYSQKEDK YEEEIKILTD KLKEAETRAE  240
FAERSVAKLE KTIDDLEDEL YAQKLKYKAI SEELDHALND MTSI                   284

SEQ ID NO: 169          moltype = AA  length = 968
FEATURE                 Location/Qualifiers
source                  1..968
                        mol_type = protein
                        organism = Meleagris gallopavo
SEQUENCE: 169
MCLQLFVFSE LPPPFSFLKN KELTFVYCLH KALMVECWCH SICSSPTLAD QIREVVGLES  60
LVSSGKCILE RIKHWLIIDI FLFLCISVWS IGESPPEETE KHDDSQRSTL FIEKPQSGSV  120
SVGGNITFIA KVEAKDLLRK PNVKWFKGKW MDLASKAGKH LQLKESFERH TKIHTFEMHI  180
IQAKENYAGN YRCEVSYKDK FDSCSFDLEV TESSQAAPSI DIRSAFKRSG DGQDDAGELD  240
FSGLLKRREI KQQEEEPEVD VWELLKNANP SEYEKIAFQY GITDLRGMLK RLKRMRREVK  300
KSAAFSKGLD PAYQVDKGGK VRFMVELADP TVELKWYKNG QEIRPSAKYI FEHKGNQRIL  360
FINNCTMTDD ARYYVTAGDE KCSTELFVRD PPILVTKGLE DTNTYVGERV ELSCEVSEDD  420
ANVKWFKNGV ELTNEPKSRY RIKVEGKKHT LIIEEAAKND NATYSVMTTG GQSEAKLSVD  480
LRPLKISLAL EDQTVRLGQE IHLKCEISEN VEGKWYKNGQ LIEASDRVKL YHKGRIHRFI  540
IASAAVDDEG EYMFVPDAYN INIPCKVHVV DPPKLHLDGL GENNTVTVVA GTKLRLEIPI  600
TGEPTPKVMW SKGDKWITDS GRIRAETYSD SSCLVIDTAE REDSGPFRIT LKNEAGEDSA  660
LINIKVVDVP DPPQAPNVTE VGEDWCVMTW EPPANDGGSP ILGYFIERKK KQSSRWMRLN  720
FELCKETTFE PKKMIEGVAY EVRVFAVNAI GTSKPSMPSK SFVPLAVTSP PTLLAVDSVT  780
DTSVTMKWRP PDHIGAAGLD GYVVEYCFEG TDEWIVANPE LTDKTKFTIN GLPTGSKILV  840
RVKAVNAAGE SEPRYHPQPI LVKEVIEPPK IRLPRHLKQT YTRRVGETVN LVIPFQVRTK  900
NTDRKTLNTH FSYATASYLL FLYVNLGYWT AFVNILHPEG KSHDKASGII HSYLGTKIGH  960
YHMSSQSN                                                           968

SEQ ID NO: 170          moltype = AA  length = 1309
FEATURE                 Location/Qualifiers
source                  1..1309
                        mol_type = protein
                        organism = Meleagris gallopavo
SEQUENCE: 170
MNVKGEISAY ASLVVKRYKG EIDASHLHAG TFSMPLSLGV TPHGFASRFE IKFIDKFEVM  60
FGREGETMSL GCTVVISPDI KRFKPDIEWY RNGVLITPSK WVQMHWTGER ASLTLTHANK  120
EDEGLYTLRV VMGEYYEEYS GYVFVRDADA EIPGAPGAPL DVQCLDANKD YVIVSWKQPA  180
VDGGNPILGY FIDRCEVGTT HWTQCNETPV KFARFPVTGL IEGRSYIFRV RAVNRAGVSL  240
PSRVSEAVAA LDPADRARLR SHPSAPWTGQ IIVTEEEPAE GVIPGPPTDL QVIEATKNYV  300
VLSWKPPGQR GHEGIMYFVE KCIAGTEDWQ RVNTEIPVKS PRFAVFDLSE GKSYCFRVRC  360
CNSAGIGEPS EATEATVVDD KLDIPKPPGR IMPTRNTDTS VVVTWTEPPD AKELVGYYIE  420
SSVVGSGHWE PCNNNPVKGT RFICHGLVNG EKYIFRVRAV NAAGLSEFSQ DSEPIEVQAA  480
IAAPSPPYDI TVLESVRDSM VLGWKQPKAT GGTEITGYYV NYREVIGGVP GKWREANIKA  540
ISERAYRIEN LKENMVYQFQ VAAANLAGVG TPSLPSKPFK CEEWTIAVPG PPHDVTYTEV  600
RKDSLVLLWK EPVYTGRSPI AGYYIDMKET EAKEEHWRSV NEKPLQKKFL KIAGLTEGVS  660
YVFRVRATNQ AGVGRPSDIT DPVVAETRPG TKEVVVNVDD DGVISLNFEC DQMSPNSKFV  720
```

-continued

```
WSKNYEPIED DSRLDIDSKG GKSKAIFKNL GEDDLGIYSC VVTDTDGVSS GYTIDEEEMK  780
RLLALSHERK FPTIPLASEL AVEILEKGEV RFWLQAEKLS GNAKANFVFN DKEIFNGEKY  840
KMKVDQRTGL VEMIMDKLED KDEGTYTFQL QDGKATNQSS LVLIGDVFKK LRDEAEFQRK  900
EWHRKQGPHF VDQLGWEVTD ECNVMLKCKV ANIKKETHIV WYKDDREIMV DEEHDFKDGV  960
CTLLISEFSK KDTGIYEVIL KDDRGKDVSE LKLTDTAFAD LMNEVCRRIA LSATDLKIQS  1020
TAEGIRLYSF VTYYVDDLRV GWVHNDTQIR YTDRVKTGVT GEQIWLQINE PTPQDKGKYT  1080
MELFDGKTGH KKTVDLSGQA FDEAFAEFQR LKQAAIAEKN RARVLGGLPD VVTIQEGKAL  1140
NLSCTVWGDP TPEVSWLKNE KSFVSDANCI LKFEAGKNVS FTISAVSTHD SGKYSLVVKN  1200
KYGTETSDVT VIVITVSGQT DTAKHTVSIS GQSQNNITVS KKKKATTMSK PQQEKEKKDI  1260
KPHVDEIIAT EETSDSKEIP RVEKEKNTEN NQQTAKKTHG KAAAARTVP              1309

SEQ ID NO: 171          moltype = AA  length = 434
FEATURE                 Location/Qualifiers
source                  1..434
                        mol_type = protein
                        organism = Meleagris gallopavo
SEQUENCE: 171
MSILKIQARE IFDSRGNPTV EVDLYTNKGL FRAAVPSGAS TGIYEALELR DNDKTRYLGK  60
GVSKAVEHVN KTIAPALISK NVNVVEQEKI DKLMLEMDGT ENKSKFGANA ILGVSLAVCK  120
AGAAEKGVPL YRHIADLAGN PEVILPVPAF NVINGGSHAG NKLAMQEFMI LPVGADNFKE  180
AMRIGAEVYH NLKNVIKEKY GKDATNVGDE GGFAPNILEN KEALELLKTA IGKAGYSDKV  240
VIGMDVAASE FYRDGKYDLD FKSPDDPSRY ISPDQLADLY KGFVKNYPLV SIEDPFDQDD  300
WAAWKKFTAS VGIQVVGDDL TVTNPKRIAK AVEEKSCNCL LLKVNQIGSV TESLQACKLA  360
QSNGWGVMVS HRSGETEDTF IADLVVGLCT GQIKTGAPCR SERLAKYNQL LRIEEELGSK  420
ARFAGRNFRN PRIN                                                    434

SEQ ID NO: 172          moltype = AA  length = 165
FEATURE                 Location/Qualifiers
source                  1..165
                        mol_type = protein
                        organism = Meleagris gallopavo
SEQUENCE: 172
MAPKKAKKRI EGANSNVFSM FEQAQIQEFK EAFTIMDQNR DGFIDKADLR DTFAALGRLN  60
VKNEEIDEMI KEAPGPINFT VFLTMFGEKL KGADPEETIL NAFKVFDPEG KGLKSAYIKE  120
MLMTQGERFS QEEIDQMFAA FPPDVSGNLD YKNLVHVITH GEEKD                  165

SEQ ID NO: 173          moltype = AA  length = 1974
FEATURE                 Location/Qualifiers
source                  1..1974
                        mol_type = protein
                        organism = Meleagris gallopavo
SEQUENCE: 173
MSSDAEMAIF GEAAPYLRKS EKERIEAQNR PFDAKAACFV VDQKQMYVKG TIQSREGGKV  60
TVKTYDDTTV TVKDDEVFPM NPPKYDKIED MAMMTHLHEP AVLYNLKERY AAWMIYTYSG  120
LFCVTVNPYK WLPVYNPEVV SGYRGKKRQE APPHIFSISD NAYQFMLTDR DNQSILITGE  180
SGAGKTVNTK RVIQYFATIA VTGEKKKDQQ PSKMQGTLED QIISANPLLE AFGNAKTVRN  240
DNSSRFGKFI RIHFGTTGKL ASADIETYLL EKSRVTFQLA SERSYHIFYQ IMSNKKPELI  300
GEYLLLITTN PYDFTYVSQG EITVASIDDS EELLATDNAV DILGFSPDEK VGMYKLTGAV  360
MHYGNMKFKQ KQREEQAEPD GTEVADKAAY LMGLNSAELL KALCYPRVKV GNEYVTKGQN  420
VQQVYNSVGA LAKSVYEKMF LWMVTRINQQ LDTKQPRQHF IGVLDIAGFE IFDFNSLEQL  480
CINFTNEKLQ QFFNHHMFVL EQEEYKKEGI DWEFIDFGMD LAACIELIEK PMGIFSILEE  540
ECMFPKATDT SFKNKLYDQH LGKSNNFQKP KPAKGKAEAH FSLIHYAGTV DYNITGWLEK  600
NKDPLNETVV GLYQKSSLKL LSFLFSNYAG AETSDSGSGK KGGKKKGGSF QTVSAVFREN  660
LNKLMTNLRS THPHFVRCLI PNETKTPGEM DHYLVMHQLR CNGVLEGIRI CRKGFPSRIL  720
YADFKQRYKI LNASAIPEGQ FIDSKKASEK LLSSIDVDHN QYRFGHTKVF FKAGLLGLLE  780
EMRDEKLVSL ITHMQAMCRG YLMRTEFKKM NARRESIYII QYNVRAFMNV KHWPWMKLYF  840
KMKPMLKSVE SEKEMANMKE EFEKTKEELA KSEAKRKELE EKMVSLLQEK NDLQLQVQSE  900
SENLADAEER CEGLIKSKIQ LEARIKELNE RMEDEEEMNA ELTAKKRKLE DECSELKKDI  960
DDLELTLAKV EKEKHATENK VKNLTEEMAA LDENISKLTK EKKALQEAHQ QTLEYLQAEE  1020
DKVSILTKIK TKLEQQVDDL EGSLEQERKQ RMDLERAKRK LDGDLKMSQD SIMDLKNDKQ  1080
QTDERLKKKD FEISHLQSKI EDEQAQSSQL QKKIKELQAR TEELEEEIEA ERANCAKTEK  1140
QRADLSRELE EISERLEEAG GATAAQIEMN KKREAEFQKM RRDLEEATLQ HEATAAALRK  1200
KHADSTAELG EQIDNLQRVK QKLEKEKSEL KMEIDDLASN MESVSKAKSN LEKMCRALED  1260
QYSEVRTKDD EHMRLINDMN TQKTRLQTEN GELTRQLEEK ESLISQLTRG KQAITQQTEE  1320
LKRQLEEENK AKNALAHGLQ SARHDCDLLR EQYEEELEAK GELQRALSKA NSEVAQWRTK  1380
YETDAIQCME ELEDAKKKLA QRLQDSEEQI EAVNSKCASL EKTKQRLQGE VDDLMIAVER  1440
SNAACAAFDK KQKNFDKVMA EWKQKYQESQ AELEAAQKES RTLSTEIFRM KNAYEEVLDQ  1500
VETVRRENKN LQQEISDLTE QLAEAGKANH GLEKAKKQTE QEKYDLQAAL EEAEVSLEHE  1560
EGKILRVQLE LNQVKSGVDR RSAEKDEEIQ QLKKNHQRVL ESMQTMLDAE IRSRNDALRL  1620
KKKMEGDLND MEIQLSRANC QIAEIQKHLK GMQGQLKDSQ LHLDDALREN DDLKEQLAVV  1680
ERRNNLMMTE LEQMRAALEQ TERARKVSEQ ELTDTSERVQ ILHSQNTSLL NTKKKLEVDI  1740
THLQNEVEDS IQEASNAEEK AKKAITDAAM MAEELKKEQD TSAHLERMKR NLEQMVKDLQ  1800
HRLDEAEQLT LKGGKKQLQK LEARIHELEN ELDIEQKRGT ESLKGARKYE RRLKELTYQS  1860
EEDKKNILRL QNLVDKLQLK VKAYKKQAEE AEEQASTNLS RCRKTQHELE EAEERADIAE  1920
CQVNKLRAKS RDTGRQKSGL SSELGDNTFL PVIWHNMCNI DRPETTPQVI LLHQ        1974
```

-continued

```
SEQ ID NO: 174          moltype = AA  length = 615
FEATURE                 Location/Qualifiers
source                  1..615
                        mol_type = protein
                        organism = Meleagris gallopavo
SEQUENCE: 174
MKWVTLISFI FLFSSATSRN LQRVARDAEH KSEIAHRYND LKEETFKAVA MITFAQYLQR  60
CSYDGLSKLV KDVVDLAQKC VANEDAPECS KSLPSIILDE ICQVEKLRDS YGAMADCCSK  120
ADPERNECFL SFKVSQPDFI QPYQRPASDV ICKEYQDNRV SLLGHFIYSV ARRNPYMYAP  180
TILSFAADYE HALQSCCQES DVGACLDAKE TAMREKAKKI SVKQQYSCGI LKKFGDRVFQ  240
AEKLVRLSQK YPKAPFSEVA KLVRESKDVH KECCEGDMVE CMDDMAEIIN NMCSKQDVFS  300
SKIKGCCEKP IMERSQCIME AEFDEKPADL PSLVEKYIQD KEVCKSFEAG HDAFMSEFVY  360
EYSRRHPEFS TQLILRIAKG YESLLEKCCK TDNPAECYAN AQDQLNQHIK ETQDVVKTNC  420
DLLSAHGEPD FLKSILIRYT KKMPQVSTET LLEIGKKMTA IGTKCCQLSE DRRMACSEGY  480
LSIVIHDMCR RQETTPINDN VSHCCSDSYA HRRPCFTAMG VDTKYVPPPF NPDMFSFDEK  540
LCSAPAEERE VGQMKLLINL IKRKPQMTEE QIKTIADGFT AMVDKCCKQS DINTCFGEEG  600
ANLIVQSRAT LGIGV                                                   615

SEQ ID NO: 175          moltype = AA  length = 381
FEATURE                 Location/Qualifiers
source                  1..381
                        mol_type = protein
                        organism = Anas platyrhynchos
SEQUENCE: 175
MPFSNSHNLL KMKHSATDEF PDLSAHNNHM AKVLTLDLYK KLRDKQTPSG FTLDDVIQTG  60
VDNPGHPFIM TVGCVAGDEE SYEVFKELFD PVIEDRHGGY KPTDQHKTDL NAGNLQGGDD  120
LDPNYVLSSR VRTGRSIRGF CLPPHCSRGE RRAIEKLSVE ALGSLGGDLK GKYYALRNMT  180
DAEQQQLIDD HFLFDKPVSP LLLASGMARD WPDARGIWHN DNKTFLVWIN EEDHLRVISM  240
QKGGNMKEVF TRFCTGLTQI ENLFKSKNHE FMWNPHLGYI LTCPSNLGTG LRAGVHIKLP  300
HLGKHEKFGE VLKKLRLQKR GTGGVDTAAV GGVFDVSNAD RLGFSEVELV QMVVDGVKLL  360
IEMEKRLEKG QSIDDLIPAQ K                                            381

SEQ ID NO: 176          moltype = AA  length = 897
FEATURE                 Location/Qualifiers
source                  1..897
                        mol_type = protein
                        organism = Anas platyrhynchos
SEQUENCE: 176
MNSMNQIETN MQYTYNYEED EYMTQEEEWD RDLLLDPAWE KQQRKTFTAW CNSHLRKAGT  60
QIENIEEDFR NGLKLMLLLE VISGERLPKP DRGKMRFHKI ANVNKALDYI ASKGVKLVSI  120
GAEEIVDGNV KMTLGMIWTI ILRFAIQDIS VEETSAKEGL LLWCQRKTAP YRNVNIQNFH  180
LSWKDGLGLC ALIHRHRPDL IDYSKLNKDD PIGNINLAME IAEKHLDIPK MLDAEDIVNT  240
PKPDERAIMT YVSCFYHAFA GAEQAETAAN RICKVLAVNQ ENERLMEEYE RLASELLEWI  300
RRTIPWLENR TPEKTMQAMQ KKLEDFRDYR RKHKPPKVQE KCQLEINFNT LQTKLRISNR  360
PAFMPSEGKM VSDIAGAWQR LEQAEKGYEE WLLNEIRRLE RLEHLAEKFR QKASTHEQWA  420
YGKEQILLQK DYESASLTEV RAMLRKHEAF ESDLAAHQDR VEQIAAIAQE LNELDYHDAA  480
SVNDRCQKIC DQWDSLGTLT QKRREALERT EKLLETIDQL HLEFAKRAAP FNNWMEGAME  540
DLQDMFIVHS IEEIQSLISA HDQFKATLPE ADGERQAILS IQNEVEKVIQ SYSMRISASN  600
PYSTVTVEEI RTKWEKVKQL VPQRDQSLQE ELARQHANER LRRQFAAQAN VIGPWIQTKM  660
EEIARSSIEM TGPLEDQMNQ LKQYEQNIIN YKHNIDKLEG DHQLIQEALV FDNKHTNYTM  720
EHIRVGWELL LTTIARTINE VETQILTRDA KGITQEQMND FRASFNHFDR RKNGLMDHDD  780
FRACLISMGY DLGEAEFARI MSLVDPNGQG TVTFQSFIDF MTRETADTDT AEQVIASFRI  840
LASDKPYILA DELRRELPPE QAQYCIKRMP PYTGPGSVPG ALDYTSFSSA LYGESDL     897

SEQ ID NO: 177          moltype = AA  length = 843
FEATURE                 Location/Qualifiers
source                  1..843
                        mol_type = protein
                        organism = Anas platyrhynchos
SEQUENCE: 177
MAAPLSDGEK RKQISVRGIA GLGDVAEVRK SFNRHLHFTL VKDRNVATPR DYFFALAHTV  60
RDHLVGRWIR TQQHYYERDP KRIYYLSLEF YMGRTLQNTM VNLGLQNACD EAIYQLGLDL  120
EELEEIEEDA GLGNGGLGRL AACFLDSMAT LGLAAYGYGI RYEFGIFNQK IVDGWQVEEA  180
DDWLRYGNPW EKARPEYMLP VHFYGRVEHT PEGVKWVDTQ VVLAMPYDTP VPGYKNNTVN  240
TMRLWSAKAP NDFNLQEFNV GDYIEAVLDR NLAENISRVL YPNDNFFEGK ELRLKQEYFV  300
VAATLQDIIR RFKSSKFGCR DPVRTSFETF PDKVAIQLND THPALSIPEL MRILVDVEKV  360
DWDKAWEITK RTCAYTNHTV LPEALERWPV SMFEKLLPRH LEIIYALNQM HLDRVAALYP  420
GDVDRLRRMS VIEEGDCKRI NMAHLCVIGS HAVNGVARIH SDIVKHSVFK DFYELEPEKF  480
QNKTNGITPR RWLLLCNPGL ADVIAEKIGE GFVTDLSQLK KLLDFIDNEM FIRDVAKVKQ  540
ENKLKFSAYL EEQYKVKINP SSMFDVQVKR IHEYKRQLLN CLHAITLYNR IRSDPSKSFV  600
PRTIMIGGKA APGYHMAKMI IKLITSIGEV INNDPYVGDK LKVIFLENYR VSLAEKVIPA  660
ADLSQQISTA GTEASGTGNM KFMLNGALTI GTMDGANVEM AEEAGEENLF IFGMRVEDVE  720
ALDRKGYNAL EYYERLPELR QAIDQISSGF FSPRDPGCFR DVVNMLMHHD RFKVFADYEA  780
YIKCQGQVDQ LFMDPREWTK KVIRNIACSG KFSSDRTITE YAQEIWGVEP SSTKIPPPNI  840
PRD                                                                843
```

-continued

