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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 June 2012 (28.06.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/087835 A2

(51) International Patent Classification:
C07K 19/00 (2006.01) *C12N 15/63* (2006.01)
C12N 15/62 (2006.01)

(21) International Application Number:
PCT/US2011/065563

(22) International Filing Date:
16 December 2011 (16.12.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/426,326 22 December 2010 (22.12.2010) US

(71) Applicant *(for all designated States except US)*: **BOSTON BIOMEDICAL RESEARCH INSTITUTE** [US/US]; 64 Grove Street, Watertown, MA 02472 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **TAKAYAMA, Shinichi** [JP/US]; 116 Orchard Street, Belmount, MA 02478 (US). **HISHIYA, Akinori** [JP/US]; 14 Leslie Road, Belmont, MA 02478 (US).

(74) Agents: **CONLIN, David, G.** et al.; Edwards Wildman Palmer LLP, P.O. Box 55874, Boston, MA 02205 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

WO 2012/087835 A2

(54) Title: COMPOSITIONS AND METHODS FOR ENHANCING PROTEIN FOLDING

(57) Abstract: The present invention features compositions and methods that enhance protein folding featuring a BAG domain.

COMPOSITIONS AND METHODS FOR ENHANCING PROTEIN FOLDING

CROSS-REFERENCE TO RELATED APPLICATION

This application claims the benefit of U.S. Provisional Application No.:61/426,326, filed December 22, 2010, the entire contents of which are incorporated herein by reference.

STATEMENT OF RIGHTS TO INVENTIONS MADE UNDER FEDERALLY SPONSORED RESEARCH

This work was supported by the following grants from the National Institutes of Health, Grant No: AR052925. The government has certain rights in the invention.

BACKGROUND OF THE INVENTION

Various different proteins can be exogenously expressed in bacteria, yeast and mammalian cells, with molecular biological techniques. For decades, methods to enhance protein expression have been extensively investigated. These efforts have brought about many new methodologies and technologies, such as expression vectors, which have strong promoters. Viral vectors have been used for high transfection efficiency and promoting the improvement of efficiency of protein production *in vivo*. Even though highly efficient expression and a huge amount of protein production can be achieved by current molecular techniques, proper protein folding remains problematic. During protein synthesis, newly synthesized peptides fold properly with the help of a molecular chaperone complex. The persistence of proteins with abnormal or premature structures turns on the degradation system, since abnormally-folded proteins are frequently toxic to cells. Because of this self-defense system against unfolded or abnormally functioning proteins, considerable numbers of proteins expressed exogenously end up in rapid degradation in cells as well. This is one of the most serious problems when a protein is expressed in cells using exogenous gene an expression system.

SUMMARY OF THE INVENTION

As described below, the present invention features compositions and methods that enhance protein folding.

The invention features a fusion polypeptide contains a BAG domain that interacts with an Hsp70 molecular chaperone linked to a heterologous polypeptide. In one embodiment, the BAG domain is derived from a mammalian BAG1 or BAG3 polypeptide.

In another embodiment, the BAG domain comprises amino acids 419-536 or 388-500 of human BAG3. In another embodiment, the heterologous polypeptide is cystic fibrosis transmembrane conductance regulator (CFTR), IL13 receptor α2 protein (IL13Rα2), or anti-trypsin (α1-AT).

In another aspect, the invention provides a polynucleotide encoding a BAG domain that interacts with an Hsp70 molecular chaperone.

In another aspect, the invention provides an expression vector contains a promoter suitable for expression in a mammalian cell operably linked to a polynucleotide sequence encoding a fusion polypeptide contains a BAG domain that interacts with an HSP70 molecular chaperone linked to a heterologous polypeptide. In one embodiment, the polynucleotide sequences encoding the BAG domain and the heterologous polypeptide are linked by a nucleic acid linker contains between about 5 and 25 nucleic acids. In another embodiment, the BAG domain is attached to the amino terminus or the carboxy terminus of the heterologous polypeptide, or wherein the BAG domain is positioned between the amino and carboxy termini of the heterologous polypeptide. In another embodiment, the BAG domain is linked to the heterologous polypeptide by a linker contains a protease-sensitive site.

In another aspect, the invention features a host cell contains the expression vector of a previous aspect.

In another aspect, the invention features a method of producing a BAG fusion protein, the method involving the steps of expressing in a host cell, a polynucleotide encoding a fusion protein of a previous aspect; and culturing the host cell under conditions appropriate for production of the fusion protein.

In another aspect, the invention features a method of treating a subject having a disease associated with the expression of a misfolded protein, the method involving expressing in a cell of the subject, a polynucleotide encoding a fusion protein of a previous aspect; and expressing the protein in the cell under conditions appropriate for production of the fusion protein, thereby treating the subject.

The invention provides compositions featuring a BAG family protein and methods of using such compositions to enhance protein folding or otherwise stabilize an exogenously expressed protein. Compositions and articles defined by the invention were isolated or otherwise manufactured in connection with the examples provided below. Other features and advantages of the invention will be apparent from the detailed description, and from the claims.

Definitions

By "agent" is meant any small molecule chemical compound, antibody, nucleic acid molecule, or polypeptide, or fragments thereof.

By "ameliorate" is meant decrease, suppress, attenuate, diminish, arrest, or stabilize the development or progression of a disease.

By "alteration" is meant a change (increase or decrease) in the expression levels or activity of a gene or polypeptide as detected by standard art known methods such as those described herein. As used herein, an alteration includes a 10% change in expression levels, preferably a 25% change, more preferably a 40% change, and most preferably a 50% or greater change in expression levels. "

By "analog" is meant a molecule that is not identical, but has analogous functional or structural features. For example, a polypeptide analog retains the biological activity of a corresponding naturally-occurring polypeptide, while having certain biochemical modifications that enhance the analog's function relative to a naturally occurring polypeptide. Such biochemical modifications could increase the analog's protease resistance, membrane permeability, or half-life, without altering, for example, ligand binding. An analog may include an unnatural amino acid.

By "BAG domain" is meant a protein interaction domain having at least about 85% amino acid sequence identity or having similarity to a BAG domain of BAG1, 2, 3, 4, 5, or 6, or a fragment thereof capable of mediating protein-protein interactions. In one particular embodiment, a BAG domain has at least 85%, 95%, or even 100% identity to a BAG domain set forth in Figure 11. In particular embodiments, a BAG domain mediates interactions with Hsp70 molecular chaperones, and includes any BAG domains recognized within members of the BAG polypeptide super family. For example, a "BAG domain" may refer to a BAG domain with a WW domain (similar to that found in the BAG3 polypeptide family), a BAG domain with a BAG1 type sequence for BAG1 N-terminal ubiquitin like domain (similar to that found in the BAG1 polypeptide family), a BAG domain with an LMBR1 domain (present only in non-mammalian BAG super family members), or a BAG domain with no known representative protein domain. An exemplary "BAG domain" is represented by amino acids 419-500 of the human BAG3 protein (GenBank Accession GI:38502170), having the sequence shown below:

HPGVLKVEAILEKVQGLEQAVDNFEGKKTDKKYLMIEEYLTKELLALDSVDPEGRADV
RQARRDGVRKVQTILEKLEQKAID.

A "BAG3 family polypeptide" is a polypeptide having at least 85% amino acid sequence identity or similarity to human BAG3 protein (GenBank Accession GI:38502170) or a fragment thereof capable of mediating protein-protein interactions.. In one embodiment, a BAG3 polypeptide has the exemplary sequence shown below:

MSAATHSPMMQVASGNGDRDPLPPGWEIKIDPQTGWPFFVDHNSRTTTWNDPRVPSEG
PKETPSSANGPSREGSRLPPAREGHPVYPQLRPGYIPIPVLHEGAENRQVHPFHVYPQ
PGMQRFRTEAAAAAPQRSQSPLRGMPETTQPDKQCGQVAAAAAAQPPASHGPERSQSP
AASDCSSSSSSASLPSSGRSSLGSHQLPRGYISIPVIHEQNVTRPAAQPSFHQAQKTH
YPAQQGEYQTHQPVYHKIQGDDWEPRPLRAASPFRSSVQGASSREGSPARSSTPLHSP
SPIRVHTVVDRPQQPMTHRETAPVSQPENKPESKPGPVGPELPPGHIPIQVIRKEVDS
KPVSQKPPPPSEKVEVKVPPAPVPCPPPSPGPSAVPSSPKSVATEERAAPSTAPAEAT
PPKPGEAEAPPKHPGVLKVEAILEKVQGLEQAVDNFEGKKTDKKYLMIEEYLTKELLA
LDSVDPEGRADVRQARRDGVRKVQTILEKLEQKAIDVPGQVQVYELQPSNLEADQPLQ
AIMEMGAVAADKGKKNAGNAEDPHTETQQPEATAAATSNPSSMTDTPGNPAAP.

A "BAG1 family polypeptide" is a polypeptide having at least 85% amino acid sequence identity or similarity to GenBank Accession GI:288915525 or a fragment thereof capable of mediating protein-protein interactions.. In one embodiment, a BAG1 polypeptide has the exemplary sequence shown below:

MAQRGGARRPRGDRERLGSRLRALRPGREPRQSEPPAQRGPPPSRRPPARSTASGHDRPTRGAAAGARR
PRMKKKTRRRSTRSEELTRSEELTLSEEATWSEEATQSEEATQGEEMNRSQEVTRDEESTRSEEVTREEMAAAGL
TVTVTHSNEKHDLHVTSQQGSSEPVVQDLAQVVEEVIGVPQSFQKLIFKGKSLKEMETPLSALGIQDGCRVMLIG
KKNSPQEEVELKKLKHLEKSVEKIADQLEELNKELTGIQQGFLPKDLQAEALCKLDRRVKATIEQFMKILEEIDT
LILPENFKDSRLKRKGLVKKVQAFLAECDTVEQNICQETERLQSTNFALAE.

Bag1 polypeptides are described, for example, in Takayama et al., Cell 80: 279-284, 1995; Takayama et al., Genomics 35: 494-498, 1996; Takayama et al., J. Biol. Chem. 274: 781-786, 1999; and Takayama et al., J Biol Chem. 1999 Jan 8;274(2):781-6.

In this disclosure, "comprises," "comprising," "containing" and "having" and the like can have the meaning ascribed to them in U.S. Patent law and can mean " includes," "including," and the like; "consisting essentially of" or "consists essentially" likewise has the meaning ascribed in U.S. Patent law and the term is open-ended, allowing for the presence of more than that which is recited so long as basic or novel characteristics of that which is recited is not changed by the presence of more than that which is recited, but excludes prior art embodiments.

"Detect" refers to identifying the presence, absence or amount of the analyte to be detected.

By "detectable label" is meant a composition that when linked to a molecule of interest renders the latter detectable, via spectroscopic, photochemical, biochemical,

immunochemical, or chemical means. For example, useful labels include radioactive isotopes, magnetic beads, metallic beads, colloidal particles, fluorescent dyes, electron-dense reagents, enzymes (for example, as commonly used in an ELISA), biotin, digoxigenin, or haptens.

By "disease" is meant any condition or disorder that damages or interferes with the normal function of a cell, tissue, or organ.

By "effective amount" is meant the amount of an active compound required to ameliorate the symptoms of a disease relative to an untreated patient. The effective amount of active compound(s) used to practice the present invention for therapeutic treatment of a disease varies depending upon the manner of administration, the age, body weight, and general health of the subject. Ultimately, the attending physician or veterinarian will decide the appropriate amount and dosage regimen. Such amount is referred to as an "effective" amount.

The invention provides a number of targets that are useful for the development of highly specific drugs to treat or a disorder characterized by the methods delineated herein. In addition, the methods of the invention provide a facile means to identify therapies that are safe for use in subjects. In addition, the methods of the invention provide a route for analyzing virtually any number of compounds for effects on a disease described herein with high-volume throughput, high sensitivity, and low complexity.

By "fragment" is meant a portion of a polypeptide or nucleic acid molecule. This portion contains, preferably, at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the entire length of the reference nucleic acid molecule or polypeptide. A fragment may contain 10, 20, 30, 40, 50, 60, 70, 80, 90, or 100, 200, 300, 400, 500, 600, 700, 800, 900, or 1000 nucleotides or amino acids.

"Hybridization" means hydrogen bonding, which may be Watson-Crick, Hoogsteen or reversed Hoogsteen hydrogen bonding, between complementary nucleobases. For example, adenine and thymine are complementary nucleobases that pair through the formation of hydrogen bonds.

By "inhibitory nucleic acid" is meant a double-stranded RNA, siRNA, shRNA, or antisense RNA, or a portion thereof, or a mimetic thereof, that when administered to a mammalian cell results in a decrease (e.g., by 10%, 25%, 50%, 75%, or even 90-100%) in the expression of a target gene. Typically, a nucleic acid inhibitor comprises at least a portion of a target nucleic acid molecule, or an ortholog thereof, or comprises at least a portion of the

complementary strand of a target nucleic acid molecule. For example, an inhibitory nucleic acid molecule comprises at least a portion of any or all of the nucleic acids delineated herein.

By "isolated polynucleotide" is meant a nucleic acid (e.g., a DNA) that is free of the genes which, in the naturally-occurring genome of the organism from which the nucleic acid molecule of the invention is derived, flank the gene. The term therefore includes, for example, a recombinant DNA that is incorporated into a vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote; or that exists as a separate molecule (for example, a cDNA or a genomic or cDNA fragment produced by PCR or restriction endonuclease digestion) independent of other sequences. In addition, the term includes an RNA molecule that is transcribed from a DNA molecule, as well as a recombinant DNA that is part of a hybrid gene encoding additional polypeptide sequence.

By an "isolated polypeptide" is meant a polypeptide of the invention that has been separated from components that naturally accompany it. Typically, the polypeptide is isolated when it is at least 60%, by weight, free from the proteins and naturally-occurring organic molecules with which it is naturally associated. Preferably, the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight, a polypeptide of the invention. An isolated polypeptide of the invention may be obtained, for example, by extraction from a natural source, by expression of a recombinant nucleic acid encoding such a polypeptide; or by chemically synthesizing the protein. Purity can be measured by any appropriate method, for example, column chromatography, polyacrylamide gel electrophoresis, or by HPLC analysis.

By “linker” is meant an amino acid sequence that joins a protein of interest and a bag domain.

By “marker” is meant any protein or polynucleotide having an alteration in expression level or activity that is associated with a disease or disorder.

As used herein, “obtaining” as in “obtaining an agent” includes synthesizing, purchasing, or otherwise acquiring the agent.

"Primer set" means a set of oligonucleotides that may be used, for example, for PCR. A primer set would consist of at least 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 30, 40, 50, 60, 80, 100, 200, 250, 300, 400, 500, 600, or more primers.

By “reduces” is meant a negative alteration of at least 10%, 25%, 50%, 75%, or 100%.

By “reference” is meant a standard or control condition.

A "reference sequence" is a defined sequence used as a basis for sequence comparison. A reference sequence may be a subset of or the entirety of a specified sequence; for example, a segment of a full-length cDNA or gene sequence, or the complete cDNA or gene sequence. For polypeptides, the length of the reference polypeptide sequence will generally be at least about 16 amino acids, preferably at least about 20 amino acids, more preferably at least about 25 amino acids, and even more preferably about 35 amino acids, about 50 amino acids, or about 100 amino acids. For nucleic acids, the length of the reference nucleic acid sequence will generally be at least about 50 nucleotides, preferably at least about 60 nucleotides, more preferably at least about 75 nucleotides, and even more preferably about 100 nucleotides or about 300 nucleotides or any integer thereabout or therebetween.

By "siRNA" is meant a double stranded RNA. Optimally, an siRNA is 18, 19, 20, 21, 22, 23 or 24 nucleotides in length and has a 2 base overhang at its 3' end. These dsRNAs can be introduced to an individual cell or to a whole animal; for example, they may be introduced systemically via the bloodstream. Such siRNAs are used to downregulate mRNA levels or promoter activity.

By "specifically binds" is meant a compound or antibody that recognizes and binds a polypeptide of the invention, but which does not substantially recognize and bind other molecules in a sample, for example, a biological sample, which naturally includes a polypeptide of the invention.

Nucleic acid molecules useful in the methods of the invention include any nucleic acid molecule that encodes a polypeptide of the invention or a fragment thereof. Such nucleic acid molecules need not be 100% identical with an endogenous nucleic acid sequence, but will typically exhibit substantial identity. Polynucleotides having "substantial identity" to an endogenous sequence are typically capable of hybridizing with at least one strand of a double-stranded nucleic acid molecule. Nucleic acid molecules useful in the methods of the invention include any nucleic acid molecule that encodes a polypeptide of the invention or a fragment thereof. Such nucleic acid molecules need not be 100% identical with an endogenous nucleic acid sequence, but will typically exhibit substantial identity. Polynucleotides having "substantial identity" to an endogenous sequence are typically capable of hybridizing with at least one strand of a double-stranded nucleic acid molecule. By "hybridize" is meant pair to form a double-stranded molecule between complementary polynucleotide sequences (e.g., a gene described herein), or portions thereof, under various

conditions of stringency. (See, e.g., Wahl, G. M. and S. L. Berger (1987) Methods Enzymol. 152:399; Kimmel, A. R. (1987) Methods Enzymol. 152:507).

For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and more preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and more preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30° C, more preferably of at least about 37° C, and most preferably of at least about 42° C. Varying additional parameters, such as hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Various levels of stringency are accomplished by combining these various conditions as needed. In a preferred: embodiment, hybridization will occur at 30° C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37° C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and 100 μg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, hybridization will occur at 42° C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50% formamide, and 200 μg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.

For most applications, washing steps that follow hybridization will also vary in stringency. Wash stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include a temperature of at least about 25° C, more preferably of at least about 42° C, and even more preferably of at least about 68° C. In a preferred embodiment, wash steps will occur at 25° C in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42 C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 68° C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will be readily apparent to those skilled in the art. Hybridization techniques are well known to those skilled in the art and are described, for

example, in Benton and Davis (Science 196:180, 1977); Grunstein and Hogness (Proc. Natl. Acad. Sci., USA 72:3961, 1975); Ausubel et al. (Current Protocols in Molecular Biology, Wiley Interscience, New York, 2001); Berger and Kimmel (Guide to Molecular Cloning Techniques, 1987, Academic Press, New York); and Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York.

By "substantially identical" is meant a polypeptide or nucleic acid molecule exhibiting at least 50% identity to a reference amino acid sequence (for example, any one of the amino acid sequences described herein) or nucleic acid sequence (for example, any one of the nucleic acid sequences described herein). Preferably, such a sequence is at least 60%, more preferably 80% or 85%, and more preferably 90%, 95% or even 99% identical at the amino acid level or nucleic acid to the sequence used for comparison.

Sequence identity is typically measured using sequence analysis software (for example, Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, Wis. 53705, BLAST, BESTFIT, GAP, or PILEUP/PRETTYBOX programs). Such software matches identical or similar sequences by assigning degrees of homology to various substitutions, deletions, and/or other modifications. Conservative substitutions typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid, asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine. In an exemplary approach to determining the degree of identity, a BLAST program may be used, with a probability score between e^{-3} and e^{-100} indicating a closely related sequence.

By "subject" is meant a mammal, including, but not limited to, a human or non-human mammal, such as a bovine, equine, canine, ovine, or feline.

Ranges provided herein are understood to be shorthand for all of the values within the range. For example, a range of 1 to 50 is understood to include any number, combination of numbers, or sub-range from the group consisting 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50.

As used herein, the terms “treat,” treating,” “treatment,” and the like refer to reducing or ameliorating a disorder and/or symptoms associated therewith. It will be appreciated that, although not precluded, treating a disorder or condition does not require that the disorder, condition or symptoms associated therewith be completely eliminated.

Unless specifically stated or obvious from context, as used herein, the term "or" is understood to be inclusive. Unless specifically stated or obvious from context, as used herein, the terms "a", "an", and "the" are understood to be singular or plural.

Unless specifically stated or obvious from context, as used herein, the term "about" is understood as within a range of normal tolerance in the art, for example within 2 standard deviations of the mean. About can be understood as within 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.5%, 0.1%, 0.05%, or 0.01% of the stated value. Unless otherwise clear from context, all numerical values provided herein are modified by the term about.

The recitation of a listing of chemical groups in any definition of a variable herein includes definitions of that variable as any single group or combination of listed groups. The recitation of an embodiment for a variable or aspect herein includes that embodiment as any single embodiment or in combination with any other embodiments or portions thereof.

Any compositions or methods provided herein can be combined with one or more of any of the other compositions and methods provided herein.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the sequence of human BAG3 protein (GenBank Accession GI:38502170), with the BAG domain underlined.

Figure 2 shows the sequence of BAG domain containing subfragments 419-536 (Panel A) and 388-500 (Panel B) of human BAG3 protein (GenBank Accession GI:38502170), with the BAG domain underlined.

Figure 3 shows the alignment of BAG domains present in exemplary BAG1 polypeptide family members.

Figure 4 shows the alignment of BAG domains present in exemplary Hsp70 regulators.

Figure 5 shows alignment data for BAG domains that display high levels of identity to the BAG domain of human BAG3 protein (GenBank Accession GI:38502170).

Figure 6 is a schematic diagram illustrating the Folding Tag system for enhancing protein folding. The gene coding the amino acid sequence of the molecular chaperone interacting domain of BAG protein (FOLDING TAG) must be attached to genes of interest, therefore, expressed proteins are fusion proteins with the FOLDING TAG. The expressed protein correctly folded in cells, to escape from the degradation by proteasome, thus leading to an enhancement of stability and an increased number of functional proteins.

Figure 7 is a Western blot showing CFTR protein expression in human cells.

Figure 8 shows the intracellular localization of CFTR in human cells. Nuclei are visualized with DAPI.

Figure 9 is a graph showing CFTR activity as measured by fluorescence intensity in response to forskolin and KSCN. The activity of the cells that recombinantly express CFTR is shown relative to wild-type and mock transfected.

Figure 10 is a Western blot showing expression of IL3Rα2 in human cells.

Figure 11 is a Western blot showing expression of alpha 1 anti-trypsin in human cells.

Figure 12 is the amino acid alignment of human BAG family protein (BAG1-6) using clustalW alignment software.

Figure 13 provides amino acid and polynucleotide sequences of N-terminal (F) and C-terminal (R) Folding Tags.

Figure 14 provides amino acid and polynucleotide sequences of CFTR, IL13 receptor α2 protein (IL13Rα2), or anti-trypsin (α1-AT).

DETAILED DESCRIPTION OF THE INVENTION

The invention features compositions and methods that are useful for enhancing protein folding. The invention is based, at least in part, on the discovery that BAG3 functions in the proper folding and stabilization of binding proteins *in vivo* through the acceleration of protein folding. Stability and functionality of interacting proteins are regulated by the C-terminus of BAG3, where Hsp70/Hsc70 interacts. Accordingly, the invention provides a fusion protein comprising a protein of interest linked to a BAG domain, wherein the BAG domain enhances both the stabilization and function of the protein. Compositions and methods described herein are useful for generating recombinant proteins that can be used for research or therapeutic purposes. In some embodiments, vectors encoding such proteins are useful in polynucleotide therapy *in vivo*.

BAG Family Proteins

BAG family proteins contain a BAG domain that interacts with Hsp70 molecular chaperones, thereby accelerating the protein-folding cycle of Hsp70 molecular chaperones. In addition to the conserved BAG domain, BAG family proteins also contain a diverse array of additional domains, which allow them to interact with specific target proteins or which target them to specific locations within cells. Based on the results reported herein, BAG family proteins likely act as bridging molecules that recruit molecular chaperones to target proteins, presumably modulating protein function through alterations in their conformation.

As reported in more detail below, the Hsp70/Hsc70 interacting motif in BAG family proteins regulates protein folding by recruiting molecular chaperone complex, Hsp70/Hsc70 and co-chaperone proteins. In addition, BAG family proteins associate with the ubiquitin proteasome system to regulate the folding or degradation decision. Thus, the tag sequence accelerates folding and/or prevents activation of the ubiquitin proteasome degradation pathway. In this technology, the amino acid sequence of BAG protein is attached to the protein of interest, and the protein is expressed as a fusion protein with a part of BAG sequence (termed a "FOLDING TAG"). The fusion protein is expressed in cells, and specifically folded with the aid of the FOLDING TAG peptides. Figure 13 provides Folding Tag sequences. One is for the N-terminal tag (F) and the other is for the C-terminal tag (R). BAG domain in nucleotides (font with blue background) and in amino acid (green background), which recruit Hsp70/Hsc70 molecular chaperone complex. Therefore, the protein escapes from the intracellular degradation system, thereby increasing expression of functional proteins.

The present invention is useful not only for in vitro methods of enhancing the expression of recombinant proteins, but is also useful for enhancing protein conformation in diseases associated with abnormal protein folding. Abnormal protein folding is observed in many conformational diseases or degenerative diseases. Misfolded or Unfolded proteins causes problems in which 1) the protein loses an original function, 2) the protein is unstable, or 3) the protein has an unexpected function, which is toxic to cells. These problems cause serious outcomes, especially when the protein is expressed exogenously. To solve these problems, molecular chaperones have been co-expressed with the target proteins. Since molecular chaperone proteins are involved in protein folding and stressed condition induces molecular chaperones, exogenous expression of molecular chaperones or pre-treatment of cells with stress (heat shock) have been used for increasing protein folding[2-6]. The protein of interest is not successfully folded, however, because of problems in the specificity of recognition by exogenously expressed chaperones or stress-induced chaperones. Many tumor cells and cancer cells express more chaperone proteins than under normal conditions, and the expression of chaperone proteins in those cells may make the cells resistant to various stresses such as apoptosis. Thus, activation of molecular chaperone itself may in some cases have harmful side effects, such as tumorigenesis when used as gene therapy[7-9]. The present invention advantageously avoids these drawbacks because the direct attachment of the FOLDING TAG to a protein of interest increases correct folding by recruiting the Hsp70/Hsc70 molecular chaperone folding complex.

Prior to the present discovery, proteasome inhibitors were sometimes used to prevent degradation of exogenously over-expressed proteins. The inhibitor of the proteasome was used in the cell culture medium[10]. Unfolded or abnormally folded proteins are degraded by a protein complex called proteasome. An inhibitor of proteasome, such as MG132, increases the stability of proteins. However, using a proteasome inhibitor to increase production of a particular protein is not very effective. Proteasome inhibitors are not specific to specific target proteins and are sometimes very toxic to cells, because of inhibition of cell division related proteins. In addition, MG132 only inhibits protein degradation, it does not increase proper protein folding. Thus, inhibition of the proteasome often increases unfolded or misfolded proteins, which don't have full biological activity. In fact, the proteasome inhibitor was used to inhibit the degradation of CFTR protein *in vitro*, but stabilized CFTR protein doesn't work properly in cells[11, 12]. By contrast, the present invention directly accesses the protein folding problem, and in fact, as reported in more detail below CFTR protein expressed with a Folding Tag was functional.

Accordingly, the invention provides compositions and methods to enhance function and stabilization of a given protein, and therefore has broad application in many fields. For example, in 1) gene therapy, 2) the production of recombinant protein, 3) experimental protein expression.