```
SEQ ID NO: 178          moltype = AA  length = 364
FEATURE                 Location/Qualifiers
source                  1..364
                        mol_type = protein
                        organism = Anas platyrhynchos
SEQUENCE: 178
MTHQFPALSP EQKKALSDIA QRIVASGKGI LAADESVGTM GNRLQRINVE NTEENRRAFR  60
EILFSSDTSI NQSIGGVILF HETLYQKDSS GKPFPAIIKE KGIVVGIKLD KGTSPLAGTN  120
GETTIEGLDG LSERCAQYKK DGVDFGKWRA VLKITSTTPS QLAIQENANT LARYASICQQ  180
NGLVPIVEPE VLPDGDHDLQ RCQYVTEKVL AAVYKALNDH HVYLEGTLLK PNMVTAGHAC  240
TKKYTPQDVA AATVTTLLRT VPAAVPGICF LSGGQSEEEA SLNLNAMNQS SLPKPWKLTF  300
SYGRALQASA LAAWLGKSEN KKAAQEAFRK RAQINSLACR GQYVVSGKTD AAATQSLFTA  360
SYTY                                                               364

SEQ ID NO: 179          moltype = AA  length = 192
FEATURE                 Location/Qualifiers
source                  1..192
                        mol_type = protein
                        organism = Anas platyrhynchos
SEQUENCE: 179
MAPKKDVKKP AAAAAPAPAP APAPAPAPAK PKEPAIDLKS IKIEFSKEQQ DDFKEAFLLF  60
DRTGDAKITL SQVGDIVRAL GQNPTNAEIN KILGNPSKEE MNAKKITFEE FLPMLQAAAN  120
NKDQGTYEDF VEGLRVFDKE GNGTVMGAEL RHVLATLGEK MTEEEVDELM KGQEDSNGCI  180
NYEAFVKHIL SV                                                      192

SEQ ID NO: 180          moltype = AA  length = 339
FEATURE                 Location/Qualifiers
source                  1..339
                        mol_type = protein
                        organism = Anas platyrhynchos
SEQUENCE: 180
MVKVGVNGFG RIGRLVTRAA ILSGKVQVVA INDPFIDLNY MVYMFKYDST HGHFRGTVKA  60
ENGKLVINGN AITIFQERDP SNIKWADAGA EYVVESTGVF TTMEKAGAHL KGGAKRVIIS  120
APSADAPMFV MGVNHEKYDK SLKIVSNASC TTNCLAPLAK VIHDNFGIVE GLMTTVHAIT  180
ATQKTVDGPS GKLWRDGRGA AQNIIPASTG AAKAVGKVIP ELNGKLTGMA FRVPTPNVSV  240
VDLTCRLEKP AKYDDIKRVV KAAADGPLKG ILAYTEDQVV SCDFNGDSHS STFDAGAGIA  300
LNDHFVKLVS WYDNEYGYSN RVLALSTLLH AANIFPSSR                         339

SEQ ID NO: 181          moltype = AA  length = 284
FEATURE                 Location/Qualifiers
source                  1..284
                        mol_type = protein
                        organism = Anas platyrhynchos
SEQUENCE: 181
MEAIKKKMQM LKLDKENALD RAEQAEAEQK QAEERSKQLE DELAAMQKKL KGTEDELDKY  60
SEALKDAQEK LELAEKKAAD AEAEVASLNR RIQLVEEELD RAQERLATAL QKLEEAEKAA  120
DESERGMKVI ENRALKDEEK MELQEIQLKE AKHIAEEADR KYEEVARKLV IIEGDLERTE  180
ERAELAESKC SELEEELKNV TNNLKSLEAQ AEKYSQKEDK YEEEIKILTD KLKEAETRAE  240
FAERSVAKLE KTIDDLEDEL YAQKLKYKAI SEELDHALND MTSI                   284

SEQ ID NO: 182          moltype = AA  length = 1247
FEATURE                 Location/Qualifiers
source                  1..1247
                        mol_type = protein
                        organism = Anas platyrhynchos
SEQUENCE: 182
MPEPVKKTVS AFTKKPKTTE VTAGSTAVFE AETEKASVKV KWQRAGTEIT DSEKYVIKAE  60
GNKHSLTINN VVKEDDVTYA VIAGSSKVKF ELKVKEPEKS ETVTPAEAAP APAASELPAP  120
PAESSQNTEV PSAGTQPEEP ADPIGLFVTR PQDGEVTVGG NITFTAKVAG ESLLKKPSVK  180
WFKGKWMDLA SKVGQHLQLH DSYDRNNKVY TFELEIIEAN MTFAGGYRCE VSTKDKFDSS  240
NFNLTVNEAP LSGDLDIRAA FRRTTEGHDE AGELNFSALL KKRGDGKTES QPDVDVWEIL  300
RKAPPSEYEK IAFQYGITDL RGMLKRLKRI KKEEKKSTAF LKKLDPAYQV DKGQKIKLMV  360
EVANPDADVK WLKNGQEIQV SGSRYIFEAV GNKRILTINH CSLADDAAYE CVVGEEKSFT  420
ELFVKEPPIL ITRPLEDQMV MVGERVEFEC EVSEEGATVK WEKDGVELTR EETFKYRFKK  480
DGKKQYLIIN ESTKEDSGHY TVKTNGGVSV AELIVQEKKL EVYQSIADLT VKARDQAVFK  540
CEVSDENVKG IWLKNGKEVV PDERIKISHI GRIHKLTIED VTPGDEADYS FIPQGFAYNL  600
SAKLQFLEVK IDFVPREEPP KIHLDCLGQS PDTIVVVAGN KLRLDVPISG DPTPTVIWQK  660
VNKKTDLLRS NDDSMTPSEN SSDLNTDSKL LFESEGRVRV EKHEDHCVFI IEGAEKEDEG  720
VYRVIVKNPV GEDKADITVK VIDVPDPPEA PKISNIGEDY CTVQWQPPKY DGGQPVLGYI  780
LERKKKKSYR WMRLNFDLLK ELTYEAKRMI EGVVYEMRIY AVNSIGMSRP SPASQPFMPI  840
APPSEPTHFT VEDVSDSTVA LKWRPPERIG AGGLDGYIVE YCKEGSTEWI PALPGLTERT  900
SALIKDLVTG DKLHFRIKAI NLAGESGAAI IKEPITVQEI LQRPKIWLPR HLRQTLVKKV  960
GETINIVIPF QGKPRPKITW TKDGQSLDAK DVGIRNSNTD TILFIRKAEL HHSGAYEITL  1020
QIENMTDKVA ITIQIIDKPG PPQNIKLTDV WGFNAALEWT PPQDDGNAQI LGYTVQKADK  1080
KTMEWYTVFD HYRRTNCVVS ELIMGNEYFF RVFSENLCGL SETAATTKNP AYIQKTGTTY  1140
KPPCYKEHDF SEPPKFTHPL VNRSVIAGYN TTLSCAVRGI PKPKIYWYKN KVDLSGDAKY  1200
```

-continued

```
RMFSKQGVLT LEIRKPTPFD GGFYTCKAVN ECGEAEIECK LDVRVPQ                1247

SEQ ID NO: 183          moltype = AA  length = 1451
FEATURE                 Location/Qualifiers
source                  1..1451
                        mol_type = protein
                        organism = Anas platyrhynchos
SEQUENCE: 183
MSSLAVPFYQ RRHKHFDQSY RNIQTRYVLE EYAARKAASR QAAYYESSGL GKTTCRLCAR  60
RARTLAHETM QESRKRTHEQ KSHASEEKRI RFASELSALE KEIHTARHRA REQLDRLAIQ  120
RMVEENMAME RHAVEERISR APEILVRLRS HTVWEKMSVS LCFTVQGFPT PVVQWYKNEE  180
LITPASDPAK YRVESKYGVH VLHINRADFD DSATYSAVAT NVHGQASTNC AVVVRRFREG  240
EEPHPAGIMP FHLPLSYDVC FTHFDVQFLE KFGVTFATEG ETLTLKCSVL VTPELKRLRP  300
RAEWYRDDVL IKDSKWTKLY FGEGQAALSF THLNKDDEGL YTLRMITKGG VNECSAYLFV  360
RDADALIAGA PGAPMDVKCH DANRDYVIVT WKPPNTTSQS PVIGYFVDKC EVGLENWVQC  420
NDAPVKICKY PVTGLYEGRS YIFRVRAVNN AGISRPSRVS EPVAALDPVD LERTQTIHVD  480
EGRKIVISKD DLEGDIQIPG PPTNVHASEI SKTYVVLSWD PPVPRGREPL TYFIEKSMVG  540
SGSWQRVNAQ VAVKSPRYAV FDLAEGKPYV FRVVSANKHG ISDPSEITEP IQPQDTVVVP  600
SAPGRVVATR NTKTSVVVQW DKPKHEEDLY GYYIDYSVVG SNHWEPSNHK PIKYNRFVVH  660
GLETGEQYIF RVKAVNAVGF SENSQESEAI KVQAALTCPS YPHGITLLNC DGHSMTLGWK  720
APKYSGGSPI LGYYMDKREA NHQNWHEVNS SLITRTIYTV EDLTEDAFYE FKIAAANIVG  780
IGQPSDPSEH FKCKAWTMPE PGPAYDLTVC EVRNTSLVLL WKAPVYEGKS PITGYLVDYK  840
EVDKEDWITA NEKPTAQRYF KVTDLHEGHT YVFKVRAVND AGVGKSSEVS EPVLVEAKPG  900
TKEIFSGVDD EGNIYLAFEC KEATDASHFA WGKSYEEIED SDKYKIETEG DHSKLYFKNP  960
DKSDLGTYSI SVSDTDGVSS SFVMDDEELE RLMAISNEIK NPTIPLKSEL AYEVLEKGEV  1020
RFWIQAESLS PDSTFRFVIN DKEVENSDRH KINCDRPNGI IEMVMDRFTI DNEGTYTVQI  1080
QDGKAKNQSS LVLIGDAFKA VLAESELQRK EFLRKQGPHF SEFLYWEVTE ECEVLLACKV  1140
ANTKKETVFK WYKDGSGIDV DEMPDLQKGE CHLSIPRLSR KDEGVYKATL SDDRGHDVST  1200
LELSGKVYDD IILALSRVSG KTASPLKILC TEEGIRLQCF LKYYNEEMKI SWSHRESKIS  1260
SGEKMKIGGG ENIAWLQITE PTEKDKGNYT FEIFGDKESF KRTLDLSGQA FDDAYAEFQR  1320
LKAAAFAEKN RGKVIGGLPD VVTIMDGKTL NLTCTVFGNP DPEVVWFKND KALELNEHYV  1380
VSLEQGKYAS LTIKGVTSED SGKFSIHVKN KYGGETVDVT VSVYRHGEKV PEVKPGQLAK  1440
PRPIPPSEDS S                                                       1451

SEQ ID NO: 184          moltype = AA  length = 408
FEATURE                 Location/Qualifiers
source                  1..408
                        mol_type = protein
                        organism = Anas platyrhynchos
SEQUENCE: 184
MPERSWIPVG TPLLRWTCTH TKACFEQRSP AVPPLVSTKH WSSETTTSPV SLEKGLSVVD  60
QEKIDNLMLE MDGTENKSKF GANAILGVSL AVCKAGAAEK DVPLYRHIAD LAGNSDLILP  120
VPAFNVINGG SHAGNKLAMQ EFMILPVGAE SFRDAMRIGA EVYHNLKSVI KEKYGKDATN  180
VGDEGGFAPN ILENSEALEL LKEAIDKAGY TDKIVIGMDV AASEFYRDGK YDLDFKSPDD  240
PSRYISADEL GDLYQSFVRD YPVVSIEDPF DQDDWEAWSK FTANVGIQIV GDDLTVTNPK  300
RIERAVEEKA CNCLLLKVNQ IGSVTEAIQA CKLAQENGWG VMVSHRSGET EDTFIADLVV  360
GLCTGQIKTG APCRSERLAK YNQLMRIEEE LGDEARFAGH NFRNPSVL               408

SEQ ID NO: 185          moltype = AA  length = 171
FEATURE                 Location/Qualifiers
source                  1..171
                        mol_type = protein
                        organism = Anas platyrhynchos
SEQUENCE: 185
MQGENHRAPK KAKKRIEGAN SNVFSMFEQA QIQEFKEAFT IMDQNRDGFI DKADLRDTFA  60
ALGRLNVKNE EIDEMIKEAP GPINFTVFLT MFGEKLKGAD PEETILNAFK VFDPEGKGLK  120
SAYIKEMLMT QGERFSQEEV DQMFAAFPPD VSGNLDYKNL VHVITHGEEK D           171

SEQ ID NO: 186          moltype = AA  length = 1626
FEATURE                 Location/Qualifiers
source                  1..1626
                        mol_type = protein
                        organism = Anas platyrhynchos
SEQUENCE: 186
FSISDNAYQF MLTDRENQSI LITGESGAGK TVNTKRVIQY FATIAASGEK KKEEQSGKMQ  60
GTLEDQIISA NPLLEAFGNA KTVRNDNSSR FGKFIRIHFG ATGKLASADI ETYLLEKSRV  120
TFQLKAERSY HIFYQIMSNK KPELIDMLLI TTNPYDYHFV SQGEITVASI NDQEELMATD  180
SAIDILGFTA DEKVAIYKLT GAVMHYGNLK FKQKQREEQA EPDGTEVADK AAYLMGLNSA  240
DLLKALCYPR VKVGNEYVTK GQTVQQVNNS VGALAKAVYE KMFLWMVVRI NQQLDTKQPR  300
QYFIGVLDIA GFEIFDFNSL EQLCINFTNE KLQQFFNHHM FVLEQEEYKK EGIEWEFIDF  360
GMDLAACIEL IEKPMGIFSI LEEECMFPKA TDTSFKNKLY DQHLGKSSNF QKPKPAKGKV  420
EAHFSLVHYA GTVDYNITGW LEKNKDPLNE TVIGLYQKSS VKTLALLFAN YGGAEINPQK  480
IIIKDTFKGS SFQTVSALFR ENLNKLMTNL RSTHPHFVRC IIPNETKTPG AMEHELVLHQ  540
LRCNGVLEGI RICRKGFPSR VLYADFKQRY KVLNASAIPE GQFIDSKKAS EKLLGSIDVD  600
HTQYKFGHTK VFFKAGLLGL LEEMRDEKLA QLITRTQARC RGFLMRVEYQ RMVERRESIF  660
CIQYNVRAFM NVKHWPWMKL FFKIKPLLKS AESEKEMANM KEEFEKTKEE LAKSEAKKKE  720
```

-continued

```
LEEKMVKLVQ EKNDLQLQVQ AEADALADAE ERCDQLIKTK IQLEAKIKEV TERAEDEEEI  780
NAELTAKKRK LEDECSELKK DIDDLELTLA KVEKEKHATE NKVKNLTEEM AALDETIAKL  840
TKEKKALQEA HQQTLDDLQA EEDKVNTLTK AKTKLEQQVD DLEGSLEQEK KLRMDLERAK  900
RKLEGDLKLS QDSIMDLEND KQQLDEKLKK KDFEISQIQS KIEDEQALGM QLQKKIKELQ  960
ARIEELEEEI EAERTSRAKA EKHRADLSRE LEEISERLEE AGGATAAQIE MNKKREAEFQ  1020
KMRRDLEEAT LQHEATAAAL RKKHADSTAE LGEQIDNLQR VKQKLEKEKS ELKMEIDDLA  1080
SNMESVSKAK ANLEKMCRTL EDQLSEIKTK EEEHQRMIND LTAQRARLQT ESGEYSRQVE  1140
EKDSQISQLS RSKQAFTQQI EELKRHLEEE IKAKNALAHA LQSARHDCDL LREQYEEEQE  1200
AKGELQRALS KANSEVAQWR TKYETDAIQR TEELEEAKKK LAQRLQDAEE HVEAVNAKCA  1260
SLEKTKQRLQ NEVEDLMIDV ERSNAACAAL DKKQKNFDKI LAEWKQKYEE TQAELEASQK  1320
ESRSLSTELF KMKNAYEESL DHLETLKREN KNLQQEISDL TEQIAEGGKA IHELEKVKKQ  1380
IEQEKSEIQA ALEEAEASLE HEEGKILRLQ LELNQVKSEI DRKIAEKDEE IDQLKRNHLR  1440
IVESMQSTLD AEIRSRNEAL RLKKKMEGDL NEMEIQLSHA NRQAAEAQKN LRNTQGVLKD  1500
TQIHLDDALR TQEDLKEQVA MVERRANLLQ AEIEELRAAL EQTERSRKVA EQELMDASER  1560
VQLLHTQNTS LINTKKKLET DITQIQSEME DAIQEARNAE EKAKKAITDA SMMAEELKKE  1620
QDTSAH                                                             1626

SEQ ID NO: 187          moltype = AA  length = 565
FEATURE                 Location/Qualifiers
source                  1..565
                        mol_type = protein
                        organism = Anas platyrhynchos
SEQUENCE: 187
MITFAQYLQR CSYDGLSKLV KDVVDLAQKC VANEDAPECS KSLPSIFLDE ICQVEKLRDS  60
YGAMADCCAK SDPERNECFL SFKIPQPDFV QPYQRPASDV ICKEYEDNRV SFLGHFIYSV  120
ARRNPFLYAP TILGVAADYE HALKSCCKES DVGACLDGKE TGIREKVKKI SVKQQYSCGI  180
LKKFGDRIFQ ADKLALLSQK YPKASFAEIS KLIHDVKDVY KECCEGDMVE CMDDRAELVN  240
YMCSKHEAFS SKIKDCCEKP IVERSQCIME AEFDEKPADL PSLVEKYIQD KEVCKSFEAG  300
HDAFMSEFVY EYSRRHPEFS TQLILRIAKG YETLLEKCCK TDNPAECYAN AQDQLNEHIK  360
ETQDVVKTNC DLLTTHGEAD FLKALLIRYT KKMPQVSTET LLEIGKKMTG VGTKCCKLPE  420
DRRLPCSEGY LSMVIQDMCR RQETSPINDN VSHCCSDSYA YRRPCFSAMG VDTKYVPPPF  480
NPDVFNFDEK LCSAPAEERA EGQMKLLINL IKRKPQMTEE QIKTIADGFT AMVDKCCKQS  540
DIETCFGEEG ANLIVQSRAT LGIGA                                        565

SEQ ID NO: 188          moltype = AA  length = 418
FEATURE                 Location/Qualifiers
source                  1..418
                        mol_type = protein
                        organism = Cairina sp.
SEQUENCE: 188
MCCLKCQQGF WDFCAPAVLA RGPCFGCGAL VTAVGSAMPF SNSHNLLKMK HSATDEFPDL  60
SAHNNHMAKV LTLDLYKKLR DKQTPSGFTL DDVIQTGVDN PGHPFIMTVG CVAGDEESYE  120
VFKELFDPVI EDRHGGYKPT DQHKTDLNAG NLQGGDDLDP NYVLSSRVRT GRSIRGFCLP  180
PHCSRGERRA IEKLSVEALG SLGGDLKGKY YALRNMTDAE QQQLIDDHFL FDKPVSPLLL  240
ASGMARDWPD ARGIWHNDNK TFLVWINEED HLRVISMQKG GNMKEVFTRF CTGLTQIENL  300
FKSKNHEFMW NPHLGYILTC PSNLGTGLRA GVHIKLPHLG KHEKFGEVLK KLRLQKRGTG  360
GVDTAAVGGV FDVSNADRLG FSEVELVQMV VDGVKLLIEM EKRLEKGQSI DDLIPAQK    418

SEQ ID NO: 189          moltype = AA  length = 885
FEATURE                 Location/Qualifiers
source                  1..885
                        mol_type = protein
                        organism = Cairina sp.
SEQUENCE: 189
RAGDKPQDEY MTQEEEWDRD LLLDPAWEKQ QRKTFTAWCN SHLRKAGTQI ENIEEDFRNG  60
LKLMLLLEVI SGERLPKPDR GKMRFHKIAN VNKALDYIAS KGVKLVSIGA EEIVDGNVKM  120
TLGMIWTIIL RFAIQDISVE ETSAKEGLLL WCQRKTAPYR NVNIQNFHLS WKDGLGLCAL  180
IHRHRPDLID YSKLNKDDPI GNINLAMEIA EKHLDIPKML DAEDIVNTPK PDERAIMTYV  240
SCFYHAFAGA EQAETAANRI CKVLAVNQEN ERLMEEYERL ASELLEWIRR TIPWLENRTP  300
EKTMQAMQKK LEDFRDYRRK HKPPKVQEKC QLEINFNTLQ TKLRISNRPA FMPSEGKMVS  360
DIAGAWQRLE QAEKGYEEWL LNEIRRLERL EHLAEKFRQK ASTHEQWAYG KEQILLQKDY  420
ESASLTEVRA MLRKHEAFES DLAAHQDRVE QIAAIAQELN ELDYHDAASV NDRCQKICDQ  480
WDSLGTLTQK RREALERTEK LLETIDQLHL EFAKRAAPFN NWMEGAMEDL QDMFIVHSIE  540
EIQSLITAHD QFKATLPEAD GERQAILSIQ NEVEKVIQSY SMRISASNPY STVTVEEIRT  600
KWEKVKQLVP QRDQSLQEEL ARQHANERLR RQFAAQANVI GPWIQTKMEE IARSSIEMTG  660
PLEDQMNQLK QYEQNIINYK HNIDKLEGDH QLIQEALVFD NKHTNYTMEH IRVGWELLLT  720
TIARTINEVE TQILTRDAKG ITQEQMNDFR ASFNHFDRRK NGLMDHDDFR ACLISMGYDL  780
GEAEFARIMS LVDPNGQGTV TFQSFIDFMT RETADTDTAE QVIASFRILA SDKPYILADE  840
LRRELPPEQA QYCIKRMPPY TGPGSVPGAL DYTSFSSALY GESDL                  885

SEQ ID NO: 190          moltype = AA  length = 843
FEATURE                 Location/Qualifiers
source                  1..843
                        mol_type = protein
                        organism = Cairina sp.
SEQUENCE: 190
```

-continued

```
MAAPLSDGEK RKQISVRGIA GLGDVAEVRK SFNRHLHFTL VKDRNVATPR DYFFALAHTV  60
RDHLVGRWIR TQQHYYERDP KRIYYLSLEF YMGRTLQNTM VNLGLQNACD EAIYQLGLDL  120
EELEEIEEDA GLGNGGLGRL AACFLDSMAT LGLAAYGYGI RYEFGIFNQK IVDGWQVEEA  180
DDWLRYGNPW EKARPEYMLP VHFYGRVEHN PEGVKWVDTQ VVLAMPYDTP VPGYKNNTVN  240
TMRLWSAKAP NDFNLQEFNV GDYIEAVLDR NLAENISRVL YPNDNFFEGK ELRLKQEYFV  300
VAATLQDIIR RFKSSKFGCR DPVRTSFETF PDKVAIQLND THPALSIPEL MRILVDVEKV  360
DWDKAWEITK RTCAYTNHTV LPEALERWPV SMFEKLLPRH LEIIYALNQM HLDRVAALYP  420
GDVDRLRRMS VIEEGDCKRI NMAHLCVIGS HAVNGVARIH SDIVKHSVFK DFYELEPEKF  480
QNKTNGITPR RWLLLCNPGL ADVIAEKIGE GFVTDLSQLK KLLDFIDNKM FIRDVAKVKQ  540
ENKLKFSAYL EEQYKVKINP SSMFDVQVKR IHEYKRQLLN CLHAITLYNR IRSDPSKSFV  600
PRTIMIGGKA APGYHMAKMI IKLITSIGEV INNDPYVGDK LKVIFLENYR VSLAEKVIPA  660
ADLSQQISTA GTEASGTGNM KFMLNGALTI GTMDGANVEM AEEAGEENLF IFGMRVEDVE  720
ALDRKGYNAL EYYERLPELR QAIDQISSGF FSPGDPGCFR DVVNMLMHHD RFKVFADYEA  780
YIKCQGQVDQ LFMDPWEWTK KVIRNIACSG KFSSDRTITE YAQEIWGVEP SSTKIPPPNI  840
PRD                                                                843

SEQ ID NO: 191          moltype = AA  length = 364
FEATURE                 Location/Qualifiers
source                  1..364
                        mol_type = protein
                        organism = Cairina sp.
SEQUENCE: 191
MTHQFPALSP EQKKALSDIA QRIVASGKGI LAADESVGTM GNRLQRINVE NTEENRRAFR  60
EILFSSDTSI NQSIGGVILF HETLYQKDSS GKPFPAIIKE KGIVVGIKLD KGTSPLAGTN  120
GETTIEGLDG LSERCAQYKK DGVDFGKWRA VLKITSTTPS QLAIQENANT LARYASICQQ  180
NGLVPIVEPE VLPDGDHDLQ RCQYVTEKVL AAVYKALNDH HVYLEGTLLK PNMVTAGHAC  240
TKKYTPQDVA AATVTTLLRT VPAAVPGICF LSGGQSEEEA SLNLNAMNQS SLPKPWKLTF  300
SYGRALQASA LAAWLGKPEN KKAAQEAFRK RAQINSLACR GQYVVSGKTD AAATQSLFTA  360
SYTY                                                               364

SEQ ID NO: 192          moltype = AA  length = 333
FEATURE                 Location/Qualifiers
source                  1..333
                        mol_type = protein
                        organism = Cairina sp.
SEQUENCE: 192
MVKVGVNGFG RIGRLVTRAA ILSGKVQVVA INDPFIDLNY MVYMFKYDST HGHFRGTVKA  60
ENGKLVINGN AITIFQERDP SNIKWADAGA EYVVESTGVF TTMEKAGAHL KGGAKRVIIS  120
APSADAPMFV MGVNHEKYDK SLKIVSNASC TTNCLAPLAK VIHDNFGIVE GLMTTVHAIT  180
ATQKTVDGPS GKLWRDGRGA AQNIIPASTG AAKAVGKVIP ELNGKLTGMA FRVPTPNVSV  240
VDLTCRLEKP AKYDDIKRVV KAAADGPLKG ILAYTEDQVV SCDFNGDSHS STFDAGAGIA  300
LNDHFVKLVS WYDNEYGYSN RVVDLMVHMA SKE                               333

SEQ ID NO: 193          moltype = AA  length = 284
FEATURE                 Location/Qualifiers
source                  1..284
                        mol_type = protein
                        organism = Cairina sp.
SEQUENCE: 193
MDAIKKKMQM LKLDKENALD RAEQAEADKK AAEERSKQLE DELVALQKKL KGTEDELDKY  60
SESLKDAQEK LELADKKATD AESEVASLNR RIQLVEEELD RAQERLATAL QKLEEAEKAA  120
DESERGMKVI ENRAQKDEEK MEIQEIQLKE AKHIAEEADR KYEEVARKLV IIESDLERAE  180
ERAELSESKC AELEEELKTV TNNLKSLEAQ AEKYSQKEDK YEEEIKVLTD KLKEAETRAE  240
FAERSVTKLE KSIDDLEDEL YAQKLKYKAI SEELDHALND MTSM                   284

SEQ ID NO: 194          moltype = AA  length = 1258
FEATURE                 Location/Qualifiers
source                  1..1258
                        mol_type = protein
                        organism = Cairina sp.
SEQUENCE: 194
YPTTKPKTTE VTAGSTAVFE AETEKAGVKV KWQRAGTEIT DSEKYVIKAE GNKHSLTINN  60
VVKEDDVTYA VIAGSSKVKF ELKVKEPEKS ETVTPAEAAP APAASELPAP PAESSQNTEV  120
PSAGTQPEEP VDPIGLFVTR PQDGEVTVGG NITFTAKVAG ESLLKKPSVK WFKGKWMDLA  180
SKVGQHLQLH DSYDRNNKVY TFELEIIEAN MTFAGGYRCE VSTKDKFDSS NFNLTVNEIY  240
HLCHDIVSFP FPSKTNYIFL LLRSTEGHDE AGELNFSALL KKRGDGKTES QPDVDVWEIL  300
RKAPPSEYEK IAFQYGITDL RGMLKRLKRI KKEEKKSTAF LKKLDPAYQV DKGQKIKLMV  360
EVANPDADVK WLKNGQEIQV SGRYIFEAVG NKRILTINHC SLADDAAYEC VVGEEKSFTE  420
LFVKEPPILI THPLEDQMVM VGERVEFECE VSEEGATVKW EKDGVELTRE ETFKYRFKKD  480
GKKQYLIINE STKEDSGHYT VKTNGGVSVA ELIVQEKKLE VYQSIADLTV KARDQAVFKC  540
EVSDENVKGI WLKNGKEVVP DERIKISHIG RIHKLTIEDV TPGDEADYSF IPQGFAYNLS  600
AKLQFLEVKI DFVPREEPPK IHLDCLGQSP DTIVVVAGNK LRLDVPISGD PTPTVIWQKV  660
NKKTDLLRSN DDSMTPSENS SDLNTDSKLL FESEGRVRVE KHEDHCVFII EGAEKEDEGV  720
YRVIVKNPVG EDKADITVKV IDVPDPPEAP KISNIGEDYC TVQWQPPKYD GGQPVLGYIL  780
ERKKKKSYRW MRLNFDLLKE LTYEAKRMIE GVVYEMRIYA VNSIGMSRPS PASQPFMPIA  840
PPSEPTHFTV EDVSDSTVAL KWRPPERIGA GGLDGYIVEY CKEGSTEWIP ALPGLTERTS  900
```

-continued

```
ALIKDLVTGD KLHFRIKAIN LAGESGAAII KEPITVQEIL QRPKIWLPRH LRQTLVKKVG  960
ETINIVIPFQ GKPRPKITWT KDGQSLDAKD VGIRNSNTDT ILFIRKAELH HSGAYEITLQ  1020
IENMTDKVAI TIQIIDKPGP PQNIKLTDVW GFNAALEWTP PQDDGNAQIL GYTVQKADKK  1080
TMEWYTVFDH YRRTNCVVSE LIMGNEYFFR VFSENLCGLS ETAATTKNPA YIQKTGTTYK  1140
PPCYKEHDFS EPPKFTHPLV NRSVIAGYNT TLSCAVRGIP KPKIYWYKNK MDLSGDAKYR  1200
MFSKQGVLTL EIRKPTPFDG GFYTCKAVNE RGEAEIECKL DVRGNHMSIM HNTNTAVK    1258

SEQ ID NO: 195          moltype = AA  length = 1451
FEATURE                 Location/Qualifiers
source                  1..1451
                        mol_type = protein
                        organism = Cairina sp.
SEQUENCE: 195
MSSLAVPFYQ RRHKHFDQSY RNIQTRYVLE EYAARKAASR QAAYYESSGL GKTTCRLCAR  60
RARTLAHETM QESRKRTHEQ KSHASEEKRI RFASELSALE KEIHTARHRA REQLDRLAIQ  120
RMVEENMAME RHAVEERISR APEILVRLRS HTVWEKMSVS LCFTVQGFPT PVVQWYKNEE  180
LITPASDPAK YRVESKYGVH VLHINRADFD DSATYSAVAT NVHGQASTNC AVVVRRFREG  240
EEPHPAGIMP FHLPLSYDVC FTHFDVQFLE KFGVTFATEG ETLTLKCSVL VTPELKRLRP  300
RAEWYRDDVL IKDSKWTKLY FGEGQAALSF THLNKDDEGL YTLRMITKGG VNECSAYLFV  360
RDADALIAGA PGAPMDVKCH DANRDYVIVT WKPPNTTSQS PVIGYFVDKC EVGLENWVQC  420
NDAPVKICKY PVTGLYEGRS YIFRVRAVNN AGISRPSRVS EPVAALDPVD LERTQTIHVD  480
EGRKIVISKD DLEGDIQIPG PPTNVHASEI SKTYVVLSWD PPVPRGREPL TYFIEKSMVG  540
SGSWQRVNAQ VAVKSPRYAV FDLAEGKPYV FRVVSANKHG ISDPSEITEP IQPQDTVVVP  600
SAPGRVVATR NTKTSVVVQW DKPKHEEDLY GYYIDYSVVG SNHWEPSNHK PIKYNRFVVH  660
GLETGEQYIF RVKAVNAVGF SENSQESEAI KVQAALTCPS YPHGITLLNC DGHSMTLGWK  720
APKYSGGSPI LGYYMDKREA NHQNWHEVNS SLITRTIYTV EDLTEDAFYE FKIAAANIVG  780
IGQPSDPSEH FKCKAWTMPE PGPAYDLTVC EVRNTSLVLL WKAPVYEGKS PITGYLVDYK  840
EVDKEDWITA NEKPTAQRYF KVTDLHEGHT YVFKVRAVND AGVGKSSEVS EPVLVEAKPG  900
TKEIFSGVDD EGNIYLAFEC KEATDASHFA WGKSYEEIED SDKYKIETEG DRSKLYFKNP  960
DKSDLGTYSI SVSDTDGVSS SFVMDDEELE RLMAISNEIK NPTIPLKSEL AYEVLEKGEV  1020
RFWIQAESLS PDSTFRFVIN DKEVENSDRH KINCDRPNGI IEMVMDRFTI DNEGTYTVQI  1080
QDGKAKNQSS LVLIGDAFKA VLAESELQRK EFLRKQGPHF SEFLYWEVTE ECEVLLACKV  1140
ANTKKETVFK WYKDGSGIDV DEMPDLQKGE CHLSIPRLSR KDEGVYKATL SDDRGHDVST  1200
LELSGKVYDD IILALSRVSG KTASPLKILC TEEGIRLQCF LKYYNEEMKI SWSHRESKIS  1260
SGEKMKIGGG ENIAWLQITE PTEKDKGNYT FEIFGDKESF KRTLDLSGQA FDDAYAEFQR  1320
LKAAAFAEKN RGKVIGGLPD VVTIMDGKTL NLTCTVFGNP DPEVVWFKND KALELNEHYV  1380
VSLEQGKYAS LTIKGVTSED SGKFSIHVKN KYGGETVDVT VSVYRHGEKV PEVKPGQLAK  1440
PRPIPPSEGS S                                                       1451

SEQ ID NO: 196          moltype = AA  length = 438
FEATURE                 Location/Qualifiers
source                  1..438
                        mol_type = protein
                        organism = Cairina sp.
SEQUENCE: 196
MVFKMSILKI HAREIFDSRG NPTVEVDLYT NKGLFRAAVP SGASTGIYEA LELRDNDKTR  60
YMGKGVSKAV EHVNKTIAPA LISKNVNVVE QEKIDKLMLD MDGSENKSKF GANAILGVSL  120
AVCKAGAAEK GVPLYRHIAD LAGNPEVILP VPAFNVINGG SHAGNKLAMQ EFMILPVGAD  180
SFKEAMRIGA EVYHNLKNVI KEKYGKDATN VGDEGGFAPN ILENKEALEL LKTAIGKAGY  240
SDKVVIGMDV AASEFYRDGK YDLDFKSPDD PSRYISPDQL ADLYKGFVKN YPVVSIEDPF  300
DQDDWAAWKK FTGSVGIQVV GDDLTVTNPK RIAKAVEEKS CNCLLLKVNQ IGSVTESLQA  360
CKLAQSNGWG VMVSHRSGET EDTFIADLVV GLCTGQIKTG APCRSERLAK YNQLLRIEEE  420
LGSKARFAGR NFRNPRIN                                                438

SEQ ID NO: 197          moltype = AA  length = 165
FEATURE                 Location/Qualifiers
source                  1..165
                        mol_type = protein
                        organism = Cairina sp.
SEQUENCE: 197
MAPKKAKKRI EGANSNVFSM FEQAQIQEFK EAFTIMDQNR DGFIDKADLR DTFAALGRLN  60
VKNEEIDEMI KEAPGPINFT VFLTMFGEKL KGADPEETIL NAFKVFDPEG KGLKSAYIKE  120
MLMTQGERFS QEEVDQMFAA FPPDVSGNLD YKNLVHVITH GEEKD                  165

SEQ ID NO: 198          moltype = AA  length = 1945
FEATURE                 Location/Qualifiers
source                  1..1945
                        mol_type = protein
                        organism = Cairina sp.
SEQUENCE: 198
MTSPDAEMAA FGEAAPYLRK SEKERIEAQN KPFDAKSSVF VAHPKESFVK GTIQSKESGK  60
VTVKTEGGET LTVKEDQIFS MNPPKYDKIE DMAMMTHLHE PAVLYNLKER YAAWMIYTYS  120
GLFCVTVNPY KWLPVYNPEV VLAYRGKKRQ EAPPHIFSIS DNAYQFMLTD RENQSILITG  180
ESGAGKTVNT KRVIQYFATI AASGEKKKEE QSGKMQGTLE DQIISANPLL EAFGNAKTVR  240
NDNSSRFGKF IRIHFGATGK LASADIETYL LEKSRVTFQL KAERSYHIFY QIMSNKKPEL  300
IDMLLITTNP YDYHFVSQGE ITVASINDQE ELMATDSAID ILGFTADEKV AIYKLTGAVM  360
```

-continued

```
HYGNLKFKQK QREEQAEPDG TEVADKAAYL MGLNSADLLK ALCYPRVKVG NEYVTKGQTV  420
QQVNNSVGAL AKAVYEKMFL WMVVRINQQL DTKQPRQYFI GVLDIAGFEI FDFNSLEQLC  480
INFTNEKLQQ FFNHHMFVLE QEEYKKEGIE WEFIDFGMDL AACIELIEKP MGIFSILEEE  540
CMFPKATDTS FKNKLYDQHL GKSSNFQKPK PAKGKAEAHF SLVHYAGTVD YNITGWLEKN  600
KDPLNETVIG LYQKSSVKTL ALLFANYGGA EAEASGGGAS GGKKGGKKKG SSFQTVSALF  660
RENLNKLMTN LRSTHPHFVR CIIPNETKTP GAMEHELVLH QLRCNGVLEG IRICRKGFPS  720
RVLYADFKQR YKVLNASAIP EGQFIDSKKA SEKLLGSIDV DHTQYKFGHT KVFFKAGLLG  780
LLEEMRDEKL AQLITRTQAR CRGFLMRVEY QRMVERRESI FCIQYNVRAF MNVKHWPWMK  840
LFFKIKPLLK SAESEKEMAN MKEEFEKTKE ELAKSEAKRK ELEEKMVKLV QEKNDLQLQV  900
QAEADALADA EERCDQLIKT KIQLEAKIKE VTERAEDEEE INAELTAKKR KLEDECSELK  960
KDIDDLELTL AKVEKEKHAT ENKVKNLTEE MASLDETIAK LTKEKKALQE AHQQTLDDLQ  1020
AEEDKVNTLT KAKTKLEQQV DDLEGSLEQE KKLRMDLERA KRKLEGDLKL SQDTIMDLEN  1080
DKQQLDEKLK KKDFEISQIQ SKIEDEQALG MQLQKKIKEL QARIEELEEE IEAERTSRAK  1140
AEKHRADLSR ELEEISERLE EAGGATAAQV EMNKKREAEF QKMRRDLEEA TLQHEATAAT  1200
LRKKHADSTA ELGEQIDNLQ RVKQKLEKEK SELKMEIDDL ASNMESVSKA KANLEKMCRT  1260
LEDQLSEIKT KEEEHQRMIN DLTAQRARLQ TESGEYSRQV EEKDSQISQL SRSKQAFTQQ  1320
IEELKRHLEE EIKAKNALAH ALQSARHDCD LLREQYEEEQ EAKGELQRAL SKANSEVAQW  1380
RTKYETDAIQ RTEELEEAKK KLAQRLQDAE EHVEAVNAKC ASLEKTKQRL QNEVEDLMID  1440
VERTNAACAA LDKKQKNFDK ILAEWKQKYE ETQAELEASQ KESRSLSTEL FKMKNAYEES  1500
LDHLETLKRE NKNLQQEISD LTEQIAEGGK AIHELEKVKK QIEQEKSEIQ AALEEAEASL  1560
EHEEGKILRL QLELNQVKSE IDRKIAEKDE EIDQLKRNHL RIVESMQSTL DAEIRSRNEA  1620
LRLKKKMEGD LNEMEIQLSH ANRQAAEAQK NLRNTQAVLK DTQIHLDDAL RTQEDLKEQV  1680
AMVERRANLL QAEIEELRAA LEQTERSRKV AEQELMDASE RVQLLHTQNT SLINTKKKLE  1740
TDIVQIQGEM EETIQEARNA EEKAKKAITD ASMMAEELKK EQDTSAHLER MKKNLDQTVK  1800
DLQHRLDEAE QLALKGGKKQ IQKLEARVRE LEGDVDAEQK RSAEAVKGVR KYERRVKELT  1860
YQSEEDRKNI LRLQDLVDKL QMKVKSYKRQ AEEAEELSNV NLSKFRKIQH ELEEAEERAD  1920
IAESQVNKLR AKSREFHGKK IEEEE                                        1945
```