Polynucleotide Therapy

Polynucleotide therapy is useful to treat not only a recessive genetic disorder, but also various other diseases such as cancer. In polynucleotide therapy, a given protein is expressed in a human tissue or in specific cells. Most of the effort has been put into developing a delivery system and high expression system in the target tissue or cells. Recently, polynucleotide therapy research has focused on the "posttranslational problem". In over-expression systems, many proteins are degraded due to inappropriate protein conformation at the post-translational level. Thus, increasing protein folding is one of the most important problems to be solved in polynucleotide therapy. The present invention enhances the expression of the given protein, since the fusion protein recruits the molecular chaperone complex to enhance proper folding. Although the expressed protein is a fusion protein and not "a native protein", the TAG sequence itself is derived from human cytoplasmic protein, BAG3. Therefore, the fusion polypeptide is unlikely to be immunogenic.

Recombinant Protein Expression of BAG3 Fusion Polypeptides

A recombinant protein is a protein, which is translated from a recombinant DNA. Molecular biology techniques enable the production of recombinant proteins in a variety of cell types (e.g., bacteria, yeast, insect cells and mammalian cells). Virtually any recombinant protein can be produced using the methods of the invention, including but not limited to antibodies, antigens, hormones, therapeutic drugs, enzymes, probes for molecular screening of interacting partners or small molecules modifying the activity of the recombinant protein *in vitro*. Since the posttranslational modification of proteins depends on each species, bacterial recombinant protein derived from mammalian DNA has some limitation of bioactivity due to physiological modification. Over the last two decades, much effort has been made to enhance the productivity of recombinant protein in mammalian cells.

As mentioned above, many proteins translated in cells are degraded due to their inappropriate conformation. The folding problem decreases the productivity of recombinant proteins. To increase proper folding of recombinant protein, many different experimental modifications have been made, such as little enhancement of protein production by decreased incubation temperature or the addition of proteasome inhibitors for preventing degradation of the protein[10, 14-20]. However, proteasome inhibitors do not solve the folding problem, since the proteasome degradation pathway is activated by inappropriate folding. In addition, proteasome inhibitors are not appropriate for "secreted proteins" and for "stable cell lines" which express the recombinant protein permanently. On the other hand, the present invention solves the protein folding problem of recombinant protein in a physiological way and enables the stabilization of the protein. After expression, the fusion protein is purified using methods known in the art. Such methods include, but are not limited to affinity purification using an antibody against bag domain or an additional tag (myc or flag etc). In one embodiment, the recombinant protein is affinity purified on a protein column conjugating ATPase domain of Hsc70 because of high affinity interaction between ATPase domain of Hsc70 and BAG domain. The TAG protein can be cleaved off with a protease recognizing the peptides sequence between the TAG and the protein of interest, for separation of only target proteins.

The ATPase domain of Hsc70 is highly stable as a bacterial recombinant protein and is purified using a GST fusion protein or equivalent methods (His tag etc). Following purification, the purified protein is cross linked on sepharose or agarose beads for column production. After affinity purification of BAG domain (FOLDING TAG) with this column, the addition of ATP will wash BAG domain (FOLDING TAG) off of the column and regenerate the Hsc70ATPase column. To make the column efficient, short fragments of

Hsc70ATPase (mini-domain), which is sufficient for interacting with BAG domain may be used [21]. The Hsc70ATPase column is an inexpensive, specific association with BAG TAG and easy to reuse. This provides significant advantages relative to antibody based purification methods.

Protein Expression

Molecular biologists rely on exogenous protein expression to assay protein function. The present invention is useful for stable protein expression, and accordingly provides expression vectors, which contain the nucleotide sequence for the TAG with the multi cloning restriction enzyme sites inserted in front or after the TAG. These vectors may be used for expression of the TAG in bacteria, yeast, insects and mammalian cells. The invention further provides methods of optimizing the expression vector constructs to ensure stable protein production. Methods for adjusting the number of nucleotides between the FOLDING TAG and the target protein are known in the art.

The present invention provides methods of treating diseases characterized by the presence of misfolded protein or a reduction in the level of correctly folded protein, and/or disorders or symptoms thereof. Methods of the invention comprise administering a therapeutically effective amount of a pharmaceutical composition comprising a recombinant polynucleotide or protein of the invention to a subject (e.g., a mammal such as a human). Thus, one embodiment is a method of treating a subject suffering from or susceptible to a disease characterized by the presence of misfolded protein or a reduction in the level of correctly folded protein or disorder or symptom thereof. The method includes the step of administering to the mammal a therapeutic amount of an amount of a polynucleotide or polypeptide described herein in an amount sufficient to treat the disease or disorder or symptom thereof, under conditions such that the disease or disorder is treated.

The methods herein include administering to the subject (including a subject identified as in need of such treatment) an effective amount of a compound described herein, or a composition described herein to produce such effect. Identifying a subject in need of such treatment can be in the judgment of a subject or a health care professional and can be subjective (e.g. opinion) or objective (e.g. measurable by a test or diagnostic method).

As used herein, the terms "treat," treating," "treatment," and the like refer to reducing or ameliorating a disorder and/or symptoms associated therewith. It will be appreciated that, although not precluded, treating a disorder or condition does not require that the disorder, condition or symptoms associated therewith be completely eliminated.

As used herein, the terms "prevent," "preventing," "prevention," "prophylactic treatment" and the like refer to reducing the probability of developing a disorder or condition in a subject, who does not have, but is at risk of or susceptible to developing a disorder or condition.

The therapeutic methods of the invention (which include prophylactic treatment) in general comprise administration of a therapeutically effective amount of the polynucleotides or polypeptides described herein, such as a compound of the formulae herein to a subject (e.g., animal, human) in need thereof, including a mammal, particularly a human. Such treatment will be suitably administered to subjects, particularly humans, suffering from, having, susceptible to, or at risk for a disease, disorder, or symptom thereof. Determination of those subjects "at risk" can be made by any objective or subjective determination by a diagnostic test or opinion of a subject or health care provider (e.g., genetic test, enzyme or protein marker, Marker (as defined herein), family history, and the like). The compounds herein may be also used in the treatment of any other disorders in which protein folding defects may be implicated.

In one embodiment, the invention provides a method of monitoring treatment progress. The method includes the step of determining a level of diagnostic marker (Marker) (e.g., any target delineated herein modulated by a compound herein, a protein or indicator thereof, etc.) or diagnostic measurement (e.g., screen, assay) in a subject suffering from or susceptible to a disorder or symptoms thereof associated with the presence of misfolded protein or a reduction in the level of correctly folded protein, in which the subject has been administered a therapeutic amount of a protein or polynucleotide herein sufficient to treat the disease or symptoms thereof. The level of Marker determined in the method can be compared to known levels of Marker in either healthy normal controls or in other afflicted patients to establish the subject's disease status. In preferred embodiments, a second level of Marker in the subject is determined at a time point later than the determination of the first level, and the two levels are compared to monitor the course of disease or the efficacy of the therapy. In certain preferred embodiments, a pre-treatment level of Marker in the subject is determined prior to beginning treatment according to this invention; this pre-treatment level of Marker can then be compared to the level of Marker in the subject after the treatment commences, to determine the efficacy of the treatment.

The practice of the present invention employs, unless otherwise indicated, conventional techniques of molecular biology (including recombinant techniques), microbiology, cell biology, biochemistry and immunology, which are well within the purview

of the skilled artisan. Such techniques are explained fully in the literature, such as, "Molecular Cloning: A Laboratory Manual", second edition (Sambrook, 1989); "Oligonucleotide Synthesis" (Gait, 1984); "Animal Cell Culture" (Freshney, 1987); "Methods in Enzymology" "Handbook of Experimental Immunology" (Weir, 1996); "Gene Transfer Vectors for Mammalian Cells" (Miller and Calos, 1987); "Current Protocols in Molecular Biology" (Ausubel, 1987); "PCR: The Polymerase Chain Reaction", (Mullis, 1994); "Current Protocols in Immunology" (Coligan, 1991). These techniques are applicable to the production of the polynucleotides and polypeptides of the invention, and, as such, may be considered in making and practicing the invention. Particularly useful techniques for particular embodiments will be discussed in the sections that follow.

The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how to make and use the assay, screening, and therapeutic methods of the invention, and are not intended to limit the scope of what the inventors regard as their invention.

EXAMPLES

C-terminal peptides of BAG3, including the BAG domain sequence (see e.g. Fig. 1, underlined sequence), increased protein stability when fused with a protein of interest. Expressed proteins are fusion proteins, part of which contain BAG3, which recruits the protein folding machinery (Hsp70/Hsc70 and co-chaperone, Hdj proteins). The distance between a BAG3 tag and the protein of interest was important for the stabilization of the fusion protein. Since all BAG domain proteins bind to the Hsc70 molecular chaperone and regulate folding activity, all BAG family proteins are likely to have similar activity. Thus, the present invention provides fusion proteins comprising a BAG family protein or fragment thereof. BAG proteins and fragments thereof may be derived from mammalian or non-mammalian BAG family proteins.

Example 1: An Hsc70 interacting domain of BAG3 for use in generating a fusion protein.

The nucleotide sequence encoding amino acids 419-536 or 388-500 of human BAG3 (see e.g. Figure 2, panels A and B, respectively) was subcloned into BamH I and EcoR I sites of Myc-pcDNA3 (Myc tag was inserted into pcDNA3 with Hind III and BamH I site) or into Xho I and Xba I sites of Flag-pcDNA3 (Flag tag was inserted into pcDNA3 with Xba I and Apa I site), respectively.

The generated vector was named Myc-(F)-pcDNA3 and (F)-Flag-pcDNA3, which are commercially available from Invitrogen. A variety of proteins were produced using this vector: (1) cystic fibrosis transmembrane conductance regulator (CFTR) The amino acid and polynucleotide sequences of CFTR are provided at Figure 14A, (2) The amino acid and polynucleotide sequences of IL13 receptor α2 protein (IL13Rα2) are provided at Figure 14B and (3) The amino acid and polynucleotide sequences of α1 anti-trypsin (α1-AT) are provided at Figure 14C. The polynucleotide sequence encoding CFTR was inserted into Myc-(F)-pcDNA3 in frame using two restricted enzyme sites, Not I and Xho I (Myc-(F)-CFTR-pcDNA3).

The nucleic acid sequence encoding IL13Rα2 and α1-AT was inserted into (F)-Flag-pcDNA3 in frame with Kpn I and Xho I or EcoR I and Xho I, respectively (IL13Rα2-(F)-Flag-pcDNA3 and α1-AT-(F)-Flag-pcDNA3). The constructed vectors expressed CFTR, IL13Rα2, or α1-AT as fusion proteins of the amino acid sequences derived from BAG3. The Myc tag or Flag tag was also conjugated to the N-terminus or C-terminus of the amino acid sequence to provide for detection of the fusion protein product. If desired, such tags (Flag and Myc) can be omitted. The antibody against the FOLDING TAG sequence can also be used to detect the expressed protein. The position of the FOLDING TAG and a linker size between the FOLDING TAG and a given protein should be considered may be varied with each protein. Figure 13 provides an exemplary sequence of a folding tag. In one embodiment, 33-38 amino acids is added to the BAG domain. In other embodiments, the linker may be between 15-20, 20-30, 30-40, 40-50, or 50-60 amino acids in length. As shown in Figure 13, a linker is an additional sequence that links an insert and a BAG domain. One of skill in the art will appreciate that the length may be varied to optimize the folding function of Hsp70 as it relates to accessibility of the protein to be folded.

If desired, the FOLDING TAG length can be varied. In one embodiment, the FOLDING TAG is shortened to generate a minimum peptide capable of recruiting a folding complex. In other embodiments, the sequence is varied to create a FOLDING TAG that can be cleaved from the fusion protein. In yet another embodiment, the position of the tag is varied. For certain applications, it may be preferred to have the FOLDING TAG at the C-terminus. In one embodiment, a C-terminal tag may be preferred for a protein with an N-terminal signal sequence or other important function at N-terminus (see example #3). In another embodiment, alternate BAG family proteins from different species may be used. For example, candidate BAG family members may include, but are not limited to, those shown in Figures 3, 4, and 5. Additional candidate BAG family members may include, but are not

limited to, those listed in Tables 1-4. In one embodiment, a BAG protein is selected from the same species as the host cell used to express the recombinant protein and different target proteins. In yet another embodiment, the amino acid sequence of the FOLDING TAG is optimized to modify the FOLDING TAG's activity. An overview of this process is shown in Figure 6.

Table 1. BAG3-like

1. PREDICTED: LOW QUALITY PROTEIN: BAG family molecular chaperone regulator 3-like [Pongo abelii] 575 aa protein XP_002821250.1 GI:297687501
2. PREDICTED: BAG family molecular chaperone regulator 3 [Macaca mulatta] 575 aa protein XP_001104160.2 GI:297301961
3. BCL2-associated athanogene 3 [Bos taurus] 585 aa protein DAA14707.1 GI:296472592
4. PREDICTED: LOW QUALITY PROTEIN: BAG family molecular chaperone regulator 3-like [Callithrix jacchus] 574 aa protein XP_002807518.1 GI:296221348
5. PREDICTED: BCL2-associated athanogene 3 [Oryctolagus cuniculus] 577 aa protein XP_002718735.1 GI:291404909
6. PREDICTED: BCL2-associated athanogene 3-like [Saccoglossus kowalevskii] 396 aa protein XP_002741303.1 GI:291242819
7. hypothetical protein BRAFLDRAFT_98022 [Branchiostoma floridae] 381 aa protein XP_002602349.1 GI:260815175
8. bcl2-associated athanogene [Schistosoma mansoni] 155 aa protein XP_002569618.1 GI:256052104
9. PREDICTED: BCL2-associated athanogene 3 [Taeniopygia guttata] 558 aa protein XP_002188317.1 GI:224052956
10. PREDICTED: similar to BCL2-associated athanogene 3 isoform 1 [Ciona intestinalis] 309 aa protein XP_002128367.1 GI:198436653
11. PREDICTED: similar to BCL2-associated athanogene 3 isoform 2 [Ciona intestinalis] 399 aa protein XP_002128391.1 GI:198436651
12. PREDICTED: BAG family molecular chaperone regulator 3 [Sus scrofa] 574 aa protein XP_001929035.1 GI:194042142
13. unknown [Schistosoma japonicum] 157 aa protein ACE06958.1 GI:189503154
14. BCL2-associated athanogene 3 [Xenopus (Silurana) tropicalis] 562 aa protein NP_001120299.1 GI:187607362
15. BCL2-associated athanogene 3 [Mus musculus] 577 aa protein AAI45766.1 GI:148877849
16. BCL2-associated athanogene 3 [Xenopus laevis] 597 aa protein NP_001079487.1 GI:147907190
17. BAG family molecular chaperone regulator 3 [Bos taurus] 585 aa protein NP_001075940.1 GI:126722829
18. PREDICTED: similar to BCL2-associated athanogene 3, [Monodelphis domestica] 647

aa protein XP_001376803.1 GI:126273363
19. BCL2-associated athanogene 3, isoform CRA_b [Homo sapiens] 575 aa protein EAW49383.1 GI:119569768
20. BCL2-associated athanogene 3, isoform CRA_a [Homo sapiens] 574 aa protein EAW49382.1 GI:119569767
21. PREDICTED: BCL2-associated athanogene 3 [Gallus gallus] 560 aa protein XP_001233435.1 GI:118093076
22. BAG family molecular chaperone regulator 3 [Mus musculus] 577 aa protein NP_038891.4 GI:115270960
23. PREDICTED: hypothetical protein [Pan troglodytes] 685 aa protein XP_508072.2 GI:114633203
24. Bcl-2-interacting death suppressor [Rattus norvegicus] 574 aa protein ABG23394.1 GI:108937164
25. unnamed protein product [Mus musculus] 577 aa protein BAE42639.1 GI:74152742
26. PREDICTED: similar to BAG-family molecular chaperone regulator-3 (BCL-2 binding athanogene-3) (BAG-3) (Bcl-2-binding protein Bis) (Docking protein CAIR-1) [Canis familiaris] 579 aa protein XP_544046.2 GI:73998917
27. BCL2-associated athanogene 3 variant [Homo sapiens] 575 aa protein BAD96520.1 GI:62897159
28. BAG family molecular chaperone regulator 3 [Rattus norvegicus] 574 aa protein NP_001011936.1 GI:58865450
29. BAG family molecular chaperone regulator 3 [Danio rerio] 459 aa protein NP_001003533.1 GI:57525959
30. unknown [Schistosoma japonicum] 128 aa protein AAW25430.1 GI:56754485
31. unnamed protein product [Tetraodon nigroviridis] 492 aa protein CAG04720.1 GI:47216542
32. hypothetical protein with C2 domain,HECT domain (ubiquitin-protein ligase, domain present in Hsp70 regulator and WW / rsp5 / WWP domain [Schistosoma japonicum] 157 aa protein AAP06461.1 GI:29841429
33. BAG family molecular chaperone regulator 3 [Homo sapiens] 575 aa protein NP_004272.2 GI:14043024
34. RecName: Full=BAG family molecular chaperone regulator 3; Short=BAG-3; AltName: Full=Bcl-2-associated athanogene 3; AltName: Full=Bcl-2-binding protein Bis577 aa protein Q9JLV1.1 GI:12643803
35. hypothetical protein DKFZp434E0610.1 - human (fragment) 616 aa protein T46292

GI:11359966
36. unnamed protein product [Mus musculus] 577 aa protein BAA95066.1 GI:7670430
37. Bcl-2-binding protein BIS [Homo sapiens] 575 aa protein AAF26839.1 GI:6724086
38. BAG-family molecular chaperone regulator-3 [Homo sapiens] 575 aa protein AAD16122.2 GI:5868898

Table 2 BAG1-like

1. unnamed protein product [Thellungiella halophila] 293 aa protein BAJ34177.1 GI:312282623
2. PREDICTED: BAG family molecular chaperone regulator 1-like [Sus scrofa] 346 aa protein XP_003130736.1 GI:311265599
3. hypothetical protein PGTG_12108 [Puccinia graminis f. sp. tritici CRL 75-36-700-3] 257 aa protein EFP86152.1 GI:309309561
4. hypothetical protein PGTG_10046 [Puccinia graminis f. sp. tritici CRL 75-36-700-3] 257 aa protein EFP84326.1 GI:309307735
5. CRE-BAG-1 protein [Caenorhabditis remanei] 249 aa protein XP_003099114.1 GI:308473781
6. hypothetical protein SELMODRAFT_103564 [Selaginella moellendorffii] 250 aa protein XP_002975739.1 GI:302787939
7. hypothetical protein SELMODRAFT_99632 [Selaginella moellendorffii] 282 aa protein XP_002973687.1 GI:302783829
8. hypothetical protein SELMODRAFT_164705 [Selaginella moellendorffii]
267 aa protein XP_002961906.1 GI:302756965
9. PREDICTED: BAG family molecular chaperone regulator 1-like [Ailuropoda melanoleuca] 279 aa protein XP_002926223.1 GI:301781606
10. hypothetical protein ARALYDRAFT_485533 [Arabidopsis lyrata subsp. lyrata] 271 aa protein XP_002877823.1 GI:297819880
11. BCL-2-associated athanogene 3 [Arabidopsis lyrata subsp. lyrata] 297 aa protein XP_002873301.1 GI:297810835
12. hypothetical protein ARALYDRAFT_357447 [Arabidopsis lyrata subsp. lyrata] 349 aa protein XP_002865891.1 GI:297796013
13. BCL-2-associated athanogene 2 [Arabidopsis lyrata subsp. lyrata] 282 aa protein XP_002864772.1 GI:297793775
14. unnamed protein product [Vitis vinifera] 238 aa protein CBI15919.3 GI:297745863
15. unnamed protein product [Vitis vinifera] 316 aa protein CBI31837.3 GI:297741106
16. unnamed protein product [Vitis vinifera] 219 aa protein CBI30967.3 GI:297740785
17. unnamed protein product [Vitis vinifera] 223 aa protein CBI18140.3 GI:297735646
18. PREDICTED: BAG family molecular chaperone regulator 1-like [Pongo abelii] 280 aa protein XP_002819746.1 GI:297684217
19. Os08g0546100 [Oryza sativa Japonica Group] 501 aa protein NP_001062419.2 GI:297608934
20. Os06g0126500 [Oryza sativa Japonica Group] 339 aa protein
NP_001056659.2 GI:297605098
21. PREDICTED: BCL2-associated athanogene-like [Bos taurus] 288 aa protein XP_002696356.1 GI:297487622

22. PREDICTED: BCL2-associated athanogene-like [Bos taurus] 269 aa protein XP_001250235.3 GI:297463172
23. PREDICTED: BAG family molecular chaperone regulator 1 isoform 1 [Macaca mulatta] 232 aa protein XP_001090384.2 GI:297270768
24. PREDICTED: BAG family molecular chaperone regulator 1 isoform 2 [Macaca mulatta] 280 aa protein XP_002800106.1 GI:297270766
25. RecName: Full=BAG family molecular chaperone regulator 1; Short=BAG-1; AltName: Full=Bcl-2-associated athanogene 1 345 aa protein Q99933.4 GI:296439462
26. PREDICTED: BAG family molecular chaperone regulator 1-like [Callithrix jacchus] 354 aa protein XP_002743090.1 GI:296190117
27. hypothetical protein LOC100382078 [Zea mays] 316 aa protein NP_001168312.1 GI:293333513
28. PREDICTED: BCL2-associated athanogene [Oryctolagus cuniculus] 354 aa protein XP_002708077.1 GI:291383091
29. predicted protein [Naegleria gruberi] 201 aa protein XP_002679429.1 GI:290993617
30. BAG family molecular chaperone regulator 1 isoform BAG-1S [Homo sapiens] 230 aa protein NP_001165886.1 GI:288915527
31. BAG family molecular chaperone regulator 1 isoform BAG-1L [Homo sapiens] 345 aa protein NP_004314.5 GI:288915525
32. BAG family molecular chaperone regulator 1 isoform 1S [Mus musculus] 219 aa protein NP_001165210.1 GI:284507286
33. hypothetical protein PANDA_015836 [Ailuropoda melanoleuca] 246 aa protein EFB29193.1 GI:281353609
34. C. briggsae CBR-BAG-1 protein [Caenorhabditis briggsae] 209 aa protein XP_002640155.1 GI:268568088
35. hypothetical protein BRAFLDRAFT_94947 [Branchiostoma floridae] 183 aa protein XP_002610811.1 GI:260831730
36. Pc13g01730 [Penicillium chrysogenum Wisconsin 54-1255] 313 aa protein XP_002558618.1 GI:255935183
37. unknown [Glycine max] 235 aa protein ACU20992.1 GI:255641432
38. unknown [Glycine max] 270 aa protein ACU18631.1 GI:255636586
39. unknown [Glycine max] 253 aa protein ACU18456.1 GI:255636230
40. conserved hypothetical protein [Ricinus communis] 296 aa protein XP_002531057.1 GI:255580463
41. protein binding protein, putative [Ricinus communis] 265 aa protein XP_002516645.1 GI:255551196
42. protein binding protein, putative [Ricinus communis] 353 aa protein XP_002510840.1 GI:255539551
43. protein binding protein, putative [Ricinus communis] 301 aa protein XP_002509707.1 GI:255537281
44. predicted protein [Micromonas sp. RCC299] 239 aa protein XP_002501226.1 GI:255075103
45. hypothetical protein SORBIDRAFT_10g001600 [Sorghum bicolor] 347 aa protein XP_002436384.1 GI:242091788
46. hypothetical protein SORBIDRAFT_03g038830 [Sorghum bicolor] 259 aa protein XP_002456583.1 GI:242054875
47. protein binding protein [Zea mays] 317 aa protein NP_001147497.1 GI:226528818
48. protein binding protein [Zea mays] 334 aa protein NP_001146867.1 GI:226502698
49. hypothetical protein LOC100273657 [Zea mays] 259 aa protein NP_001141543.1 GI:226502648

50. protein binding protein [Zea mays] 324 aa protein NP_001149344.1 GI:226500512
51. hypothetical protein LOC100276280 [Zea mays] 242 aa protein NP_001143580.1 GI:226499124
52. protein binding protein [Zea mays] 320 aa protein NP_001151280.1 GI:226495281
53. PREDICTED: hypothetical protein [Vitis vinifera] 343 aa protein XP_002279827.1 GI:225455457
54. PREDICTED: hypothetical protein [Vitis vinifera] 256 aa protein XP_002280537.1 GI:225443974
55. PREDICTED: hypothetical protein [Vitis vinifera] 263 aa protein XP_002271076.1 GI:225439501
56. PREDICTED: hypothetical protein [Vitis vinifera] 279 aa protein XP_002278599.1 GI:225434520
57. predicted protein [Populus trichocarpa] 235 aa protein XP_002323006.1 GI:224139206
58. predicted protein [Populus trichocarpa] 349 aa protein XP_002321874.1 GI:224134647
59. predicted protein [Populus trichocarpa] 308 aa protein XP_002318848.1 GI:224122484
60. predicted protein [Populus trichocarpa] 265 aa protein XP_002313979.1 GI:224105919
61. predicted protein [Populus trichocarpa] 276 aa protein XP_002304467.1 GI:224074851
62. predicted protein [Populus trichocarpa] 241 aa protein XP_002300336.1 GI:224061055
63. predicted protein [Populus trichocarpa] 263 aa protein XP_002298491.1 GI:224055346
64. predicted protein [Populus trichocarpa] 283 aa protein XP_002298065.1 GI:224053973
65. PREDICTED: BCL2-associated athanogene [Taeniopygia guttata] 210 aa protein XP_002187538.1 GI:224045892
66. predicted protein [Thalassiosira pseudonana CCMP1335] 207 aa protein XP_002296269.1 GI:224005236
67. unknown [Zea mays] 322 aa protein ACN25625.1 GI:223943083
68. hypothetical protein OsJ_19959 [Oryza sativa Japonica Group] 339 aa protein EEE65009.1 GI:222634877
69. hypothetical protein OsJ_16195 [Oryza sativa Japonica Group] 213 aa protein EEE61706.1 GI:222629574
70. hypothetical protein OsJ_03960 [Oryza sativa Japonica Group] 237 aa protein EEE55622.1 GI:222619490
71. unknown [Zea mays] 259 aa protein ACL53357.1 GI:219885965
72. hypothetical protein OsI_32096 [Oryza sativa Indica Group] 415 aa protein EEC84905.1 GI:218202478
73. hypothetical protein OsI_04304 [Oryza sativa Indica Group] 262 aa protein EEC71739.1 GI:218189312
74. BAG domain protein [Penicillium marneffei ATCC 18224] 421 aa protein XP_002144054.1 GI:212527794
75. protein binding protein [Zea mays] 334 aa protein ACG37214.1 GI:195635491
76. protein binding protein [Zea mays] 336 aa protein ACG35882.1 GI:195628104
77. unknown [Zea mays] 241 aa protein ACF85598.1 GI:194703028
78. PREDICTED: similar to BAG family molecular chaperone regulator 1 (Bcl-2-associated athanogene 1) (BAG-1) (Glucocorticoid receptor-associated protein RAP46) [Equus caballus] 320 aa protein XP_001917762.1 GI:194224914
79. ATBAG4 (ARABIDOPSIS THALIANA BCL-2-ASSOCIATED ATHANOGENE 4); protein binding [Arabidopsis thaliana] 269 aa protein NP_190746.2 GI:186510943
80. predicted protein [Laccaria bicolor S238N-H82] 178 aa protein XP_001877763.1 GI:170093083
81. hypothetical protein CC1G_03189 [Coprinopsis cinerea okayama7#130] 193 aa protein XP_001830652.1 GI:169847888

82. predicted protein [Physcomitrella patens subsp. patens] 259 aa protein XP_001763281.1 GI:168021504
83. predicted protein [Physcomitrella patens subsp. patens] 231 aa protein XP_001752788.1 GI:168000168
84. hypothetical protein MGL_3498 [Malassezia globosa CBS 7966] 307 aa protein XP_001729463.1 GI:164656671
85. Bcl2-associated athanogene isoform 1L [Rattus norvegicus] 358 aa protein NP_001100117.2 GI:157952208
86. BAG family molecular chaperone regulator 1 isoform 1L [Mus musculus] 355 aa protein NP_033866.4 GI:157952206
87. BCL2-associated athanogene [Gallus gallus] 209 aa protein NP_001103162.1 GI:157841180
88. Bcl2-associated athanogene 1 (predicted), isoform CRA_b [Rattus norvegicus] 181 aa protein EDL98649.1 GI:149045649
89. Bcl2-associated athanogene 1 (predicted), isoform CRA_a [Rattus norvegicus] 219 aa protein EDL98648.1 GI:149045648
90. hypothetical protein LOC558108 [Danio rerio] 206 aa protein NP_001092206.1 GI:148922861
91. Bcl2-associated athanogene 1, isoform CRA_b [Mus musculus] 181 aa protein EDL05423.1 GI:148673476
92. hypothetical protein [Vitis vinifera] 319 aa protein CAN67094.1 GI:147809934
93. hypothetical protein [Vitis vinifera] 343 aa protein CAN63593.1 GI:147799221
94. ATBAG1 (ARABIDOPSIS THALIANA BCL-2-ASSOCIATED ATHANOGENE 1); protein binding [Arabidopsis thaliana] 342 aa protein NP_200019.2 GI:145359142
95. hypothetical protein CNBB2190 [Cryptococcus neoformans var. neoformans B-3501A] 276 aa protein XP_777418.1 GI:134108072
96. PREDICTED: hypothetical protein [Monodelphis domestica] 243 aa protein XP_001364319.1 GI:126333887
97. hypothetical protein OsJ_28168 [Oryza sativa Japonica Group] 317 aa protein EAZ43546.1 GI:125604221
98. hypothetical protein OsI_30134 [Oryza sativa Indica Group] 305 aa protein EAZ07876.1 GI:125562428
99. BCL2-associated athanogene, isoform CRA_d [Homo sapiens] 297 aa protein EAW58517.1 GI:119578921
100. unknown [Picea sitchensis] 254 aa protein ABK26055.1 GI:116791648
101. H0714H04.4 [Oryza sativa (indica cultivar-group)] 272 aa protein CAH66977.1 GI:116309946
102. BCL2-associated athanogene [Bos taurus] 237 aa protein NP_001069759.1 GI:115495295
103. Os09g0524800 [Oryza sativa Japonica Group] 334 aa protein NP_001063716.1 GI:115480245
104. Os04g0619900 [Oryza sativa Japonica Group] 272 aa protein NP_001053903.1 GI:115460606
105. Os01g0831200 [Oryza sativa Japonica Group] 262 aa protein NP_001044700.1 GI:115440841
106. PREDICTED: similar to BAG1 protein [Pan troglodytes] 353 aa protein XP_520528.2 GI:114624364
107. BAG1 protein [Homo sapiens] 310 aa protein AAH14774.2 GI:110611785
108. BAG domain-containing protein [Oryza brachyantha]
332 aa protein

ABG73438.1 GI:110430648
109. BAG family molecular chaperone regulator 1 [Suberites domuncula] 258 aa protein CAJ65915.1 GI:110162114
110. hypothetical protein [Trifolium pratense] 347 aa protein BAE71291.1 GI:84468416
111. PREDICTED: similar to BAG-family molecular chaperone regulator-1 (BCL-2 binding athanogene-1) (BAG-1) (Glucocorticoid receptor-associated protein RAP46) [Canis familiaris] 371 aa protein XP_854585.1 GI:73971741
112. BCL2-associated athanogene isoform 1L variant [Homo sapiens] 345 aa protein BAD96469.1 GI:62897057
113. BCL2-associated athanogene [synthetic construct] 275 aa protein AAX36941.1 GI:60830701
114. hypothetical protein CNB03510 [Cryptococcus neoformans var. neoformans JEC21] 277 aa protein XP_569186.1 GI:58264060
115. BCL-2 binding anthanogene-1 [Hordeum vulgare subsp. vulgare] 259 aa protein CAI39214.1 GI:56797970
116. BCL2-associated athanogene [Homo sapiens] 207 aa protein CAH72518.1 GI:55661642
117. putative BAG domain containing protein [Oryza sativa Japonica Group] 316 aa protein BAD09231.1 GI:42408090
118. ATBAG2 (ARABIDOPSIS THALIANA BCL-2-ASSOCIATED ATHANOGENE 2); protein binding [Arabidopsis thaliana] 285 aa protein NP_568950.2 GI:30697630
119. unknown [Arabidopsis thaliana] 268 aa protein AAM63329.1 GI:21554254
120. unknown [Arabidopsis thaliana] 300 aa protein AAM61448.1 GI:21537107
121. BAG family molecular chaperone regulator Bag101 [Schizosaccharomyces pombe 972h-] 195 aa protein NP_596760.1 GI:19113552
122. BAG1 (human) homolog family member (bag-1) [Caenorhabditis elegans] 210 aa protein NP_491893.1 GI:17507755
123. ATBAG3 (ARABIDOPSIS THALIANA BCL-2-ASSOCIATED ATHANOGENE 3); protein binding [Arabidopsis thaliana] 303 aa protein NP_196339.1 GI:15240726
124. hypothetical protein [Cicer arietinum] 262 aa protein CAC10210.1 GI:10334497
125. unnamed protein product [Arabidopsis thaliana] 326 aa protein BAB11054.1 GI:10177741
126. unnamed protein product [Arabidopsis thaliana] 302 aa protein BAB10172.1 GI:10176928
127. BAG domain containing protein-like [Oryza sativa Japonica Group] 321 aa protein BAA90810.1 GI:6983875
128. hypothetical protein [Oryza sativa Indica Group] 268 aa protein CAB51831.2 GI:6624711
129. RecName: Full=BAG family molecular chaperone regulator 1; Short=BAG-1; AltName: Full=Bcl-2-associated athanogene 1
355 aa protein Q60739.3 GI:5915764
130. Bag1 protein variant [Homo sapiens] 240 aa protein AAD25045.1 GI:4583371
131. Bcl-2 associating athanogene-1 protein [Homo sapiens] 274 aa protein AAD11467.1 GI:4204712
132. unknown [Arabidopsis thaliana] 269 aa protein AAC14405.1 GI:3068705
133. glucocortoid receptor-associated protein RAP46 [Homo sapiens 274 aa protein CAA84624.1 GI:1143476

Table 3 LMBR1 and BAG-like

1. LMBR1 domain-containing protein [Paracoccidioides brasiliensis Pb01] 1035 aa protein XP_002795957.1 GI:295670820

2. conserved hypothetical protein [Uncinocarpus reesii 1704] 1098 aa protein XP_002582789.1 GI:258564088
3. LMBR1 domain protein [Aspergillus flavus NRRL3357] 771 aa protein XP_002383137.1 GI:238503810
4. conserved hypothetical protein [Ajellomyces capsulatus NAm1] 878 aa protein XP_001543103.1 GI:154284616
5. conserved hypothetical protein [Aspergillus terreus NIH2624] 1061 aa protein XP_001212198.1 GI:115389386
6. hypothetical protein AN4211.2 [Aspergillus nidulans FGSC A4] 1124 aa protein XP_661815.1 GI:67527933

Table 4. BAG containing

1. hypothetical protein [Leptosphaeria maculans] 484 aa protein CBX91750.1 GI:312211665
2. hypothetical protein MGYG_01124 [Arthroderma gypseum CBS 118893] 566 aa protein EFQ98087.1 GI:311338885
3. hypothetical protein PTT_14740 [Pyrenophora teres f. teres 0-1] 1082 aa protein EFQ89119.1 GI:311321631
4. PREDICTED: BAG family molecular chaperone regulator 2-like [Sus scrofa] 210 aa protein XP_003128384.1 GI:311260217
5. BAG domain-containing protein [Glomerella graminicola M1.001] 323 aa protein EFQ25840.1 GI:310790307
6. hypothetical protein PGTG_13832 [Puccinia graminis f. sp. tritici CRL 75-36-700-3] 521 aa protein EFP88028.1 GI:309311437
7. CRE-UNC-23 protein [Caenorhabditis remanei] 450 aa protein XP_003097664.1 GI:308470863
8. BAG domain-containing protein Samui [Harpegnathos saltator] 737 aa protein EFN89500.1 GI:307214463
9. BAG domain-containing protein Samui [Camponotus floridanus] 590 aa protein EFN75007.1 GI:307191466
10. BAG domain containing protein [Coccidioides posadasii C735 delta SOWgp] 377 aa protein XP_003068556.1 GI:303317108
11. predicted protein [Micromonas pusilla CCMP1545] 197 aa protein XP_003062309.1 GI:303286039
12. hypothetical protein NECHADRAFT_101954 [Nectria haematococca mpVI 77-13-4] 305 aa protein XP_003047677.1 GI:302897597
13. hypothetical protein TRV_06834 [Trichophyton verrucosum HKI 0517] 482 aa protein XP_003019150.1 GI:302654699
14. hypothetical protein ARB_05573 [Arthroderma benhamiae CBS 112371] 550 aa protein

XP_003016176.1 GI:302508431
15. BAG domain-containing protein [Verticillium albo-atrum VaMs.102] 381 aa protein XP_003000558.1 GI:302405443
16. PREDICTED: BAG family molecular chaperone regulator 2-like [Ailuropoda melanoleuca] 224 aa protein XP_002918970.1 GI:301767100
17. PREDICTED: BAG family molecular chaperone regulator 5-like [Ailuropoda melanoleuca] 569 aa protein XP_002918888.1 GI:301766948
18. PREDICTED: BAG family molecular chaperone regulator 4-like [Ailuropoda melanoleuca] 457 aa protein XP_002917278.1 GI:301763725
19. PREDICTED: BAG family molecular chaperone regulator 3-like [Ailuropoda melanoleuca] 545 aa protein XP_002915457.1 GI:301759215
20. PREDICTED: BAG family molecular chaperone regulator 1 [Xenopus (Silurana) tropicalis] 224 aa protein XP_002939543.1 GI:301620355
21. PREDICTED: BAG family molecular chaperone regulator 4-like [Xenopus (Silurana) tropicalis] 405 aa protein XP_002932772.1 GI:301606422
22. hypothetical protein CC1G_09698 [Coprinopsis cinerea okayama7#130] 352 aa protein XP_001834198.2 GI:299741066
23. BCL-2-associated athanogene 5 [Arabidopsis lyrata subsp. lyrata] 212 aa protein XP_002889912.1 GI:297844062
24. BCL-2-associated athanogene 6 [Arabidopsis lyrata subsp. lyrata] 1050 aa protein XP_002880227.1 GI:297824689
25. unnamed protein product [Vitis vinifera] 908 aa protein CBI40461.3 GI:297745381
26. PREDICTED: BAG family molecular chaperone regulator 5-like isoform 1 [Pongo abelii] 447 aa protein XP_002825189.1 GI:297695954
27. PREDICTED: BAG family molecular chaperone regulator 4-like, partial [Pongo abelii] 367 aa protein XP_002819057.1 GI:297682712
28. PREDICTED: BAG family molecular chaperone regulator 2-like [Pongo abelii] 211 aa protein XP_002817076.1 GI:297678430
29. PREDICTED: BAG family molecular chaperone regulator 4 isoform 2 [Macaca mulatta] 421 aa protein XP_002805355.1 GI:297299241
30. PREDICTED: BAG family molecular chaperone regulator 4 isoform 1 [Macaca mulatta] 457 aa protein XP_001092290.2 GI:297299239
31. PREDICTED: BAG family molecular chaperone regulator 5-like isoform 3 [Macaca mulatta] 488 aa protein XP_002805271.1 GI:297298674
32. PREDICTED: BAG family molecular chaperone regulator 5-like isoform 1 [Macaca

mulatta] 447 aa protein XP_002805269.1 GI:297298670
33. conserved hypothetical protein [Arthroderma otae CBS 113480] 493 aa protein XP_002850741.1 GI:296824976
34. hypothetical protein [Tuber melanosporum Mel28] 266 aa protein XP_002840054.1 GI:296420998
35. PREDICTED: BAG family molecular chaperone regulator 4-like [Callithrix jacchus] 457 aa protein XP_002757004.1 GI:296222015
36. PREDICTED: BAG family molecular chaperone regulator 5-like, partial [Callithrix jacchus] 546 aa protein XP_002754360.1 GI:296215959
37. PREDICTED: BAG family molecular chaperone regulator 2-like [Callithrix jacchus] 211 aa protein XP_002746721.1 GI:296198483
38. Chain C, Crystal Structure Of The Complex Between The Bag5 Bd5 And Hsp70 Nbd 142 aa protein 3A8Y_C GI:292659563
39. PREDICTED: BCL2-associated athanogene 4-like [Oryctolagus cuniculus] 459 aa protein XP_002720829.1 GI:291409094
40. PREDICTED: BCL2-associated athanogene 2 [Oryctolagus cuniculus] 211 aa protein XP_002714564.1 GI:291396428
41. unnamed protein product [Sordaria macrospora] 456 aa protein CBI61348.1 GI:289617625
42. hypothetical protein PANDA_007430 [Ailuropoda melanoleuca] 457 aa protein EFB29769.1 GI:281354185
43. hypothetical protein PANDA_007521 [Ailuropoda melanoleuca] 174 aa protein EFB16610.1 GI:281341026
44. hypothetical protein PANDA_003439 [Ailuropoda melanoleuca] 519 aa protein EFB13668.1 GI:281338084
45. hypothetical protein TcasGA2_TC005382 [Tribolium castaneum] 118 aa protein EFA07817.1 GI:270011369
46. C. briggsae CBR-UNC-23 protein [Caenorhabditis briggsae] 457 aa protein XP_002637426.1 GI:268558870
47. hypothetical protein, conserved [Trypanosoma brucei gambiense DAL972] 415 aa protein CBH12270.1 GI:261329289
48. BAG domain-containing protein [Ajellomyces dermatitidis SLH14081] 482 aa protein XP_002622007.1 GI:261191198
49. hypothetical protein CLUG_00229 [Clavispora lusitaniae ATCC 42720] 198 aa protein XP_002619070.1 GI:260949547

50. putative Uncoordinated protein 23 [Angiostrongylus cantonensis] 98 aa protein CAR63696.1 GI:256251586
51. conserved hypothetical protein [Candida tropicalis MYA-3404] 201 aa protein XP_002549008.1 GI:255728165
52. KLTH0E04840p [Lachancea thermotolerans] 178 aa protein XP_002553694.1 GI:255714825
53. unknown [Glycine max] 179 aa protein ACU24252.1 GI:255647578
54. unknown [Glycine max] 192 aa protein ACU22784.1 GI:255644563
55. protein binding protein, putative [Ricinus communis] 152 aa protein XP_002530491.1 GI:255579292
56. conserved hypothetical protein [Ricinus communis] 182 aa protein XP_002527845.1 GI:255573852
57. hypothetical protein RCOM_1078930 [Ricinus communis] 389 aa protein XP_002514871.1 GI:255547628
58. hypothetical protein RCOM_1617200 [Ricinus communis] 1170 aa protein XP_002511942.1 GI:255541756
59. ZYRO0D15576p [Zygosaccharomyces rouxii] 169 aa protein XP_002497107.1 GI:254582244
60. hypothetical protein [Pichia pastoris GS115] 217 aa protein XP_002490201.1 GI:254566181
61. BAG domain protein [Talaromyces stipitatus ATCC 10500] 326 aa protein XP_002480434.1 GI:242784668
62. predicted protein [Postia placenta Mad-698-R] 188 aa protein XP_002477173.1 GI:242222979
63. predicted protein [Postia placenta Mad-698-R] 698 aa protein XP_002476215.1 GI:242220912
64. hypothetical protein SORBIDRAFT_05g024730 [Sorghum bicolor] 432 aa protein XP_002451133.1 GI:242071713
65. hypothetical protein SORBIDRAFT_05g018040 [Sorghum bicolor] 373 aa protein XP_002450781.1 GI:242071009
66. hypothetical protein SORBIDRAFT_05g018030 [Sorghum bicolor] 405 aa protein XP_002450780.1 GI:242071007
67. hypothetical protein SORBIDRAFT_05g024770 [Sorghum bicolor] 424 aa protein XP_002449877.1 GI:242069201
68. Involucrin, putative [Pediculus humanus corporis] 556 aa protein XP_002429335.1

GI:242017720
69. conserved hypothetical protein [Candida dubliniensis CD36] 199 aa protein XP_002417633.1 GI:241949821
70. hypothetical protein IscW_ISCW012456 [Ixodes scapularis] 137 aa protein XP_002414382.1 GI:241749054
71. BAG protein [Ajellomyces capsulatus H143] 492 aa protein EER45017.1 GI:240281514
72. BAG domain-containing protein [Ajellomyces dermatitidis ER-3] 464 aa protein EEQ83829.1 GI:239606842
73. protein binding protein [Zea mays] 200 aa protein NP_001151301.1 GI:226508350
74. protein binding protein [Zea mays] 213 aa protein NP_001152525.1 GI:226508022
75. hypothetical protein LOC100279586 [Zea mays] 397 aa protein NP_001146055.1 GI:226494841
76. conserved hypothetical protein [Paracoccidioides brasiliensis Pb18] 573 aa protein EEH44237.1 GI:226288725
77. conserved hypothetical protein [Paracoccidioides brasiliensis Pb03] 493 aa protein EEH19866.1 GI:225681582
78. BAG domain-containing protein [Ajellomyces capsulatus G186AR] 488 aa protein EEH04941.1 GI:225556653
79. PREDICTED: similar to IQ calmodulin-binding region; Apoptosis regulator Bcl-2 protein, BAG [Vitis vinifera] 220 aa protein XP_002278460.1 GI:225470016
80. PREDICTED: hypothetical protein [Vitis vinifera] 1221 aa protein XP_002279584.1 GI:225454410
81. BCL-2 associated athanogene 3-like protein [Piriformospora indica] 279 aa protein ACN76856.1 GI:224995894
82. BAG family molecular chaperone regulator 5 [Salmo salar] 252 aa protein ACN58723.1 GI:224587836
83. predicted protein [Populus trichocarpa] 211 aa protein XP_002323161.1 GI:224139542
84. PREDICTED: similar to BCL2-associated athanogene 4, partial [Taeniopygia guttata] 148 aa protein XP_002194339.1 GI:224080885
85. predicted protein [Populus trichocarpa] 530 aa protein XP_002303771.1 GI:224072528
86. predicted protein [Populus trichocarpa] 1227 aa protein XP_002301387.1 GI:224064112
87. PREDICTED: BCL2-associated athanogene 5 [Taeniopygia guttata] 450 aa protein XP_002200616.1 GI:224051771

88. hypothetical protein OsJ_06145 [Oryza sativa Japonica Group] 881 aa protein EEE56689.1 GI:222622557
89. novel protein similar to vertebrate BCL2-associated athanogene 5 (BAG5, zgc:114117) [Danio rerio] 494 aa protein CAX12188.1 GI:220678541
90. hypothetical protein OsI_06641 [Oryza sativa Indica Group] 881 aa protein EEC72871.1 GI:218190444
91. Chain C, Chaperone Complex 88 aa protein 3CQX_C GI:215261183
92. BAG family molecular chaperone regulator 1A [Schizosaccharomyces japonicus yFS275] 203 aa protein XP_002175892.1 GI:213410244
93. BAG family molecular chaperone regulator 1B [Schizosaccharomyces japonicus yFS275] 238 aa protein XP_002172980.1 GI:213404416
94. GA28307 [Drosophila pseudoobscura pseudoobscura] 623 aa protein XP_002135463.1 GI:198463237
95. BCL2-associated athanogene 5 [Xenopus (Silurana) tropicalis] 450 aa protein NP_001124516.1 GI:196259972
96. hypothetical protein TRIADDRAFT_53782 [Trichoplax adhaerens] 411 aa protein XP_002109596.1 GI:195999456
97. BAG domain containing protein [Zea mays] 397 aa protein ACG47983.1 GI:195657031
98. GD14453 [Drosophila simulans] 709 aa protein XP_002084784.1 GI:195590098
99. GE21975 [Drosophila yakuba] 611 aa protein XP_002094712.1 GI:195494144
100. GK17074 [Drosophila willistoni] 602 aa protein XP_002061647.1 GI:195427161
101. GJ13715 [Drosophila virilis] 615 aa protein XP_002047962.1 GI:195378380
102. GM25424 [Drosophila sechellia] 683 aa protein XP_002030411.1 GI:195327408
103. GL20748 [Drosophila persimilis] 623 aa protein XP_002025601.1 GI:195169585
104. GI13926 [Drosophila mojavensis] 603 aa protein XP_002009228.1 GI:195129569
105. GH14821 [Drosophila grimshawi] 651 aa protein XP_001984817.1 GI:195018621
106. GG15647 [Drosophila erecta] 617 aa protein XP_001972663.1 GI:194870498
107. GF23468 [Drosophila ananassae] 638 aa protein XP_001958531.1 GI:194752443
108. unnamed protein product [Homo sapiens] 421 aa protein BAG64979.1 GI:194384412
109. unnamed protein product [Homo sapiens] 178 aa protein BAG63354.1 GI:194377982
110. PREDICTED: BCL2-associated athanogene 5 [Equus caballus] 501 aa protein

XP_001917330.1 GI:194225389
111. PREDICTED: BAG family molecular chaperone regulator 2-like [Sus scrofa] 228 aa protein XP_001927915.1 GI:194040264
112. unnamed protein product [Homo sapiens] 211 aa protein BAG51221.1 GI:193785786
113. unnamed protein product [Homo sapiens] 211 aa protein BAG50848.1 GI:193785482
114. PREDICTED: similar to starvin CG32130-PE [Acyrthosiphon pisum] 413 aa protein XP_001946335.1 GI:193713619
115. BAG family molecular chaperone regulator 2 [Rattus norvegicus] 210 aa protein NP_001121667.1 GI:189491875
116. BAG domain containing protein [Pyrenophora tritici-repentis Pt-1C-BFP] 411 aa protein XP_001930427.1 GI:189188176
117. unnamed protein product [Homo sapiens] 457 aa protein BAG36731.1 GI:189054211
118. hypothetical protein [Podospora anserina S mat+] 453 aa protein XP_001911221.1 GI:171692593
119. predicted protein [Laccaria bicolor S238N-H82] 298 aa protein XP_001880002.1 GI:170097565
120. conserved hypothetical protein [Culex quinquefasciatus] 345 aa protein XP_001844965.1 GI:170034205
121. hypothetical protein [Aspergillus oryzae RIB40] 299 aa protein XP_001816931.1 GI:169764919
122. hypothetical protein SNOG_11736 [Phaeosphaeria nodorum SN15] 349 aa protein XP_001801975.1 GI:169617121
123. predicted protein [Physcomitrella patens subsp. patens] 335 aa protein XP_001778220.1 GI:168051556
124. predicted protein [Physcomitrella patens subsp. patens] 1483 aa protein XP_001774151.1 GI:168043356
125. predicted protein [Physcomitrella patens subsp. patens] 161 aa protein XP_001772412.1 GI:168039855
126. predicted protein [Physcomitrella patens subsp. patens] 184 aa protein XP_001757454.1 GI:168009521
127. predicted protein [Physcomitrella patens subsp. patens] 1028 aa protein XP_001752778.1 GI:168000148
128. starvin, isoform E [Drosophila melanogaster] 635 aa protein NP_001097600.1 GI:161083666

129. BAG family molecular chaperone regulator 4 [Bos taurus] 457 aa protein NP_001104002.1 GI:160333615
130. Chain A, Solution Structure Of The Bag Domain (275-350) Of Bag-Family Molecular Chaperone Regulator-5 89 aa protein 2D9D_A GI:159164011
131. Chain A, Solution Structure Of The Murine Bag Domain Of Bcl2- Associated Athanogene 3 111 aa protein 1UK5_A GI:159163160
132. Chain A, Solution Structure Of The First Murine Bag Domain Of Bcl2- Associated Athanogene 5 99 aa protein 1UGO_A GI:159163137
133. Chain A, Solution Structure Of The Sodd Bag Domain 99 aa protein 1M7K_A GI:159162676
134. Chain A, Solution Structure Of The Bag Domain From Bag4SODD 87 aa protein 1M62_A GI:159162675
135. Chain A, Bag Domain Of Bag1 Cochaperone 135 aa protein 1I6Z_A GI:159162460
136. BAG domain protein [Aspergillus fumigatus A1163] 393 aa protein EDP55944.1 GI:159130831
137. unnamed protein product [Homo sapiens] 447 aa protein BAF84001.1 GI:158256060
138. hypothetical protein AaeL_AAEL000883 [Aedes aegypti] 328 aa protein XP_001651673.1 GI:157111666
139. hypothetical protein AaeL_AAEL000883 [Aedes aegypti] 336 aa protein XP_001651672.1 GI:157111664
140. hypothetical protein Kpol_1063p21 [Vanderwaltozyma polyspora DSM 70294] 165 aa protein XP_001643268.1 GI:156839147
141. PREDICTED: similar to Samui [Nasonia vitripennis] 751 aa protein XP_001607528.1 GI:156549088
142. predicted protein [Nematostella vectensis] 237 aa protein XP_001641534.1 GI:156407404
143. predicted protein [Nematostella vectensis] 51 aa protein XP_001626567.1 GI:156364870
144. predicted protein [Nematostella vectensis] 414 aa protein XP_001622761.1 GI:156352438
145. predicted protein [Nematostella vectensis] 418 aa protein XP_001622759.1 GI:156352434
146. hypothetical protein SS1G_07244 [Sclerotinia sclerotiorum 1980] 297 aa protein XP_001591798.1 GI:156051674
147. hypothetical protein [Leishmania braziliensis MHOM/BR/75/M2904] 291 aa protein