```
SEQ ID NO: 199          moltype = AA  length = 1945
FEATURE                 Location/Qualifiers
source                  1..1945
                        mol_type = protein
                        organism = Cairina sp.
SEQUENCE: 199
MTSPDAEMAA FGEAAPYLRK SEKERIEAQN KPFDAKSSVF VAHPKESFVK GTIQSKESGK  60
VTVKTEGGET LTVKEDQIFS MNPPKYDKIE DMAMMTHLHE PAVLYNLKER YAAWMIYTYS  120
GLFCVTVNPY KWLPVYNPEV VLAYRGKKRQ EAPPHIFSIS DNAYQFMLTD RENQSILITG  180
ESGAGKTVNT KRVIQYFATI AASGEKKKEE QSGKMQGTLE DQIISANPLL EAFGNAKTVR  240
NDNSSRFGKF IRIHFGATGK LASADIETYL LEKSRVTFQL KAERSYHIFY QIMSNKKPEL  300
IDMLLITTNP YDYHFVSQGE ITVASINDQE ELMATDSAID ILGFTADEKV AIYKLTGAVM  360
HYGNLKFKQK QREEQAEPDG TEVADKAAYL MGLNSADLLK ALCYPRVKVG NEYVTKGQTV  420
QQVNNSVGAL AKAVYEKMFL WMVVRINQQL DTKQPRQYFI GVLDIAGFEI FDFNSLEQLC  480
INFTNEKLQQ FFNHHMFVLE QEEYKKEGIE WEFIDFGMDL AACIELIEKP MGIFSILEEE  540
CMFPKATDTS FKNKLYDQHL GKSSNFQKPK PAKGKAEAHF SLVHYAGTVD YNITGWLEKN  600
KDPLNETVIG LYQKSSVKTL ALLFANYGGA EAEASGGGAS GGKKGGKKKG SSFQTVSALF  660
RENLNKLMTN LRSTHPHFVR CIIPNETKTP GAMEHELVLH QLRCNGVLEG IRICRKGFPS  720
RVLYADFKQR YKVLNASAIP EGQFIDSKKA SEKLLGSIDV DHTQYKFGHT KVFFKAGLLG  780
LLEEMRDEKL AQLITRTQAR CRGFLMRVEY QRMVERRESI FCIQYNVRAF MNVKHWPWMK  840
LFFKIKPLLK SAESEKEMAN MKEEFEKTKE ELAKSEAKRK ELEEKMVKLV QEKNDLQLQV  900
QAEADALADA EERCDQLIKT KIQLEAKIKE VTERAEDEEE INAELTAKKR KLEDECSELK  960
KDIDDLELTL AKVEKEKHAT ENKVKNLTEE MASLDETIAK LTKEKKALQE AHQQTLDDLQ  1020
AEEDKVNTLT KAKTKLEQQV DDLEGSLEQE KKLRMDLERA KRKLEGDLKL SQDTIMDLEN  1080
DKQQLDEKLK KKDFEISQIQ SKIEDEQALG MQLQKKIKEL QARIEELEEE IEAERTSRAK  1140
AEKHRADLSR ELEEISERLE EAGGATAAQV EMNKKREAEF QKMRRDLEEA TLQHEATAAT  1200
LRKKHADSTA ELGEQIDNLQ RVKQKLEKEK SELKMEIDDL ASNMESVSKA KANLEKMCRT  1260
LEDQLSEIKT KEEEHQRMIN DLTAQRARLQ TESGEYSRQV EEKDSQISQL SRSKQAFTQQ  1320
IEELKRHLEE EIKAKNALAH ALQSARHDCD LLREQYEEEQ EAKGELQRAL SKANSEVAQW  1380
RTKYETDAIQ RTEELEEAKK KLAQRLQDAE EHVEAVNAKC ASLEKTKQRL QNEVEDLMID  1440
VERTNAACAA LDKKQKNFDK ILAEWKQKYE ETQAELEASQ KESRSLSTEL FKMKNAYEES  1500
LDHLETLKRE NKNLQQEISD LTEQIAEGGK AIHELEKVKK QIEQEKSEIQ AALEEAEASL  1560
EHEEGKILRL QLELNQVKSE IDRKIAEKDE EIDQLKRNHL RIVESMQSTL DAEIRSRNEA  1620
LRLKKKMEGD LNEMEIQLSH ANRQAAEAQK NLRNTQAVLK DTQIHLDDAL RTQEDLKEQV  1680
AMVERRANLL QAEIEELRAA LEQTERSRKV AEQELMDASE RVQLLHTQNT SLINTKKKLE  1740
TDIVQIQGEM EETIQEARNA EEKAKKAITD ASMMAEELKK EQDTSAHLER MKKNLDQTVK  1800
DLQHRLDEAE QLALKGGKKQ IQKLEARVRE LEGDVDAEQK RSAEAVKGVR KYERRVKELT  1860
YQSEEDRKNI LRLQDLVDKL QMKVKSYKRQ AEEAEELSNV NLSKFRKIQH ELEEAEERAD  1920
IAESQVNKLR AKSREFHGKK IEEEE                                        1945
```

```
SEQ ID NO: 200          moltype = AA  length = 615
FEATURE                 Location/Qualifiers
source                  1..615
                        mol_type = protein
                        organism = Cairina sp.
SEQUENCE: 200
MKWVTLISFI FLFSSATSRN LQRVARDADH KSEIAHRYND LKEETFKAVA MITFAQYLQR  60
CSYDGLSKLV KDVVDLAQKC VANEDAPECS KSLPSIFLDE ICQVEKLRDS YGAMADCCAK  120
```

-continued

```
SDPERNECFL SFKIPQPDFV QPYQRPASDV ICKEYEDNRV SFLGHFIYSV ARRNPFLYAP  180
TILGVAADYE HALKSCCKES DVGACLDGKE TAIREKVKKI SVKQQYSCGI LQKFGDRIFQ  240
ADKLAFLSQK YPKASFTEIS KLIHDVKDVY KECCEGDMVE CMDDRAELVN YICSKHESFS  300
SKIKDCCEKP IVERSQCIME AEFDEKPADL PSLVEKYIQD KEVCKSFEAG HDAFMSEFVY  360
EYSRRHPEFS TQLILRIAKG YETLLEKCCK TDNPAECYAN AQDQLNEHIK ETQDVVKTNC  420
DLLTTHGEAD FLKALLIRYT KKMPQVSTET LLEIGKKMTG VGTKCCKLPE DRRLPCSEGY  480
LSMVIQDMCR RQETSPINDN VSHCCSDSYA YRRPCFSAMG VDTKYVPPPF NPEMFNFDEK  540
LCSAPAEERA AGQMKLLVNL IKRKPQMTEE QIKTIADGFT AMVDKCCKQS DIETCFGEEG  600
ANLIVQSRAT LGIGA                                                   615

SEQ ID NO: 201          moltype = AA  length = 381
FEATURE                 Location/Qualifiers
source                  1..381
                        mol_type = protein
                        organism = Salmo salar
SEQUENCE: 201
MPFGNTHNNF KLNFKVEEEY PDLTKHNNHM AKVLTKDMYA KLRDKQTPSG FTLDDVIQTG  60
VDNPGHPFIM TVGCVAGDEE SYEIFKDLLD PIISDRHSGY KPTDKHKTDL NFENLKGGDD  120
LDPNYVLSSR VRTGRSIKGY TLPPHNSRGE RRAVEKLSVE ALNTLDGEFK GKYYPLNKMT  180
DAEQEQLIAD HFLFDKPVSP LLLGAGMARD WPDARGIWHN DAKSFLVWVN EEDHLRVISM  240
EKGGNMKEVF RRFCVGLKRI EETFKKHNHG FMWNEHLGYV LTCPSNLGTG LRGGVHVKLP  300
KLSTHPKFEE ILTRLRLQKR GTGGVDTASV GGVFDISNAD RLGSSEVDQV QMVVDGVKLM  360
VEMEKKLEKG EAIDGMIPAQ K                                            381

SEQ ID NO: 202          moltype = AA  length = 895
FEATURE                 Location/Qualifiers
source                  1..895
                        mol_type = protein
                        organism = Salmo salar
SEQUENCE: 202
MTAIETQVQY GSYMTSEQVY MTQEDDWDRD LLLDPAWEKQ QRKTFTAWCN SHLRKAGTQI  60
ENIEEDFRNG LKLMLLLEVI SGERLPKPDK GKMRFHKIAN VNKALDFICS KGVKLVSIGA  120
EEIVDGNVKM TLGMIWTIIL RFAIQDISVE ETSAKEGLLL WCQRKTAPYR NVNVQNFHIS  180
WKDGLALCAL IHRHRPDLID YSKLRKDDPM GNLNTAFEVA EKYLDIPKML DAEDIVNTPK  240
PDEKAIMTYV SCFYHAFAGA EQAETAANRI CKVLAVNQEN EKLMEEYEKL ASELLEWIRR  300
TIPWLENRVA EQTMRAMQQK LEDFRDYRRV HKPPKVQEKC QLEINFNTLQ TKLRLSNRPA  360
FMPSEGKMVS DIANAWKGLE QVEKGYEEWL LTEIRRLERL DHLAEKFKQK SSLHESWTSG  420
KVELLSMKDY ESASLMEIRA LMRKHEAFES DLAAHQDRVE QIAAIAQELN ELDYHDAVTI  480
NARCQGICDQ WDNLGTLTQK RRDSLERVEK LWETIDQLYL EFAKRAAPFN NWMDGAMEDL  540
QDMFIVHSIE EIQSLITAHD QFKATLPEAD KERVATMGIQ NEIVKIAQTY GIKLSGVNPY  600
TNLSPQDITD KWDAVKHLVP LRDQMLQEEV ARQQSNERLR RQFAAQANII GPWIQTKMEE  660
ISHVSVDIAG SLEEQMSNLK QYEQNIINYK CNIDKLEGDH QLSQESLIFD NKHTNYTMEH  720
VRVGWEQLLT TIARTINEVE NQILTRDAKG ISQEQLNEFR ASFNHFDRKR NGMMDPDDFR  780
ACLISMGYDL GEVEFARIMT LVDSNNTGVV TFQAFIDFMT RETAETDTAD QVMASFKILA  840
SDKTYITVEE LRRELPPEQA DYCISRMTSY IGSGAPPGAL DYISFSSALY GESDL       895

SEQ ID NO: 203          moltype = AA  length = 843
FEATURE                 Location/Qualifiers
source                  1..843
                        mol_type = protein
                        organism = Salmo salar
SEQUENCE: 203
MPKPLTDQDK RKQISVRGLA GVENVSDLKT NFNRHLHFTL VKDRNVATKR DYYFALANTV  60
RDHLVGRWIR TQQHYYEKDP KRVYYLSLEF YMGRTLQNTM VNLALENACD EATYQLGLDM  120
EELQDIEEDA GLGNGGLGRL AACFLDSMAS LGLAAYGYGI RYEFGIFNQK IVNGWQVEEA  180
DDWLRYGNPW EKARPEYMRP VHFYGRVEHT PEGVKWVDTQ VVLALPYDTP VPGYRNNVVN  240
TMRLWSAKAP CDFNLKDFNV GGYIQAVLDR NLAENISRVL YPNDNFFEGK ELRLKQEYFV  300
VAATLQDIVR RFKSSKFGST EVVRVDLSTL PDKVAIQLND THPAMAIPEL MRILLDTEHQ  360
TWEKAWDICT RTCAYTNHTV LPEALERWPT DLLQNLLPRH LEIIYEINRR HLERIAKLYP  420
GDHDRLRRMS LVEEGDQKKI NMAHLCIVGS HAVNGVARIH SDILKATLFK DFYEVDPHKF  480
QNKTNGITPR RWLVMCNPGL AEVIAERIGE DYIRDLDQLK KLLTFVDDDS LIHDIAKVKQ  540
ENKLKFSAYL EEHYKVKINP NSMFDVQVKR IHEYKRQLLN CLHIITLYNR IKKEPNKNWT  600
PRTIMIGGKA APGYHTAKLI IRLITAIGEI VNHDTVVGDR LKVIFLENYK VTLAEKIIPA  660
SDLSEQISTA GTEASGTGNM KFMLNGALTI GTMDGANVEM AEEAGEENLF IFGMRVDDVD  720
AMDKSGYDAM DYYNRIPELK QAMDQIAGGF FSPDQHDLFK DIVNMLLHHD RFKVFADYEA  780
YIKCQEKVSA LYKNPKEWTK MVIHNIAGCG KFSSDRTISQ YAREIWGMEP SLERIPAPDD  840
PRQ                                                                843

SEQ ID NO: 204          moltype = AA  length = 363
FEATURE                 Location/Qualifiers
source                  1..363
                        mol_type = protein
                        organism = Salmo salar
SEQUENCE: 204
MPHSFPFLST EQKKELSDIA QKIVAPGKGI LAADESTGSV AKRFQSINAE NTEENRRLYR  60
QLLFTADRVE PCIGGVIFFH ETLYQKTDDG KLFPQLIKER DMVVGIKVDK GVVPLAGTNG  120
```

-continued

```
ETTTQGLDGL YERCAQYKKD GADFAKWRCV LKITDTTPSE LAIKENANVL ARYASICQMH  180
GIVPIVEPEI LPDGDHDLKR CQYVTEKVLA AVYKALSDHH VYLEGTLLKP NMVTPGHSCS  240
QKYSNQEIAM ATVTALRRTV PPAVPGITFL SGGQSEEEAS INLNAMNQCP LHRPWAITFS  300
YGRALQASAL KAWGGKPENG KACQEEFIKR ALNNSKACQG KYVSSGDKGA AGGESLYVAN  360
HAY                                                                363

SEQ ID NO: 205          moltype = AA  length = 193
FEATURE                 Location/Qualifiers
source                  1..193
                        mol_type = protein
                        organism = Salmo salar
SEQUENCE: 205
MAPKKDAKAP AKKAEPAKAA APAPAPEPDV VAAPPPLDLS TVKVEFTPDQ QEDYKEAFGL  60
FDRVGDAKVA YNQVADIMRA LGQNPTNKEV RKVLGNPSDE DMAAKRLEFE AFLPMLQHIV  120
NDPNKGTFDD YVEGLRVFDK EGNGTVMGAE LRIVLGTLGE KMSEAEIDAL MQGQEDENGS  180
INFEAFVKHI MSI                                                     193

SEQ ID NO: 206          moltype = AA  length = 334
FEATURE                 Location/Qualifiers
source                  1..334
                        mol_type = protein
                        organism = Salmo salar
SEQUENCE: 206
MVKVGVNGFG RIGRLVTRAA FHSKKGVEIV AINDPFIDLD YMVYMFKYDS THGRFHGEVK  60
AEGGKLVIDG HKITVFHERD PANIKWGDAG ATYVVESTGV FTTIEKASTH LKGGAKRVVI  120
SAPSADAPMF VMGVNHEKYD NSLKVVSNAS CTTNCLAPLA KVIHDNYHII EGLMSTVHAV  180
TATQKTVDGP SGKLWRDGRG ASQNIIPAST GAAKAVGKVI PELNGKITGM AFRVPTPNVS  240
VVDLTVRLEK PASYDAIKKV VKAAADGPMK GILGYTEQQV VSSDFNGDTH SSIFDAGAGI  300
ALNDHFVKLV TWYDNEFGYS NRVIDLMAHM ATKE                              334

SEQ ID NO: 207          moltype = AA  length = 284
FEATURE                 Location/Qualifiers
source                  1..284
                        mol_type = protein
                        organism = Salmo salar
SEQUENCE: 207
MDAIKKKMQM LKLDKENALD RAEGAEGDKK AAEDKSKQLE DDLVALQKKL KGTEDELDKY  60
SESLKDAQEK LEVAEKTATD AEADVASLNR RIQLVEEELD RAQERLATAL TKLEEAEKAA  120
DESERGMKVI ENRASKDEEK MELQDIQLKE AKHIAEEADR KYEEVARKLV IIESDLERTE  180
ERAELSEGKC SELEEELKTV TNNLKSLEAQ AEKYSQKEDK YEEEIKVLTD KLKEAETRAE  240
FAERSVAKLE KTIDDLEDEL YAQKLKYKAI SEELDNALND MTSI                   284

SEQ ID NO: 208          moltype = AA  length = 1138
FEATURE                 Location/Qualifiers
source                  1..1138
                        mol_type = protein
                        organism = Salmo salar
SEQUENCE: 208
MPEAKKPAAK TDKAAAAPAD KEARSAIEEG SECRDELPDD ETVPGGQSEL TGLFVERPES  60
TTVIKGKNVT FVAKVDSSDL LRKPNMKWLK GKWLDLGSKA GKHVQFKEAY DRNSKVYTYE  120
MSIIKVVEGD AGGYRCEVTS KDKCDSCTFE VTVEAVQEEQ PANILDAFKR SGDAGEDAGD  180
LDFSALLKKR EKKARDEPKE DVDVWEILKD AKPCDYEKIA FDYGITDLRG LLKRLKKMKK  240
VEPKKSDAFL KKLEQCYSVD KGKRTQMHVE LHDPNVQVKW LKNGVEIKPS AKYVFECVGN  300
KRTLTINKCN LSDDAAYECV VGEEKSFTEV FVKEPPVTIT KLLDDVHVVV GEKVEFECEV  360
SEEGANVKWM KDGVELTKDG KIRIKKDGKK HTLVINEATI EDIGMYYVYT NGGESKGELE  420
VEAKELEVLQ SIADLSVKAC EQAMFKCEVS DEKVVGKWFK DGVEVKPGNR IKMSHIGRIH  480
KLTIDDVKPQ DEGNYTFVPE GYALSLSAKL NFIEIKIEYV PRQDPPKIHL DTSSTGSKNT  540
IVVVAGNKLR LDVEITGEPV PTVCWMKGDT VISEAEGRVR VETRTTLSSF VIEGAERPDE  600
GQYSIIVTNP AGEDRAELTV KIVDVPNPPE NVRCMGVGED TATITWDPPK FDGGAPIKGY  660
LMERKKQGSS RWTKLNFEVF ESTTYEAKKM IEGVFYEMRV FAVNGIGISQ PSGNSKPFMP  720
IAPTSEPTRL TVEDVTDSTC ALKWRPPERV GAGGVDGYII EWCKEGEDNW VAANKEPVDK  780
NTYRVKGLPT GEKLLFRVVA MNIAGRSPPC TMKQSVTIRE IMEYPKIRLP RQLRTKFIRK  840
VGEKINLVIP FQGKPRPVVN WLKDGEPLEN KSVGIRTSDF DTILFIRSAE RDHSGTYTLS  900
VQIDNMQDKA DIHIQIVDKP GPPINVMVTD VWGFNAALEW KPPKDTGNTD ITGYTIQKAD  960
KKTQGWFTVY EHNRRPSCTA SDLVMGNEYS FRVFSENICG LSDEVAFSKN TAIIGKTDLE  1020
YNKPAFKERD MSSSPKFTAP LVDRVVVAGY STAISCAVRA YPKAKIVWMK NKMIIGEDPK  1080
FLMQNNQGVL TLNIRKPGQF DGGKYSCKAI NDLGEDEVEC RLEVRVLKEK KEGDEEKK    1138