XP_001564807.1 GI:154337148
148. hypothetical protein BC1G_05107 [Botryotinia fuckeliana B05.10] 298 aa protein XP_001555733.1 GI:154312812
149. PREDICTED: similar to BCL2-associated athanogene 4 [Equus caballus] 456 aa protein XP_001493229.1 GI:149742573
150. PREDICTED: similar to BAG family molecular chaperone regulator 2 (Bcl-2-associated athanogene 2) (BAG-2) [Equus caballus] 265 aa protein XP_001499778.1 GI:149732651
151. PREDICTED: similar to BAG family molecular chaperone regulator 3 (Bcl-2-associated athanogene 3) (BAG-3) (Bcl-2-binding protein Bis) (Docking protein CAIR-1) [Equus caballus] 578 aa protein XP_001496329.1 GI:149690028
152. PREDICTED: similar to Bcl2-associated athanogene 2 [Ornithorhynchus anatinus] 206 aa protein XP_001505275.1 GI:149639355
153. PREDICTED: similar to BCL2-associated athanogene 3 [Ornithorhynchus anatinus] 530 aa protein XP_001514849.1 GI:149634658
154. PREDICTED: similar to BCL2-associated athanogene 4 [Ornithorhynchus anatinus] 96 aa protein XP_001506271.1 GI:149634106
155. PREDICTED: hypothetical protein [Ornithorhynchus anatinus] 447 aa protein XP_001511029.1 GI:149441871
156. conserved hypothetical protein [Lodderomyces elongisporus NRRL YB-4239] 224 aa protein XP_001527802.1 GI:149246754
157. BCL2-associated athanogene 4 [Rattus norvegicus] 457 aa protein EDM09068.1 GI:149057825
158. rCG22435 [Rattus norvegicus] 152 aa protein EDL99317.1 GI:149046424
159. hypothetical protein PGUG_02352 [Meyerozyma guilliermondii ATCC 6260] 222 aa protein XP_001484623.1 GI:146417308
160. BAG domain protein [Aspergillus fumigatus Af293] 393 aa protein XP_001481711.1 GI:146322491
161. hypothetical protein MGG_05448 [Magnaporthe oryzae 70-15] 508 aa protein XP_360073.2 GI:145615221
162. hypothetical protein An18g04750 [Aspergillus niger] 313 aa protein XP_001398933.1 GI:145255335
163. PREDICTED: similar to BAG-family molecular chaperone regulator-2; BAG-2 [Monodelphis domestica] 210 aa protein XP_001371890.1 GI:126310295
164. PREDICTED: similar to BCL2-associated athanogene 5 [Monodelphis domestica] 505 aa protein XP_001373299.1 GI:126290435

165. hypothetical protein PICST_30847 [Scheffersomyces stipitis CBS 6054] 202 aa protein XP_001383495.1 GI:126133941
166. PREDICTED: si:dkey-220o5.3 [Danio rerio] 399 aa protein XP_001337521.1 GI:125824741
167. BCL2-associated athanogene 2 [Homo sapiens] 108 aa protein CAM28256.1 GI:123230635
168. BAG domain protein [Aspergillus clavatus NRRL 1] 430 aa protein XP_001269553.1 GI:121702577
169. BCL2-associated athanogene 4, isoform CRA_b [Homo sapiens] 243 aa protein EAW63331.1 GI:119583735
170. BAG domain protein [Neosartorya fischeri NRRL 181] 393 aa protein XP_001265061.1 GI:119496575
171. hypothetical protein CIMG_03702 [Coccidioides immitis RS] 383 aa protein XP_001244261.1 GI:119187309
172. AGAP011762-PA [Anopheles gambiae str. PEST] 347 aa protein XP_320748.3 GI:118793232
173. PREDICTED: similar to silencer of death domains [Gallus gallus] 369 aa protein XP_424388.2 GI:118101378
174. hypothetical protein CHGG_07434 [Chaetomium globosum CBS 148.51] 358 aa protein XP_001225090.1 GI:116198557
175. PREDICTED: hypothetical protein [Strongylocentrotus purpuratus] 475 aa protein XP_001199380.1 GI:115647102
176. Os02g0719700 [Oryza sativa Japonica Group] 213 aa protein NP_001047945.1 GI:115448331
177. PREDICTED: similar to silencer of death domains isoform 1 [Pan troglodytes] 506 aa protein XP_001170419.1 GI:114619701
178. PREDICTED: similar to BAG-family molecular chaperone regulator-2; BAG-2 isoform 1 [Pan troglodytes] 175 aa protein XP_001158291.1 GI:114607983
179. BAG domain-containing protein Samui [Bombyx mori] 677 aa protein NP_001036843.1 GI:112983960
180. PREDICTED: similar to CG32130-PA, isoform A, partial [Apis mellifera] 655 aa protein XP_001123195.1 GI:110776578
181. PREDICTED: BAG family molecular chaperone regulator 2 isoform 3 [Macaca mulatta] 211 aa protein XP_001110620.1 GI:109071591
182. BAG-domain protein 1 / regulator of cell death [Plantago major] 159 aa protein

CAJ38401.1 GI:106879635
183. BCL2-associated athanogene 4 [Bos taurus] 447 aa protein ABF57351.1 GI:95768382
184. BAG family molecular chaperone regulator 5 [Bos taurus] 447 aa protein NP_001035634.1 GI:94966879
185. PREDICTED: similar to starvin CG32130-PE [Tribolium castaneum] 514 aa protein XP_970462.1 GI:91089057
186. unnamed protein product [Macaca fascicularis] 191 aa protein BAE91540.1 GI:90085599
187. IQ calmodulin-binding region; Apoptosis regulator Bcl-2 protein, BAG [Medicago truncatula] 187 aa protein ABD33007.1 GI:87241149
188. hypothetical protein NCU01220 [Neurospora crassa OR74A] 357 aa protein XP_961586.1 GI:85103726
189. BAG family molecular chaperone regulator 2 [Bos taurus] 211 aa protein NP_001029436.1 GI:77735483
190. BAG family molecular chaperone regulator 5 [Danio rerio] 471 aa protein NP_001029355.1 GI:77681460
191. PREDICTED: similar to BAG-family molecular chaperone regulator-4 (BCL2-associated athanogene 4) (BAG-4) (Silencer of death domains) [Canis familiaris] 457 aa protein XP_849918.1 GI:73979313
192. PREDICTED: similar to BAG-family molecular chaperone regulator-2 [Canis familiaris] 211 aa protein XP_853118.1 GI:73973426
193. PREDICTED: similar to BCL2-associated athanogene 5 isoform a [Canis familiaris] 576 aa protein XP_537562.2 GI:73964517
194. hypothetical protein [Trypanosoma brucei TREU927] 415 aa protein XP_845835.1 GI:72391082
195. UNCoordinated family member (unc-23) [Caenorhabditis elegans] 457 aa protein NP_001024011.1 GI:71997311
196. BAG family molecular chaperone regulator 5 [Gallus gallus] 450 aa protein NP_001026382.1 GI:71894943
197. hypothetical protein CaO19.8612 [Candida albicans SC5314] 199 aa protein XP_711238.1 GI:68489890
198. BAG family molecular chaperone regulator 4 [Rattus norvegicus] 457 aa protein NP_001020301.1 GI:68341993
199. BAG family molecular chaperone regulator 5 isoform a [Homo sapiens] 488 aa protein NP_001015049.1 GI:62548856

200. BAG family molecular chaperone regulator 5 [Mus musculus] 447 aa protein NP_081680.1 GI:58037205
201. BAG family molecular chaperone regulator 5 [Rattus norvegicus] 447 aa protein NP_001008526.1 GI:56606102
202. BCL2-associated athanogene 2 [Xenopus (Silurana) tropicalis] 213 aa protein NP_001006729.1 GI:55741948
203. BCL2-associated athanogene [Homo sapiens] 60 aa protein CAH72520.1 GI:55661644
204. YALI0D10769p [Yarrowia lipolytica] 327 aa protein XP_502669.1 GI:50550393
205. DEHA2B14982p [Debaryomyces hansenii CBS767] 206 aa protein XP_457599.1 GI:50416966
206. hypothetical protein [Candida glabrata CBS 138] 153 aa protein XP_447805.1 GI:50290745
207. Chain A, Structural Genomics Of Caenorhabditis Elegans: Structure Of Bag-1 Protein 137 aa protein 1T7S_A GI:49259265
208. BAG2 [Homo sapiens] 211 aa protein CAG38527.1 GI:49065418
209. unknown protein [Oryza sativa Japonica Group] 410 aa protein AAT44147.1 GI:48475078
210. unnamed protein product [Tetraodon nigroviridis] 323 aa protein CAG03774.1 GI:47216770
211. unnamed protein product [Tetraodon nigroviridis] 93 aa protein CAG12283.1 GI:47215891
212. hypothetical protein FG11061.1 [Gibberella zeae PH-1] 570 aa protein XP_391237.1 GI:46139093
213. hypothetical protein FG02098.1 [Gibberella zeae PH-1] 305 aa protein XP_382274.1 GI:46110433
214. hypothetical protein [Ashbya gossypii ATCC 10895] 213 aa protein NP_986905.1 GI:45201335
215. UNCoordinated family member (unc-23) [Caenorhabditis elegans] 399 aa protein NP_872142.1 GI:32566728
216. BCL2-associated athanogene 5 [Homo sapiens] 447 aa protein AAH50551.1 GI:29792008
217. hypothetical protein [Macaca fascicularis] 60 aa protein BAC41231.1 GI:25815212
218. starvin, isoform B [Drosophila melanogaster] 516 aa protein NP_729912.1 GI:24663868

219. starvin, isoform C [Drosophila melanogaster] 542 aa protein NP_729911.1 GI:24663865
220. starvin, isoform A [Drosophila melanogaster] 609 aa protein NP_729910.1 GI:24663861
221. Bag4 protein [Mus musculus] 469 aa protein AAH37239.1 GI:23273683
222. BAG family molecular chaperone regulator 2 [Mus musculus] 210 aa protein NP_663367.1 GI:21703784
223. Bag5 protein [Mus musculus] 205 aa protein AAH27827.1 GI:20379946
224. BAG family molecular chaperone regulator Bag102 [Schizosaccharomyces pombe 972h-] 206 aa protein NP_595316.1 GI:19112108
225. BAG family molecular chaperone regulator 4 [Mus musculus] 457 aa protein NP_080397.1 GI:17975504
226. UNCoordinated family member (unc-23) [Caenorhabditis elegans] 458 aa protein NP_505307.1 GI:17564906
227. BAG6 (BCL-2-ASSOCIATED ATHANOGENE 6); calmodulin binding / protein binding [Arabidopsis thaliana] 1043 aa protein NP_182147.1 GI:15225945
228. ATBAG5 (ARABIDOPSIS THALIANA BCL-2-ASSOCIATED ATHANOGENE 5); protein binding [Arabidopsis thaliana] 215 aa protein NP_172670.1 GI:15221182
229. Bag4 protein [Mus musculus] 147 aa protein AAH09102.1 GI:14318608
230. Chain B, Crystal Structure Of A Bag Domain In Complex With The Hsc70 Atpase Domain 114 aa protein 1HX1_B GI:13399492
231. unnamed protein product [Mus musculus] 147 aa protein BAB28930.1 GI:12851044
232. unnamed protein product [Mus musculus] 414 aa protein BAB24105.1 GI:12838157
233. BAG family molecular chaperone regulator 5 isoform b [Homo sapiens] 447 aa protein NP_004864.1 GI:6631077
234. BAG family molecular chaperone regulator 4 [Homo sapiens] 457 aa protein NP_004865.1 GI:6631075
235. Snl1p [Saccharomyces cerevisiae S288c] 159 aa protein NP_012248.1 GI:6322173
236. BAG family molecular chaperone regulator 2 [Homo sapiens] 211 aa protein NP_004273.1 GI:4757834
237. KIAA0873 protein [Homo sapiens] 466 aa protein BAA74896.1 GI:4240235
238. F12F1.7 [Arabidopsis thaliana] 215 aa protein AAC17606.1 GI:3157923

Example 2: Construction of CFTR expressing plasmid and CFTR expression in human cells

A pcDNA3 with a Myc tag (Myc-pcDNA3) was generated. The gene for Myc tag was inserted into pcDNA3 (Invitrogen) using Hind III and BamH I sites. The *cftr* gene or its mutant (Δ508F) was inserted into Myc-pcDNA3 using Not I and Xho I sites (Myc-CFTR-pcDNA3 or Myc-CFTR Δ508F -pcDNA3). In order to express the protein, the gene for the amino acids (human *bag3* gene amino acid 419-536: indicated by (S) in the constructs) was inserted into Myc-pcDNA3 with two restriction sites (BamH I and EcoR I), then the *cftr* gene was inserted into it using Not I and Xho I sites (Myc-(S)-CFTR-pcDNA3 or Myc-(S)-CFTR □508F -pcDNA3).

To verify the expression of CFTR protein in human cells, the plasmids for Myc-CFTR, Myc-(S)-CFTR, Myc-CFTR Δ508F, and Myc-(S)-CFTR Δ508F were transfected to HEK293 cells with pEGFP-N2 plasmid (Clontech), and cultured for two days. Cells were lysed in RIPA buffer (50 mM Tris-HCl, pH7.5, 150 mM NaCl, 0.1% NP-40, 0.5% sodium deoxycholate, and 0.1% SDS) containing protease inhibitor cocktail and PMSF. After brief sonication, the sample was incubated with SDS-sample buffer at 37 ℃ for 30 minutes, and immunoblotting was performed. CFTR protein was detected with anti-Myc antibody (Figure 7). GFP expression was used for the transfection efficiency, and actin as a loading control.

Example 3: Intracellular localization of CFTR in human cells

The plasmid for Myc-CFTR-pcDNA3 or Myc-(S)-CFTR-pcDNA3 was transfected into HEK293 as described above. For staining, cells were washed with PBS, and fixed with 4% paraformaldehyde in PBS for 5 minutes. After three 5-minute washes with PBS, the fixed samples were incubated in 0.2% Triton X-100 in PBS for 3 minutes and blocked with 2% BSA in PBS for 1 hour. The antibody against Myc tag was incubated with the samples for 1 hour at room temperature. After washing with PBS, the samples were incubated with Alexa Fluor™ conjugated antibodies (Alexa Fluor™ 594 goat anti-mouse IgG (Molecular Probes, Eugene, OR)). After washing with PBS, samples were mounted using Vectashield™ with Dapi (4', 6-diamidino-2-phenylindole: Vector Laboratories, Burlingame, CA) (Figure 8).

Example 4: CFTR function in human cells

HEK293 cells were cultured onto 96-well plate. The cultured cells were transfected with the plasmid for Myc-CFTR-pcDNA3 or Myc-(S)-CFTR-pcDNA3, and incubated for an additional two days. Cells were loaded with SPQ (6-methoxy-*N*-(3-sulfopropyl) quinolinium) by incubating the cells with 5 mM SPQ in hypotonic medium for 15 minutes at

37 °C. The cells were then allowed to recover in HBSS for 15 minutes at 37 °C. Cells were incubated with quenching buffer (10 mM HEPES, pH7.4, 135 mM NaI, 2.4 mM K_2HPO_4, 0.6 mM KH_2PO_4, 1 mM $MgSO_4$, 1 mM $CaSO_4$, 10 mM dextrose) for 30 minutes at 37 °C, and the medium was washed with assay buffer (10 mM HEPES, pH7.4, 135 mM $NaNO_3$, 2.4 mM K_2HPO_4, 0.6 mM KH_2PO_4, 1 mM $MgSO_4$, 1 mM $CaSO_4$, 10 mM dextrose). To activate the CFTR protein, the medium was replaced with assay buffer containing 10 μM forskolin. Baseline activity was measured by adding an equal volume of 300 mM KSCN onto the well directly (final concentration; 150 mM). The fluorescence intensity from SPQ was measured with Tecan Safire II™ (Figure 9).

Soluble IL13Rα2 may be used to inhibit inflammation (e.g., in connection with asthma), or to reduce fibrosis (pulmonary fibrosis, hepatic fibrosis etc).

Example 5: Construction of IL13Rα2 expressing plasmid

Generating recombinant proteins in mammalian cells offers several advantages over other expression systems. Most of recombinant proteins require complicated post-translational modifications such as glycosylation for their function, which cannot be attained or are not accurate in other host cells, such as E. coli or yeast. Despite their slower growth rates and requirements of complex nutrition, mammalian cells have grown as a major host cell for generating complex recombinant proteins[25]. However, mammalian cells have strict intracellular quality control systems for protein production, in which misfolded or unfolded proteins are rapidly degraded in cells[26]. This system could lead to extremely low efficiency of recombinant protein generation. Therefore, attempting enhancement of protein folding has been one of the major issues in this field. For example, cell cultures at low temperature to express molecular chaperones with recombinant proteins have been performed, but production and recovery of recombinant protein from mammalian cells are still not efficient enough for mass production. Interleukin 13 (IL13) mediates allergic inflammation. Therefore, its decoy receptor, IL13Rα2 protein is used in the treatment of asthma. However, the production of soluble IL-13Rα2 using human cells is inefficient because of its instability in cells. The present invention provides efficient methods and compositions for the production of IL-13Rα2.

The gene sequence for Flag tag was made by annealing two single strand nucleotides, and inserted into pcDNA3 with two restriction enzyme sites (Xba I and Apa I; pcDNA3-Flag). cDNA of Jurkat cells (immortalized cell line of human T lymphocyte) was used as a template to generate IL13Rα2 gene by PCR. The gene for IL13Rα2 was inserted into

pcDNA3-Flag with two restriction sites (Kpn I and Xho I). The cDNA for the FOLDING TAG was inserted between Xho I and Xba I sites.

Example 6: IL13Rα2 expression in human cells

To verify the expression of IL13Rα2 protein in human cells, the plasmid for IL13Rα2-Flag-pcDNA3 or IL13Rα2-(S)-Flag-pcDNA3 was transfected to HEK293 cells with pEGFP-N2 plasmid (Clontech), and cultured cells for two days. Cells were lysed in RIPA buffer (50 mM Tris-HCl, pH7.5, 150 mM NaCl, 0.1% NP-40, 0.5% sodium deoxycholate, and 0.1% SDS) containing protease inhibitor cocktail and PMSF. After brief sonication, the sample was boiled in SDS-sample buffer, and immunoblotting was performed. IL13Rα2 protein was detected with anti-Flag antibody (Figure 10). GFP expression was used for the transfection efficiency, and actin as a loading control.

The activity of IL13Rα2 protein is measured by ELISA, which is commercially available.

Example 7: Construction of α1 anti-trypsin expressing plasmid

Most of the proteins currently engineered are secreted proteins. Mammalian cells have become the dominant system for the production of recombinant proteins because of their abilities for post-translational modifications, protein folding, and efficient secretion[25]. Soon after translation commences, secreted proteins with a signal sequence are inserted into the ER for further translation, transferred to the Golgi apparatus, then moved to cell membrane for secretion via vesicle transportation[27]. In the ER, protein folding is strictly controlled since abnormally folded proteins may acquire unfavorable functions, which could have systematic effects if they were to be secreted and circulated in the blood. Mal-folded proteins are typically retained in the ER and are unable to be secreted[28]. The accumulation of unfolded proteins in the ER elicits the unfolded protein response. The typical response has three different part, (i) suppression of translation, (ii) enhanced expression of molecular chaperones, and (iii) degradation of unfolded proteins[29]. Accumulation of unfolded protein in ER induces cell death unless the unfolded proteins are cleared[29]. This explains why enhancing the efficiency of secretion is as important as enhancing the expression level.

To verify the effect of the folding tag on protein secretion, α1 anti-trypsin was used as an example. α1 anti-trypsin is secreted from liver, and circulates in blood vessels. The function of α1 anti-trypsin is to protect tissues, particularly lung tissue, from excess proteases. The lung is damaged by proteases in a disease called α 1 anti-trypsin deficiency. Subjects affected by α 1 anti-trypsin deficiency may develop emphysema, asthma and/or

chronic obstructive pulmonary disease (COPD)[30]. Currently α 1 anti-trypsin (human serum) is used for the treatment of patients with α 1 anti-trypsin deficiency. This treatment is expensive and subjects that receive the human serum are at risk of contracting pathogens present in the human serum[31]. Therefore, the present invention provides recombinant protein technology useful in generating α 1 anti-trypsin, which is not subject to the expense and risks [32].

The construction of an α1 anti-trypsin expressing plasmid is similar to that described above (refer "Construction of IL13Rα2 expressing plasmid"). cDNA of α1 anti-trypsin was generated by PCR using cDNA of Jurkat cells (immortalized cell line of human T lymphocyte) as a template. The generated cDNA for α1 anti-trypsin was inserted into pcDNA3-Flag with two restriction sites (EcoR I and Xho I), and cDNA for FOLDING TAG was inserted into Xho I and Xba I sites.

Example 8: Expression and secretion of α1 anti-trypsin

The plasmid for α1 anti-trypsin was transfected to HEK293 cells with pEGFP-N2 plasmid (Clontech), and cultured cells for two days. Cell medium was harvested to monitor the secretion of α1 anti-trypsin, and cells were lysed in RIPA buffer (50 mM Tris-HCl, pH7.5, 150 mM NaCl, 0.1% NP-40, 0.5% sodium deoxycholate, and 0.1% SDS) containing protease inhibitor cocktail and PMSF. Cell lysate was briefly sonicated, and the sample was boiled in SDS-sample buffer, and applied onto SDS-PAGE. After transfer to PVDF membrane, α1 anti-trypsin was detected with anti-Flag antibody (Figure 11). GFP expression was used for the transfection efficiency. α1 anti-trypsin activity is carried out using a commercially available kit.

In particular embodiments, the invention provides compositions and methods for generating BAG3 fusion polypeptides comprising the proteins delineated in the following tables, as well as vectors encoding these BAG3 fusion polypeptides.

Polynucleotide Therapy

Protein	Application
Cystic fibrosis transmembrane conductance regulator (CFTR)	Cystic Fibrosis [12]
Alpha-1-antitrypsin	Alpha-1-antitrypsin deficiency (COPD)[33]
Survival motor neuron-1 (SMN1)	Spinal muscular atrophy [34]
Aspartoacylase	Canavan disease [35]
Calpain 3 Dysferlin Sarcoglycan	Limb girdle muscular dystrophy [36]
Dystrophin	Muscular dystrophy (DMD) [37, 38]
Utrophin	Muscular dystrophy (DMD) [39]
HIV1 gag	HIV vaccine [40]
NGF	Alzheimer's disease [41]
GAD65 (glutamic acid decarboxylase)	Parkinson's disease [42]
GAD67	Parkinson's disease [42]
AADC (dopa decarboxylase)	Parkinson's disease [43]
Aspartoacylase	Canavan's disease [44]
GM-CSF	Prostate cancer [45]
Low-density lipoprotein receptor	Familial hypercholesterolaemia [46]
Glucocerebrosidase	Gaucher disease [47-49]
FANCC (Fanconi anemia complementation C)	Fanconi anaemia [50]
Factor VIII	Hemophilia A [51, 52]
Factor IX	Hemophilia B [53]
CD86(B7-2)	Malignant melanoma[54]
IL-12	Malignant melanoma [55]
p53 (wild type)	Most type of Cancer (expressing wild type p53) [56]
Adenosine deaminase	Adenosine deaminase deficiency (severe combined immune deficiency) [57]

Recombinant proteins

Protein	**Application**
Factor VIII	Hemophilia A [58]
Factor IX	Hemophilia B [59]
Activated protein C	Sepsis [60]
laronidase	Mucopolysaccharidosis I [61]
Erythropoietin	Anemia [15, 16]
GM-CSF	Hematopoiesis [19]
Iduronate-2-sulfatase	Hunter syndrome [62]
alpha-galactosidase A	Fabry disease [63, 64]
Interferon-beta	Relapsing multiple sclerosis [65]
glycodelin	Immunosuppression [66]
sIL13 receptor alpha 2	Asthma, Pulmonary fibrosis [67]
Tissue plasminogen activator	Myocardial infraction [68]
Alpha-glucosidase	Pompe disease [69-71]
Luteinizing hormone	Infertility [72]
Follicle stimulating hormone	Infertility [73]
TNFalpha receptor fusion protein	Rheumatoid arthritis [74-77]
Beta-glucocerebrosidase	Gaucher's disease [78]
Deoxyribonuclease I	Cystic fibrosis [79]
Antibodies	[80]
Soluble cytokine receptors	Biological therapy [81]
Cytokines	

Experimental / Industrial application

Protein	**Application**
Enzymes	
Cytokines	
Antibodies	
Other proteins	

References

1. Takayama S, Reed JC. Molecular chaperone targeting and regulation by BAG family proteins. *Nat Cell Biol.* Oct 2001;3(10):E237-241.
2. Hsu TA, Betenbaugh MJ. Coexpression of molecular chaperone BiP improves immunoglobulin solubility and IgG secretion from Trichoplusia in insect cells. *Biotechnol Prog.* Jan-Feb 1997;13(1):96-104.
3. Whiteley EM, Hsu TA, Betenbaugh MJ. Thioredoxin domain non-equivalence and anti-chaperone activity of protein disulfide isomerase mutants in vivo. *J Biol Chem.* Sep 5 1997;272(36):22556-22563.
4. Ailor E, Betenbaugh MJ. Overexpression of a cytosolic chaperone to improve solubility and secretion of a recombinant IgG protein in insect cells. *Biotechnol Bioeng.* Apr 20-May 5 1998;58(2-3):196-203.
5. Yokoyama N, Hirata M, Ohtsuka K, Nishiyama Y, Fujii K, Fujita M, Kuzushima K, Kiyono T, Tsurumi T. Co-expression of human chaperone Hsp70 and Hsdj or Hsp40 co-factor increases solubility of overexpressed target proteins in insect cells. *Biochim Biophys Acta.* Sep 7 2000;1493(1-2):119-124.
6. Smales CM, Dinnis DM, Stansfield SH, Alete D, Sage EA, Birch JR, Racher AJ, Marshall CT, James DC. Comparative proteomic analysis of GS-NS0 murine myeloma cell lines with varying recombinant monoclonal antibody production rate. *Biotechnol Bioeng.* Nov 20 2004;88(4):474-488.
7. Jaattela M. Escaping cell death: survival proteins in cancer. *Exp Cell Res.* Apr 10 1999;248(1):30-43.
8. Volloch VZ, Sherman MY. Oncogenic potential of Hsp72. *Oncogene.* Jun 17 1999;18(24):3648-3651.
9. Seo JS, Park YM, Kim JI, Shim EH, Kim CW, Jang JJ, Kim SH, Lee WH. T cell lymphoma in transgenic mice expressing the human Hsp70 gene. *Biochem Biophys Res Commun.* Jan 17 1996;218(2):582-587.
10. Martensen PM, Justesen J. Specific inhibitors prevent proteolytic degradation of recombinant proteins expressed in High Five cells. *Biotechniques.* Apr 2001;30(4):782-784, 786, 788 passim.
11. Ward CL, Omura S, Kopito RR. Degradation of CFTR by the ubiquitin-proteasome pathway. *Cell.* Oct 6 1995;83(1):121-127.
12. Jensen TJ, Loo MA, Pind S, Williams DB, Goldberg AL, Riordan JR. Multiple proteolytic systems, including the proteasome, contribute to CFTR processing. *Cell.* Oct 6 1995;83(1):129-135.
13. Palomares LA, Estrada-Mondaca S, Ramirez OT. Production of recombinant proteins: challenges and solutions. *Methods Mol Biol.* 2004;267:15-52.
14. Schlaeger EJ, Lundstrom K. Effect of temperature on recombinant protein expression in Semliki Forest virus infected mammalian cell lines growing in serum- free suspension cultures. *Cytotechnology.* Nov 1998;28(1-3):205-211.
15. Fogolin MB, Wagner R, Etcheverrigaray M, Kratje R. Impact of temperature reduction and expression of yeast pyruvate carboxylase on hGM-CSF-producing CHO cells. *J Biotechnol.* Apr 8 2004;109(1-2):179-191.
16. Yoon SK, Hong JK, Lee GM. Effect of simultaneous application of stressful culture conditions on specific productivity and heterogeneity of erythropoietin in Chinese hamster ovary cells. *Biotechnol Prog.* Jul-Aug 2004;20(4):1293-1296.
17. Fox SR, Patel UA, Yap MG, Wang DI. Maximizing interferon-gamma production by Chinese hamster ovary cells through temperature shift optimization: experimental and modeling. *Biotechnol Bioeng.* Jan 20 2004;85(2):177-184.
18. Rodriguez J, Spearman M, Huzel N, Butler M. Enhanced production of monomeric interferon-beta by CHO cells through the control of culture conditions. *Biotechnol Prog.* Jan-Feb 2005;21(1):22-30.
19. Bollati-Fogolin M, Forno G, Nimtz M, Conradt HS, Etcheverrigaray M, Kratje R. Temperature reduction in cultures of hGM-CSF-expressing CHO cells: effect on productivity and product quality. *Biotechnol Prog.* Jan-Feb 2005;21(1):17-21.

20. Kumar N, Gammell P, Meleady P, Henry M, Clynes M. Differential protein expression following low temperature culture of suspension CHO-K1 cells. *BMC Biotechnol.* 2008;8:42.
21. Brive L, Takayama S, Briknarova K, Homma S, Ishida SK, Reed JC, Ely KR. The carboxyl-terminal lobe of Hsc70 ATPase domain is sufficient for binding to BAG1. *Biochem Biophys Res Commun.* Dec 21 2001;289(5):1099-1105.
22. Koch C, Hoiby N. Pathogenesis of cystic fibrosis. *Lancet.* Apr 24 1993;341(8852):1065-1069.
23. Kopito RR. Biosynthesis and degradation of CFTR. *Physiol Rev.* Jan 1999;79(1 Suppl):S167-173.
24. Griesenbach U, Alton EW. Gene transfer to the lung: lessons learned from more than 2 decades of CF gene therapy. *Adv Drug Deliv Rev.* Feb 27 2009;61(2):128-139.
25. Wurm FM. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat Biotechnol.* Nov 2004;22(11):1393-1398.
26. Hebert DN, Molinari M. In and out of the ER: protein folding, quality control, degradation, and related human diseases. *Physiol Rev.* Oct 2007;87(4):1377-1408.
27. Sakaguchi M. Eukaryotic protein secretion. *Curr Opin Biotechnol.* Oct 1997;8(5):595-601.
28. Anelli T, Sitia R. Protein quality control in the early secretory pathway. *Embo J.* Jan 23 2008;27(2):315-327.
29. Schroder M, Kaufman RJ. The mammalian unfolded protein response. *Annu Rev Biochem.* 2005;74:739-789.
30. Greene CM, Miller SD, Carroll T, McLean C, O'Mahony M, Lawless MW, O'Neill SJ, Taggart CC, McElvaney NG. Alpha-1 antitrypsin deficiency: a conformational disease associated with lung and liver manifestations. *J Inherit Metab Dis.* Feb 2008;31(1):21-34.
31. Petrache I, Hajjar J, Campos M. Safety and efficacy of alpha-1-antitrypsin augmentation therapy in the treatment of patients with alpha-1-antitrypsin deficiency. *Biologics.* 2009;3:193-204.
32. Karnaukhova E, Ophir Y, Golding B. Recombinant human alpha-1 proteinase inhibitor: towards therapeutic use. *Amino Acids.* Jun 2006;30(4):317-332.
33. Cruz PE, Mueller C, Flotte TR. The promise of gene therapy for the treatment of alpha-1 antitrypsin deficiency. *Pharmacogenomics.* Sep 2007;8(9):1191-1198.
34. Passini MA, Bu J, Roskelley EM, Richards AM, Sardi SP, O'Riordan CR, Klinger KW, Shihabuddin LS, Cheng SH. CNS-targeted gene therapy improves survival and motor function in a mouse model of spinal muscular atrophy. *The Journal of clinical investigation.* Apr 1;120(4):1253-1264.
35. Leone P, Janson CG, Bilaniuk L, Wang Z, Sorgi F, Huang L, Matalon R, Kaul R, Zeng Z, Freese A, McPhee SW, Mee E, During MJ. Aspartoacylase gene transfer to the mammalian central nervous system with therapeutic implications for Canavan disease. *Ann Neurol.* Jul 2000;48(1):27-38.
36. Daniele N, Richard I, Bartoli M. Ins and outs of therapy in limb girdle muscular dystrophies. *Int J Biochem Cell Biol.* 2007;39(9):1608-1624.
37. Romero NB, Braun S, Benveniste O, Leturcq F, Hogrel JY, Morris GE, Barois A, Eymard B, Payan C, Ortega V, Boch AL, Lejean L, Thioudellet C, Mourot B, Escot C, Choquel A, Recan D, Kaplan JC, Dickson G, Klatzmann D, Molinier-Frenckel V, Guillet JG, Squiban P, Herson S, Fardeau M. Phase I study of dystrophin plasmid-based gene therapy in Duchenne/Becker muscular dystrophy. *Hum Gene Ther.* Nov 2004;15(11):1065-1076.
38. Wells DJ. Therapeutic restoration of dystrophin expression in Duchenne muscular dystrophy. *J Muscle Res Cell Motil.* 2006;27(5-7):387-398.
39. Sonnemann KJ, Heun-Johnson H, Turner AJ, Baltgalvis KA, Lowe DA, Ervasti JM. Functional substitution by TAT-utrophin in dystrophin-deficient mice. *PLoS Med.* May 26 2009;6(5):e1000083.
40. Shiver JW, Fu TM, Chen L, Casimiro DR, Davies ME, Evans RK, Zhang ZQ, Simon AJ, Trigona WL, Dubey SA, Huang L, Harris VA, Long RS, Liang X, Handt L, Schleif WA, Zhu L, Freed DC, Persaud NV, Guan L, Punt KS, Tang A, Chen M, Wilson KA, Collins KB, Heidecker GJ, Fernandez VR, Perry HC, Joyce JG, Grimm KM, Cook JC, Keller PM, Kresock DS, Mach H, Troutman RD, Isopi LA, Williams DM, Xu Z, Bohannon KE, Volkin DB, Montefiori DC, Miura A, Krivulka GR, Lifton MA, Kuroda MJ, Schmitz JE, Letvin NL,

Caulfield MJ, Bett AJ, Youil R, Kaslow DC, Emini EA. Replication-incompetent adenoviral vaccine vector elicits effective anti-immunodeficiency-virus immunity. *Nature.* Jan 17 2002;415(6869):331-335.

41. Tuszynski MH, Thal L, Pay M, Salmon DP, U HS, Bakay R, Patel P, Blesch A, Vahlsing HL, Ho G, Tong G, Potkin SG, Fallon J, Hansen L, Mufson EJ, Kordower JH, Gall C, Conner J. A phase 1 clinical trial of nerve growth factor gene therapy for Alzheimer disease. *Nat Med.* May 2005;11(5):551-555.

42. Kim J, Yoon YS, Lee H, Chang JW. AAV-GAD gene for rat models of neuropathic pain and Parkinson's disease. *Acta Neurochir Suppl.* 2008;101:99-105.

43. Christine CW, Starr PA, Larson PS, Eberling JL, Jagust WJ, Hawkins RA, VanBrocklin HF, Wright JF, Bankiewicz KS, Aminoff MJ. Safety and tolerability of putaminal AADC gene therapy for Parkinson disease. *Neurology.* Nov 17 2009;73(20):1662-1669.

44. Arun P, Madhavarao CN, Moffett JR, Hamilton K, Grunberg NE, Ariyannur PS, Gahl WA, Anikster Y, Mog S, Hallows WC, Denu JM, Namboodiri AM. Metabolic acetate therapy improves phenotype in the tremor rat model of Canavan disease. *J Inherit Metab Dis.* Jun;33(3):195-210.

45. Harzstark AL, Small EJ. Immunotherapeutics in development for prostate cancer. *Oncologist.* Apr 2009;14(4):391-398.

46. Grossman M, Rader DJ, Muller DW, Kolansky DM, Kozarsky K, Clark BJ, 3rd, Stein EA, Lupien PJ, Brewer HB, Jr., Raper SE, et al. A pilot study of ex vivo gene therapy for homozygous familial hypercholesterolaemia. *Nat Med.* Nov 1995;1(11):1148-1154.

47. McEachern KA, Nietupski JB, Chuang WL, Armentano D, Johnson J, Hutto E, Grabowski GA, Cheng SH, Marshall J. AAV8-mediated expression of glucocerebrosidase ameliorates the storage pathology in the visceral organs of a mouse model of Gaucher disease. *J Gene Med.* Jun 2006;8(6):719-729.

48. Kim EY, Hong YB, Lai Z, Kim HJ, Cho YH, Brady RO, Jung SC. Expression and secretion of human glucocerebrosidase mediated by recombinant lentivirus vectors in vitro and in vivo: implications for gene therapy of Gaucher disease. *Biochem Biophys Res Commun.* May 28 2004;318(2):381-390.

49. Enquist IB, Nilsson E, Ooka A, Mansson JE, Olsson K, Ehinger M, Brady RO, Richter J, Karlsson S. Effective cell and gene therapy in a murine model of Gaucher disease. *Proc Natl Acad Sci U S A.* Sep 12 2006;103(37):13819-13824.

50. Liu JM, Kim S, Read EJ, Futaki M, Dokal I, Carter CS, Leitman SF, Pensiero M, Young NS, Walsh CE. Engraftment of hematopoietic progenitor cells transduced with the Fanconi anemia group C gene (FANCC). *Hum Gene Ther.* Sep 20 1999;10(14):2337-2346.

51. Roth DA, Tawa NE, Jr., O'Brien JM, Treco DA, Selden RF. Nonviral transfer of the gene encoding coagulation factor VIII in patients with severe hemophilia A. *N Engl J Med.* Jun 7 2001;344(23):1735-1742.

52. Powell JS, Ragni MV, White GC, 2nd, Lusher JM, Hillman-Wiseman C, Moon TE, Cole V, Ramanathan-Girish S, Roehl H, Sajjadi N, Jolly DJ, Hurst D. Phase 1 trial of FVIII gene transfer for severe hemophilia A using a retroviral construct administered by peripheral intravenous infusion. *Blood.* Sep 15 2003;102(6):2038-2045.

53. Qiu X, Lu D, Zhou J, Wang J, Yang J, Meng P, Hsueh JL. Implantation of autologous skin fibroblast genetically modified to secrete clotting factor IX partially corrects the hemorrhagic tendencies in two hemophilia B patients. *Chin Med J (Engl).* Nov 1996;109(11):832-839.

54. Westerman LE, Sund SC, Selvaraj P, Jensen PE. Induction of tumor-specific immunity in mice by immunization with reconstituted tumor membrane liposomes containing recombinant B7-2 (CD86). *J Immunother.* Jul-Aug 2000;23(4):456-463.

55. Komita H, Zhao X, Katakam AK, Kumar P, Kawabe M, Okada H, Braughler JM, Storkus WJ. Conditional interleukin-12 gene therapy promotes safe and effective antitumor immunity. *Cancer Gene Ther.* Dec 2009;16(12):883-891.

56. Bouchet BP, de Fromentel CC, Puisieux A, Galmarini CM. p53 as a target for anti-cancer drug development. *Crit Rev Oncol Hematol.* Jun 2006;58(3):190-207.

57. Qasim W, Gaspar HB, Thrasher AJ. Progress and prospects: gene therapy for inherited immunodeficiencies. *Gene Ther.* Nov 2009;16(11):1285-1291.

58. Roberts HR, Monroe DM, White GC. The use of recombinant factor VIIa in the treatment of bleeding disorders. *Blood.* Dec 15 2004;104(13):3858-3864.
59. White G, Shapiro A, Ragni M, Garzone P, Goodfellow J, Tubridy K, Courter S. Clinical evaluation of recombinant factor IX. *Semin Hematol.* Apr 1998;35(2 Suppl 2):33-38.
60. Levi M, van der Poll T. Recombinant human activated protein C: current insights into its mechanism of action. *Crit Care.* 2007;11 Suppl 5:S3.
61. Kakkis ED, Muenzer J, Tiller GE, Waber L, Belmont J, Passage M, Izykowski B, Phillips J, Doroshow R, Walot I, Hoft R, Neufeld EF. Enzyme-replacement therapy in mucopolysaccharidosis I. *N Engl J Med.* Jan 18 2001;344(3):182-188.
62. Muenzer J, Beck M, Eng CM, Escolar ML, Giugliani R, Guffon NH, Harmatz P, Kamin W, Kampmann C, Koseoglu ST, Link B, Martin RA, Molter DW, Munoz Rojas MV, Ogilvie JW, Parini R, Ramaswami U, Scarpa M, Schwartz IV, Wood RE, Wraith E. Multidisciplinary management of Hunter syndrome. *Pediatrics.* Dec 2009;124(6):e1228-1239.
63. Brady RO, Tallman JF, Johnson WG, Gal AE, Leahy WR, Quirk JM, Dekaban AS. Replacement therapy for inherited enzyme deficiency. Use of purified ceramidetrihexosidase in Fabry's disease. *N Engl J Med.* Jul 5 1973;289(1):9-14.
64. Mapes CA, Anderson RL, Sweeley CC, Desnick RJ, Krivit W. Enzyme replacement in Fabry's disease, an inborn error of metabolism. *Science.* Sep 4 1970;169(949):987-989.
65. Sottini A, Capra R, Serana F, Chiarini M, Caimi L, Imberti L. Interferon-beta therapy monitoring in multiple sclerosis patients. *Endocr Metab Immune Disord Drug Targets.* Mar 2009;9(1):14-28.
66. Seppala M, Koistinen H, Koistinen R. Glycodelins. *Trends Endocrinol Metab.* Apr 2001;12(3):111-117.
67. Donaldson DD, Whitters MJ, Fitz LJ, Neben TY, Finnerty H, Henderson SL, O'Hara RM, Jr., Beier DR, Turner KJ, Wood CR, Collins M. The murine IL-13 receptor alpha 2: molecular cloning, characterization, and comparison with murine IL-13 receptor alpha 1. *J Immunol.* Sep 1 1998;161(5):2317-2324.
68. Martin U, Sponer G, Strein K. Evaluation of thrombolytic and systemic effects of the novel recombinant plasminogen activator BM 06.022 compared with alteplase, anistreplase, streptokinase and urokinase in a canine model of coronary artery thrombosis. *J Am Coll Cardiol.* Feb 1992;19(2):433-440.
69. Raben N, Danon M, Gilbert AL, Dwivedi S, Collins B, Thurberg BL, Mattaliano RJ, Nagaraju K, Plotz PH. Enzyme replacement therapy in the mouse model of Pompe disease. *Mol Genet Metab.* Sep-Oct 2003;80(1-2):159-169.
70. Raben N, Fukuda T, Gilbert AL, de Jong D, Thurberg BL, Mattaliano RJ, Meikle P, Hopwood JJ, Nagashima K, Nagaraju K, Plotz PH. Replacing acid alpha-glucosidase in Pompe disease: recombinant and transgenic enzymes are equipotent, but neither completely clears glycogen from type II muscle fibers. *Mol Ther.* Jan 2005;11(1):48-56.
71. Zhu Y, Li X, McVie-Wylie A, Jiang C, Thurberg BL, Raben N, Mattaliano RJ, Cheng SH. Carbohydrate-remodelled acid alpha-glucosidase with higher affinity for the cation-independent mannose 6-phosphate receptor demonstrates improved delivery to muscles of Pompe mice. *Biochem J.* Aug 1 2005;389(Pt 3):619-628.
72. Durnerin CI, Erb K, Fleming R, Hillier H, Hillier SG, Howles CM, Hugues JN, Lass A, Lyall H, Rasmussen P, Thong J, Traynor I, Westergaard L, Yates R. Effects of recombinant LH treatment on folliculogenesis and responsiveness to FSH stimulation. *Hum Reprod.* Feb 2008;23(2):421-426.
73. Loumaye E, Campbell R, Salat-Baroux J. Human follicle-stimulating hormone produced by recombinant DNA technology: a review for clinicians. *Hum Reprod Update.* Mar 1995;1(2):188-199.
74. Loetscher H, Gentz R, Zulauf M, Lustig A, Tabuchi H, Schlaeger EJ, Brockhaus M, Gallati H, Manneberg M, Lesslauer W. Recombinant 55-kDa tumor necrosis factor (TNF) receptor. Stoichiometry of binding to TNF alpha and TNF beta and inhibition of TNF activity. *J Biol Chem.* Sep 25 1991;266(27):18324-18329.
75. Ashkenazi A, Marsters SA, Capon DJ, Chamow SM, Figari IS, Pennica D, Goeddel DV, Palladino MA, Smith DH. Protection against endotoxic shock by a tumor necrosis factor receptor immunoadhesin. *Proc Natl Acad Sci U S A.* Dec 1 1991;88(23):10535-10539.

76. Van Zee KJ, Kohno T, Fischer E, Rock CS, Moldawer LL, Lowry SF. Tumor necrosis factor soluble receptors circulate during experimental and clinical inflammation and can protect against excessive tumor necrosis factor alpha in vitro and in vivo. *Proc Natl Acad Sci U S A.* Jun 1 1992;89(11):4845-4849.

77. Digel W, Porzsolt F, Schmid M, Herrmann F, Lesslauer W, Brockhaus M. High levels of circulating soluble receptors for tumor necrosis factor in hairy cell leukemia and type B chronic lymphocytic leukemia. *The Journal of clinical investigation.* May 1992;89(5):1690-1693.

78. Friedman B, Vaddi K, Preston C, Mahon E, Cataldo JR, McPherson JM. A comparison of the pharmacological properties of carbohydrate remodeled recombinant and placental-derived beta-glucocerebrosidase: implications for clinical efficacy in treatment of Gaucher disease. *Blood.* May 1 1999;93(9):2807-2816.

79. Suri R. The use of human deoxyribonuclease (rhDNase) in the management of cystic fibrosis. *BioDrugs.* 2005;19(3):135-144.

80. Hoogenboom HR. Selecting and screening recombinant antibody libraries. *Nat Biotechnol.* Sep 2005;23(9):1105-1116.

81. Fernandez-Botran R, Crespo FA, Sun X. Soluble cytokine receptors in biological therapy. *Expert Opin Biol Ther.* Aug 2002;2(6):585-605.

Other Embodiments

From the foregoing description, it will be apparent that variations and modifications may be made to the invention described herein to adopt it to various usages and conditions. Such embodiments are also within the scope of the following claims.

The recitation of a listing of elements in any definition of a variable herein includes definitions of that variable as any single element or combination (or subcombination) of listed elements. The recitation of an embodiment herein includes that embodiment as any single embodiment or in combination with any other embodiments or portions thereof.

All patents and publications mentioned in this specification are herein incorporated by reference to the same extent as if each independent patent and publication was specifically and individually indicated to be incorporated by reference.

What is claimed is:

1. A fusion polypeptide comprising or consisting essentially of a BAG domain, or fragment thereof, that interacts with an Hsp70 molecular chaperone linked to a heterologous polypeptide.

2. The fusion polypeptide of claim 1, wherein the BAG domain is derived from a mammalian BAG3 polypeptide.

3. The fusion polypeptide of claim 1, wherein the BAG domain comprises amino acids 419-536 or 388-500 of human BAG3.

4. The fusion polypeptide of claim 1, wherein the heterologous polypeptide is cystic fibrosis transmembrane conductance regulator (CFTR), IL13 receptor α2 protein (IL13Rα2), or anti-trypsin (α1-AT).

5. A polynucleotide encoding the fusion polypeptide of any of claims 1-4.

6. An expression vector comprising a promoter suitable for expression in a mammalian cell operably linked to a polynucleotide sequence encoding a fusion polypeptide comprising a BAG domain that interacts with an HSP70 molecular chaperone linked to a heterologous polypeptide.

7. The expression vector of claim 6, wherein the polynucleotide sequences encoding the BAG domain and the heterologous polypeptide are linked by a nucleic acid linker comprising between about 5 and 25 nucleic acids.

8. The expression vector of claim 6, wherein the BAG domain is attached to the amino terminus or the carboxy terminus of the heterologous polypeptide, or wherein the BAG domain is positioned between the amino and carboxy termini of the heterologous polypeptide.

9. The fusion protein of claim 17, wherein said BAG domain is linked to said heterologous polypeptide by a linker comprising a protease-sensitive site.

10. A host cell comprising the expression vector of claim 6.

11. A method of producing a BAG fusion protein, said method comprising the steps of:

(a) expressing in a host cell, a polynucleotide encoding a fusion protein of claim 1;

and (b) culturing said host cell under conditions appropriate for production of said fusion protein.

12. A method of treating a subject having a disease associated with the expression of a misfolded protein, the method comprising:

(a) expressing in a cell of the subject, a polynucleotide encoding a fusion protein of claim 1; and

(b) expressing said protein in said cell under conditions appropriate for production of said fusion protein, thereby treating said subject.

FIG. 1

MSAATHSPMMQVASGNGDRDPLPPGWEIKIDPQTGWPFFVDHNSRTTTWNDPRVPSEGPKETPSSANGPSREGSRLPPAR
EGHPVYPQLRPGYIPIPVLHEGAENRQVHPFHVYPQPGMQRFRTEAAAAAPQRSQSPLRGMPETTQPDKQCGQVAAAAAA
QPPASHGPERSQSPAASDCSSSSSSSASLPSSGRSSLGSHQLPRGYISIPVIHEQNVTRPAAQPSFHQAQKTHYPAQQGEY
QTHQPVYHKIQGDDWEPRPLRAASPFRSSVQGASSREGSPARSSTPLHSPSPIRVHTVVDRPQQPMTHRETAPVSQPENK
PESKPGPVGPELPPGHIPIQVIRKEVDSKPVSQKPPPPSEKVEVKVPPAPVPCPPPSPGPSAVPSSPKSVATEERAAPST
APAEATPPKPGEAEAPPKHPGVLKVEAILEKVQGLEQAVDNFEGKKTDKKYLMIEEYLTKELLALDSVDPEGRADVRQAR
RDGVRKVQTILEKLEQKAIDVPGQVQVYELQPSNLEADQPLQAIMEMGAVAADKGKKNAGNAEDPHTETQQPEATAAATS
NPSSMTDTPGNPAAP

FIG. 2A

HPGVLKVEAILEKVQGLEQAVDNFEGKKTDKKYLMIEEYLTKELLALDSVDPEGRADVRQARRDGVRKVQTILEKLEQKA
IDVPGQVQVYELQPSNLEADQPLQAIMEMGAVAADKGK

FIG. 2B

SVATEERAAPSTAPAEATPPKPGEAEAPPKHPGVLKVEAILEKVQGLEQAVDNFEGKKTDKKYLMIEEYLTKELLALDSV
DPEGRADVRQARRDGVRKVQTILEKLEQKAID

FIG. 3

BAG[*smart00264*], BAG domains, present in Bcl-2-associated athanogene 1 and silencer of death domains

```
1M7K_A          22 . [2] .KKIIHVLE    KV. [1] .YLEQEVEE    F. [2] .KKTDKAYWLLEEMLTKELLELDS    VE    T 71
gi 194384412   343 . [2] .KKIIHVLE    KV. [1] .YLEQEVEE    F. [2] .KKTDKAYWLLEEMLTKELLELDS    VE    T 392
gi 62548856    406 . [2] .KAVWNVLG    NL. [1] .EIQGEVLS    F. [2] .NRTDKNYIRLEELLTKQLLALDA    VD    P 455
gi 62548856     50 . [2] .SRLQEIQK    EV. [1] .SVEQQVIG    F. [2] .LSDDKNYKKLERILTKQLFEIDS    VD    T 99
gi 62548856    223 . [2] .AKINFVMC    EV. [2] .ARGVLIAL    L. [2] .VNNNETCRHLSCVLSGLIADLDA    LD    V 273
gi 62548856    316 . [2] .LKIEKVLK    RM. [1] .EIKNELLQ    A      QNPSELYLSSKTELQGLIGQLDE    VS    L 363
gi 32566728    311 . [2] .EDATLMID    EV. [2] .MMHSNIEK    A. [2] .CLQTYMNACSYEETAGATCQNFL    KI    I 361
gi 15225945    595 . [2] .ATVREQMG    DV. [2] .RIEALEAS    T. [2] .HIEEKEIVVNGELVMNLLIKLDA    VE    G 645
gi 15221182     72 . [2] .SSINREAN    RV. [2] .IIQRQETV    D. [2] .RSDEKERLRMNETLMALLIKLDS    VP    G 122
gi 19112108    122 . [2] .AYIDELQQ    DL. [2] .KIEAFCQS. [2] .A. [2] .QDVQDLHTRLSETLLARMIKLDA    VN    V 174
gi 194377982    76 . [2] .KHATRIID    EV. [2] .KFLDDLGN    A. [2] .HLMSLYSACSSEVPHGPVDQKFQ    SI    V 126
gi 3068705     138 . [2] .AAVNAVTG    EV. [6] .VVALEVAV    N. [3] .QVAVREFDMAAELLMRQLLKLDG    IE    A 193
gi 124494251   246 . [2] .EELNKELT    GI. [3] .FLPKDLQA    E. [2] .CKLDRRVKATIEQFMKTLEEIDT    LI. [1] .P 298
gi 114633203   531 . [2] .LKVEAILE    KV. [1] .GLEQAVDN    F. [2] .KKTDKKYLMIEEYLTKELLALDS    VD    P 580
gi 17507755    108 . [2] .DLNLRDVA    DL. [3] .FLEKPKQV    E. [2] .KKLEKKVKYFNEEAERHLETLDG. [2] .II    T 161
gi 19113552    109 . [2] .KAIDQYVD    KE. [3] .MYDNYVNK. [2] .N. [2] .KQKNKQKLMISELLLQQLIKLDG    VD    V 162
gi 6321173      73 . [2] .DNVSLRYG. [1] .EL. [2] .RSKDLINR    F. [4] .EKDIYERNYCNEMLLKLLIELDS    ID    L 126
1UGO_A          16 . [2] .SRLQEIQR    EV. [1] .AIEPQVVG    F. [2] .LSDDKNYKRLERILTKQLFEIDS    VD    T 65
3COX_C           8 . [2] .KHATRIID    EV. [2] .KFLDDLGN    A. [2] .HLMSLYSACSSEVPPGPVDQKFQ    SI    V 58
2D9D_A           8 . [2] .LKIEKVLK    RM. [1] .EIKNELLQ    A      QNPSELYLSSKTELQGLIGQLDE    VS    L 55
1T7S_A          35 . [2] .DLNLRDVA    DL. [3] .FLEKPKQV    E. [2] .KKLEKKVKYFNEEAERHLETLDG. [2] .II    T 88
1HX1_B          26 . [2] .EELNKELT    GI. [3] .FLPKDLQA    E. [2] .CKLDRRVKATIEQFMKTLEEIDT    LI. [1] .P 78
```

```
1M7K_A          72 GG. [3] .VRQARKEAVCKIQAILE    KLEKK. [1] . 99
gi 194384412   393 GG. [3] .VRQARKEAVCKIQAILE    KLEKK. [1] . 420
gi 62548856    456 QG. [3] .CKAARKQAVRLAQNILS    YLDLK. [1] . 483
gi 62548856    100 EG. [3] .IQQARKRAAQETERLLK    ELEQN. [1] . 127
gi 62548856    274 CG. [3] .IRNYRREVVEDINKLLK    YLDLE. [1] . 301
gi 62548856    364 EK. [3] .IREARRRAVIEVQTLIT    YTDLK. [1] . 391
gi 32566728    362 IQ. [5] .QKRIKRRLENLMSQIEN    AERTK. [1] . 391
gi 15225945    646 LH. [2] .IREFRKALATELSSIQD    KLDSL. [1] . 672
gi 15221182    123 LD. [2] .IREARRKVSRKIVGMQE    YLDSI. [1] . 149
gi 19112108    175 ED. [3] .ARLKRKEAIRLSQQYLS    KLDST. [1] . 202
gi 194377982   127 IG. [5] .QKKIKRRLETLIRNIEN    SDKAI. [1] . 156
gi 3068705     194 EG. [1] .AKVQRKAEVRRIQNLQE    AVDKL. [1] . 219
gi 124494251   292 EN. [3] .SRLKRKGLVKKVQAFLA    ECDTV. [1] . 326
gi 114633203   581 EG. [7] .RRDGVRKVQTILEKLEQ. [1] .AIDVP. [1] . 613
gi 17507755    162 ET. [8] .NREKRKTLVNGIQTLLN    QNDAL. [1] . 194
gi 19113552    163 LG. [3] .LRFERKQLVSKIQKMLD    HVDQT. [1] . 190
gi 6322173     127 IN. [8] .LKEKRKGVIKEIQAMLK    SLDSL. [1] . 159
1UGO_A          66 EG. [3] .IQQARKRAAQETERLLK    ELEQN. [1] . 93
3COX_C          59 IG. [5] .QKKIKRRLETLIRNIDN    SDKAI. [1] . 88
2D9D_A          56 EK. [3] .IREARRRAVIEVQTLIT    YTDLK. [1] . 83
1T7S_A          89 ET. [8] .NREKR TLVNGIQTLLN    QNDAL. [1] . 121
1HX1_B          79 EN. [3] .SRLKRKGLVKKVQAFLA    ECDTV. [1] . 106
```

FIG. 3 (Cont.)