SEQ ID NO: 209          moltype = AA  length = 1502
FEATURE                 Location/Qualifiers
source                  1..1502
                        mol_type = protein
                        organism = Salmo salar
SEQUENCE: 209
MATKVMHFNQ KKHVSHVSHH SSHTTKHVVK EQSSRKSSSK SVNQVQHTHA TEAMAEPAYT  60
VPAFRQRSAD EIEEYQRVSS HVGKGLASIQ KELHRMRVVT KTQVENIAIK REVQEMMHKK  120
```

-continued

```
MTLSTETDKA PDFMVALRPH TVWEKTPVKL FCTVDGHPRP IVKWYKGGKP VDPLSAPGKY  180
KIESKYGVHS LIISRCMVSD TAEYSAVATN SHGSTTSKAS VIVKRPAGGA YGSCQLFGQV  240
PHLTVIHHSK LEITMLDRFG VSFGTEGGSI SLVCTMVVVP DLPNVPPLAQ WYRDDKLLKA  300
GKLAEIKVGG GAATLTLPHL AKDDEGLYTL RMWTKDGTTE HSAYLFVKDA APSVAGAPGA  360
PISVKAFDIN SDYVLVAWKP PNTTNEAAIT GYFVDKRESG SATWSQCNDA PVQICKYPVH  420
GLNVGHAYHF RVRAVNSAGI SRPSRESDKV TALDPAERER LQVIKLDGKH EVVIKDDDLE  480
GYVTLPGEPT DMHASEIFKS YIVLSWKPPS PRGRAPLWYI IEKSVSGTGA WQRINTAVKL  540
QSPRYPVFDL EAGKKYQFRV YSENLYGASE ASKPTDPIEK VDEHGVPSAP GQVVATRNTK  600
SSVFVQWDPP KHPNHLMGYY IDASLVGSKT WAPCNHKPYK NTRFVVHGLT PGETYVFRVQ  660
AVNVYGLSDE SQESTPIAVE PALTTPSAPH GITMLSCDGA SMIIAWNSPK HRGGSKINAY  720
YVDKRDADSL AWKEVNSAPT NTRTYTVDGL TEGTFYEFKV QAGNLAGVGV PSAPSTPLKC  780
EAWTSAVPGP AYDLAFREVR GDSLVILWKA PVYTGASAVT GYIVEMAKKG HHYHPLNEEA  840
IGHCYLQVTG LEAGAEYTFK VKAVNAEGVG KASQTSEPVF AKALPGTQEI QCGVDEDTGD  900
IFLSFEACEI SETSNFAWSK NYKEISDCTR VSIETKGKHS KLTFVNPDVS DLGAFSVHVT  960
DTDGVSASHT VTEDALNTML ELSYNIRHPI VPLKHQLNYE VLEKGHVRFW LQCLKMSPAV  1020
NYRFIVNDKE VTSSDSHKIS HDVNTGVIQM TVDHFSKASE ATYTVQIHDG RAKNQSSLVL  1080
VGDVFKAALK EAEFQRSEHI RKQGPHFAEY LSYHVDDDCT VLLVCKVANV KKETVFHWFK  1140
DEEEIVPETP PNVMSGSVAL PISLFSRKDQ GHYKAKLSDD RGKDTSVFDI SGQVFDDIIN  1200
AIAHIAGSSA SELVLQCTPE GIRLQCYMKY YTEEIRTAWL HKDSKISSSE KMRIAGTAEM  1260
AWMQIREPSE KEKGHYSIQI MDATKSHTRT FDLSGQAYTD AYEEYLRLKA AAFAEKNRGR  1320
VVGGLPDVVT IMEQKTLSLT CTVWGDPSPE VTWFKNENEV VSTEHAKITL EANKFASLTI  1380
TAVTSEDSGK YSINVRNKYG GEFVEITVSV YKQGESPPEP KMGGRGATPA PLASKTPAKT  1440
PTPAQTPTPS LKSPTPSLKS PTPSLRSPAK SPTPTPKSPT PPRRVKSPSP ARSLKSPTPP  1500
RK                                                                 1502
```

```
SEQ ID NO: 210          moltype = AA  length = 434
FEATURE                 Location/Qualifiers
source                  1..434
                        mol_type = protein
                        organism = Salmo salar
SEQUENCE: 210
MSITKIHARE ILDSRGNPTV EVDLYTAKGR FRAAVPSGAS TGVHEALELR DGDKSRYLGK  60
GTVKAVDHVN KDIAAKLIEK KFSVVDQEKI DHFMLELDGT ENKSKFGANA ILGVSLAVCK  120
AGAAEKGVPL YRHIADLAGH KDVILPCPAF NVINGGSHAG NKLAMQEFMI LPIGASNFHE  180
AMRIGAEVYH NLKNVIKAKY GKDATNVGDE GGFAPNILEN NEALELLKTA IEKAGYPDKI  240
IIGMDVAASE FYKAGKYDLD FKSPDDPARY ITGDQLGDLY KSFIKGYPVQ SIEDPFDQDD  300
WAAWTKFTAA VDIQVVGDDL TVTNPKRIQQ AVEKKACNCL LLKVNQIGSV TESIKACKLA  360
QSNGWGVMVS HRSGETEDTF IADLVVGLCT GQIKTGAPCR SERLAKYNQL MRIEEELGAK  420
AKFAGKDYRH PKIN                                                    434
```

```
SEQ ID NO: 211          moltype = AA  length = 170
FEATURE                 Location/Qualifiers
source                  1..170
                        mol_type = protein
                        organism = Salmo salar
SEQUENCE: 211
MAPKKAKRRG AAAEGGSSNV FSMFEQSQIQ EYKEAFTIID QNRDGIISKD DLRDVLASMG  60
QLNVKNEELE AMVKEASGPI NFTVFLTMFG EKLKGADPED VIVSAFKVLD PEATGFIKKE  120
FLQELLTTQC DRFSAEEMKN LWAAFPPDVA GNVDYKQICY VITHGEEKEE             170
```

```
SEQ ID NO: 212          moltype = AA  length = 1979
FEATURE                 Location/Qualifiers
source                  1..1979
                        mol_type = protein
                        organism = Salmo salar
SEQUENCE: 212
MTTLGLTEGH PAKYQHLTNL VPAISYELGQ KKERVIMSTD AEMAAFGPAA IYLRKPERER  60
IASQAAPFDA KTAFFVVEPK EMYLKGVLQS KEGGKATVKT LCNKVLTVKE DDIHPMNPPK  120
YDKIEDMAMM THLNEATVLY NLKERYAAWM IYTYSGLFCV TVNPYKWLPV YDSVVVNAYR  180
GKKRIEAPPH IFSISDNAYQ FMLTDHENQS ILITGESGAG KTVNTKRVIQ YFATIAVAGG  240
KKEQTAAATS GKIKGSLEDQ IIAANPLLEA YGNAKTVRND NSSRFGKFIR IHFGTTGKLA  300
SADIETYLLE KSRVTFQLSA ERSYHIFYQL MTGHQPQLLE ALLITTNPYD YPIISHGEIA  360
VKSIDDTEEF IATDTAIDIL GFTAEEKIGI YKLTGSVMHH GAMKFKQKQR EEQAEPDGTE  420
EADKISYLMG LNSADLLKAL CYPRVKVGNE MVTKGQTVPQ VNNSTMALCK SVYEKMFLWM  480
VVRINEMLDT KQSRQFFIGV LDIAGFEIFD YNSLEQLCIN FTNEKLQQFF NHHMFVLEQE  540
EYKKEGIEWA FIDFGMDLAA CIELIEKPMG IFSILEEECM FPKASDTTFK SKLYDQHIGK  600
TKAFEKPKPG KGKAEAHFAL VHYAGTVDYN VTGWLDKNKD PLNDSVVQLY QKSSVKLLAM  660
LYVGAAASQA DDAARGGGGK KKKGGSFQTV SALFRENLGK LMTNLRSTHP HFVRCLIPNE  720
SKTPGLMENF LVIHQLRCNG VLEGIRICRK GFPSRIQYGD FKQRYRVLNA SVIPDGQFID  780
NKKASEKLLG SIDVDKSQYK FGHTKVFFKA GLLGTLEEMR DEKLATLVTM TQALCRGYVM  840
RKEFVKMMAR REAIYSIQYN IRSFMNVKHW PWMKLYFKIK PLLKSAETEK EMAAMKENFE  900
KMKEDLAKAL AKIKGLEEKI VSLLQENNDL QLQIAAEENT LCDAEERCEG LIKSKIQMEA  960
KLKETFERLE DEEELNAELT AKKRKLEDEC SELKKDIDDL ELTLAKVEKE KHATENKVKN  1020
LTEEMSSQDE SLIKLTKEKK ALQEVHQQVL DDLQAEEDKV NNLTKAKTKL EQQVDDLEGS  1080
LEQEKKVRMD LERAKRKLEG DLKLALESLM DLENDKQQSD EKIKKKDFET SQLLSKIEDE  1140
QALGAQLQKK IKELQARIEE LEEEIEAERA ARAKVEKQRS DLSRELEEIS ERLEEAGGAT  1200
```

-continued

```
SVQIEMSKKR EAEFQKLRRD LEESTLQHEA TAAALRRKQA DTVAELGEQI DNLQRVKQKL  1260
EKEKSEYKME IDDLSSNMEA IAKAKGNLEK MCRSLEDQLS ELKTKNDEHV RQLNDINVQR  1320
ARLLTENGEL GRQMEEKESL VSQLTRAKQT FTQQIEDLKR QIEEEVKAKN ALAHGLQSAR  1380
HDCDLLREQY EEEQEAKAEL QRALSKANSE VAQWRSRYET DAIQRTEELE DAKKKLTQRL  1440
QDAQEAVEAT NAKCASLEKT KQRLLGEVED LMIDVERANA SAAQLDKKQR NFDKVLSEWK  1500
QKYEEGQAEL EGSLKEARSL STELFKMKNS YEEALDHLET LKRENKNLQH EISDLTELVG  1560
ESGKSIHELD KAKKTVENEK AEIQAALEEA EGTLEHEESK ILRIQLELNQ IKGEVDRKLA  1620
EKDEETEQIK RNSQRMMDSM QSTLDSEIRS RNDAMRVKKK MEGDLNEMEI QLSHANRQAA  1680
ESQKQLRNVQ GALKDAQLQL DDAIRVADDM KEQVAMVERR NNLMMAEIEE LRAALEQTER  1740
GRKVAEQELV DASERVGLLH SQNTSLINTK KKLEVDLTQV QGEMEDTVQE ARNAEEKAKK  1800
AITDAAMMAE ELKKEQDTSS HLERMKKNLE VTVKDLQHRL DEAENLAMKG GKKQLQKLEA  1860
RVRELESEVD AEQKRGVEAI KGVRKYERRV KELTYQTEED NKNVSRLQDL VDRLQLKMKA  1920
YKRAAEDAEE QSNIHLSRYR KVQHELEEAQ ERADISESQV NKLRAKSREI GSKGNVAEE   1979

SEQ ID NO: 213          moltype = AA  length = 5447
FEATURE                 Location/Qualifiers
source                  1..5447
                        mol_type = protein
                        organism = Salmo salar
SEQUENCE: 213
MDENEEFEEE FEEEEIIEEE YFEETVTSTT SHQTTSRAAA VTQQPGVPVR RVVRKKVKVD  60
TSKFMTPYLA HSQKMLDLFS YNKYRRAYDK AKGKPYAITA DTPEMMRIKK AQEQLSEVKY  120
RMEGNKVRCT SLYDGEAREI AHVKHVSELI SKVLYRAKWD ETKDRYLLPP DAPELVLAVK  180
NAANYSKKLY TEQWDEDKTM CYPYSDSPEL RRVAKAQEVL SDILYKKGHA ERTAKYTSLA  240
VPNDMELAKK NFEHRSDLKY HEDYNKNVKG HWCETPYIEV AVARVAMDNL SNKMYTAAYE  300
DIKDQINFMQ TETPEYKQHK KAGHAASSIK YKKDYENSKS VADYNVLPAT DNPLLRQLRY  360
AGTILSDKVY KASYEKSRGS SINYCDTPKF KMDSVVQQFS DAHYKDKYEN EVKGHYIGSY  420
EDIYMQHCKK MEEIKSEQNY KADYEDIKTR CFYPQTMSPE YEAHKKVAQC KDKVYRQHPD  480
TVKFTQVTDS PVQLQAAINA KQLSDLNYHA AYELDKFKCA LPPDYPFFIQ SRANAYNLSD  540
SCYRYDWEKV KAKQFEVKGD AISILAARAH TNIASDVKYK KEFEKNKGKM VGALSIHDDP  600
KILHSMHVAK IQSDREYKKN YEKTKTKYIT TLDMMSNTLA KKSQGIASMT GYRTIHNRYL  660
LPHDAMSLDL AKKANIIQSD NEYHSDYNNS TKGTGWVAFG SMEVEKAKKA GDILNERKYR  720
QHPDTVPFTS IADHPVMMQA QVNQIMRSDL HYKKGLEEIH QKYSLPHNDP QFLQAKCNAY  780
NISKNYYKQS WEETIAKGYD MRPDAISVIH AKHGRHIASD IQYKKDYVKG KGKQIGCRNL  840
QDDPLLVHYM EVARLQSDKN YKKDYHKSKL KYHSPVDMMS VVHAKHASTV QTFAGYRKIQ  900
HHYSLLPGAM NLELARTMNT NASDVNYKAD WENSVKGTGW IPIGSIPVEL AKQAGAVLDE  960
HKYRQHPENF KFSSKMDGMN MELCKHNQKI MNVQEYKASG EKFLHTYNLP ADAPEFMQAK  1020
CNAVNISKNY YTYAHRQDIL KGHQVKADAI PILAARACTN IASNYKYKLA YEKGRGKHVG  1080
FRNVKDDPLL VHYMEVAKMQ SEQNYKKDYH KSKLKYHSPV DMMSLVHAKH ASAAQTFAGY  1140
RKSLHNYTLL PEAMPVQLAR TMMEIQSNNA YKSDYNNMCR GSGWVPIGSM QVELAKTAMA  1200
ARDEQGYRVH PSTLKFTSLT GQMNMALAMT NTKLLDDASY RASGEKFKHT YNLPADSPQF  1260
LQAKCNAQTM SESHYKHKWV EDIAKGYDIR PDAISVIHAK HGRHIASDIQ YKKDYVKGKG  1320
KQIGCRNLQD DPLLVHYMEV AKMQSDKNYK KDYHKSKLKY HSPVDMMSLE HAKHASTVQT  1380
FAGYRKIQHH YTLTADSMNL ALARTMNASS SDYEYKSDWN NSVRGTGWIP IGSLPVELAK  1440
VGGDIQSEQK YRTHPSKFQF KKLMDSMDIV LATDNNKIMN KQEYTKQWHN DKLTVHVMPD  1500
APEIVLAKAN AINMSQKTYK AGLAKMYKNG YDLKVDAIAI QAAKQGTKIA SDYKYKLSHR  1560
MNRGKQIGCR NLQDDPLLVH YMEVARLQSD KNYKKDYHKS KLKYHSPVDM MSVVHAKHAS  1620
TVQTFAGYRK IQHHYTVLAD AMNVELAKQM QTIASDNLYK SDYETSVKGT GWIPIGSLDV  1680
EKAKTAGLIG SEKLYRQHPS SFKFTKDMHS MDVVLANANN AIMNKQTYVD AWNKDKTSVH  1740
IMPDSMEVVL ARANKINCSA TLYKLANEQD KKKGYDLRND AIAILAAKQG TSIASDFKYR  1800
AGYRKQVGHH IGARCVEEDP LLMLAINSAK IASDALYKKD FHQSKTKFHL PLDMIAFELA  1860
KKNQIQVNHA KYVTRLHNWT CLPDSTDVVH ARHVYDIASD AVYKADLKCL QGIGWVPIGS  1920
LDVEKAKKAG EVLSEKLYRQ HPSNFKFTSS VQDIGMVLAK QNNDNANKQA YTEAWETDKT  1980
KNHIMPDAMD IVLANANKSN YSWNQYRKAM VEDKKKGYDM RNDAISIIAA KHGREIISDH  2040
KYKSSYRKQV GHHIGARCVE DDPLLVLALN SAKIASDALY KKDFHQSKTK FHLPLDMIAF  2100
ELCKKNQIQV NDFNYRTHLH NWTCLPDSTD VVHARHVYDI ASEAVYRADM KWLRGIGWVP  2160
TGSLDVEKAK KAGEILSEKK YRQHPSTLDF KSPLDGMDIV LAKHNGQIMN KRQYVEQWDA  2220
DKIKIHVGNE LPELVLAKAN AITMSNKLYK QGFDETIKKG YYLPVDAISV KAAKHGREII  2280
SDYKYKAGYS KQVGHHIGAR CVEEDPLLML AINSAKIASD ALYKKDFHQS KTKFHLPLDM  2340
IAFELCKKNQ IQVNDFNYRT HLHNWTCLPD STDVVHARRV YDLHSDAVYR ADMEFIRGTG  2400
WVPIGSLDVL KAKKAAEILS DRLYKQKPDS LKYTTDMHSM PMVLAKANAE IINKGNYVAQ  2460
WEADKIKIHV GNELPELVLA KANAINNSYK LYKAELEELY KKGYNLKSDC IAIKAAKHGR  2520
EIISDYKYKA GYSEQVGHHI GARCVEEDPL IMLALNSAKI ASDALYKKDF HQSKTKFHLP  2580
VDMLTIELAK KCQMQVNDFN YRTRLHNWTC LPDSSDVVQA RKVYALHSDA VYRADMEWLK  2640
GTGWVPIGSI DVEKAKKAGE ILSERKYREH PSNFKFTATT DDMPYALATA NAQIMDKRQY  2700
VQQWEADKIQ IHVGNELPEL VLAKANAITM SNKLYREGFN EMIGQGYHLD KDAIAVQHAK  2760
KSREIISDYK YKAGFRNQCG HHIGAHCVED DPLLVLAANA ARISSDINYK KDFHLTKTKF  2820
SIPLDMMSFD LAKKSQFQIQ EFTYRTKLHN WTCMPDNTDV VRARKVYSLH SDAVYRADLE  2880
WLRGCGWLPQ GSVDVIKAKN AQTILNERLY RQPPSSIKFT SVVDSPEVVL SKQNADILSD  2940
RKYREVWDTE KTNIHLNNDL PSLVLSKANA IAISNKLYTK DWDADKAKGY HIKEDAIAVL  3000
KARKSRDIIS DYKYHEGYRK QVGHHIGARC VEEDPLIMLA LNSAKIASDA LYKKDFNKSK  3060
TKFHLPVDML TIELAKKCQV QVNDFNYRTH LHNWTCLPDS TDVVQARKTY DLQSDAVYKA  3120
DLEWLRGCGW NPHESVEVLK VKHAQAVLAD RGYRVKLDQQ KFTAPVDRID MVCAKNAAEI  3180
LNEANYRALW HTEKTKYTIT DTPGLATARE VAKNLHPKLY TKDWDKVKAT GFFMPGDAVP  3240
IKQCIHNNKV QSDHKYKADY RKQAGHHIGA RCVEDDPLIM LAINSAKIAS DALYKKDFNK  3300
SKTKFHLPVD MLTIELAKKC QVQVNDDHYR TRLHNWTCLP DSTDVTQARK AYDLQSDVVY  3360
```

-continued

```
KADLEWLRGC GWVAADCVDH VKVRNAQKIM NERLYKKDAK DNFAKYTHIV DRPEIILATI  3420
NAFNLSDLKY KETFNAEKGK YIGSEDTPQL AHSREVAKVI SEKYYKLAWD ESKATGYQLD  3480
AHYIPVVGAK KTREIVSDVK YHDTHEKAKG HYLAGTLVDF PEVMRCGEQE KQKGLRDYHK  3540
EYNTNKAKNH IPADIIPNMV AKRCQDMLSD VLYRTLLHNW TCHPDQQDII RARKNNEILS  3600
DVYYKDDLNW IKGIGCYAWD TPELVRAKNA HLLQSDLKYK EEGKKEFNNY SIVTDTPVYV  3660
TALLGHTWAS DLNYREAYHK DKHMYTTILD TYNYAHNATL KDIYSTKTYT SAWDKIKAKD  3720
YTLPSDSHAL KHAKQQKVIL SNVKYREDYE KFKSLYTLPK SLEDDPQTLR CLKAGKMGLD  3780
RLYKADYEKT KAKNHIPPDM LDVLSARQTQ NYVSEIGYRK HLHQWICMPD MQVYVQARKV  3840
NEQLSDIFYK DDLNWLRGIG CYAWDTPEIL RVKAADKLQS DNKYRAKGIE AFKDYSVVID  3900
TPTYETAKLS SQNLSDLHYR NDYATNVKGT NSSPAVTVDT ERARLANFNQ SDNLYKEANK  3960
NFMPTGYSLP HDNPLMKQAK QNSIVASSVK YKNDYEKSKA TNYTMDPDGV SFVNIRNANK  4020
VTNPRLYREL YEKGKGKIHS TYDTPEIRQV KMTQEAVSDL CYKEKYFANR GTMIPMGVTP  4080
QMIHCNHVNE ITSDLRYKED LLWLRGIGCF LYDTPEMSTV RNITKFRSNY EVEAKKNHAN  4140
FTVVLDTPEY QRVTELKTHM SDLMYRAQGN VEKTKCTTTL DSLDVKRVKW AQQLSNKYMY  4200
TDLASKERAH FTPEKDTPIM SHARAMKVVF SEIKYKDQYE KMKHKYTSIH DTPILVRAKK  4260
AYLQSSDLRY KESFERSKGH YHSVKDALDI VCHRRVTDDI SEVKYREKYI NSLGTWRSMP  4320
DRPEFFHSRI IGENVSDIKY KEDLDWLKGI GCYVWDRPEM VQAEKNKVLF SERHYKETFE  4380
KNRHRFKYTC DTPFFQAAKK ASVIINDRHY RAAYEKTKDK YSSVLDDPRN LLAKENAKNS  4440
QWKYREQYER AKNKFTSILE TPEHESHKRS KAIGDNVYKM DYNKNKAKGY TLGHDTPHST  4500
HMKKVKEITS VLKYKELYEK SKAQINMDPE SFEIRAAKEA YKNISNLDYR KKYQATKNKW  4560
IWTVDRPDFI QAAKANFQQS DIEYTYDKEM MKGCVISIVD DKLAVLAKMN ADLSNVKYRE  4620
KFEKAKGQYK TVLDTPTNLH AKSVRILASE SKYKLAGKME QQTGQFTTLP TTRDTIHAKD  4680
MGKLVSEKQY KQKFEKEKGK SKYNSMRVPP DVQHAMDVAK NQSNFAYRQE AKAKLNYTSV  4740
AMRPDIMRAT QASKLASNIG YREKARQEAA HGGSLVHRPD IALATAVSKL NSQVQYKAKY  4800
DKEKGLRPKY DMKESVVYKT QKDAHQLASE VKYKGDLKKL HKPVTDMAES LSMQHNLSTS  4860
KLSSDYQYKK KHEDSKGHYL MIADSPEQLH HKEASELQSH VKYKEKYEKA KGKAMLDFET  4920
PTYVTAKEAQ HMQSAKDYHK EYEESIKGKN LSGLEVTPAM LHVRHATKIA SERDYRRDLE  4980
EGVKGKGLTV FEETPELLRA RNATAILSQK EYKKTLETEI KGKGMLALAT DTPDFMRARN  5040
ATDILSQQKY KQTAEADRAS YTSVIDTPDI LHAQNIRNMV SQKKYREEAE KCMSHYVPVL  5100
DTPEMMRVRE NQKNFSTIKY KDLLCRCTAI SDLPEVKRVR EMQKNISSVL YKDSSAKGTP  5160
QVFTPEMERA KRNQEQISSV LYSDSFRKQV QGKAAFVLDT PEMRRVKETQ RNISGVKYHQ  5220
DFEKSKGSFT PTTSDPVMDR VKKNTQEFSD INYRGIQRRV VEMEKRRVVE HEQETVTDLR  5280
VWRTNPGSVF DYDPAEDNIQ SRSLHMMSVQ AQRRSKEHSR STSAMSGVGD EKSEVSEDPA  5340
HHVSLYSNGF ATSSMGYTHT KTVELQQRSS SVATQQTTVS SVPSHPSTTG KTVRAMYDYS  5400
SADMDEVSFK DGDVIVNVQS IDEGWMYGTV QRSGKTGMLP ANYVEAI                5447

SEQ ID NO: 214          moltype = AA  length = 377
FEATURE                 Location/Qualifiers
source                  1..377
                        mol_type = protein
                        organism = Salmo salar
SEQUENCE: 214
MCDDDETTAL VCDNGSGLVK AGFAGDDAPR AVFPSIVGRP RHQGVMVGMG QKDSYVGDEA  60
QSKRGILTLK YPIEHGIITN WDDMEKIWHH TFYNELRVAP EEHPTLLTEA PLNPKANREK  120
MTQIMFETFN VPAMYVAIQA VLSLYASGRT TGIVLDSGDG VTHNVPIYEG YALPHAIMRL  180
DLAGRDLTDY LMKILTERGY SFVTTAEREI VRDIKEKLCY VALDFENEMA TAASSSSLEK  240
SYELPDGQVI TIGNERFRCP ETLFQPSFIG MESAGIHETA YNSIMKCDID IRKDLYANNV  300
LSGGTTMYPG IADRMQKEIT ALAPSTMKIK IIAPPERKYS VWIGGSILAS LSTFQAMWIT  360
KQEYDEAGPS IVHRKCF                                                 377

SEQ ID NO: 215          moltype = AA  length = 608
FEATURE                 Location/Qualifiers
source                  1..608
                        mol_type = protein
                        organism = Salmo salar
SEQUENCE: 215
MQWLSVCSLL VLLSVLSRSQ AQNQICTIFT EAKEDGFKSL ILVGLAQNLP DSTLGDLVPL  60
IAEALAMGVK CCSDTPPEDC ERDVADLFQS AVCSSETLVE KNDLKMCCEK TAAERTHCFV  120
DHKAKIPRDL SLKAELPAAD QCEDFKKDHK AFVGRFIFKF SKSNPMLPPH VVLAIAKGYG  180
EVLTTCCGEA EAQTCFDTKK ATFQHAIAKR VAELKSLCIV HKKYGDRVVK AKKLVQYSQK  240
MPQASFQEMA GMVDKIVATV APCCSGDMVT CMKERKTLVD EVCADESVLS RAAGLSACCK  300
EDAVHRGSCV EAMKPDPKPD GLSEHYDVHA DIAAVCQTFT KTPDVAMGKL VYEISVRHPE  360
SSQQVILRFA KEAEQALLQC CDMEDHAECV KTALAGSDID KKITDETDYY KKMCAAEAAV  420
SDDNFEKSMM VYYTRIMPQA SFDQLHMVSE TVHDVLHACC KDEPGHFVLP CAEEKLTDAI  480
DATCDDYDPS SINPHIAHCC NQSYSMRRHC ILAIQPDTEF TPPELDASSF HMGPELCTKD  540
SKDLLLSGKK LLYGVVRHKT TITEDHLKTI STKYHTMKDK CCAAEDQAAC FTEEAPKLVS  600
ESAELVKV                                                           608

SEQ ID NO: 216          moltype = AA  length = 608
FEATURE                 Location/Qualifiers
source                  1..608
                        mol_type = protein
                        organism = Salmo salar
SEQUENCE: 216
MQWLSVCSLL VLLSVLSRSQ AQNQICTIFT EAKEDGFKSL ILVGLAQNLP DSTLGDLVPL  60
IAEALAMGVK CCSDTPPEDC ERDVADLFQS AVCSSETLVE KNDLKMCCEK TAAERTHCFV  120
```

-continued

```
DHKAKIPRDL SLKAELPAAD QCEDFKKDHK AFVGRFIFKF SKSNPMLPPH VVLAIAKGYG  180
EVLTTCCGEA EAQTCFDTKK ATFQHAVMKR VAELRSLCIV HKKYGDRVVK AKKLVQYSQK  240
MPQASFQEMG GMVDKIVATV APCCSGDMVT CMKERKTLVD EVCADESVLS RAAGLSACCK  300
EDAVHRGSCV EAMKPDPKPD GLSEHYDIHA DIAAVCQTFT KTPDVAMGKL VYEISVRHPE  360
SSQQVILRFA KEAEQALLQC CDMEDHAECV KTALAGSDID KKITDETDYY KKMCAAEAAV  420
SDDSFEKSMM VYYTRIMPQA SFDQLHMVSE TVHDVLHACC KDEQGHFVLP CAEEKLTDAI  480
DATCDDYDPS SINPHIAHCC NQSYSMRRHC ILAIQPDTEF TPPELDASSF HMGPELCTKD  540
SKDLLLSGKK LLYGVVRHKT TITEDHLKTI STKYHTMKEK CCAAEDQAAC FTEEAPKLVS  600
ESAELVKV                                                           608

SEQ ID NO: 217          moltype = AA  length = 380
FEATURE                 Location/Qualifiers
source                  1..380
                        mol_type = protein
                        organism = Gadus morhua
SEQUENCE: 217
MAKNCHNDYK MKFSGDEEFP DLSLHNNHMA KVLTKDIYNK LRGKSTPSGF TVDDIIQTGV  60
DNPGHPFIMT VGCVAGDEES YEVFKDLLDP IISDRHSGYK PTDKHRTDLN FENLKGGDDL  120
DPAYVLSSRV RTGRSIKGFT LPPHNSRGER RTIEKLSIEA LATLSGEFKG KYYPLNGMTD  180
KEQDQLINDH FLFDKPVSPL LTCAGMARDW PDARGIWHND AKTFLVWVNE EDHLRVISMQ  240
QGGNMKEVFR RFCTGLLKIE ETFKKHNHGF MWNEHLGYVL TCPSNLGTGL RGGVHVKLPK  300
LSTHAKFEEI LTRLRLQKRG TGGVDTASVG GVFDISNADR LGSSEVDQVQ MVVDGVKLMV  360
AMEKKLEKSE SIDDMIPAQK                                              380

SEQ ID NO: 218          moltype = AA  length = 892
FEATURE                 Location/Qualifiers
source                  1..892
                        mol_type = protein
                        organism = Gadus morhua
SEQUENCE: 218
FHKLLQTNGK GMIMLAYMTQ EEGWDRDLLL DPAWEKQQRK TFTAWCNSHL RKAGTQIENI  60
EEDFRNGLKL MLLLEVISGE RLPKPDKGKM RFHKIANVNK ALDYICSKGV KLVSIGAEEI  120
VDGNAKMTLG MIWTIILRFA IQDISVEETS AKEGLLLWCQ RKTAPYRNVN VQNFHISWKD  180
GLALCALIHR HRPDLIDYSK LRKDDPIGNL NTAFEVAEKF LDIPKMLDAE DIVNTPKPDE  240
KAIMTYVSCF YHAFAGAEQA ETAANRICKV LAVNQENEKL MEEYEKLASE LLEWIRRTIP  300
WLENRAAEQT MRAMQQKLED FRDYRRVHKP PRVQEKCQLE INFNTLQTKL RLSNRPAFMP  360
SEGKMVSDIA NAWKGLEQVE KGYEEWLLTE IRRLERLDHL AEKFKQKCSL HETWTSGKEE  420
LLSQKDFESA SLMEIRALMR KHEAFESDLA AHQDRVEQIA AIAQELNELD YHDAVTVNAR  480
CQGICDQWDN LGTLTQKRRD SLERVEKLWE TIDQLYLEFA KRAAPFNNWM DGAMEDLQDM  540
FIVHSIEEIQ SLITAHDQFK ATLPEADKER MATMGIQSEI LKIAQTYGIK LSGINPYTNL  600
SPQDISNKWD AVKHLVPLRD QMLQEEVARQ QANERLRRQF AAQANIIGPW IQTKMEEIGH  660
VSADITGSLE DQMNNLKQYE QNIINYKSNI DKLEGDHQLS QESLIFDNKH TNYTMEHIRV  720
GWEQLLTTIA RTINEVENQI LTRDAKGISQ EQLNEFRASF NHFDRKRNGM MDPDDFRACL  780
ISMGYDLGEV EFARIMTLVD ANNTGVVTFQ AFIDFMTRET AETDTAEQVM ASFKILASDK  840
SYITVDELRR ELPLEQAEYC ISRMTRYVGA DGTTGALDYI SFSSALYGES DL          892

SEQ ID NO: 219          moltype = AA  length = 843
FEATURE                 Location/Qualifiers
source                  1..843
                        mol_type = protein
                        organism = Gadus morhua
SEQUENCE: 219
MAKPLTDQEK RKQISVRGLA GVENVTDIKS NFNRHLHYTL VKDRNVATKR DYYFALANTV  60
RDHLVGRWIR TQQHYYEKDP KRVYYLSLEF YMGRTLQNTM VNLALENACD EATYQLGLDM  120
EELQDIEEDA GLGNGGLGRL AACFLDSMAS LGLAAYGYGI RYEFGIFNQK IVDGWQVEEA  180
DDWLRYGNPW EKARPEYMRP VHYYGRVEHT PGGVKWVDTQ VVLALPYDTP VPGYRNNVVN  240
TMRLWSAKAP CDFNLKDFNV GGYIQAVLDR NLAENISRVL YPNDNFFEGK ELRLKQEYFV  300
VAATLQDIIR RFKSSKFGST EFVRMDLSTL PDKVAIQLND THPALAIPEL MRILMDLEEL  360
PWEKAWDIVV RTCAYTNHTV LPEALERWPI DLFQNLLPRH LDIVYEINRR HLERVAQQFP  420
GDEGRLRRMS LIEEGDHKKI NMAHLCIVGS HAVNGVARIH SEIIKDTVFK DFYEMDPEKF  480
QNKTNGITPR RWLVMSNPGL AEVIAERIGD DYIRDLDQLK NLLEFVDDDS LIRDVAKVKQ  540
ENKMKFSAYL EEHYKVKINP ESMFDVQVKR IHEYKRQLLN CLHIITFYNR IKKQPNKAWT  600
PRTIMIGGKA APGYHTAKMI IKLITAIGDI VNHDAVVGDR LKVIFLENYK VTLAEKIIPA  660
SDLSEQISTA GTEASGTGNM KFMLNGALTI GTMDGANVEM AEEAGEENLF IFGMRVDDVK  720
AMDQTGYDAK SYYDRIPELK QAMDQISGGF FSPGQPDLFK DIVNMLMNHD RFKVFADYEE  780
YIKCQDKVSA LYKDPKEWTK MVIRNIGSSG KFSSDRTIAQ YAREIWGMEP SLEKIPAPDD  840
PRD                                                                843

SEQ ID NO: 220          moltype = AA  length = 367
FEATURE                 Location/Qualifiers
source                  1..367
                        mol_type = protein
                        organism = Gadus morhua
SEQUENCE: 220
MPHAYPFLST EQKKELSDIA KKIVAPGKGI LAADESTGSV AKRFQGINAE NTEENRRLYR  60
QLLFTADDRV APCIGGVIFF HETLYQKTDD GKLFPQLIRE RGQVVGIKVD KGVMPLAGTD  120
```

-continued

```
GETTTQGLDG LYERCAQYKK DGADFAKWRC VLKITENTPS ELAIMENANV LARYASICQM  180
VRHHGIVPIV EPEILPDGDH DLQRTQYVTE KVLAAVYKAL SDHHVYLEGT LLKPNMVTAG  240
HSCTRKYSAQ EVAMATVTAL RRTVPPAVPG VTFLSGGQSE EQASINLNAI NQCPLGRPWA  300
LTFSYGRALQ ASALKAWGGK KENGKACQEE YIKRALNNNL ASQGKYVSSG DKGAAGGESL  360
YVANHAY                                                            367

SEQ ID NO: 221          moltype = AA  length = 194
FEATURE                 Location/Qualifiers
source                  1..194
                        mol_type = protein
                        organism = Gadus morhua
SEQUENCE: 221
MAPKKDVKAP AAAAKKAEPA KKAEPAPEPV AVPAPKTVDL SAVKVEFTPD QMEDYREAFG  60
LFDRVGDNKV CYNQIADIMR ALGQNPTNKE VKAILGNPSD EDMNSKRVDF EGFLPMMQTI  120
VNSPNKGTLD DYVEGLRVFD KEGNGTVMGA ELRIVLSTLG EKMTEAEIDA LMQGQEDENG  180
CINYESFVKH IMSI                                                    194

SEQ ID NO: 222          moltype = AA  length = 333
FEATURE                 Location/Qualifiers
source                  1..333
                        mol_type = protein
                        organism = Gadus morhua
SEQUENCE: 222
MVKVGINGFG RIGRLVTRAG FQSKKVEIVA INDPFIDLEY MVYMFKYDST HGRFHGEVKI  60
EGDKLVIDGH KISVFHERDP ANIKWGDAGA QYVVESTGVF TTIEKASAHL KGGAKRVIIS  120
APSADAPMFV MGVNHEKYEN SLQVVSNASC TTNCLAPLAK VIHDNFGIVE GLMSTVHAIT  180
ATQKTVDGPS GKLWRDGRGA SQNIIPASTG AAKAVGKVIP ELNGKITGMA FRVPTPNVSV  240
VDLTVRLEKP AKYDDIKKVI KAASEGPMKG ILGYTEDQVV STDFNGDIRS SIFDAGAGIA  300
LNDHFVKLVT WYDNEFGYSN RVIDLMAHMS CKE                               333

SEQ ID NO: 223          moltype = AA  length = 284
FEATURE                 Location/Qualifiers
source                  1..284
                        mol_type = protein
                        organism = Gadus morhua
SEQUENCE: 223
MDAIKKKMQM LKLDKENALD RAEQAEGDKK AAEDRSKQLE DDLVALQKKL KGTEDELDKY  60
SEALKDAQEK LELAEKKAAD AESDVASLNR RIQLVEEELD RAQERLATAL TKLEEAEKAA  120
DESERGMKVI ENRAMKDEEK MELQEIQLKE AKHIAEEADR KYEEVARKLV IIESDLERTE  180
ERAELSEGKC SELEEELKTV TNNLKSLEAQ AEKYSQKEDK YEEEIKVLTD KLKEAETRAE  240
FAERSVAKLE KTIDDLEDEL YAQKLKYKAI SEELDHALND MTSI                   284