FIG. 4

BAG[*pfam02179*], Domain present in Hsp70 regulators.

Domain present in Hsp70 regulators.

1T7S_B	38	.[4].D.[1].ADL.[1].R.[1].FLEKPKQV.[4].KL	EKKV.[4].EE	AERH	LETLDGX	84
gi 121937237	299	.[4].L.[1].DYV.[1].K.[1].LKPLIRQY.[9].AL	EMEH.[4].ET	AMTQ.[1].LLKADGV	351
gi 121924795	264	.[4].I.[1].SKF.[1].T.[1].FVPVAIQY.[9].KR	EFEY.[4].ES	ILTQ.[1].IFKLDGV	316
gi 121781258	273	.[4].I.[1].ADF.[1].S.[1].WLPLCLEY.[9].KR	VDEH.[4].ES	VMQQ.[1].LLKLDAV	325
gi 121807061	216	.[4].L.[1].AYF.[1].R.[1].LLPLCNEY.[9].SR	EFEH.[4].ET	ILAQ.[1].ILRADGI	268
gi 74691723	125	.[4].V.[1].ARY.[1].K.[1].FRERAERV.[9].KD	EWEC.[4].EM	LLKL	LIELDGI	176
gi 731783	73	.[4].V.[1].LRY.[1].N.[1].LEGRSKDL.[9].KD	IYER.[4].EM	LLKL	LIELDSL	124
gi 74612815	67	.[4].I.[1].LRY.[1].N.[1].FEDRVNNI.[9].KS	VYEK.[4].EM	LLKL	LIELDGV	118
gi 74603384	119	.[4].I.[1].DFV.[1].E.[1].YVPEIDEY.[10].DL	EYKY.[4].EM	LLKQ	LMKLDGI	171
gi 74584351	112	.[4].I.[1].DYV.[1].S.[1].YVPEIDKF.[10].DV	EYKF.[4].EM	LLKE	LMKLDAI	164
gi 75004161	372	.[4].M.[1].DEV.[1].E.[1].MESNIEKA.[2].CL	QTYM.[5].EE	TAGA.[1].CQNELKI	418
gi 82183433	117	.[4].I.[1].DEI.[1].K.[1].VMDNLENG.[2].QL	MGLY.[5].EV	PAGP.[1].DQKFQSI	163
gi 12229698	112	.[4].I.[1].DEV.[1].N.[1].FLDDLGNA.[2].HL	MSLY.[5].EV	PHGP.[1].DQKFQSI	158
gi 75220240	598	.[4].Q.[1].GDV.[1].K.[1].IEALEAST.[2].HI	EEKE.[5].EL	VMNL	LIKLDAV	643
gi 122221715	218	.[4].E.[1].EGV.[1].R.[1].VAAEAEAL.[1].RD	ARGR.[4].EA	LMRL	LIRLDAV	261
gi 75219931	75	.[4].E.[1].NRV.[1].S.[1].IQRQETVD.[3].SD	EKER.[4].ET	LMAL	LIKLDSV	120
gi 122199176	60	.[4].E.[1].DQI.[1].R.[1].IQLQDTVD.[3].TN	HLEK.[4].EA	LMKL	LIKLDSI	105
gi 26391464	109	.[4].I.[1].QYV.[1].K.[1].LSPMYDNY.[9].QK	NKQK.[4].EL	LLQQ	LLKDLGV	160
gi 122092851	104	.[4].E.[1].TGI.[1].K.[1].FVQSELVA.[3].RG	LAKR.[5].EF	LMQN	LEKLDAI	150
gi 123797691	259	.[4].E.[1].SGI.[1].Q.[1].FLAKELQA.[3].CK	LDRK.[5].EQ	FMKI	LEEIDTM	305
gi 122143593	141	.[4].D.[1].AGI.[1].Q.[1].FLAKDLQA.[3].CK	LDRR.[5].EQ	FMKI	LEEIDTL	187

FIG. 4 (Cont.)

```
gi 122168666  262 . [4] . V. [1] . AEV. [1] . K    LAPKVAAL. [ 9] . VA    ENDV. [4] . EL    LMNE    LLKLDAV 312
gi 122222012  274 . [4] . V. [1] . LDV. [1] . K    LASKVSAL. [ 9] . VV    DADV. [4] . EA    LMNE    LVKLDSI 324
gi 75170982   144 . [4] . I. [1] . LEV. [1] . R    LGGRVSAF. [ 9] . IA    EKDL. [4] . EL    LMNE    LIKLDAI 194
gi 122202289  142 . [4] . I. [1] . LEV. [1] . R    LAGQVSAM. [ 9] . VV    ETDL. [4] . EK    LMNQ    LIKLDGI 192
gi 122213973  132 . [4] . I. [1] . FQV. [1] . R    LAGQLSAF. [ 9] . VE    EKNL. [4] . EM    LMNQ    LVKLDAI 182
gi 75181093   138 . [4] . I. [1] . FEV. [1] . R    LAGQVSAF. [ 9] . VE    EKSL. [4] . EM    LMNQ    LIRLDAI 188
gi 75107612   131 . [4] . V. [1] . SEV. [1] . K    LSAKVCEL. [ 9] . VE    DKEF. [4] . EL    LMVQ    LLKLDGI 181
gi 75158813   141 . [4] . V. [1] . GEV. [1] . K    LSDRVVAL. [ 9] . VA    VREF. [4] . EL    LMRQ    LLKLDGI 191
gi 12643895   185 . [4] . V. [1] . CEV. [1] . K. [1] . RGVLIALL. [ 3] . NN    NETC. [4] . CV    LSGL    IADLDAL 230
gi 82261118   183 . [4] . V. [1] . VKV. [1] . V. [1] . RGQLVALL. [ 3] . SG    RDSC. [4] . RI    LTEM    QVELDAL 228
gi 82225998   215 . [4] . V. [1] . SQV. [1] . G. [1] . RCQEVALL. [ 3] . SG    RDSC. [4] . RI    LTEL    LVKLDAL 260
gi 82225998   308 . [4] . I. [1] . GRV. [1] . V    LREEVLRH    GA    SGQG    IE    LQGL    LNHLDQV 345
gi 75075382   278 . [4] . V. [1] . TRM. [1] . E    IKNELLQA    QN    PPEL. [5] . TE    LQGL    IGQLDEV 320
gi 68052044   278 . [4] . I. [1] . KKL. [1] . E    VNSLLIKT    EN    ASDL. [5] . AE    LQGL    IAQLDEV 320
gi 74634667   191 . [4] . T. [1] . RDT. [1] . A. [1] . LGPLIKKY. [12] . AP    KEQR. [4] . EM    YLQR    QFTLIDV 245
gi 74812596    71 . [4] . V. [1] . SKA. [1] . S    LQPEIDRF. [ 3] . PH    SKEF. [4] . EN    LEQL    ILSLDNL 115
gi 75075382    12 . [4] . I. [1] . KEV. [1] . S    VEQQVVGF. [ 3] . SD    DKNY. [4] . RI    LTKQ    LTEIDSV 56
gi 82225998    36 . [4] . V. [1] . KEI. [1] . S    LGPQVCSY. [ 3] . QN    DREY. [4] . RE    LTRL    LLEVDKV 80
gi 82261118     7 . [4] . V. [1] . REV. [1] . S    LGPQVCTF. [ 3] . QN    DRDY. [4] . RE    LTQM    LLQVDQV 51
gi 122069375  196 . [4] . I. [1] . RDV. [1] . Q    IFDQVEKF. [ 6] . KK    DKTY. [4] . EM    LTQN    LLKLDSI 243
gi 74870980   376 . [4] . I. [1] . RDV. [1] . E    LMGKVEQF. [ 3] . RE    EKEY. [4] . EM    LTRN    LLKLDTI 420
gi 27151475   383 . [4] . I. [1] . TDV. [1] . N    LMTDVENF. [ 3] . KK    DKRY. [4] . EM    LTRN    LIKLDNI 427
gi 81889881   382 . [4] . V. [1] . EKV. [1] . Y    LEQEVEEF. [ 3] . KT    DKAY. [4] . EM    LTKE    LLELDSV 426
gi 82182154   358 . [4] . I. [1] . ERV. [1] . K    LAQNVKGF. [ 3] . KN    DKRY. [4] . EM    LTKE    LLALDSV 402
gi 82176811   482 . [4] . I. [1] . ERV. [1] . A    LEQAVTGF. [ 3] . KN    EKNY. [4] . ED    LTKV    LLALDSV 526
```

FIG. 4 (Cont.)

```
1UK5_A          31 .[4].I.[1].EKV.[1].G      LEQAVDSF.[ 3].KT  DKKY.[4].EY  LTKE  LLALDSV 75
gi 82225998    392 .[4].V.[1].GSL.[1].D      LQAQVLGF.[ 3].RA  DKSY.[4].EL  LTKQ  LLALDAV 436
gi 75075382    368 .[4].V.[1].GNL.[1].E      IQGEVLSF.[ 3].RT  DKNY.[4].EL  LTKQ  LLALDAV 412
gi 154312812   212 .[4].I.[1].DTF.[1].N.[1].ILPLAEEF.[ 9].KK  DFEH.[4].ET  IMGQ.[1].LLKLDAV 264
gi 146417308   135 .[4].V.[1].DYV.[1].D.[1].YVPQIDEY.[10].DV  QYKY.[4].EM  LLKE  LMKLDGI 187
gi 145615221   422 .[4].I.[1].ESF.[1].E.[1].LRPMCEQF.[ 9].KC  VEEY.[4].ET.[1].LQKV  LLKLDEI 474
gi 145255335   229 .[4].L.[1].SYL.[1].R.[1].LVPLCEEY.[ 9].SR  EFEY.[4].ET  ILAQ.[1].LINADNI 281
gi 149246754   129 .[4].I.[1].DFV.[1].L.[1].YVPEIDKY.[10].DA  EFKF.[4].EM  LLKE  TMRLDGI 181
gi 225455457   149 .[4].I.[1].LEV.[1].R      LAGQVSAL.[ 9].VA  EKVM.[4].EL  LMNQ  LINLDSI 199
gi 126133941   115 .[4].I.[1].DFV.[1].S.[1].YVPQIDTY.[10].EL  EYKY.[4].EM  LLKE  LMKLDEV 167
gi 154284616   793 .[4].L.[1].SYF.[1].D.[1].LEPLIREF.[ 9].TR  DIEH.[4].ET  TMMQ.[1].LLKVDGI 845
gi 156051674   213 .[4].I.[1].HYL.[1].S.[1].IVPLAEAF.[ 9].KR  DFEH.[4].ET  IMGQ.[1].LLKLDAI 265
gi 156407404   161 .[4].E.[1].SQK.[1].R.[1].LADRVLEF.[ 3].KT  DKEY.[4].EM  LLRG  ILTLDEI 206
1UGO_A          19 .[4].I.[1].REV      K.[1].IEPQVVGF.[ 3].SD  DKNY.[4].EI  LTKQ  LFEIDSV 63
1M7K_A          25 .[4].Y.[1].EKV.[1].Y      LEQEVEEF.[ 3].KT  DKAY.[4].EM  LTKE  LLELDSV 69
gi 115448331    78 .[4].E.[1].TRL.[1].R.[1].LRRQETVD.[ 3].GD  ERER.[4].EA  LMAV  LLRLDAV 123
gi 156352434   335 .[4].I      EKI.[1].E.[1].LQGEVEGL.[ 6].KN  SNKY.[4].EM  LTRC  MLDLDQI 382
gi 157357345   125 .[4].Y.[1].AEV.[1].K      LLEKVVAL.[ 3].VN.[6].NKEF.[4].EL  LMRQ  LLKLDGI 175
gi 225425196    77 .[4].E.[1].DEV.[1].R.[1].ISREEIVD.[ 3].RD  SKER.[4].ET  LMSL  LLRLDSV 122
gi 110430648   154 .[4].I.[1].LDV.[1].K      IAAKVTAL.[ 9].VV  DADV.[4].EA  LMNE  IVKLDAI 204
gi 157337073   139 .[4].E.[1].DKL.[1].E.[1].IVALERAV.[ 5].VV  DKEF.[4].EL  LMVQ  LLQLDTI 186
gi 156839147    79 .[4].V.[1].NKY.[1].K.[1].YKDGILKL.[ 9].SE  VYQR.[4].EM  LIKL  MIELDGI 130
gi 148922861   110 .[4].E.[1].TGL.[1].N.[1].FLAKELQA.[ 3].NK  LDQR.[5].EQ  EMKI  LEEIDGM 156
```

FIG. 4 (Cont.)

1T7S B	85	.[14].	NREKRKTLVNGIQTLLNQND	A.	[2].	121
gi 1219372371	352	.[7].	ARKQRKALILKVQSLLKSLD	A.	[2].	381
gi 121924795	317	.[7].	ARLQRKALVKEVQGMLNKLD	E.	[2].	346
gi 121781258	326	.[7].	ARAVRKALVIEVQGVLNKMD	A.	[2].	355
gi 121807061	269	.[7].	ARNARKALVKEAQSTLTKLD	Q.	[2].	298
gi 74691723	177	.[12].	LKKRRKAAILKIQEQLKRLD	E.	[2].	211
gi 731783	125	.[12].	LKEKRKGVIKEIQAMLKSLD	S.	[2].	159
gi 74612815	119	.[12].	LKAKRKAVIKEIQTQLKKLD	T.	[2].	153
gi 74603384	172	.[7].	LRDNRKKVIKFIQDHQRRLD	A.	[2].	201
gi 74584351	165	.[7].	LRDNRRKVIKFIQDHQKRLD	K.	[2].	194
gi 75004161	419	.[4].	AADDQKRIKRRLENLMSQIE	.[4].	T.	[2]. 449
gi 82183433	164	.[4].	AIEDQKRIKRRLETLIRNID	.[4].	S.	[2]. 194
gi 12229698	159	.[4].	ALEDQKKIKRRLETLLRNIE	.[4].	A.	[2]. 189
gi 75220240	644	.[6].	IREFRKALATELSSIQDKLD	S.	[2].	672
gi 122221715	262	.[2].	AREYRRRVTKRVLALQDAVD	A.	[2].	286
gi 75219931	121	.[6].	IREARRKVSRKIVGMQEILD	S.	[2].	149
gi 122199176	106	.[6].	VREARRKVTRRIVGLQEILD	S.	[2].	134
gi 26391464	161	.[7].	LRFERKQLVSKIQKMLDHVD	Q.	[2].	190
gi 122092851	151	.[8].	IRGKRKSAIVRIQALLKRND	D.	[2].	181
gi 123797691	306	.[8].	SRLKRKNLVKKVQVFLAECD	T.	[2].	336
gi 122143593	188	.[8].	SRMKRKGLVKRIQAFLAECD	T.	[2].	218
gi 122168666	313	.[5].	VKAQRRLQVKRVQKYVETLD	A.	[2].	340
gi 122222012	325	.[5].	VKEQRRVQEKRVQKYVEALD	A.	[2].	352
gi 75170982	195	.[5].	VKLQRKMQVKRVQNYVETLD	A.	[2].	222
gi 122202289	193	.[6].	VKLQRKMQVKRVQKYVETLD	M.	[2].	221
gi 122213973	183	.[5].	VKLKKKMQEERLHKYVEALD	L.	[2].	210
gi 75181093	189	.[5].	VKLMRKMQVQRVQKYVEALD	L.	[2].	216
gi 75107612	182	.[5].	ARAQRKAEVRRVQNLVETLD	K.	[2].	209
gi 75158813	192	.[5].	AKVQRKAEVRRIQNLQEAVD	K.	[2].	219
gi 12643895	231	.[7].	IRNYRREVVEDINKLLKYLD	L.	[2].	260
gi 82261118	229	.[7].	IRNYRKQVVEEINGLLKHLD	L.	[2].	258
gi 82225998	261	.[7].	VRNYRKQVVEEINGLLKHLD	L.	[2].	290
gi 82225998	346	.[7].	IREARRRAVLEVQALITFLD	L.	[2].	375
gi 75075382	321	.[7].	IREARRRAVIEVQTLITYID	L.	[2].	350
gi 68052044	321	.[7].	IREARRRAVIEVQTLITYID	L.	[2].	350
gi 74634667	246	.[7].	IRALRKQAVNKMHQYHESLD	E.	[2].	275
gi 74812596	116	.[7].	FRTMRRDAVKEIQQLMEMLD	Y.	[2].	145
gi 75075382	57	.[7].	IQQARKRAAQETERLLKELE	Q.	[2].	86
gi 82225998	81	.[7].	LQLPRKRAAGEVEGLLHYLE	S.	[2].	110
gi 82261118	52	.[7].	LQGARKRAAQEVEGLLRYLE	E.	[2].	81

FIG. 4 (Cont.)

gi 122069375	244	[7]	IKSARKEAIKSINHCIAVLE	A.[2].	273
gi 74870980	421	[7]	IRLARKEAIKCIQASINVLE	A.[2].	450
gi 27151475	428	[7]	IRQARKEAIKCIQKCIAVLE	A.[2].	457
gi 81889881	427	[7]	VRQARKEAVCKIQAILEKLE	K.[2].	456
gi 82182154	403	[7]	VRQARRDGVRRVQNILDELE	M.[2].	432
gi 82176811	527	[7]	VRQARKDGVRKVQKILETLE	Q.[2].	556
1UK5 A	76	[7]	VRQARRDGVRKVQTILEKLE	Q.[2].	105
gi 82225998	437	[7]	TKTARKQAVKHAQNILSYLD	M.[2].	466
gi 75075382	413	[7]	CKAARKQAVKLAQNILSYLD	L.[2].	442
gi 154312812	265	[7]	ARETRKALVREVQNVLNELD	S.[2].	294
gi 146417308	188	[7]	LRENRRKVIKFVQDHQKRLD	K.[2].	217
gi 145615221	475	[7]	IRMRRKELVKYVQEVLKPVD	K.[2].	504
gi 145255335	282	[7]	IRDARRALVKSAQNALNSLD	Q.[2].	311
gi 149246754	182	[7]	LRENRKKVIKFIQDHQKRLD	K.[2].	211
gi 225455457	200	[5]	VKLQRKMQVRRVQKYVETLD	M.[2].	227
gi 126133941	168	[7]	LRENRKKVIKFVQDHQKRLD	R.[2].	197
gi 154284616	846	[7]	ARQQRRTLIVTVQNLLKLLD	E.[2].	875
gi 156051674	266	[5]	ARETRKALVKEAYAVLAGID	S.[2].	293
gi 156407404	207	[7]	VKQARRGAVREIQGYLDKLE	E.[2].	236
1UGO A	64	[7]	IQQARKRAAQETERLLKELE	Q.[2].	93
1M7K A	70	[7]	VRQARKEAVCKIQAILEKLE	K.[2].	99
gi 115448331	124	[6]	VREARRAVTRRVVGLQEVED	A.[2].	152
gi 156352434	383	[7]	VRDARKSAVNHCQDALDKLE	A.[2].	412
gi 157357345	176	[5]	AKVQRRAEVRRVQSLVEMLD	T.[2].	203
gi 225425196	123	[6]	VRDCRKAVIRRAIALQEKVD	A.[2].	151
gi 110430648	205	[5]	VKVQRRMQEKRVQKYVESLD	A.[2].	232
gi 157337073	187	[5]	AKVQRRIEVCRIQSFVDTLD	N.[2].	214
gi 156839147	131	[12]	LKQSRKSAIKLIQGQLSKLD	S.[2].	165
gi 148922861	157	[8]	CRMKKKGLVKTVQGYLAQCD	K.[2].	187

FIG. 5

Sequences producing significant alignments:

	Score (Bits)	E Value
ref|NP_001075940.1| BAG family molecular chaperone regulator ...	159	9e-38
ref|XP_508072.2| PREDICTED: hypothetical protein [Pan troglod...	159	9e-38
gb|DAA14707.1| BCL2-associated athanogene 3 [Bos taurus]	159	1e-37
ref|XP_002821250.1| PREDICTED: LOW QUALITY PROTEIN: BAG famil...	158	3e-37
ref|XP_001104160.2| PREDICTED: BAG family molecular chaperone...	158	3e-37
pir||T46292 hypothetical protein DKFZp434E0610.1 - human (fra...	157	3e-37
gb|EAW49382.1| BCL2-associated athanogene 3, isoform CRA_a [H...	157	4e-37
gb|AAD16122.2| BAG-family molecular chaperone regulator-3 [Ho...	157	4e-37
ref|NP_004272.2| BAG family molecular chaperone regulator 3 [...	157	4e-37
dbj|BAD96520.1| BCL2-associated athanogene 3 variant [Homo sa...	157	4e-37
gb|EAW49383.1| BCL2-associated athanogene 3, isoform CRA_b [H...	157	4e-37
ref|XP_002807518.1| PREDICTED: LOW QUALITY PROTEIN: BAG famil...	157	5e-37
gb|AAF26839.1|AF127139_1 Bcl-2-binding protein BIS [Homo sapi...	157	5e-37
gb|ABG23394.1| Bcl-2-interacting death suppressor [Rattus nor...	157	5e-37
ref|NP_001011936.1| BAG family molecular chaperone regulator ...	156	6e-37
dbj|BAE42639.1| unnamed protein product [Mus musculus]	155	1e-36
sp|Q9JLV1.1|BAG3_MOUSE RecName: Full=BAG family molecular cha...	155	2e-36
gb|AAI45766.1| BCL2-associated athanogene 3 [Mus musculus]	55	2e-36
ref|XP_002718735.1| PREDICTED: BCL2-associated athanogene 3 [...	154	2e-36
ref|XP_001929035.1| PREDICTED: BAG family molecular chaperone...	154	2e-36
ref|NP_038891.4| BAG family molecular chaperone regulator 3 [...	154	3e-36
dbj|BAA95066.1| unnamed protein product [Mus musculus]	154	3e-36
ref|XP_001496329.1| PREDICTED: similar to BAG family molecula...	153	6e-36
pdb|1UK5|A Chain A, Solution Structure Of The Murine Bag Doma...	153	7e-36

FIG. 5 (Cont.)

ref|XP_001376803.1| PREDICTED: similar to BCL2-associated ath...	150	5e-35
ref|XP_544046.2| PREDICTED: similar to BAG-family molecular c...	150	6e-35
ref|XP_002915457.1| PREDICTED: BAG family molecular chaperone...	147	5e-34
gb|EFB13668.1| hypothetical protein PANDA_003439 [Ailuropoda ...	146	7e-34
ref|XP_002188317.1| PREDICTED: BCL2-associated athanogene 3 [...	141	3e-32
ref|XP_001514849.1| PREDICTED: similar to BCL2-associated ath...	137	4e-31
ref|XP_001233435.1| PREDICTED: BCL2-associated athanogene 3 [...	135	1e-30
ref|NP_001079487.1| BCL2-associated athanogene 3 [Xenopus lae...	116	8e-25
ref|NP_001120299.1| BCL2-associated athanogene 3 [Xenopus (Si...	112	1e-23
ref|NP_001003533.1| BAG family molecular chaperone regulator ...	107	3e-22
ref|XP_002819057.1| PREDICTED: BAG family molecular chaperone...	92.4	2e-17
ref|XP_002805355.1| PREDICTED: BAG family molecular chaperone...	92.4	2e-17
ref|XP_001092290.2| PREDICTED: BAG family molecular chaperone...	92.4	2e-17
ref|XP_002757004.1| PREDICTED: BAG family molecular chaperone...	92.4	2e-17
dbj|BAG64979.1| unnamed protein product [Homo sapiens]	92.4	2e-17
ref|XP_001493229.1| PREDICTED: similar to BCL2-associated ath...	92.4	2e-17
ref|NP_001104002.1| BAG family molecular chaperone regulator ...	92.4	2e-17
ref|XP_001170419.1| PREDICTED: similar to silencer of death d...	92.4	2e-17
ref|NP_001020301.1| BAG family molecular chaperone regulator ...	92.4	2e-17
ref|XP_849918.1| PREDICTED: similar to BAG-family molecular c...	92.4	2e-17
ref|NP_004865.1| BAG family molecular chaperone regulator 4 [...	92.4	2e-17
gb|AAH37239.1| Bag4 protein [Mus musculus]	92.4	2e-17
ref|NP_080397.1| BAG family molecular chaperone regulator 4 [...	92.4	2e-17
ref|XP_002720829.1| PREDICTED: BCL2-associated athanogene 4-l...	92.0	2e-17
ref|XP_003133402.1| PREDICTED: BAG family molecular chaperone...	91.3	3e-17