SEQ ID NO: 224          moltype = AA  length = 1119
FEATURE                 Location/Qualifiers
source                  1..1119
                        mol_type = protein
                        organism = Gadus morhua
SEQUENCE: 224
MPEAADAGEL PTAEEDSGDE EGQGTELTGL FVERPQENVV IIAGADVVIT AKVDSTTLTR  60
KPTVKWLKGK WLDLASKAGK HLQFKETYDR ETKIYTYEMK VIKAKVGAAG AAGAYRCEVT  120
SKDKFDSASF EITVEAAQQS DDVDILSSFK RADAGEDDGN LDFSALLKAA KKKKRPEDDK  180
PVIDVWELLK SAHPSEYEKI AFEYGITDLR GMLKRLKKMK AEPKPSEAFL KRMESCYSVE  240
KGKKIVLTCE VFDPNVQVKW LKNGQEIKGS AKYIMETIGN IRKLTIMKTT LADDAAYECV  300
VGEDKCFTEV FVKEPPVTIT KLMDDYHVVV GERVEFEVEV SEEGANAMWF FEDIEIFKDT  360
NPAKYRIKKD GRKHMLIILE AGLEDIGMYH CWTNGGHTKG ELEVEEKELL VLQDIADLTV  420
KATDQAIFKC EVSDEKVTGK WFKDGVEVLP DERHKMTHIG RFHKLLIDNV KPQDAGDYTF  480
IPNGYALSLS AKLNFLEIKI DYVPRQDPPK IHLDTTGNMV SQNTIIVVAG NKLRLDVEIT  540
GEPAPTVVWS KGEEVITATE GRVRVDNRKT LSSFVIEGAE REDEANYTIC VTNPAGEDKA  600
KLFVKIVDVP DPPQNVKCLS VGEDSSSMVW EPPAFDGGAP LKGYLMERKK KGSSRWTKLN  660
FDVYEATTYE AKRMIEGVYY EMRVFAINSI GMSQPSITSK PFMPLAPTSE PTRLLVHDVT  720
DGTCTLKWMA PERIGAGGLD GYIIEYCKEG ETEWVVANKE LCERQGYVVR GLPVGEKINF  780
RVFAKNLAGV SPPATLAQPV TVREIVDLPK IRLPRDLRTK YIRKVGDKIN LVIPFQGKPR  840
PVVKWYKDGA PVDEKIVGVR NSTVDSILFI RSAERDHSGK YELVVQIENM EDRATIDLRV  900
VERPGPPTKV KVTEVWGFNA ALEWEAPKDD GNCEITGYTI QKADKKTMEW YTIYEHNRRP  960
SCTSSDLIMG NEYMFRVYSE NLVGPSEEPG QSKNTAVIGK TGLAYNPPEF KEKDVSCIPT  1020
FTQPLLDRSV VAGYSTAMSC SVKGFPKPKI VWMKNKMILG GDPKYIMLNN QGVLTLNIRK  1080
PSTFDGGRYS CMAVNDLGQA LVECKLEIRI PTDPTAEVK                         1119

SEQ ID NO: 225          moltype = AA  length = 1413
FEATURE                 Location/Qualifiers
source                  1..1413
                        mol_type = protein
                        organism = Gadus morhua
SEQUENCE: 225
MAMRSGNYNH EYYHKQSKYV VKEYSSSESV VDDYQHKAAV RIQGVVRKKM VEKYVREEPM  60
IRGPKFVVRL RSHTVFENSP IKLFCTVEGF PVPIVTWYKD GVILEHGSAG RYIVDDNVGM  120
```

-continued

```
RSLEIPRCGT DDTAQYSAVA SNIHGQASSQ ASIIVKRFRQ EDESCLYAWL PYTVPMLPSI  180
QYTKINITFT EKFGPVFASE GGSASLTATM HLEPNLANLQ PDSQWYRDDT RLVDSKWVKI  240
NTGRGTSTLT IPNLFKDDEG LYTLRLVTKG GTAEHSAFIS VSDGPPPVVG APGAPMDIRI  300
HDANRDYVIV SWKPPNTTTE GPILGYFVDK CEVGTENWAQ CNDHPIKICK YPVPGLFEGH  360
SYYFRVRAVN ARGISKPSRM SDPIAALDPT EFERLHALKL GGKLDVVTYD DDMEAEGKPP  420
GVSSGIYASE TDRTYVVLSW KAPAYSSKAP MWYYIEKCQV GTAAWQRVNT QVPIRSPRYA  480
VFDLAEGKEY QFRVLSANMY GTSEPSEPSK LIQTQALKGV PSAPGQVIST RETDTSLLIQ  540
WAPPKEPNNL IGYYIDQCVK GSKDWTSANH KPHSSTKFVV SGLTTGETYV FRVQAINELG  600
LSDESQESSP LSVKAALTSP TAPYDVALLG CDGQSMVLNW KKPLHTGGTA INAYFVDKRR  660
SGATMWREVH IPAVTERVYQ VEGLTQGAIY QFRVYAANLA GLGKSSKASA DFSCEAWTMP  720
EPGPAYDLTA SEVRDTSLVL EWKKPVYYGS GPITGYHVEY AKKGSSDWTR ANETAVSHRY  780
LKVEGLEEGA KYVFRVRGVN EAGVGMASLP SDSVVAKALA GAQEVSCSVD EKSGDIVFAF  840
ESCQMNVGSE FLWKKNYKEI TEFSSSIVVH TEGNLCKLMF KNAVKEDFGT YSVSVSHTDG  900
VSSSCSISAE ELQKMLELSY SIRHPIIPLK SELAYKILER GRMRFWLQAE EISSKVTYKF  960
FANNKELSGA DPNKMAHDVS TGVIEFILDH FTEDSEGTYT CQITDGGGKA QSSLVLIGDA  1020
FKAALAEADY QRREYIRVKE GPHFSEFLTV KVGEDCSVTL VCKVANLKKE SEFHWYKEDT  1080
EIIPDVKPDL GSGVCKLPIH LFSRKTTGVY KATISDDRGK DMSQVDISGK VFDDVINKIN  1140
QLAGASAAEL VINCTATGIQ LQCHMKYYTS EMNIHWMHGE SKVGHSEKCV IGGTPAMATM  1200
EVIEPNDKDK GVYSIVITDL ENTNKRTLDL SGDVYDKAYA QFQKLKAEAY AEMNRGKVMG  1260
GLPDVVTIME KKTLSLTCTV CGDPKPLVSW LKNGNDVEPD NQYVVSLDQG KFACLTINGI  1320
GMDDSGKYTM IVQNKYGGEC VDIVVSVYRH GDTIPVAKPN LTAKTIIPPK MPIELPQPKT  1380
AAPSPAPPAS PSPAPPKAVG GRGMKSPTPT RRK                               1413

SEQ ID NO: 226          moltype = AA  length = 457
FEATURE                 Location/Qualifiers
source                  1..457
                        mol_type = protein
                        organism = Gadus morhua
SEQUENCE: 226
MKSNPLTSNT GEHASYLLHV NGVMSITKIH AREILDSRGN PTVEVDLWTA KGLFRAAVPS  60
GASTGIHEAL ELRDGDKTRY LGKGTQKAVD HVNKDISPKL IEKNFSVVDQ EKIDKFMLEL  120
DGTENKSKFG ANAILGVSLA VCKAGAAEKG VPLYRHIADL AGNKDVILPV PAFNVINGGS  180
HAGNKLAMQE FMILPVGAAN FHEAMRIGAE VYHNLKNVIK AKYGKDATNV GDEGGFAPNI  240
LENNEALELL KSAIEKAGYP DKIIIGMDVA ASEFYRSGKY DLDFKSPDDP ARHISGKELG  300
DLYKTFIANY PVQSIEDPFD QDDWANWADF TAAVDIQVVG DDLTVTNPKR IQQAVEKKAC  360
NCLLLKVNQI GSVTESIQAC KLAQSNGWGV MVSHRSGETE DTFISDLVVG LCTGQIKTGA  420
PCRSERLSKY NQLMRIEEEL GDKAKFAGKD YRHPKIN                           457

SEQ ID NO: 227          moltype = AA  length = 170
FEATURE                 Location/Qualifiers
source                  1..170
                        mol_type = protein
                        organism = Gadus morhua
SEQUENCE: 227
MAPKKAKRRA GGGDSGSSNV FSMFEQSQIQ EYKEAFTIID QNRDGIISKD DLRDVLASMG  60
QINVKNEELE AMVKEASGPI NFTVFLTMFG EKLKGADPED VIVAAFKILD PEATGSIKKE  120
FLQELLTTQC DRFSPEEIKN MWSAFPPDVA GNVDYKNICY VITHGEEKEE             170

SEQ ID NO: 228          moltype = AA  length = 1935
FEATURE                 Location/Qualifiers
source                  1..1935
                        mol_type = protein
                        organism = Gadus morhua
SEQUENCE: 228
IMGDGEMQCF GEAAVYLRKP ERERLEAQTS PFDAKTAFFV SEPKEMYLKG VLQSRDGGKA  60
TVKTLCDKTL TVKEEEIYPM NPPKYDKIED MAMMTHLNEP TVLYNLKERY AAWMIYTYSG  120
LFCVTVNPYK WLPVYDSSVV NAYRGKKRIE APPHIFSISD NAYQAMLTDR ENQSILITGE  180
SGAGKTVNTK RVIQYFATIA VSGPKKTEQV AGKMQGSLED QIIAANPLLE AYGNAKTVRN  240
DNSSRFGKFI RIHFATSGKL ASADIETYLL EKSRVTFQLS AERSYHIFYQ LMTGHKPELI  300
DALLITKNPF DYHMISQGEI TVKSIDDVEE FIATDDAIDI LGFTADEKIG IYKLTGAVLH  360
HGNMKFKQKQ REEQAEPDGN EEADKIGYLM GLNSADVLKA LCYPRVKVGN EYVTKGQTCP  420
QVHNSVMALC KSVYEKMFLW MVVRINEMLD TKQPRQFFIG VLDIAGFEIF NYNSLEQLCI  480
NFTNEKLQQF FNHHMFVLEQ EEYKKEGIEW EFIDFGMDLA ACIELIEKPM GIFSILEEEC  540
MFPKASDTTF KNKLHDQHLG KTKAFEKPKP AKGKAEAHFS LVHYAGTVDY NITGWLDKNK  600
DPLNDSVIQL YQKSSVKLLA LLYASHALPD ADAASKKGAK KKGGSFQTVS ALFRENLGKL  660
MTNLRSTHPH FVRCLIPNES KTPGLMENFL VIHQLRCNGV LEGIRICRKG FPSRVLYGDF  720
KQRYKVLNAS VIPEGQFIDN KKASEKLLGS IDVDSTQYKF GHTKVFFKAG LLGALEEMRD  780
EKLAALVTLT QALCRGFNMR KEFVKMMERR EAVFTIQYNI RSFMNVKTWP WMKVYFKIKP  840
LLKSAETEKE MAQMKETFGK TKEDLAKALS KKKELEEKMV SLLQEKNDLL LQVQSEAETL  900
SDSEERCEGL IKSKIQLEAR LKETSERLED EEEINSELTA KKRKLEDECS ELKKDIDDLE  960
LTLAKVEKEK HATENKVKNL TEEMSSQDEI IAKLTKEKKA LQEAHQQTLD DLQAEEDKVN  1020
TLSKAKCKLE QQVDDLEGSL EQEKKLRMDL ERAKRKLEGD LKLAQESIMD LENDKQQSDE  1080
KIKKRDFETS QLLNKIEDEQ TLSAQFQKKI KELQARIEEL EEEIEAERAA RAKVEKQRAD  1140
LSRELEEISE RLEEAGGATS LQIEMNKKRE AEFQKLRRDL EESTLQHEAT SAALRKKQAD  1200
SVAELGEQID NLQRVKQKLE KEKSECKMEI DDLSSNMEAV AKSKANLEKL ARTLEDQLSE  1260
LKSKNDEHVR QLNDLSVQRA RLQTENGEFG RQQEEKEALI SQLTRGKQAY TQQIEELKRH  1320
```

-continued

```
IEEEVKAKNA LAHSVQSARH DCDLLREQFE EEQEAKAELQ RGMSKANSEV AQWRSKYETD  1380
AIQRTEELEE AKKKLAQRLQ DAEEAVEATN SKCASLEKTK QRLLGEVEDL MIDVERANAL  1440
AANLDKKQRN FDKVLAEWKQ KYEESQAELE GAQKEARSLS TEMFKLKNSY EESLDQLETL  1500
KRENKNLQQE ISDLSEQMGE TGKTIHELEK GKKTVEVEKS ELQTSLEEAE ATLEHEESKI  1560
LRVQLELTQI KGEVDRKMAE KDEEMEQIKR NSLRVIDSMQ SNLDSEIRSR NDALRIKKKM  1620
EGDLNEMEIQ LSHANRQASE AQKQLRNVQG QLKDAQLHLD DALRGQEDMK EQVAMVERRN  1680
NLMLAEIEEL RAALEQTERG RKVAEQELVD ASERLGLLHS QNTSLINTKK KLETDLIQIQ  1740
GEVEDTVQEA RNAEEKAKKA ITDAAMMAEE LKKEQDTSSH LERMKKNLEL TVKDLQHRLD  1800
EAESLAMKGG KKQLQKLEAR VRELETECEA EQRRGSEAIK GVRKYERKVK ELSYQCDEDK  1860
KNICRLQDLV DKLQLKMKAY KRQAEESEEQ ANTHLSRYRK VQHELEEAQE RADVSESQVN  1920
KLRTKSRNGI TKIIN                                                   1935

SEQ ID NO: 229          moltype = AA  length = 6118
FEATURE                 Location/Qualifiers
source                  1..6118
                        mol_type = protein
                        organism = Gadus morhua
SEQUENCE: 229
MATAGMPTRK VRKRVAVDTS KFMTPYLAHS MKMLEQYSHN MYRKAYDLSR GTPPVHIADT  60
PEMVRIRKAQ EQLSEVKYRM EGNKARAGSL YDGEAKEIAH VKHMSELLSK VLYRQKWDET  120
KDRYLLPPDA PELVLAVKNA ANYSKKLYTE DWDLEKASFL PYSDSPELRR VAKAQEALSD  180
IQYKKGYDQR KLKFTSLADR PDVELAKKVF TQRSDLKYKE DYNKNVKGQW CETPYFDVAT  240
ARVAMENLSS RRYTQDWEAA KDQINFMQTE TPVYDTNKKA GTSASSVKYR KDYEKTKDKS  300
DYNTLPATEN PLLRQLRYAG TILSDKVYKK NYEKSRGFSI NYCDTPKFKM DSVLKKFTDT  360
HYQDKYKNEV QGHYIGSYED IYMLHCQKVE GFKSEQNYKA EYEDIKTRCF FPQTITPEYE  420
VDKKLAQCKD KVYRQHPDKV KFTQVTDSPV QVQAALNAIQ LSDINYKAKY EQTKTKNSLP  480
ADYPFFIQSR VNAFNLSENC YKYNWEKEKA KRFEMKGDAI SILAARAHTN IASDVKYKKE  540
YEKNRGRMVG ALSIHDDPKI LHSVHVAKIV SNREYTKDYE KVKTKYHTSL DMMSVTLAKK  600
SQAIASMTGY KKISNRYLMP YDAILVDLAK KANSIQSDNE YKSDYNQYTK GSPWITFGSM  660
DVELAKKAGE ILSEKKYRKH PDTVPFTAID DHPVMLQARV NQLQRSDLFY KKGLEEIHQK  720
YSLPIDMPEL QLARANSYNI SKNYYKAAWE ETISKGYDLK LDAISIKAAK AARHAASDIQ  780
YKKAYEKARG HHVGFRNIKD DPLLVHYMQV AKMQSEKNYK KDYHKSKLKY HSPVDMMSLV  840
HAKQASKAQT YAGYRKIQHH YSLMPDAMNL VLARKMNTDA SDCEYKSEWN NYVRGTGWVP  900
IGSIGVETAK IGGQIQSENK YRTHPSTFKF SKLMDSMDLE LATQAYTSAW EKDKRTIHVM  960
PDSPEIVLAK ANAITMSDVR EYYYKLHQNG YHLKGDAIPI VAAKLGTKIA SNYKYKLDYE  1020
KARGHHVGFR NIKDDPLLVH YMQVAKMQSE KNYKKDYEKA KLKYHSPVDM MSLVHAKEAS  1080
KAQTYAGYRK IPHTYFLLPD NMQLQLCRKM QEIQSDNEYK SEYNNLYKGT GWVPIGSIDV  1140
ELAKTANSAL REHGYRQHPS SFKFTSKTDS MDMVLALANS KIIDQAAYKA SGEKVKHTYH  1200
LPADAPEFIQ AKFNAQTMSK SHYKQKWVEH IAKGYDMRPD AISILHAKQG RHIASDIQYK  1260
KAYEKARGHH VGFRSIKDDP LLVHYMQVAK MQSEKNYKKD YEKAKLKYHS PVDMMSLVHA  1320
KEASKAQTYA GYRKIQHHYN LMPDAMNLVL ARKMNTDASD CEYKSEWNNY VRGTGWVPIG  1380
SIGVETAKIG GEIQSENKYR THPSTFKFSK LMDSMDLELA TANNKIMNKD KRTIHVMPDS  1440
PEIVLAKANA ITMSDKLYKA GVMEMAKKGY HLKGDAIPIV AAKLGTKIAS NYKYKLDYEK  1500
ARGHHVGFRN IKDDPLLVHY MQVAKMQSER NYTKDYQKSK LKYHSPVDMM SLVHAKEASK  1560
AQTYAGYRKI EHHYNLMPDA MNLVLARKMN TDASDCEYKS EWNNYVRGTG WVPIGSIGVE  1620
TAKIGGEIQS ENKYRTHPST FKFSKLMDSM DLELATANNK IMNKQAYTSA WEKDKRTIHV  1680
MPDSPEIVLA KANAITMSDK LYKAGVMEMA KKGYHLKGDA IPIVAAKLGT KIASNYKYKL  1740
DYEKARGHHV GFRCIQDDPL LVHYMQLAKQ QSEKNYKKDY EKAKLKYHSP VDMMSLVHAK  1800
DATKAQTYAG YRKIQHHYSL LPDAMNLVLA RNMQVNISDN QYKSDYESYC KGIGWIPIGS  1860
LDVEKAKIAG KIQSEKLYRQ HPSNLKFTKD MFSMDLTLAT ANNQIMNKVL LALYCVDYTL  1920
MNILSSGNNL TNNFVVSFSK LYKLDNELSK KKGHNVRKDA IAILAAKNCT TIASDYKYKA  1980
GYRKQVGHHI GARCVEDDPL LVLALNSARI ASDALYKKDF NKSKTKFNLP VDMMAFELAK  2040
KNQIQVNHAN YITRLHNWTC LPDSTDVVNA RRAYEIQSDN YYKEDQKLMQ GIGWVPIGSL  2100
DVEKAKKAGE ILSEHLYRKH PSNYKFTSKC DAMPLVLAKA NAQITNKKAY HDAWDKEKSN  2160
IHIMPDAMDV VLAKANRANY SIKRYKQANE DAKKKGYDLR NDAIGIIAAR AQRNIISDIK  2220
YKAGYRKQVG HHIGARCVED DPLLMLALNS ARIASDYLYK KDFNKSKTKF NLPVDMMAFE  2280
LAKKNQIQVN HANYITRLHN WTCLPDSTDV VNARRAYEIQ SDNYYKEDQK LMQGIGWVPI  2340
GSLDVEKAKK AGEILSEHLY RKHPSNYKFT SKCDAMPLVL AKANAQITNK KAYHDAWDKE  2400
KSNIHIMPDA MDVVLAKANR ANYSIKGYKQ ANEDAKKKGY DLRNDAIGII AARAQRNIIS  2460
DIKYKAGYRK QVGHHIGARC VEDDPLLVLA LNSARIASDY LYKKDFNKSK TKFNLPVDMM  2520
AFELAKKNQI QVNHANYITR LHNWTCLPDS TDVVNARRAY EIQSDNYYKE DQKLMQGIGW  2580
VPIGSLDVEK AKKAGQILSE HLYRKHPSNY KFTSKCDAMP LVLAKANAQI TNKKAYHEAW  2640
DKEKSNIHIM PDAMDVVLAK ANRANYSIKK YKEANEESKK KGYDLRNDAI SIIAARAQRN  2700
IVSDYKYKAG YRKQTGHHIG ARCVEDDPLL VLALNSARIA SDALYKKDFN KSKTKFNLPV  2760
DMMAFELAKK NQIQVNHANY ITRLHNWTCL PDSTDVVHAR HVYEIQSDAV YKADLKWLQG  2820
MGWVPIGSLD VEKTKKAGEI LSERKYRQPA SSHKFTSPLD AMEIVLAKNN ALTMNQRLYT  2880
EAWEKDKTKL HITSETPEII LSQQNAINMS NKKYKQGYED TIKKGYYLPG DAISVKSAKH  2940
SRNIISDYKY KAGYRTQVGH HIGARCVEDD PLLMLALNSA RIASDYLYKK DFNKSKTKFN  3000
LPVDMMAFEL AKKNQIQVNH ANYITRLHNW TCLPDSTDVV NARKAYDLQS DAVYKADMEF  3060
MKGIGWVPIG SLDVVKAKHA GKILSDHLYK QKTDKLKYTS DTKAMPMLLA KANAEVMNKK  3120
SYVAAWEKDK IKTHLIPDLP EILLSRSNAY NLSNKMYKED VVKMFRKGYH LKGDAIAIVA  3180
AKKGREIISD IKYKAGYRKQ VGHHIGALSV RDDPLIMLAL NSARIASDAL YKKDFNQSKT  3240
KFHLPVDMLT IELAKKCQVQ VNDFNYRTRL HNWTCLPDST DVVQARKAYD LQSDAVYKAD  3300
MEWIKGQGWV PIGSLDVEKA KKAAEILSER KYREHPSKFK FTAKTDDMPY ALAKANAEIM  3360
NKQSYVAAWE SDKTKLHIMP DTPEIILGQQ NKLNNSRKYY REAYEETLKK GYYLPKDAIS  3420
VLAARASRNI ISDYKYKAGF RKQCGHHIGA LSVQDDPLLM MAANAAKIFS ANLYKQDFHK  3480
```

-continued

```
TKTKYKLPVD MVAFDLAKTA QSHMKEVTYR TRLHNWTCMP DNTDVVHARK VYDLHSDAVY  3540
RADMEWIRGC GWSPQDSVDV VKARNAQKIL NENLYRQAPS TVKFTSVVDL PAIVLSKQNA  3600
ENISEMKYTA AWNKDKLTFH LKPDIPSVQL AKANAINISN KLYTKAWDNQ KAKAYHIKED  3660
TIAVLKARAS RDIISDYKYK TGYRKQVGHH IGALSVRDDP LIMLALNSAR IASDALYKKD  3720
FNQSKTKFHL PVDMLTIELA KKCQIQVNDF NYRTRLHNWT CLPDSTDVVQ ARKAYDLQSD  3780
AVYKADMEWI RGCGWVPHEC VDVVKVKHAQ KILADRGYRV KLDTQKFTAP VDRMDFVCAK  3840
NAAEVLNDAK YREAWHTDKS KFSLVDTPGL ATSREVAKIM HPKLYTKDWD KVKATGYFMP  3900
GDAVPIKQCT QTTKVQSSHK YLEGYRKQVG HHIGALSVFD DPLIMLAMNS AKIASEALYK  3960
KDFNKSKTKF HLPVDMLTIE LAKKCQIQVN DFNYRTRLHN WTCLPDSTDV VNARKAYDLQ  4020
SDFVYKADLE WIRGCGWVPA ESVHHVKVRK AQEILNERLY KKGAIENFSK YTHVVDRPEL  4080
ILARVAAANL SDLKYKETFN NEKGHYIGSE DTPQLAHSRE CAKNISEAKY HESHEKAKGH  4140
YMANTLVDFP EVVRCGEQEK AKALRVYQKE YHESKTNNHI PSDIIANQVA KRCQDMLSDV  4200
IYRTHLHQWT CPPDQLDTVR ARKGNEILSD VFYKEDLTWM KGMGCYVWDT PEIVRAKKSY  4260
ELQSELKYKE AGKKEFNNYS IVTDTPVYVT AILGSTWASE LNYREAYHKE KHLYTTILDT  4320
HDYARCHDFK HFFSNKIYTA AWDKIKAKGY KIPHDSHAMM HAKQQKVILS GVKYKEDYEK  4380
FKSLYSLPKC LQDDPATARC IKAGKLNLDR LYKEDYEKTK AKNHIPADML DIVSGREVQR  4440
NVSEINYRKR LHQWVCLPDM QVFVQARKVN DQLSDIFYKD DLNYLRGIGC YAWDTPEILR  4500
VKTAGDQQSE NKYKAKAVEA YKDYTVVLDT PTYETAKQNA VNLSDLHYRL DYNANVKGTN  4560
TAPAVTMETE RARLANYNQS DNLYKEANKN SMATGYLMTN DTPVVKQVKN NSLVASSVKY  4620
KAAYEMMKAT GYTLHPEGVT FVNTRKVNQV INPRLYREQY EKDKDKIHST LDTPDIRQVK  4680
MNQEHLSELC YREKYNNTKG QLINMPMTPQ LMHCQHVNQI TSDLRYKEDL LWLRGLGCFM  4740
YDTPEMVNVR NITKFKNNYQ VEAKKNHSKF SVVLDTPEYQ RVTELKSHMS MVSPTSNTSL  4800
NTLDSVDIQR SKWAQQLSNK FLYTDLASKE RAHFSQEVET PAMGLARTMK VTNSYNKYKE  4860
QYEKMKHRYT SIADTPILIR AKKAYLQSSD LRYKETFERS KGHYHTVKDA LDILCHRRVT  4920
DDISEVKYRE KYINSLGTWK SIPDRPEFFF NRIVGENVSS VKYKEDLEWM KGIGCFVWDT  4980
PEMVQADINK TLYSGRLYKA SFEKNRYNFK YTCDTPFFHA AKNASSLINE RSYRSSYEKN  5040
KDKFTITTDD PRYQLARENN KMSQAKYKSR AKELLKSGCN ELHRPDILTA LYQSVLGSKW  5100
KYREQYERVK SKFTSILETP EYESHKRSKK IGDIVYRMEY NKNKAKGYSL PYDTPHQNHM  5160
KKVQALTSNL KYREVYEKAK SQINMDPDAF EIRAAKEAYK NTTNLDYKKK YEATKNKWIW  5220
TVDRPDFISH SRNTLQQSDV EYKYDKEMMK GCVIPIVDDK ITLLAKKGCE MASDIKYKEK  5280
YEKAKGHYNP VADSPQILHA KAVRTLASES KYKEASKKEM QSGSFTTLPE TRDTVHSKEV  5340
NKLISGKLYK SKFEKEKGQS KYNRMIVPPD VQHAMDVAKS QSSYAYKRDA KDNLHYTTTA  5400
DRPDIRKATQ AAKLISDVGY REKARAEASR GGSLTYRPDI ALATEVSKLT SLVDAKSALH  5460
GAASYDTLQM RHSKKMSALT SNVKYHEKFD KEMKGKKPKY DLKDSKIYQT LKDATNLASE  5520
VKYKGDLKKL HKPVTDMAES LSMQHNLSTS KLSSDYRYKK KFKDSKGHYL VIADTPEQLH  5580
HKEATELQSS VSGPSLCGVH VCKDYKKDYE EGMRGKNLCG LEVTPAMIHV RHATKIASER  5640
EYRKDLEEGV KGKGLSAPEE TPEHLRAKNA THILSEREYK KSLELEIKGK GMLALATDTP  5700
DFMRARNATD ILSQTKYKHT AELDRASFTS VIDTPDIIHA QQMRNIISQN KYKAEAGKTM  5760
SHYVPVFDTP EMQRVRENQK NFSTLQYQAD HKIIKGKVLT VKDTPEMLRV KENTKNFSSV  5820
LYKDSKAKGT PVVFTPEMER VKRNQEHISS VLYSDSFRKQ VQGKAAFVLD TPEMRRVKES  5880
QLNISGVRYH QDHEKSKGSF TAMASYPEME RVKKNTQEFS DINYRGIQRR VVEMERRRVV  5940
EHEQETTTDL RVWRTNPGSV FDYDPAEDNI QSRSLHMMSV HAQRRSKEHS RSTSAMSGPA  6000
DEKSEMSHEG DHHMSTYSNG FGASSVGYQT TKTVELQQRS SSVATQQTTV SSIPSHPSTT  6060
GKTVRAMYDY GAADNDEVTF KDGDVIVNVQ SIDEGWMYGT VQRTGKTGML PANYVEAV    6118

SEQ ID NO: 230          moltype = AA  length = 1935
FEATURE                 Location/Qualifiers
source                  1..1935
                        mol_type = protein
                        organism = Gadus morhua
SEQUENCE: 230
MSTDAEMAVY GQAAIYLRKP EKERLEAQSK PFDAKSAAYV TDVKELYLKC TILKKEGGKV  60
TVKVHDTEEE KTVKDDDVTP MNPPKYDKIE DMAMMTHLNE ASVLYNLKER YAAWMIYTYS  120
GLFCATVNPY KWLPVYDQHV VCAYRGKKRM EAPPHIFSVS DNAYQFMLTD RENQSVLITG  180
ESGAGKTVNT KRVIQYFATI SVGGDKKKDI GAGKIQGSLE DQIIAANPLL EAYGNAKTVR  240
NDNSSRFGKF IRIHFGTSGK LSSADIETYL LEKSRVTFQL PAERGYHIFY QMMTNHKPEI  300
IEMTLITTNP YDFPTCSMGQ IAVASIDDKE ELVATDTAID ILGFNNDEKM GIYKLTGAVL  360
HHGNMKFKQK QREEQAEPEG TEDADKVAYL LGLNSADMLK GLCYPRVKVG NEYVTKGQTV  420
PQVLNSVSAL AKSIYERMFL WMVIRINQML DTKQPRSSFI GVLDIAGFEI FDYNSMEQLC  480
INFTNEKLQQ FFNHHMFVLE QEEYKKEGII WEFIDFGMDL AACIELIEKP MGIFSILEEE  540
CMFPKASDVT FKAKLYDQHL GKTKAFEKPK PAKGKAEAHF SLVHYAGTVD YNITGWLDKN  600
KDPLNDSVVQ LYQKSGVKLL PVLYPPVASE DSTKKGGKKK GGSMQTVSSQ FRENLGKLMT  660
NLRSTHPHFV RCLIPNESKT PGLMENFLVI HQLRCNGVLE GIRICRKGFP SRILYADFKQ  720
RYKVLNASVI PEGQFIDNKK ASEKLLGSID VNHEEYKFGH TKVFFKAGLL GTLEEMRDEK  780
LATLVGMIQA LCRGFQMRKE FKKMMERRHA VYTIQYNIRS FMNVKTWPWM KVYYKIKPLL  840
KSAETEKELS AMKENYEKMK KDLATALAKK KELEEKMVSL LQEKNDLSLQ VASEGENLND  900
AEERCEGLIK SKIQLEARLK ETTERLEDEE EMNADLTAKK RKLEDECSEL KKDIDDLELT  960
LAKVEKEKHA TENKVKNLTE EMASQDESVA KLSKEKKALQ EAHQQTLDDL QAEEDKVNTL  1020
SKAKTKLEQQ VDDLEGSLEQ EKKLRMDLER SKRKLEGDLK LAQESIMDLE NDKQQSDEKI  1080
KKKDFETSQL LSKIEDEQSL AIQLQKKIKE LQARIEELEE EIEAERAARA KVEKQRADLS  1140
RELEEISERL EEAGGATSAQ IEMNKKREAE FQKLRRDLEE STLQHEATAA ALRKKQADSV  1200
AELGEQIDNL QRVKQKLEKE KSEYKMEIDD LSSNMETVAK AKGNLEKICR TLEDQLSEIK  1260
AKSDENSRQI NDIGAQRARL LTENGEFGRQ LEEKEALVSQ LTRGKQAYTQ QIEELKRQIE  1320
EEVKAKNALA HGVQSARHDC DLLREQFEEE QEAKAELQRG MSKANGEVAQ WRSKYETDAI  1380
QRTEELEESK KKLAQRLQEA EEQIEAVNSK CASLEKTKQR LQGEVEDLMV DVERANGLAA  1440
NLDKKQRNFD KVLAEWKQKY EEGQAELEGT QKEARSLSTE LFKMKNSYEE ALDQLETLKR  1500
```

-continued

```
ENKNLQQEIS DLSEQIGETG KSIHELEKSK KQVETEKTEI QTALEEAEGT LEHEESKILR  1560
VQLELNQIKG EVDRKLAEKD EEMEQIKRNS LRITDSMQST LDSEVRSRND ALRIKKKMEG  1620
DLNEMEIQLS HANRQAAEAQ KQLRNVQGQL KDAQLHLDDA VRGGEDLKEQ AAMVDRRNGL  1680
MVAEIEELRA ALEQTERGRK VAEQELIDAS ERVGLLHSQN TSLLNTKKKL ETDLVQVQGE  1740
VDDTVQEARN AEEKAKKAIT DAAMMAEELK KEQDTSSHLE RMKKNLEVTV KDLQHRLDEA  1800
ENLAMKGGKK QLQKLESRVR ELETEVEGEQ RRGVDAVKGV RKYERRVKEL TYQTEEDKKN  1860
GARLQDLVDK LQLKVKAYKR QAEEAEEQAN SYLSKCRKVQ HELEEAEERA DIAETQVNKL  1920
RSKSRDGGKG KDPAE                                                   1935

SEQ ID NO: 231          moltype = AA  length = 377
FEATURE                 Location/Qualifiers
source                  1..377
                        mol_type = protein
                        organism = Gadus morhua
SEQUENCE: 231
MCDDEETTAL VCDNGSGLVK AGFAGDDAPR AVFPSIVGRP RHQGVMVGMG QKDSYVGDEA  60
QSKRGILTLK YPIEHGIITN WDDMEKIWHH TFYNELRVAP EEHPTLLTEA PLNPKANREK  120
MTQIMFETFN VPAMYVAIQA VLSLYASGRT TGIVLDSGDG VTHNVPVYEG YALPHAIMRL  180
DLAGRDLTDY LMKILTERGY SFVTTAEREI VRDIKEKLCY VALDFENEMA TAASSSSLEK  240
SYELPDGQVI TIGNERFRCP ETLFQPSFIG MESAGIHETA YNSIMKCDID IRKDLYANNV  300
LSGGTTMYPG IADRMQKEIT ALAPSTMKIK IIAPPERKYS VWIGGSILAS LSTFQQMWIS  360
KQEYDEAGPS IVHRKCF                                                 377

SEQ ID NO: 232          moltype = AA  length = 377
FEATURE                 Location/Qualifiers
source                  1..377
                        mol_type = protein
                        organism = Thunnus maccoyii
SEQUENCE: 232
MCDDDETTAL VCDNGSGLVK AGFAGDDAPR AVFPSIVGRP RHQGVMVGMG QKDSYVGDEA  60
QSKRGILTLK YPIEHGIITN WDDMEKIWHH TFYNELRVAP EEHPTLLTEA PLNPKANREK  120
MTQIMFETFN VPAMYVAIQA VLSLYASGRT TGIVLDAGDG VTHNVPVYEG YALPHAIMRL  180
DLAGRDLTDY LMKILTERGY SFVTTAEREI VRDIKEKLCY VALDFENEMA TAASSSSLEK  240
SYELPDGQVI TIGNERFRCP ETLFQPSFIG MESAGIHETA YNSIMKCDID IRKDLYANNV  300
LSGGTTMYPG IADRMQKEIT ALAPSTMKIK IIAPPERKYS VWIGGSILAS LSTFQQMWIS  360
KQEYDEAGPS IVHRKCF                                                 377
```

1.-**41**. (canceled)

42. A method for producing a pet food product, the method comprising:

(a) contacting a plurality of host cells with a recombinant nucleic acid, wherein the recombinant nucleic acid encodes an albumin protein of an animal selected from the group consisting of a mammal, a bird, and a fish;

(b) expressing the albumin protein encoded by the recombinant nucleic acid in the plurality of host cells; and

(c) generating a pet food product comprising the albumin protein expressed in the plurality of host cells.

43. The method of claim **42**, wherein the albumin protein is a rabbit albumin protein.

44. The method of claim **42**, wherein host cells of the plurality of host cells comprise one or more members selected from the group consisting of *Pichia* cells, *Saccharomyces* cells, *Schizosaccharomyces* cells, *Trichoderma* cells, *Aspergillus* cells, *Escherichia* cells, *Bacillus* cells, and *Lactococcus* cells.

45. The method of claim **42**, wherein the albumin protein comprises an amino acid sequence that is at least 70% identical to any one of SEQ ID NOs: 16, 78, 93, 108, 123, 138, 161, 174, 187, 200, 215, or 216.

46. The method of claim **45**, wherein the amino acid sequence is at least 70% identical to SEQ ID NO 78.

47. The method of claim **45**, wherein the amino acid sequence is at least 90% identical to any one of SEQ ID NOs: 16, 78, 123, 161, 174, 187, 200, 215, or 216.

48. The method of claim **45**, wherein the amino acid sequence is at least 95% identical to any one of SEQ ID NOs: 16, 78, 108, 123, 161, 174, 187, 200, 215, or 216.

49. The method of claim **42**, further comprising isolating the albumin protein from the plurality of host cells.

50. A method for producing a pet food product, the method comprising:

(a) contacting a plurality of host cells with a first recombinant nucleic acid encoding an albumin protein and a second recombinant nucleic acid that encodes a protein different from the albumin protein;

(b) expressing the albumin protein encoded by the first recombinant nucleic acid and the protein encoded by the second recombinant nucleic acid in the plurality of host cells; and

(c) generating a pet food product comprising the albumin protein and the protein expressed in the plurality of host cells.

51. The method of claim **50**, further comprising isolating the albumin protein and the protein from the plurality of host cells.

52. The method of claim **50**, wherein the albumin protein or the protein is a rabbit protein.

53. The method of claim **52**, wherein the albumin protein and the protein are rabbit proteins.

54. The method of claim **50**, wherein the albumin protein comprises an amino acid sequence that is at least 70% identical to any one of SEQ ID NOs: 16, 78, 93, 108, 123, 138, 161, 174, 187, 200, 215, or 216.

55. The method of claim **54**, wherein the amino acid sequence is at least 70% identical to SEQ ID NO 78.

56. The method of claim **50**, wherein the protein comprises an amino acid sequence that is at least 70% identical to any one of SEQ ID NOs: 1-15, 17-77, 79-92, 94-107, 109-122, 124-137, 139-160, 162-173, 175-186, 188-199, 201-204, or 217-232.

57. The method of claim **56**, wherein the amino acid sequence is at least 70% identical to any one of SEQ ID NOs: 4, 54, 57, 58, 59, 60, 63, 64, 67, 74, or 77.

58. A method for producing a pet food product, the method comprising:

(a) contacting a plurality of host cells with a recombinant nucleic acid;

(b) expressing a protein encoded by the recombinant nucleic acid; and

(c) generating a pet food product comprising the protein expressed in the plurality of host cells,

wherein the protein comprises an amino acid sequence that is at least 70% identical to any one of SEQ ID NOs: 5, 8, 9, 11, 14, 16, 19-21, 23, 24, 26-29, 32-35, 37-41, 44-49, 51-54, 56, 57, 59, 65-79, 81, 82, 84, 93, 94, 96, 97, 99, 108, 109, 111, 112, 114, 117, 123, 124, 126, 127, 129, 138-141, 143, 147, 148, 150, 151, 153, 161, 162, 164, 165, 167, 169, 170, 172, 174, 175, 177, 178, 180, 184, 185, 187, 188, 190-192, 197, 200, 201, 203-206, 208, 209, 213, 215, 216, 217, 219-222, 224, 225, or 229.

59. The method of claim **58**, wherein the protein is a rabbit protein.

60. The method of claim **58**, wherein the protein is an albumin protein.

61. The method of claim **58**, wherein the amino acid sequence is any one of SEQ ID NOs: 5, 8, 9, 11, 14, 16, 19-21, 23, 24, 26-29, 32-35, 37-41, 44-49, 51-54, 56, 57, 59, 65-79, 81, 82, 84, 93, 94, 96, 97, 99, 108, 109, 111, 112, 114, 117, 123, 124, 126, 127, 129, 138-141, 143, 147, 148, 150, 151, 153, 161, 162, 164, 165, 167, 169, 170, 172, 174, 175, 177, 178, 180, 184, 185, 187, 188, 190-192, 197, 200, 201, 203-206, 208, 209, 213, 215, 216, 217, 219-222, 224, 225, or 229.

* * * * *