FIG. 5 (Cont.)

gb|EDM09068.1| BCL2-associated athanogene 4 [Rattus norvegicus]	90.5	6e-17
dbj|BAG36731.1| unnamed protein product [Homo sapiens]	90.5	6e-17
ref|XP_002917278.1| PREDICTED: BAG family molecular chaperone...	88.2	3e-16
gb|EAW63331.1| BCL2-associated athanogene 4, isoform CRA_b [H...	87.8	4e-16
gb|AAH09102.1| Bag4 protein [Mus musculus]	86.7	9e-16
ref|XP_002932772.1| PREDICTED: BAG family molecular chaperone...	86.7	1e-15
pdb|1M7K|A Chain A, Solution Structure Of The Sodd Bag Domain	86.3	1e-15
pdb|1M62|A Chain A, Solution Structure Of The Bag Domain From...	85.9	1e-15
emb|CAG04720.1| unnamed protein product [Tetraodon nigroviridis]	84.3	4e-15
ref|XP_424388.2| PREDICTED: similar to silencer of death doma...	83.6	8e-15
ref|XP_001506271.1| PREDICTED: similar to BCL2-associated ath...	83.2	1e-14
gb|ABF57351.1| BCL2-associated athanogene 4 [Bos taurus] >gb|...	82.4	2e-14
ref|XP_002194339.1| PREDICTED: similar to BCL2-associated ath...	72.8	1e-11
ref|XP_002741303.1| PREDICTED: BCL2-associated athanogene 3-l...	67.0	8e-10
ref|XP_001337521.1| PREDICTED: si:dkey-220o5.3 [Danio rerio] ...	66.6	9e-10
ref|NP_001026382.1| BAG family molecular chaperone regulator ...	65.5	2e-09
pdb|3A8Y|C Chain C, Crystal Structure Of The Complex Between ...	65.1	2e-09
ref|NP_001015049.1| BAG family molecular chaperone regulator ...	65.1	3e-09
dbj|BAA74896.1| KIAA0873 protein [Homo sapiens]	64.7	3e-09
gb|AAH50551.1| BCL2-associated athanogene 5 [Homo sapiens]	64.7	3e-09
dbj|BAF84001.1| unnamed protein product [Homo sapiens]	64.7	4e-09
ref|NP_004864.1| BAG family molecular chaperone regulator 5 i...	64.7	4e-09
emb|CAX12188.1| novel protein similar to vertebrate BCL2-asso...	64.3	5e-09
ref|NP_001029355.1| BAG family molecular chaperone regulator ...	63.9	6e-09
ref|XP_001511029.1| PREDICTED: hypothetical protein [Ornithor...	63.9	6e-09

FIG. 5 (Cont.)

ref|XP_002754360.1| PREDICTED: BAG family molecular chaperone...	63.9	7e-09
ref|XP_002805271.1| PREDICTED: BAG family molecular chaperone...	63.5	9e-09
ref|XP_002825189.1| PREDICTED: BAG family molecular chaperone...	63.2	1e-08
ref|XP_002805269.1| PREDICTED: BAG family molecular chaperone...	63.2	1e-08
ref|XP_001917330.1| PREDICTED: BCL2-associated athanogene 5 [...	62.8	1e-08
ref|XP_002200616.1| PREDICTED: BCL2-associated athanogene 5 [...	62.8	1e-08
ref|NP_001035634.1| BAG family molecular chaperone regulator ...	62.8	1e-08
gb|EFB29769.1| hypothetical protein PANDA_007430 [Ailuropoda ...	62.4	2e-08
ref|XP_537562.2| PREDICTED: similar to BCL2-associated athano...	62.4	2e-08
gb|ACN58723.1| BAG family molecular chaperone regulator 5 [Sa...	62.4	2e-08
ref|XP_002918888.1| PREDICTED: BAG family molecular chaperone...	62.4	2e-08
ref|XP_001373299.1| PREDICTED: similar to BCL2-associated ath...	60.8	5e-08
ref|XP_001199380.1| PREDICTED: hypothetical protein [Strongyl...	60.5	6e-08
dbj|BAB24105.1| unnamed protein product [Mus musculus]	60.5	6e-08
ref|NP_081680.1| BAG family molecular chaperone regulator 5 [...	60.5	7e-08
gb|AAH27827.1| Bag5 protein [Mus musculus]	60.1	8e-08
ref|XP_001958531.1| GF23468 [Drosophila ananassae] >gb|EDV413...	60.1	9e-08
ref|NP_001008526.1| BAG family molecular chaperone regulator ...	59.3	1e-07
ref|NP_001036843.1| BAG domain-containing protein Samui [Bomb...	58.9	2e-07
ref|XP_001641534.1| predicted protein [Nematostella vectensis...	58.9	2e-07
pdb|1UGO|A Chain A, Solution Structure Of The First Murine Ba...	58.2	4e-07
ref|XP_002061647.1| GK17074 [Drosophila willistoni] >gb|EDW72...	58.2	4e-07
ref|XP_002047962.1| GJ13715 [Drosophila virilis] >gb|EDW70304...	58.2	4e-07
gb|EFR27392.1| hypothetical protein AND_05926 [Anopheles darl...	57.8	4e-07
ref|XP_002128391.1| PREDICTED: similar to BCL2-associated ath...	57.8	4e-07
ref|XP_002135463.1| GA28307 [Drosophila pseudoobscura pseudoo...	57.8	5e-07

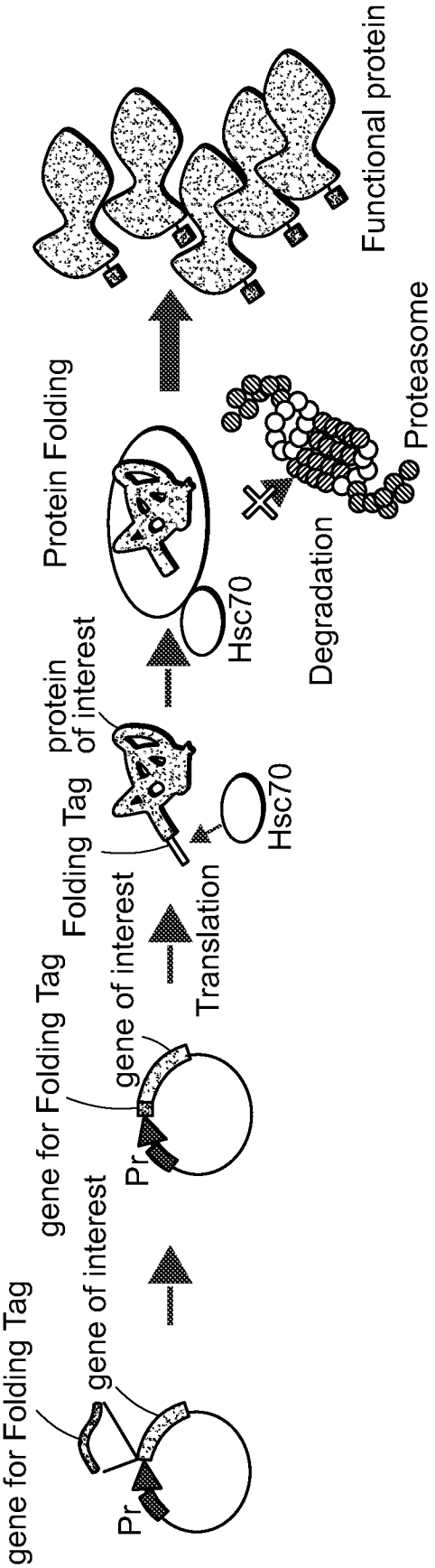
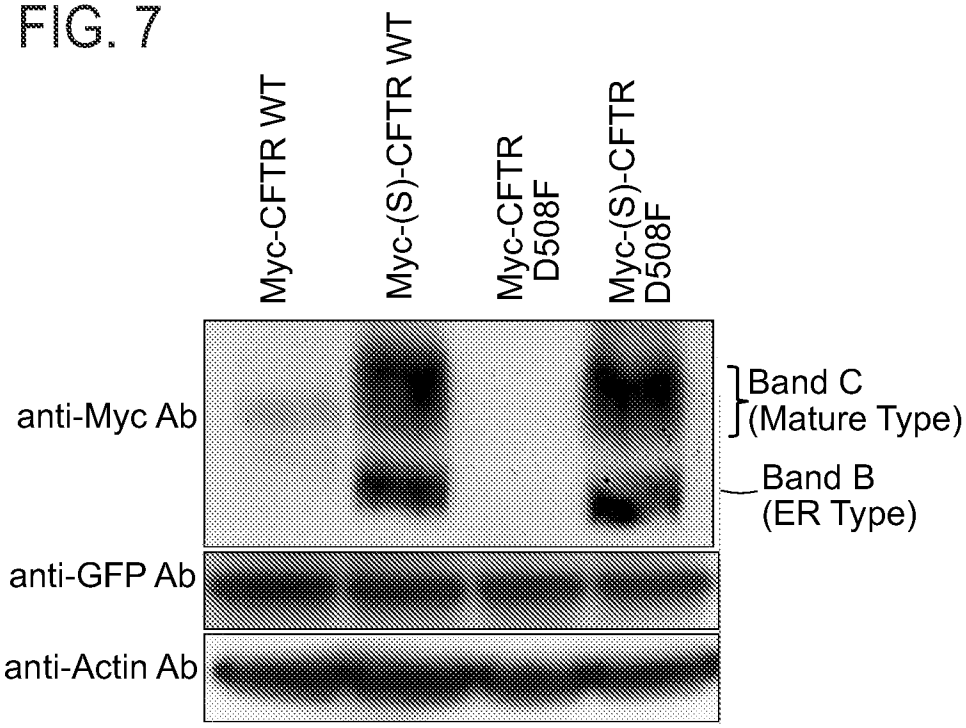

gene for Folding Tag
gene of interest
Pr
gene for Folding Tag
gene of interest
Pr
Translation
Folding Tag
protein of interest
Hsc70
Protein Folding
Hsc70
Degradation
Proteasome
Functional protein

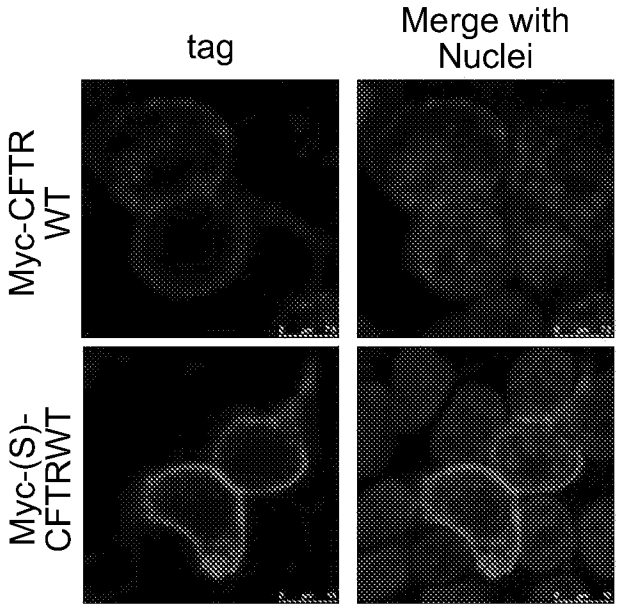
FIG. 6

FIG. 7



CFTR expression in human cells

FIG. 8



Intracellular localization of CFTR in human cells

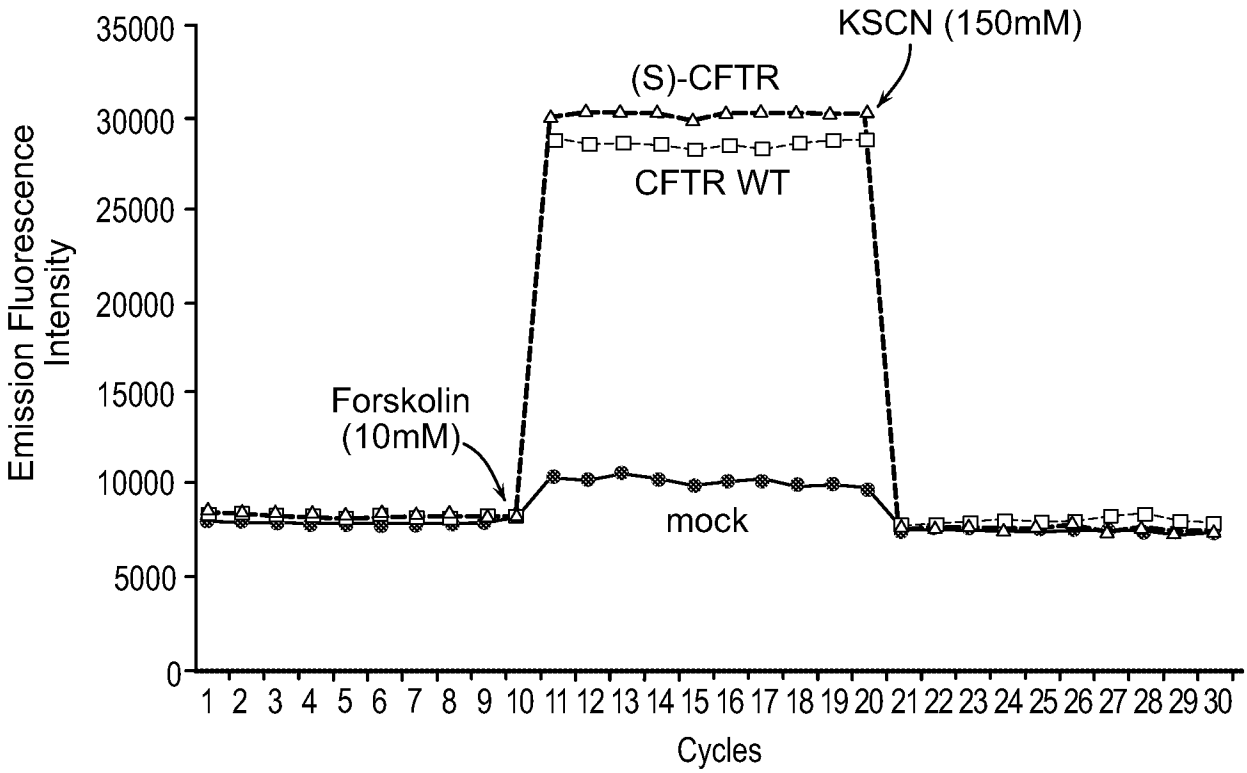
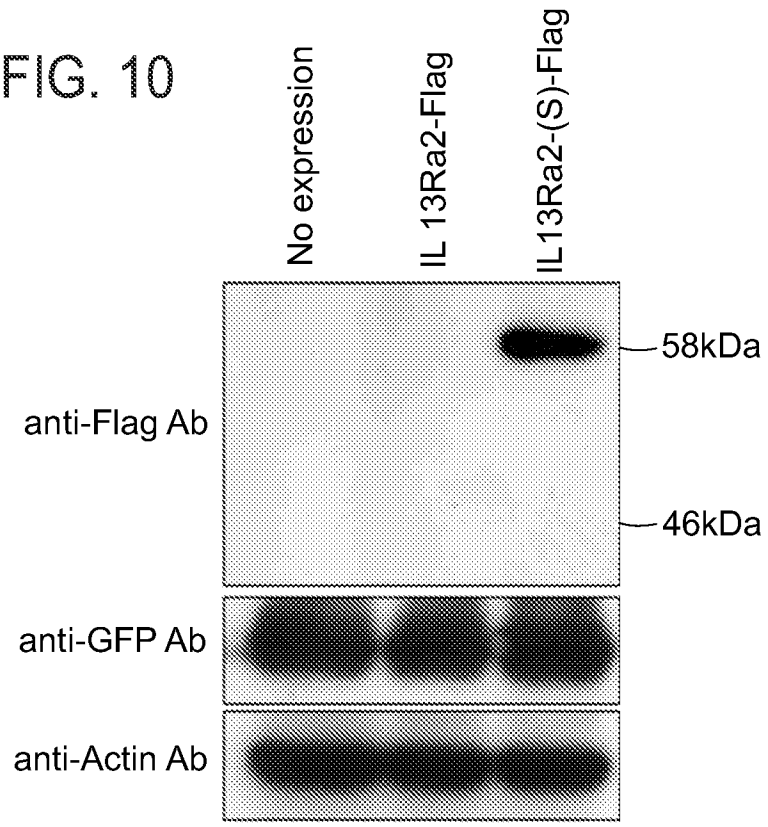


FIG. 9

CFTR function in human cells

FIG. 10



IL13Ra2 expression in human cells

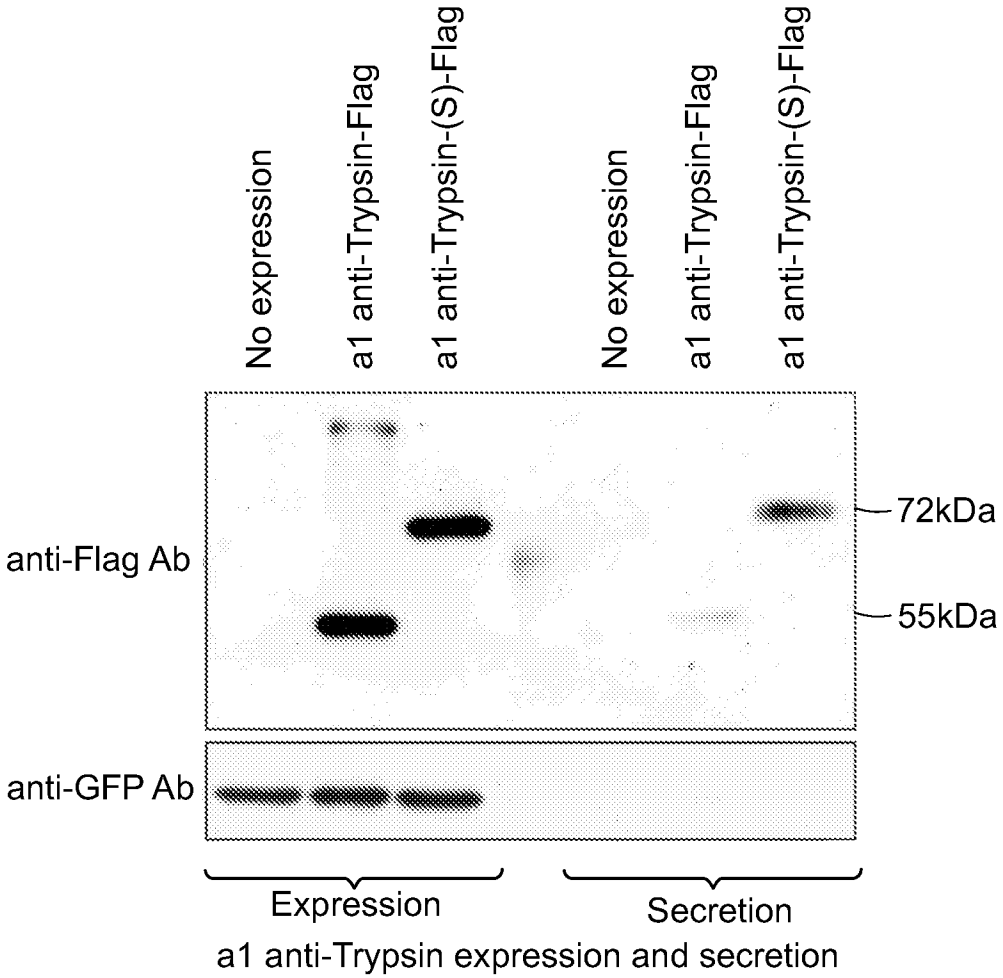

No expression
a1 anti-Trypsin-Flag
a1 anti-Trypsin-(S)-Flag
No expression
a1 anti-Trypsin-Flag
a1 anti-Trypsin-(S)-Flag
anti-Flag Ab
72kDa
55kDa
anti-GFP Ab
Expression
Secretion
a1 anti-Trypsin expression and secretion

FIG. 11

FIG. 12

```
hBAG3    --------------GVLKVEAILEKVQ-GLEQAVDN-FEGKKTDKKYLMIEEYLTKELLA 44
hBAG4    --------------SIKKIIHVLEKVQ-YLEQEVEE-FVGKKTDKAYWLLEEMLTKELLE 44
hBAG5-5  --------------SHKAVWNVLGNLS-EIQGEVLS-FDGNRTDKNYIRLEELLTKQLLA 44
hBAG5-1  --------------SISRLQEIQKEVK-SVEQQVIG-FSGLSDDKNYKKLERILTKQLFE 44
hBAG5-4  --------------SILKIEKVLKRMR-EIKNEL---LQAQNPSELYLSSKTELQGLIGQ 42
hBAG5-3  --------------SVAKINFVMCEVN-KARGVLIALLMGVNNNETCRHLSCVLSGLIAD 45
hBAG5-2  --------------HRIEIQNIFEEAQSLVREKIVPFYNGGNCVTD--EFEEGIQDIILR 44
hBAG1    QEEVELKKLKHLEKSVEKIADQLEELNKELTGIQQGFLPKDLQAEALCKLDRRVKATIEQ 60
hBAG2    ------------QESLKHATRIIDEVVNKFLDDLGN--AKSHLMSLYSACSSEVPHGPVD 46
hBAG6    --------------QGEGPQLLLSEAVSRAAKAA-----GARPLTSPESLSRDLEAPEVQ 41

hBAG3    ----LDSVD----PEGRADVRQARRDGVRKVQTILEKLEQKA---- 78
hBAG4    ----LDSVE----TGGQDSVRQARKEAVCKIQAILEKLEKKG---- 78
hBAG5-5  ----LDAVD----PQGEEKCKAARKQAVRLAQNILSYLDLKS---- 78
hBAG5-1  ----IDSVD----TEGKGDIQQARKRAAQETERLLKELEQNA---- 78
hBAG5-4  ----LDEVS----LEKNPCIREARRRAVIEVQTLITYIDLKE---- 76
hBAG5-3  ----LDALD----VCGRTEIRNYRREVVEDINKLLKYLDLEEE--- 80
hBAG5-2  ----LTHVK----TGGKISLRKARYHTLTKICAVQEIIEDCM---- 78
hBAG1    FMKILEEIDTLILPENFKDSRLKRKGLVKKVQAFLAECDTVEQN-- 104
hBAG2    -----QKFQ----SIVIGCALEDQKKIKRRLETLLRDIENSDKAIK 83
hBAG6    -----ESYR----QQLRSDIQKRLQEDPNYSPQRFPNAQR------ 72
```

FIG. 13A

Sequence

BAG3

Folding Tag (R)

```
1   AAGAGTGTGGCTACAGAAGAGAGGGCAGCCCCCAGCACTGCCCCCTGCAGAAGCTACACCT 60
1   K  S  V  A  T  E  E  E  R  A  A  P  P  S  T  A  P  A  E  A  T  P  20
61  CCAAAACCAGGAGAAGCCGAGGCTCCCCCAAAACATCCAGGAGTGCTGAAAGTGGAAGCC 120
21  P  K  P  G  E  A  E  A  P  P  P  K  H  P  G  V  L  K  V  E  A  40
121 ATCCTGGAGAAGGTACAGGGGCTGGAGCAGGCTGTAGACAACTTTGAAGGCAAGAAGACT 180
41  I  L  E  K  V  Q  G  L  E  Q  A  V  D  N  F  E  G  K  K  T  60
181 GACAAAAAGTACCTGATGATCGAAGAGTATTTGACCAAAGAGCTGCTGGCCCTGGATTCA 240
61  D  K  K  Y  L  M  I  E  E  Y  L  T  K  E  L  L  A  L  D  S  80
241 GTGGACCCCGAGGGACGAGCCGATGTGCGTCAGGCCAGGAGACGGTGTCAGGAAGGTT 300
81  V  D  P  E  G  R  A  D  V  R  Q  A  R  R  D  G  V  R  K  V  100
301 CAGACCATCTTGGAAAACTTGAACAGAAAGCCATTGAT 339
101 Q  T  I  L  E  K  L  E  Q  K  A  I  D  113
```

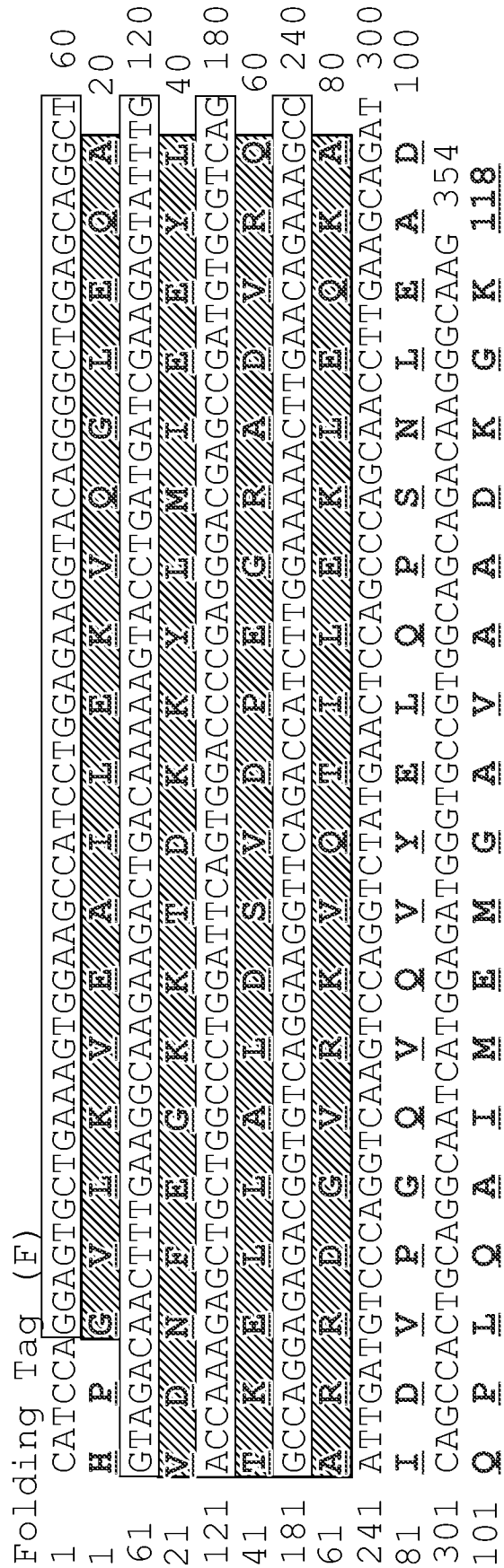

Folding Tag (F)
1 CATCCAGGAGTGCTGAAAGTGGAAGCCATCCTGGAGAAGGTACAGGGGCTGGAGCAGGCT 60
1 H P G V L K V E A I L L E K V Q G L E Q A 20
61 GTAGACAACTTTGAAGGCAAGAAGACTGACAAAAAGTACCTGATGATCGAAGAGTATTTG 120
21 V D N F E G K K T D K K Y L M I E E Y L 40
121 ACCAAAGAGCTGCTGGCCCTGGATTCAGTGGACCCCGAGGGACGAGCCGATGTGCGTCAG 180
41 T K E L L A L D S V D P E G R A D V R Q 60
181 GCCAGGAGACGGTGTCAGGAAGGTTCAGACCATCTTGGAAAAACTTGAACAGAAAGCC 240
61 A R R D G V R K V Q T I L E K L E Q K A 80
241 ATTGATGTCCCAGGTCAAGTCCAGGTCTATGAACTCCAGCCCAACCTTGAAGCAGAT 300
81 I D V P G Q V Q V Y E L Q P S N L E A D 100
301 CAGCCACTGCAGGCAATCATGGAGATGGGTGCCGTGGCAGACAAGGGCAAG 354
101 Q P L Q A I M E M G A V A A D K G K 118

FIG. 13A (Cont.)

FIG. 13B

```
DNA
Folding Tag (R)
1 AAGAGTGTGGCTACAGAAGAGAGGGCAGCCCCCAGCACTGCCCCTGCAGAAGCTACACCT 60
61 CCAAAACCAGGAGAAGCCGAGGCTCCCCCAAAACATCCAGGAGTGCTGAAAGTGGAAGCC 120
121 ATCCTGGAGAAGGTACAGGGGCTGGAGCAGGCTGTAGACAACTTTGAAGGCAAGAAGACT 180
181 GACAAAAAGTACCTGATCGAAGAGTATTTGACCAAAGAGCTGCTGGCCCTGGATTCA 240
241 GTGGACCCCCGAGGGACGAGCCGATGTGCGTCAGGCCAGGAGACGGTGTCAGGAAGGTT 300
301 CAGACCATCTTGGAAAAACTTGAACAGAAAGCCATTGAT 339

Folding Tag (F)
1 CATCCAGGAGTGCTGAAAGTGGAAGCCATCCTGGAGAAGGGTACAGGGGCTGGAGCAGGCT 60
61 GTAGACAACTTTGAAGGCAAGAAGACTGACAAAAAAGTACCTGATCGAAGAGTATTTG 120
121 ACCAAAGAGCTGCTGGCCCTGGATTCAGTGGACCCCGAGGGACGAGCCGATGTGCGTCAG 180
181 GCCAGGAGACGGTGTCAGGAAGGTTCAGACCATCTTGGAAAAACTTGAACAGAAAGCC 240
241 ATTGATGTCCCAGGTCAAGTCCAGGTCTATGAACTCCAGCCCAGCAACCTTGAAGCAGAT 300
301 CAGCCACTGCAGGCAATCATGGGAGATGGGTGCCGTGGCAGCAGACAAGGGCAAG 354
```

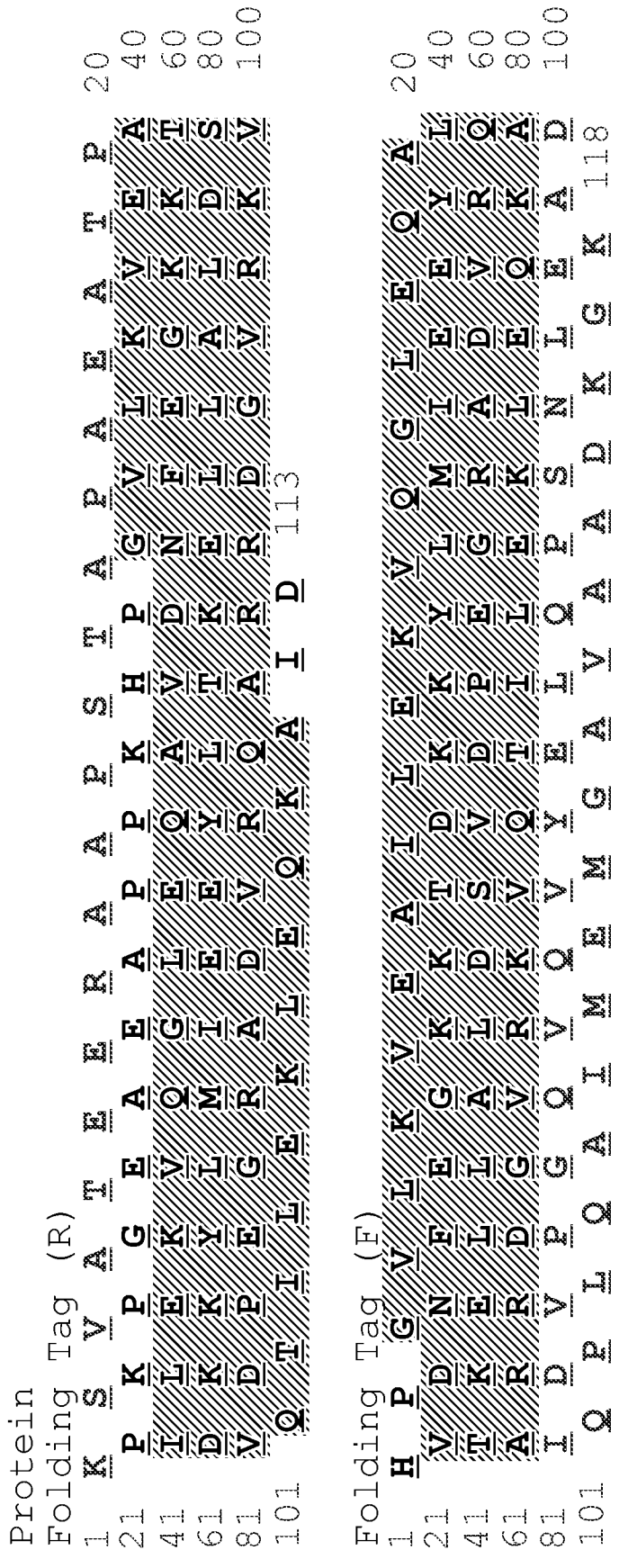

Protein
Folding Tag (R)
1 K S V A T E E E R A A P P K S T A P A E A T P
21 P K P G E A E E R A P P K H P G V L K V E A
41 T L E K V Q G L E Q A V D N E L L G K K T
61 D K K Y L M T E E Y L T K E L A L D S
81 V D P E G R A D V R Q A R R D G V R K V
101 Q T I L E K L E Q K A I D 113
Folding Tag (F)
1 H P G V L K V E A I L E K V Q G L E Q A
21 V D N E E G K K T D K K Y L M I E E V L
41 T K E L L A L D S V D P E G R A D V R Q
61 A R R D G V R K V Q T L L K L E Q K A
81 I D V P G Q Q V Q V Y E L Q P S N L E A D
101 Q P L Q A I M E M G A V A A D K G K 118
20
40
60
80
100

FIG. 13B (Cont.)

FIG. 14A

Sequence file

CFTR (NM_000492)

ATGCAGAGGTCGCCTCTGGAAAAGGCCAGCGTTGTCTCCAAACTTTTTTTCAGCTGGACCAGACCAATTTGAGGAAA
GGATACAGACAGCGCCTGGAATTGTCAGACATATACCAAATCCCTTCTGTTGATTCTGCTGACAATCTATCTGAAAAA
TTGGAAAGAGAATGGGATAGAGAGCTGGCTTCAAAGAAAAATCCTAAACTCATTAATGCCCTTCGGCGATGTTTTTC
TGGAGATTTATGTTCTATGGAATCTTTTTATATTTAGGGGGAAGTCACCAAAGCAGTACAGCCTCTCTTACTGGGAAGAA
TCATAGCTTCCTATGACCCGGATAACAAGGAGGAACGCTCTATCGCGATTTATCTAGGCATAGGCTTATGCCTTCTCTT
TATTGTGAGGACACTGCTCCTACACCCAGCCATTTTTGGCCTTCATCACATTGGAATGCAGATGAGAATAGCTATGTTT
AGTTTGATTTATAAGAAGACTTTAAAGCTGTCAAGCCGTGTTCTAGATAAAATAAGTATTGGACAACTTGTTAGTCTCC
TTTCCAACAACCTGAACAAATTTGATGAAGGACTTGCATTGGCACATTTCGTGTGGATCGCTCCTTTGCAAGTGGCACT
CCTCATGGGGCTAATCTGGGAGTTGTTACAGGCGTCTGCCTTCTGTGGACTTGGTTTCCTGATAGTCCTTGCCCTTTTTC
AGGCTGGGCTAGGGAGAATGATGATGAAGTACAGAGAGATCAGTGGGAAGAGTCAGTGAAAGACTTGTGATTACC
TCAGAAATGATTGAAAATATCCAATCTGTTAAGGCATACTGCTGGGAAGAAGCAATGGAAAAAATGATTGAAAACTT
AAGACAAACAGAACTGAAACTGACTCGGAAGGCAGCCTATGTGAGATACTTCAATAGCTCAGCCTTCTTCTTCTCAGG
GTTCTTTGTGGTGTTTTTATCTGTGCTTCCCTATGCACTAATCAAAGGAATCATCCTCCGGAAAATATTCACCACCATCT
CATTCTGCATTGTTCTGCGCATGGCGGTCACTCGGCAATTTCCCTGGGCTGTACAAACATGGTATGACTCTCTTGGAGC
AATAAACAAAATACAGGATTTCTTACAAAAGCAAGAAATATAAGACATTGGAAATATAACTTAACGACTACAGAAGTAG
TGATGGAGAATGTAACAGCCTTCTGGGAGGAGGGATTTGGGGAATTATTTGAGAAAGCAAAACAAAACAATAACAAT
AGAAAAACTTCTAATGGTGATGACAGCCTCTTCTTCACTTCTTGGTACTCCTGTCCTGAAAGATATTA
ATTTCAAGATAGAAAGAGGACAGTTGTTGGCGGTTGCTGGATCCACTGGAGCAGGCAAGACTTCACTTCTAATGGTGA
TTATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTTCATTCTGTTCTCAGTTTTCCTGGA
TTATGCCTGGCACCATTAAAGAAAATATCATCTTTGGTGTTTCCTATGATGAATATAGATACAGAAGCGTCATCAAAG
CATGCCAACTAGAAGAGGACATCTCCAAGTTTGCAGAGAAAGACAATATAGTTCTTGGAGAAGGTGGAATCACACTG
AGTGGAGGTCAACGAGCAAGAATTTCTTTAGCAAGAGCAGTATACAAAGATGCTGATTTGTATTTATTAGACTCTCCT
TTTGGATACCTAGATGTTTTAACAGAAAAAGAAATATTTGAAAGCTGTGTCTGTAAACTGATGGCTAACAAAAACTAGG

FIG. 14A (Cont.)

ATTTTGGTCACTTCTAAAATGGAACATTTAAAGAAAGCTGACAAAATATTAATTTGCATGAAGGTAGCAGCTATTTTT
ATGGGACATTTTCAGAACTCCAAAATCTACAGCCAGACTTTAGCTCAAAACTCATGGGATGTGATTCTTTCGACCAATT
TAGTGCAGAAAGAAGAAATTCAATCCTAACTGAGACCTTACACCGTTTCTCATTAGAAGGAGATGCTCCTGTCTCCTG
GACAGAAACAAAAAAAACAATCTTTTAAACAGACTGGAGAGTTTGGGGAAAAAAGGAAGAATTCTATTCTCAATCCAA
TCAACTCTATACGAAAATTTTCCATTGTGCAAAAGACTCCCTTACAAATGAATGGCATCGAAGAGGATTCTGATGAGC
CTTTAGAGAGAAGGCTGTCCTTAGTACCAGATTCTGAGCAGGGAGAGGCGATACTGCCTCGCATCAGCGTGATCAGCA
CTGGCCCCACGCTTCAGGCACGAAGGAGGCAGTCTGTCCTGAACCTGATGACACACTCAGTTAACCAAGGTCAGAAC
ATTCACCGAAAGACAACAGCATCCACACGAAAAAGTGTCACTGGCCCCTCAGGCAAACTTGACTGAACTGGATATATAT
TCAAGAAGGTTATCTCAAGAAACTGGCTTGGAAATAAGTGAAGAAATTAACGAAGAAGACTTAAAGGAGTGCTTTTTT
GATGATATGGAGAGCATACCAGCAGTGACTACATGGAACACATACCTTCGATATATTACTGTCCACAAGAGCTTAATT
TTTGTGCTAATTTGGTGCTTAGTAATTTTTCTGGCAGAGGTGGCTGCTTCTTTGGTTGTGCTGTGGCTCCTTGGAAACAC
TCCTCTTCAAGACAAAGGGAATAGTACTCATAGTAGAAATAACAGCTATGCAGTGATTATCACCAGCACCAGTTCGTA
TTATGTGTTTTACATTTACGTGGGAGTAGCCGACACTTTGCTTGCTATGGGATTCTTCAGAGGTCTACCACTGGTGCAT
ACTCTAATCACAGTGTCGAAAATTTTACACCACAAAATGTTACATTCTGTTCTTCAAGCACCTATGTCAACCCTCAACA
CGTTGAAAGCAGGTGGGATTCTTAATAGATTCTCCAAAGATATAGCAATTTTGGATGACCTTCTGCCTCTTACCATATT
TGACTTCATCCAGTTGTTATTAATTGTGATTGGAGCTATAGCAGTTGTCGCAGTTTTACAACCCTACATCTTTGTTGCAA
CAGTGCCAGTGATAGTGGCTTTTATTATGTTGAGAGCATATTTCCTCCAAACCTCACAGCAACTCAAACAACTGGAATC
TGAAGGCAGGAGTCCAATTTTCACTCATCTTGTTACAAGCTTAAAAGGACTATGGACACTTCGTGCCTTCGGACGGCA
GCCTTACTTTGAAACTCTCTGTTCCACAAAGCTCTGAATTTACATACTGCCAACTGGTTCTTGTACCTGTCAACACTGCGC
TGGTTCCAAATGAGAATAGAAATGATTTTTGTCATCTTCTTCATTGCTGTTACCTTCATTTCCATTTTAACAACAGGAGA
AGGAGAAGGAAGAGTTGGTATTATCCTGACTTTAGCCATGAATATCATGAGTACATTGCAGTGGGCTGTAAACTCCAG
CATAGATGTGGATAGCTTGATGCGATCTGTGAGCCGAGTCTTTAAGTTCATTGACATGCCAACAGAAGGTAAACCTAC
CAAGTCAACCAAACCATACAAGAATGGCCAACTCTCGAAAGTTATTATGATTATTGAGAATTCACACGTGAAGAAAGATG
ACATCTGGCCCTCAGGGGGCCAAATGACTGTCAAAGATCTCACAGCAAAATACACAGAAGGTGGAAATGCCATATTA
GAGAACATTTCCTTCTCAATAAGTCCTGGCCAGAGGGTGGGCCTCTTGGGAAGAACTGGATCAGGGAAGAGTACTTTG

FIG. 14A (Cont.)

TTATCAGCTTTTTTTGAGACTACTGAACACTGAAGGAGAAATCCAGATCGATGGTGTGTCTTGGGATTCAATAACTTTGC
AACAGTGGAGGAAAGCCTTTGGAGTGATACCACACAGAAAGTATTTATTTTTTCTGGAACATTTAGAAAAAACTTGGATC
CCTATGAACAGTGGAGTGATCAAGAAATATGGAAAGTTTGCAGATGAGGTTGGGCTCAGATCTGTGATAGAACAGTTTC
CTGGGAAGCTTGACTTTGTCCTTGTGGATGGGGGCTGTGTCCTAAGCCATGGCCACACAAGCAGTTGATGTGCTTGGCTA
GATCTGTTCTCAGTAAGGCGAAGATCTTGCTGCTTGATGAACCCAGTGCTCATTTGGATCCAGTAACATACCAAATAAT
TAGAAGAACTCTAAAACAAGCATTTGCTGATTGCACAGTAATTCTCTGTGAACACAGGATAGAAGCAATGCTGGAAATG
CCAACAATTTTTGGTCATAGAAGAGAACAAAGTGCGGCAGTACGATTCCATCCAGAAACTGCTGAACGAGAGGAGCC
TCTTCCGGCAAGCCATCAGCCCCTCCGACAGGGTGAAGCTCTTTCCCCACCGGAACTCAAGCAAGTGCAAGTCTAAGC
CCCAGATTGCTGCTCTGAAAGAGGAGACAGAAGAGGTGCAAGATACAAGGCTTTAG

FIG. 14A (Cont.)

NP_000483
MQRSPLEKASVVSKLFFSWTRPILRKGYRQRLELSDIYQIPSVDSADNLSEKLEREWDRELASKKNPKLI
NALRRCFFWRFMFYGIFLYLGEVTKAVQPLLLGRIIASYDPDNKEERSIAIYLGIGLCLLFIVRTLLLHP
AIFGLHHIGMQMRIAMFSLIYKKTLKLSSRVLDKISIGQLVSLLSNNLNKFDEGLALAHFVWIAPLQVAL
LMGLIWELLQASAFCGLGFLIVLALFQAGLGRMMMKYRDQRAGKISERLVITSEMIENIQSVKAYCWEEA
MEKMIENLRQTELKLTRKAAYVRYFNSSAFFFSGFFVVFLSVLPYALIKGIILRKIFTTISFCIVLRMAV
TRQFPWAVQTWYDSLGAINKIQDFLQKQEYKTLEYNLTTTEVVMENVTAFWEEGFGELFEKAKQNNNNRK
TSNGDDSLFFSNFSLLGTPVLKDINFKIERGQLLAVAGSTGAGKTSLLMVIMGELEPSEGKIKHSGRISF
CSQFSWIMPGTIKENIIFGVSYDEYRYRSVIKACQLEEDISKFAEKDNIVLGEGGITLSGGQRARISLAR
AVYKDADLYLLDSPFGYLDVLTEKEIFESCVCKLMANKTRILVTSKMEHLKKADKILILHEGSSYFYGTF
SELQNLQPDFSSKLMGCDSFDQFSAERRNSILTETLHRFSLEGDAPVSWTETKKQSFKQTGEFGEKRKNS
ILNPINSIRKFSIVQKTPLQMNGIEEDSDEPLERRLSLVPDSEQGEAILPRISVISTGPTLQARRRQSVL
NLMTHSVNQGQNIHRKTTASTRKVSLAPQANLTELDIYSRRLSQETGLEISEEINEEDLKECFFDDMESI
PAVTTWNTYLRYITVHKSLIFVLIWCLVIFLAEVAASLVVLWLLGNTPLQDKGNSTHSRNNSYAVIITST
SSYYVFYIYVGVADTLLAMGFFRGLPLVHTLITVSKILHHKMLHSVLQAPMSTLNTLKAGGILNRFSKDI
AILDDLLPLTIFDFIQLLLIVIGAIAVVAVLQPYIFVATVPVIVAFIMLRAYFLQTSQQLKQLESEGRSP
IFTHLVTSLKGLWTLRAFGRQPYFETLFHKALNLHTANWFLYLSTLRWFQMRIEMIFVIFFIAVTFISIL
TTGEGEGRVGIILTLAMNIMSTLQWAVNSSIDVDSLMRSVSRVFKFIDMPTEGKPTKSTKPYKNGQLSKV
MIIENSHVKKDDIWPSGGQMTVKDLTAKYTEGGNAILENISFSISPGQRVGLLGRTGSGKSTLLSAFLRL
LNTEGEIQIDGVSWDSITLQQWRKAFGVIPQKVFIFSGTFRKNLDPYEQWSDQEIWKVADEVGLRSVIEQ
FPGKLDFVLVDGGCVLSHGHKQLMCLARSVLSKAKILLLDEPSAHLDPVTYQIIRRTLKQAFADCTVILC
EHRIEAMLECQQFLVIEENKVRQYDSIQKLLNERSLFRQAISPSDRVKLFPHRNSSKCKSKPQIAALKEE
TEEEVQDTRL

FIG. 14B

IL13Rα2 (NM_000640)

ATGGCTTTCGTTTGCTTGGCTATCGGATGCTTATATACCTTTCTGATAAGCACAACATTTGGCTGTACTTCATCTTCAGA
CACCGAGATAAAAGTTAACCCTCCTCAGGATTTTGAGATAGTGGATCCCGGATACTTAGGTTATCTCTATTTGCAATGG
CAACCCCCACTGTCTCTGGATCATTTTAAGGAATGCACAGTGGAATATGAACTAAAATACCGAAACATTGGTAGTGAA
ACATGGAAGACCATCATTACTACATTACAAAGATGGGTTTGATCTTAACAAGGGCATTGAAGCGAAGATA
CACACGCTTTTACCATGGCAATGCACACAAATGGATCAGAAGTTCCTGGGCAGAAACTACTTATTGGATATCA
CCACAAGGAATTCCAGAAACTAAAGTTCAGGATATGCGTATATTACAATTGGCAATATTTACTCTGTTCTTGGA
AACCTGGCATAGGTGTACTTCTTGATACCAATTACAACTTGTTTTACTGGTATGAGGGCTTGGATCATGCATTACAGTG
TGTTGATTACATCAAGGCTGATGGACAAAATATAGGATGCAGATTTCCCTATTTGGAGGCATCAGACTATAAAGATTT
CTATATTTGTGTTAATGGATCATCAGAGAACAAGCCTATCAGATCCAGTTATTTCACTTTTCAGCTTCAAAATATAGTT
AAACCTTTGCCGCCAGTCTATCTTACTTTTACTCGGGAGAGTTCATGTGAAATTAAGCTGAAATGGAGCATACCTTTGG
GACCTATTCCAGCAAGGTGTTTTGATTATGAAATTGAGATCAGAGAAGATGATACTACCTTGGTGACTGCTACAGTTG
AAAATGAAACATACACCTTGAAAACAACAAATGAAACCCGACAATTATGCTTTGTAGTAAGAAGCAAAGTGAATATT
TATTGCTCAGATGACGGAATTTGGAGTGAGTGGATAAACAATGCTGGGAAGGTGAAGACCTATCGAAGAAAAC
TTTGCTACGTTTCTGGCTACCATTTGGTTTCATCTTAATATTAGTTATATTTGTAACCGGTCTGCTTTTGCGTAAGCCAA
ACACCTACCCAAAAATGATTCCAGAATTTTTCTGTGATACATGA

NP_000631

MAFVCLAIGCLYTFLISTTFGCTSSSDTEIKVNPPQDFEIVDPGYLGYLYLQWQPPLSLDHFKECTVEYE
LKYRNIGSETWKTIITKNLHYKDGFDLNKGIEAKIHTLLPWQCTNGSEVQSSWAETTYWISPQGIPETKV
QDMDCVYYNWQYLLCSWKPGIGVLLDTNYNLFYWYEGLDHALQCVDYIKADGQNIGCRFPYLEASDYKDF
YICVNGSSENKPIRSSYFTFQLQNIVKPLPPVYLTFTRESSCEIKLKWSIPLGPIPARCFDYEIEIREDD
TTLVTATVENETYTLKTTNETRQLCFVVRSKVNIYCSDDGIWSEWSDKQCWEGEDLSKKTLLRFWLPFGF
ILILVIFVTGLLLRKPNTYPKMIPEFFCDT

FIG. 14C

α1AT (NM_001127707)

ATGCCGTCTTCTGTCTCGTGGGGCATCCTCCTGCTGGCAGGCCTGTGCTGCCTGGTCCCTGTCTCCCTGGCTGAGGATC
CCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCACCCAACCTTCAACAAGATCACCCCC
AACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAGCTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAG
TGAGCATCGCTACAGCCTTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTGA
ATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTCCGTACCCTCAACCAGCCAG
ACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTCAGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGG
ATGTTAAAAAGTTGTACCACTCAGAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAAC
GATTACGTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACAGTTTTTGCTCTG
GTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTCAAGGACACCGAGGAAGAGGACTTCCACGT
GGACCAGGTGACCACCGTGAAGGTGCCTATGATGAAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTC
CAGCTGGGTGCTGCTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTACAGCA
CCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGACAGAAGGTCTGCCAGCTTACATTT
ACCCAAACTGTCCATTACTGGAACCTATGATCTGAAGAGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAA
TGGGGCTGACCTCTCCGGGGTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCAT
CGACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATCCCCCCCGAGGTCAAGTT
CAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAGTCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACC
CAAAAATAA

NP_001121179

MPSSVSWGILLLAGLCCLVPVSLAEDPQGDAAQKTDTSHHDQDHPTFNKITPNLAEFAFSLYRQLAHQSN
STNIFFSPVSIATAFAMLSLGTKADTHDEILEGLNFNLTEIPEAQIHEGFQELLRTLNQPDSQLQLTTGN
GLFLSEGLKLVDKFLEDVKKLYHSEAFTVNFGDTEEAKKQINDYVEKGTQGKIVDLVKELDRDTVFALVN
YIFFKGKWERPFEVKDTEEEDFHVDQVTTVKVPMMKRLGMFNIQHCKKLSSWVLLMKYLGNATAIFFLPD
EGKLQHLENELTHDIITKFLENEDRRSASLHLPKLSITGTYDLKSVLGQLGITKVFSNGADLSGVTEEAP
LKLSKAVHKAVLTIDEKGTEAAGAMFLEAIPMSIPPEVKFNKPFVFLMIEQNTKSPLFMGKVVNPTQK

FIG. 14D

pcDNA3 (EF437956.1)
GACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGAT
GCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTG
CGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAAT
CTGCTTAGGGTTAGGCGTTTTGCGCTTCGCGATGTACGGGCCAGATATACGCGTTTG
ACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGGTCATTAGTTCATAGCC
CATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC
AACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAG
GGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTA
CATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCC
CGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCT
ACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGT
GGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGA
GTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCA
TTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCT
GGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGG
GAGACCCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAGTGTGCTGGA
ATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCTATT
CTATAGTGTCACCTAAATGCTAGAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTG
CCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCC
CACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATT
CTATTCTGGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATA
GCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTG
GGGCTCTAGGGGGTATCCCCACGCGCCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTG
GTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGC

FIG. 14D (Cont.)

TTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGG
CATCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATT
AGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACG
TTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCC
TATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGGGGATTTCGGCCTATTGGTTAAA
AAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGGAATGTGTGTCAGTT
AGGGTGTGGAAAGTCCCCAGGCTCCCCAGGCAGGCAGAAGTATGCAAAGCATGCATCT
CAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATG
CAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCC
GCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTAT
TTATGCAGAGGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGC
TTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCCCGGGAGCTTGTATATCCATTTTCGG
ATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCAC
GCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGA
CAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTT
TTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGC
TATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAA
GCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTC
ACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATAC
GCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCA
CGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGG
GGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGGA
TCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCT
TTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCG
TTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGT
GCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACG

FIG. 14D (Cont.)

AGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCCAACCTG
CCATCACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGT
TTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCG
CCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACA
AATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCAT
CAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATG
GTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAG
CCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAAT
TGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAAT
GAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTC
GCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAA
AGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAG
CAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA
TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA
AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCT
CTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGC
GTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTC
CAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTA
ACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCAC
TGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGG
TGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCC
AGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGT
AGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAG
AAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAA
GGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAA

FIG. 14D (Cont.)

ATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAAT
GCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCT
GACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC
TGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAG
CCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGT
CTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAAC
GTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTC
AGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAAG
CGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCA
CTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTT
TCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGA
GTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAA
AGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGT
TGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACT
TTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGA
ATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAG
CATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATA
AACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC