

RECOMBINANT MILK PROTEINS AND FOOD COMPOSITIONS COMPRISING THE SAME

CROSS-REFERENCE TO RELATED APPLICATIONS

[1] This application claims priority to U.S. Application Nos. 63/240,621 filed on September 3, 2021; 63/221,642, filed on July 14, 2021; 63/189,547, filed on May 17, 2021; 63/174,244, filed on April 13, 2021; 63/152,694, filed on February 23, 2021; 63/138,089, filed on January 15, 2021; 63/129,720, filed on December 23, 2020; 63/121,468, filed on December 4, 2020; 63/116,528, filed on November 20, 2020; and 63/085,899, filed on September 30, 2020, each of which is incorporated by reference herein in its entirety for all purposes.

DESCRIPTION OF THE TEXT FILE SUBMITTED ELECTRONICALLY

[2] The contents of the text file submitted electronically herewith are incorporated herein by reference in their entirety: A computer readable format copy of the Sequence Listing, filename: ALRO_008_06WO_SeqList_ST25.txt, date created: September 29, 2021, file size: $\approx$ 1 MB.

FIELD OF THE DISCLOSURE

[3] The present disclosure generally relates to recombinant milk proteins, and methods of production, extraction, and purification thereof. The present disclosure also relates to food compositions (e.g., cheese compositions) comprising one or more recombinant milk proteins.

BACKGROUND

[4] Globally, more than 7.5 billion people consume milk and milk products, and it is estimated that cow milk accounts for 83% of global milk production. Demand for cow milk and dairy products is expected to continue rising due to increased reliance on these products in developing countries as well as growth in the human population, which is expected to exceed 9 billion people by 2050. Relying on animal agriculture to meet the growing demand for food is not a sustainable solution. According to the Food & Agriculture Organization of the United Nations, animal agriculture is responsible for 18% of all greenhouse gases, more than the entire transportation sector combined. Dairy cows alone account for 3% of this total.

[5] In addition to impacting the environment, animal agriculture poses a serious risk to human health. A startling 80% of antibiotics used in the United States go towards treating animals, resulting in the development of antibiotic resistant microorganisms, also known as superbugs. For years, food companies and farmers have administered antibiotics not only to sick animals, but also to healthy animals, to prevent illness. In September 2016, the United Nations announced the use of antibiotics in the food system as a crisis on par with Ebola and HIV.

[6] For at least these reasons, alternative dairy compositions produced without the use of mammalian milk have become increasingly popular. However, such compositions often fail to achieve the same organoleptic properties as their milk-derived counterparts. For example, some alternative dairy compositions on the market today are known to have an off-putting texture or taste, poor melt characteristics and/or lack of stretch. These compositions may also have reduced nutritional value compared to their milk-derived counterparts. For example, some “vegan” cheeses, typically made from oils and starch, may contain little to no protein.

[7] Accordingly, there is an urgent need to provide bovine milk and/or essential high-quality proteins from bovine milk in a more sustainable and humane manner, instead of solely relying on animal farming. Also, there is a need for selectively producing the specific milk proteins that confer nutritional and clinical benefits, and/or do not provoke allergic responses, and a need to prepare improved alternative dairy compositions which comparable nutritional value and similar organoleptic properties as their milk-derived counterparts.

BRIEF SUMMARY

[8] Provided herein are recombinant fusion proteins comprising (i) a first milk protein, and (ii) a second milk protein. At least one of the first milk protein and the second milk protein may be, for example, α -S1 casein, α -S2 casein, β -casein, κ -casein, para- κ -casein, β -lactoglobulin, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, serum albumin, or an immunoglobulin. In some embodiments, at least one of the first milk protein and the second milk protein is β -lactoglobulin. In some embodiments, at least one of the first milk protein and the second milk protein is α -S1 casein, α -S2 casein, β -casein, κ -casein, or para- κ -casein. In some embodiments, i) the first milk protein is α -S1 casein, α -S2 casein, β -casein, κ -casein, or para- κ -casein; and ii) the second milk protein is α -S1 casein, α -S2 casein, β -casein, κ -casein, or para- κ -casein. In some

embodiments, at least one of the first milk protein and the second milk protein is κ -casein and comprises the sequence of SEQ ID NO: 4, or a sequence at least 90% identical thereto. In some embodiments, at least one of the first milk protein and the second milk protein is para- κ -casein and comprises the sequence of SEQ ID NO: 2, or a sequence at least 90% identical thereto. In some embodiments, at least one of the first milk protein and the second milk protein is β -casein and comprises the sequence of SEQ ID NO: 6, or a sequence at least 90% identical thereto. In some embodiments, at least one of the first milk protein and the second milk protein is α -S1 casein and comprises the sequence SEQ ID NO: 8, or a sequence at least 90% identical thereto. In some embodiments, at least one of the first milk protein and the second milk protein is α -S2 casein and comprises the sequence SEQ ID NO: 84, or a sequence at least 90% identical thereto. In some embodiments, the first milk protein and the second milk protein are different proteins. In some embodiments, the first milk protein and the second milk protein are the same proteins. In some embodiments, the fusion protein is plant-expressed. In some embodiments, the fusion protein is expressed in soybean plant. In some embodiments, the fusion protein comprises a protease cleavage site. In some embodiments, the protease cleavage site is a chymosin cleavage site.

[9] Also provided herein are nucleic acids encoding one or more of the recombinant fusion proteins of the disclosure, and expression vectors comprising the same. In some embodiments, the nucleic acids are codon-optimized for expression in a plant, such as a soybean.

[10] Additionally, provided herein are host cells comprising a nucleic acid or an expression vector of the disclosure; i.e., a nucleic acid or expression vector encoding a fusion protein. The host cells may be, for example, plant cells, bacterial cells, fungal cells, or mammalian cells. In some embodiments, the host cells are soybean cells.

[11] Also provided herein are plants stably transformed with a nucleic acid or an expression vector of the disclosure. In some embodiments, the fusion protein is expressed in the plant in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant.

[12] Also provided herein are methods for making a fusion protein, the methods comprising: (a) transforming a host cell with a nucleic acid or an expression vector described herein; and (b) growing the transformed host cell under conditions wherein the fusion protein is expressed. In some embodiments, the method comprises co-expressing in the host cell a protein capable of forming a protein body, such as a prolamin selected from a gliadin, a hordein, a secalin, a zein, a

kafirin, or an avenin. In some embodiments, the method comprises expressing a kinase in the host cell. In some embodiments, expression of one or more proteases is knocked down or knocked out in the cell.

[13] Also provided herein are transgenic plants comprising a recombinant fusion protein, or a nucleic acid or expression vector comprising the same. In some embodiments, the transgenic plant is a soybean plant. In some embodiments, the fusion protein is expressed in the plant in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant.

[14] Also provided herein are methods for stably expressing a recombinant fusion protein in a plant, the methods comprising: (i) transforming a plant with a plant transformation vector comprising an expression cassette comprising a nucleic acid molecule encoding the fusion protein; and (ii) growing the transformed plant under conditions wherein the recombinant fusion protein is expressed. In some embodiments, the fusion protein is expressed in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant.

[15] Also provided herein are seed processing compositions comprising a fusion protein of the disclosure.

[16] Also provided herein are food compositions comprising a fusion protein of the disclosure. In some embodiments, the food composition is selected from the group consisting of cheese and processed cheese products, yogurt and fermented dairy products, directly acidified counterparts of fermented dairy products, cottage cheese dressing, frozen dairy products, frozen desserts, desserts, baked goods, toppings, icings, fillings, low-fat spreads, dairy-based dry mixes, soups, sauces, salad dressing, geriatric nutrition, creams and creamers, analog dairy products, follow-up formula, baby formula, infant formula, milk, dairy beverages, acid dairy drinks, smoothies, milk tea, butter, margarine, butter alternatives, growing up milks, low-lactose products and beverages, medical and clinical nutrition products, protein/nutrition bar applications, sports beverages, confections, meat products, analog meat products, meal replacement beverages, weight management food and beverages, cultured buttermilk, sour cream, yogurt, skyr, leben, lassi, kefir, powder containing a milk protein, and low-lactose products. In some embodiments, the food composition comprises a total amount of casein protein; wherein about 32% to 100% by weight of the total amount of casein protein in the food composition is beta-casein. In some embodiments, the food composition is a cheese composition. In some embodiments, the cheese composition has the ability to stretch to at

least 3 cm in length without breaking, as determined by heating a 100 gram mass of the composition at a temperature of 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[17] Also provided herein is method of making a food composition, comprising combining a fusion protein disclosed herein into a food composition.

[18] Also provided herein is an alternative dairy food composition comprising i) a recombinant fusion protein described herein; and ii) at least one lipid. In some embodiments, the recombinant fusion protein confers on the alternative dairy food composition one or more characteristics of a dairy food product selected from the group consisting of: taste, aroma, appearance, handling, mouthfeel, density, structure, texture, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess and emulsification. In some embodiments, the alternative dairy food composition does not comprise any other milk proteins. In some embodiments, the alternative dairy food composition comprises calcium at a concentration of about 0.01 to about 2% by weight. In some embodiments, the alternative dairy food composition comprises a total amount of casein protein; wherein about 32% to 100% by weight of the total amount of casein protein in the food composition is beta-casein. In some embodiments, the alternative dairy food composition has a pH of about 5.2 to about 5.9. In some embodiments, the alternative dairy food composition is selected from the group consisting of cheese and processed cheese products, yogurt and fermented dairy products, directly acidified counterparts of fermented dairy products, cottage cheese dressing, frozen dairy products, frozen desserts, desserts, baked goods, toppings, icings, fillings, low-fat spreads, dairy-based dry mixes, soups, sauces, salad dressing, geriatric nutrition, creams and creamers, analog dairy products, follow-up formula, baby formula, infant formula, milk, dairy beverages, acid dairy drinks, smoothies, milk tea, butter, margarine, butter alternatives, growing up milks, low-lactose products and beverages, medical and clinical nutrition products, protein/nutrition bar applications, sports beverages, confections, meat products, analog meat products, meal replacement beverages, weight management food and beverages, cultured buttermilk, sour cream, yogurt, skyr, leben, lassi, kefir, powder containing a milk protein, and low-lactose products. In some embodiments, the alternative dairy food composition is a cheese composition.

[19] Also provided herein are solid phase, protein-stabilized emulsions comprising a fusion protein described herein, wherein the emulsions have the ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the emulsion to a temperature of about 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[20] Also provided herein are colloidal suspensions comprising a fusion protein described herein, wherein the colloidal suspension has at least one, at least two, or at least three characteristics that are substantially similar to bovine milk selected from taste, appearance, mouthfeel, structure, texture, density, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess, and emulsification.

[21] These and other embodiments are described in detail below.

BRIEF DESCRIPTION OF THE DRAWINGS

[22] The accompanying figures, which are incorporated herein and form a part of the specification, illustrate some, but not the only or exclusive, example embodiments and/or features. It is intended that the embodiments and figures disclosed herein are to be considered illustrative rather than limiting.

[23] FIGs. 1A, 1B, 1C, 1D, 1E, 1F, 1G, 1H, 1I, 1J, 1K, 1L, 1M, 1N, 1O, and 1P show expression cassettes having different combinations of fusions between sequences encoding structured and intrinsically unstructured proteins (not to scale). Coding regions and regulatory sequences are indicated as blocks (not to scale). As used in the figures, “L” refers to linker; “Sig” refers to a signal sequence that directs foreign proteins to protein storage vacuoles, “5’ UTR” refers to the 5’ untranslated region, and “KDEL” refers to an endoplasmic reticulum retention signal.

[24] FIGs. 2A, 2B, 2C, 2D, 2E, 2F, 2G, 2H, 2I, 2J, 2K, 2L, 2M, 2N, 2O, and 2P show expression cassettes having different combinations of fusions between sequences encoding a first protein and a second protein (not to scale), wherein the first and/or second protein is a milk protein (not shown). Coding regions and regulatory sequences are indicated as blocks (not to scale). As used in the figures, “L” refers to linker; “Sig” refers to a signal sequence that directs foreign proteins to protein storage vacuoles, “5’ UTR” refers to the 5’ untranslated region, and “KDEL” refers to an endoplasmic reticulum retention signal.

[25] **FIG. 3** shows the modified pAR15-00 binary vector containing a selectable marker cassette conferring herbicide resistance. Coding regions and regulatory sequences are indicated as blocks (not to scale).

[26] **FIG. 4** shows an example expression cassette comprising a OKC1-T:OLG1 fusion (Optimized Kappa Casein version 1:beta-lactoglobulin version 1, SEQ ID NOs: 71-72), expression of which is driven by PvPhas promoter fused with arc5'UTR:sig10, followed by the ER retention signal (KDEL) and the 3'UTR of the arc5-1 gene, "arc-terminator". "arc5'UTR" refers to the 5' untranslated region of the arc5-1 gene. "Sig10" refers to the lectin 1 gene signal peptide. "RB" refers to ribosomal binding site. Coding regions and regulatory sequences are indicated as blocks (not to scale).

[27] **FIG. 5** shows an example expression cassette comprising a OBC-T2:FM:OLG1 fusion (Optimized Beta Casein Truncated version 2:Chymosin cleavage site:beta-lactoglobulin version 1, SEQ ID NOs: 73-74), expression of which driven by PvPhas promoter fused with arc5'UTR:sig10, followed by the 3'UTR of the arc5-1 gene, "arc-terminator". "arc5'UTR" refers to the 5' untranslated region of the arc5-1 gene. "Sig10" refers to the lectin 1 gene signal peptide. "RB" refers to ribosomal binding site. Coding regions and regulatory sequences are indicated as blocks (not to scale). The Beta Casein is "truncated" in that the bovine secretion signal is removed and replaced with a plant targeting signal.

[28] **FIG. 6** shows an example expression cassette comprising a OaS1-T:FM:OLG1 fusion (Optimized Alpha S1 Casein Truncated version 1:Chymosin cleavage site:beta-lactoglobulin version 1, SEQ ID NOs: 75-76), expression of which is driven by PvPhas promoter fused with arc5'UTR:sig10, followed by the 3'UTR of the arc5-1 gene, "arc-terminator". "arc5'UTR" refers to the 5' untranslated region of the arc5-1 gene. "Sig10" refers to the lectin 1 gene signal peptide. "RB" refers to ribosomal binding site. Coding regions and regulatory sequences are indicated as blocks (not to scale). The Alpha S1 Casein is "truncated" in that the bovine secretion signal is removed and replaced with a plant targeting signal.

[29] **FIG. 7** shows an example expression cassette comprising a para-OKC1-T:FM:OLG1:KDEL fusion (Optimized paraKappa Casein version 1:Chymosin cleavage site:beta-lactoglobulin version 1, SEQ ID NOs: 77-78), expression of which is driven by PvPhas promoter

fused with arc5'UTR:sig 10, followed by the ER retention signal (KDEL) and the 3'UTR of the arc5-1 gene, "arc-terminator". "arc5'UTR" refers to the 5' untranslated region of the arc5-1 gene. "Sig10" refers to the lectin 1 gene signal peptide. "RB" refers to ribosomal binding site. Coding regions and regulatory sequences are indicated as blocks (not to scale).

[30] **FIG. 8** shows an example expression cassette comprising a para-OKC1-T:FM:OLG1 fusion (Optimized paraKappa Casein version 1:Chymosin cleavage site:beta-lactoglobulin version 1, SEQ ID NOs: 79-80), expression of which is driven by PvPhas promoter fused with arc5'UTR:sig 10, followed by the 3'UTR of the arc5-1 gene, "arc-terminator." "arc5'UTR" refers to the 5' untranslated region of the arc5-1 gene. "Sig10" refers to the lectin 1 gene signal peptide. "RB" refers to ribosomal binding site. Coding regions and regulatory sequences are indicated as blocks (not to scale).

[31] **FIG. 9** shows an example expression cassette comprising a OKC1-T:OLG1 fusion (Optimized Kappa Casein version 1:beta-lactoglobulin version 1, SEQ ID NOs: 81-82), expression of which is driven by the promoter and signal peptide of glycinin 1 (GmSeed2:sig2) followed by the ER retention signal (KDEL) and the nopaline synthase gene termination sequence (nos term). Coding regions and regulatory sequences are indicated as blocks (not to scale).

[32] **FIGs. 10A, 10B, 10C, and 10D** show protein detection by western blotting. **FIG. 10A** shows detection of the fusion protein using a primary antibody raised against κ -casein (kCN). The kCN commercial protein is detected at an apparent MW of ~26 kDa (theoretical: 19 kDa - arrow). The fusion protein is detected at an apparent MW of ~40 kDa (theoretical: 38 kDa - arrowhead). **FIG. 10B** shows detection of the fusion protein using a primary antibody raised against β -lactoglobulin (LG). The LG commercial protein is detected at an apparent MW of ~18kDa (theoretical: 18 kDa - arrow). The fusion protein is detected at an apparent MW of ~40 kDa (theoretical: 38 kDa - arrowhead). **FIG. 10C** and **FIG. 10D** show protein gels as control for equal lane loading (image is taken at the end of the SDS run).

[33] **FIG. 11A-11E** provides a series of illustrations showing potential mechanisms by which casein proteins may be degraded in plant cells, and how fusion of a casein protein with a second protein (i.e., a fusion partner) may lead to accumulation thereof. KCN stands for kappa-casein, BC

stands for beta casein, aS1 stands for alpha-S1 casein, aS2 stands for alpha-S2 casein, PTM stands for post-translational modification.

[34] **FIGs. 12A** and **12B** show two illustrative fusion proteins. In **FIG. 12A**, a κ -casein protein is fused to a β -lactoglobulin protein. The κ -casein comprises a natural chymosin cleavage site (arrow 1). Cleavage of the fusion protein with rennet (or chymosin) yields two fragments: a para-kappa casein fragment, and a fragment comprising a κ -casein macropeptide fused to β -lactoglobulin. In some embodiments, a second protease cleavage site may be added at the C-terminus of the κ -casein protein (i.e., at arrow 2), in order to further allow separation of the κ -casein macropeptide and the β -lactoglobulin. The second protease cleavage site may be a rennet cleavage site (e.g., a chymosin cleavage site), or it may be a cleavage site for a different protease. In **FIG. 12B**, a para- κ -casein protein is fused directly to β -lactoglobulin. A protease cleavage site (e.g., a chymosin cleavage site) is added between the para- κ -casein and the β -lactoglobulin to allow for separation thereof. By fusing the para- κ -casein directly to the β -lactoglobulin, no κ -casein macropeptide is produced upon cleavage of the fusion by chymosin (or other protease).

[35] **FIG. 13** is a flow-chart showing an illustrative process for producing a food composition comprising an unstructured milk protein, as described herein. Initially, an expression construct for expression of a fusion protein in a plant cell is designed. The construct is transformed into a plant, and the plant is regenerated. Seeds are collected from the plant, and processed (e.g., by seed hulling and grinding) to produce a seed processing composition. Protein is extracted, and optionally enriched and/or concentrated (i.e., to produce a protein concentrate composition). The extracted fusion protein may optionally be cleaved or used directly to produce a food composition.

[36] **FIGs. 14A** and **14B** are images of a western blot used to detect kappa-casein protein (kCN) in samples comprising soybean total protein extracts (WT) and soybean total protein extracts spiked with 100 ng of KCN in the presence (WT + kCN + Halt) or absence (WT + KCN) of protease inhibitors. 5 μ g of total protein was loaded in each lane. **FIG. 14A** shows protein detected using a primary antibody raised against KCN. **FIG. 14B** shows total protein, as a loading control (Stain-Free detection by Bio-Rad®).

[37] **FIGs. 15A** and **15B** are images that show protein detection by western blotting. **FIG. 15A** shows detection of a fusion protein comprising β -casein and β -lactoglobulin using a primary

antibody raised against β -casein (B-CN). Commercial protein was detected at an apparent MW of ~30 kDa (arrowhead; theoretical: 23.5 kDa). The fusion protein was detected at an apparent MW of ~40 kDa (arrow; theoretical: 42 kDa). **FIG. 15B** shows a protein gel as a control for equal lane loading, visualized using stain-free detection by Bio Rad® (image is taken at the end of the SDS run). 5 μ g of total protein extracts were loaded per lane.

[38] **FIG. 16A** shows molecular weight of various proteins, and levels of kappa-casein expression observed in transformed soybeans when those proteins are fused to the kappa-casein. **FIG. 16B** shows hydrophobicity of various proteins, and levels of kappa-casein expression observed in transformed soybeans when those proteins are fused to the kappa-casein. **FIG. 16C** shows flexibility of various proteins (i.e., number of disulfide bonds), and levels of kappa-casein expression observed in transformed soybeans when those proteins are fused to the kappa-casein. Expression levels shown in **FIG. 16A-16C** are relative to kappa casein expressed alone (i.e., not as a fusion, KCN only). The values for % KCN only are presented as a $\log_{10}$ scale. Values above 100% indicate that kappa-casein was stabilized by the fusion.

[39] **FIG. 17** is a schematic showing an illustrative process for producing a food composition. The food composition produced according to this method may comprise one or more of: (i) one or more constituent proteins derived from a fusion protein, (ii) the fusion protein itself, or (ii) other protein extracted from the seed that was used to produce the fusion protein.

[40] **FIG. 18** is a schematic that shows how knocking-down or knocking-out the expression and/or activity of one or more proteases in a plant seed may prevent degradation of a casein protein expressed therein. As shown in the schematic, the casein accumulates in the seed at a higher level than in a seed with wildtype levels of protease expression and/or activity.

[41] **FIG. 19** is a schematic demonstrating how the properties of a seed processing composition, or a food composition comprising the same, may be improved if the composition comprises one or more casein proteins. These properties may be improved if the composition comprises a casein protein monomer (i.e., a casein protein that is not part of a fusion protein), or a fusion protein comprising one or more caseins.

[42] **FIG. 20** is a schematic demonstrating an illustrative mechanism that may be used to protect one or more proteins (e.g., casein proteins) from degradation in a host cell, leading to accumulation

thereof. The protein (e.g., a casein protein) is fused to one or more proteins that is capable of forming a protein body (e.g., a prolamin). After the fusion protein is synthesized and retained in the endoplasmic reticulum (ER), a protein body is formed (PB). The fusion protein (including, for example, the casein protein) is contained within the PB. Proteases that would degrade the caseins, do not have access to the fusion protein inside the PB. In this figure, the term “PSV” refers to protein storage vacuole.

[43] FIG. 21 shows protein detection by western blotting. The top panel shows detection of a fusion protein comprising β -casein and cancin using a primary antibody raised against β -casein (B-CN). Commercial protein was detected at an apparent MW of ~30 kDa (arrowhead; theoretical: 23.5 kDa). The fusion protein was detected at an apparent MW of ~50 kDa (arrow; theoretical: 44.3 kDa). The first lane shows molecular weight markers. The second lane shows protein from T1 seed from recombinant plant line KV7. Lane 3-7 shows soybean wildtype seed extracts spiked with 0%, 1%, 2%, 4%, or 6% TSP commercially available β -casein. The bottom panel shows a protein gel as a control for equal lane loading, visualized using stain-free detection by Bio Rad® (image is taken at the end of the SDS run). 2.5 μ g of total protein extracts were loaded per lane.

[44] FIG. 22 shows protein detection by western blotting. The top panel shows detection of a fusion protein comprising β -casein and a partial zein (amino acids 17-112) using a primary antibody raised against β -casein (B-CN). Commercial protein was detected at an apparent MW of ~30 kDa (arrowhead; theoretical: 23.5 kDa). The fusion protein was detected at an apparent MW of ~30 kDa (arrow; theoretical: 23.5 kDa). The first four lanes show protein from T1 seed from a recombinant plant. The fifth lane shows molecular weight markers. Lanes 6-9 shows soybean wildtype seed extracts spiked with 0%, 1.5%, 2.5%, or 5% TSP commercially available β -casein. The bottom panel shows a protein gel as a control for equal lane loading, visualized using stain-free detection by Bio Rad® (image is taken at the end of the SDS run). 2.5 μ g of total protein extracts were loaded per lane.

[45] FIG. 23 shows a binary *Agrobacterium* vector used to co-express a Gene of Interest (GOI, e.g., a casein protein) and a kinase (e.g., a Fam20C kinase) in a plant cell.

[46] FIG. 24A-24E shows expression constructs used to co-express a Gene of Interest (GOI, e.g., a casein protein) and a kinase (e.g., a Fam20C kinase) in a plant cell.

[47] FIG. 25A-25F show expression constructs used to express a Gene of Interest (GOI, e.g., a casein protein) in a plant cell, wherein the GOI is fused to a glycoprotein tag, such as a (SP)11 tag.

[48] FIG. 26A-26G shows expression constructs used to co-express a Gene of Interest (GOI, e.g., a casein protein) and a protein capable of inducing a protein body (e.g., a prolamin, zein, canein, hydrophobin, or elastin-like protein) in a plant cell.

[49] FIG. 27 shows a binary *Agrobacterium* vector used to co-express a Gene of Interest (GOI, e.g., a casein protein) and a protein capable of inducing a protein body in a plant cell.

[50] FIG. 28 is a photograph which depicts the melting properties of various cheese compositions made with isolated kappa and beta-caseins. Top left: composition A (75% kappa-casein, 25% beta-casein); top right: composition B (100% kappa-casein); bottom left: composition C (50% kappa-casein, 50% beta-casein), bottom right: composition A (100% beta-casein).

[51] FIG. 29 is a line graph showing cheese stretch with increasing contribution of protein from beta-casein (see also Tables 23-28).

[52] FIG. 30 is a line graph showing melt scores of cheese compositions comprising one or more of beta-casein, kappa-casein and alpha casein (see also Tables 23-28).

[53] FIG. 31 is a line graph showing stretch of cheese compositions comprising one or more of beta-casein, kappa-casein and alpha casein (see also Tables 23-28).

[54] FIG. 32 is a graph showing estimated apparent viscosity (in centipoise (cP)) at shear rates in the range of 0.01 to 1000 sec^{-1} for a milk composition comprising beta-casein as the only casein (BC milk), a yogurt composition comprising beta-casein as the only casein (BC yogurt), and an ice cream mix composition comprising beta-casein as the only casein (BC IC mix).

[55] FIG. 33 is a western blot showing expression of a beta-casein tetramer (BC4) in *E. Coli*. Commercial beta-casein, in monomeric form, was detected at an apparent molecular weight of ~30 kDa (theoretical: 23.5 kDa - arrowhead). The BC4 fusion protein was detected at an apparent MW of ~100 kDa (theoretical: 94 kDa - arrow).

[56] FIG. 34 is a western blot showing expression of a fusion protein comprising beta-casein and beta-lactoglobulin in tobacco leaves. Commercial beta-casein, in monomeric form, was

detected at an apparent molecular weight of ~30 kDa (theoretical: 23.5 kDa - arrowhead). The fusion protein was detected at an apparent MW of ~48 kDa (theoretical: 42 kDa - arrow).

DETAILED DESCRIPTION

[57] Provided herein are compositions and methods for producing milk proteins, which allow for safe, sustainable and humane production of milk proteins for commercial use, such as use in food compositions. The disclosure provides recombinant fusion proteins comprising at least first protein and a second protein, wherein at least one of the first protein and the second protein is a milk protein, or fragment thereof. The disclosure also provides methods for producing the recombinant fusions proteins, and food compositions comprising the same.

[58] Also provided herein are alternative dairy compositions, solid phase protein-stabilized emulsions, cheese compositions, and colloidal suspensions, comprising one or more casein proteins, wherein the casein proteins are isolated or recombinant, and are selected from the group consisting of kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein and alpha-S2-casein. The compositions, emulsions, or suspensions may be used to produce food compositions that have organoleptic properties similar to traditional dairy compositions.

[59] The following description includes information that may be useful in understanding the present disclosure. It is not an admission that any of the information provided herein is prior art or relevant to the presently claimed disclosures, or that any publication specifically or implicitly referenced is prior art.

Definitions

[60] While the following terms are believed to be well understood by one of ordinary skill in the art, the following definitions are set forth to facilitate explanation of the presently disclosed subject matter.

[61] All technical and scientific terms used herein, unless otherwise defined below, are intended to have the same meaning as commonly understood by one of ordinary skill in the art. References to techniques employed herein are intended to refer to the techniques as commonly understood in the art, including variations on those techniques and/or substitutions of equivalent techniques that would be apparent to one of skill in the art.

[62] Any ranges listed herein are intended to be inclusive of endpoints. For example, a range of 2-4 includes 2 and 4.

[63] As used herein, the singular forms “a,” “an,” and “the” include plural referents unless the content clearly dictates otherwise.

[64] The term “about” or “approximately” when immediately preceding a numerical value means a range (e.g., plus or minus 10% of that value). For example, “about 50” can mean 45 to 55, “about 25,000” can mean 22,500 to 27,500, etc., unless the context of the disclosure indicates otherwise, or is inconsistent with such an interpretation. For example, in a list of numerical values such as “about 49, about 50, about 55, ...”, “about 50” means a range extending to less than half the interval(s) between the preceding and subsequent values, e.g., more than 49.5 to less than 52.5. Furthermore, the phrases “less than about” a value or “greater than about” a value should be understood in view of the definition of the term “about” provided herein. Similarly, the term “about” when preceding a series of numerical values or a range of values (e.g., “about 10, 20, 30” or “about 10-30”) refers, respectively to all values in the series, or the endpoints of the range.

[65] As used herein, “mammalian milk” can refer to milk derived from any mammal, such as bovine, human, goat, sheep, camel, buffalo, water buffalo, dromedary, llama and any combination thereof. In some embodiments, a mammalian milk is a bovine milk.

[66] As used herein, “structured” refers to those proteins having a well-defined secondary and tertiary structure, and “unstructured” refers to proteins that do not have well defined secondary and/or tertiary structures. An unstructured protein may also be described as lacking a fixed or ordered three-dimensional structure. “Disordered” and “intrinsically disordered” are synonymous with unstructured.

[67] As used herein, “rennet” refers to a set of enzymes typically produced in the stomachs of ruminant mammals. Chymosin, its key component, is a protease enzyme that cleaves κ -casein (to produce para- κ -casein and a macropeptide (see e.g., **FIG. 12**)). In addition to chymosin, rennet contains other enzymes, such as pepsin and lipase. Rennet is used to separate milk into solid curds (for cheesemaking) and liquid whey. Rennet or rennet substitutes are used in the production of many cheeses.

[68] As used herein "whey" refers to the liquid remaining after milk has been curdled and strained, for example during cheesemaking. Whey comprises a collection of globular proteins, typically a mixture of β -lactoglobulin, α -lactalbumin, bovine serum albumin, and immunoglobulins.

[69] The term "plant" includes reference to whole plants, plant organs, plant tissues, and plant cells and progeny of same, but is not limited to angiosperms and gymnosperms such as Arabidopsis, potato, tomato, tobacco, alfalfa, lettuce, carrot, strawberry, sugar beet, cassava, sweet potato, soybean, lima bean, pea, chick pea, maize (corn), turf grass, wheat, rice, barley, sorghum, oat, oak, eucalyptus, walnut, palm and duckweed as well as fern and moss. Thus, a plant may be a monocot, a dicot, a vascular plant reproduced from spores such as fern or a nonvascular plant such as moss, liverwort, hornwort and algae. The word "plant," as used herein, also encompasses plant cells, seeds, plant progeny, propagule whether generated sexually or asexually, and descendants of any of these, such as cuttings or seed. Plant cells include suspension cultures, callus, embryos, meristematic regions, callus tissue, leaves, roots, shoots, gametophytes, sporophytes, pollen, seeds and microspores. Plants may be at various stages of maturity and may be grown in liquid or solid culture, or in soil or suitable media in pots, greenhouses or fields. Expression of an introduced leader, trailer or gene sequences in plants may be transient or permanent.

[70] The term "vascular plant" refers to a large group of plants that are defined as those land plants that have lignified tissues (the xylem) for conducting water and minerals throughout the plant and a specialized non-lignified tissue (the phloem) to conduct products of photosynthesis. Vascular plants include the clubmosses, horsetails, ferns, gymnosperms (including conifers) and angiosperms (flowering plants). Scientific names for the group include Tracheophyta and Tracheobionta. Vascular plants are distinguished by two primary characteristics. First, vascular plants have vascular tissues which distribute resources through the plant. This feature allows vascular plants to evolve to a larger size than non-vascular plants, which lack these specialized conducting tissues and are therefore restricted to relatively small sizes. Second, in vascular plants, the principal generation phase is the sporophyte, which is usually diploid with two sets of chromosomes per cell. Only the germ cells and gametophytes are haploid. By contrast, the principal generation phase in non-vascular plants is the gametophyte, which is haploid with one set of chromosomes per cell. In these plants, only the spore stalk and capsule are diploid.

[71] The term "non-vascular plant" refers to a plant without a vascular system consisting of xylem and phloem. Many non-vascular plants have simpler tissues that are specialized for internal transport of water. For example, mosses and leafy liverworts have structures that look like leaves, but are not true leaves because they are single sheets of cells with no stomata, no internal air spaces and have no xylem or phloem. Non-vascular plants include two distantly related groups. The first group are the bryophytes, which is further categorized as three separate land plant Divisions, namely Bryophyta (mosses), Marchantiophyta (liverworts), and Anthocerotophyta (hornworts). In all bryophytes, the primary plants are the haploid gametophytes, with the only diploid portion being the attached sporophyte, consisting of a stalk and sporangium. Because these plants lack lignified water-conducting tissues, they can't become as tall as most vascular plants. The second group is the algae, especially the green algae, which consists of several unrelated groups. Only those groups of algae included in the Viridiplantae are still considered relatives of land plants.

[72] The term "plant part" refers to any part of a plant including but not limited to the embryo, shoot, root, stem, seed, stipule, leaf, petal, flower bud, flower, ovule, bract, trichome, branch, petiole, internode, bark, pubescence, tiller, rhizome, frond, blade, ovule, pollen, stamen, and the like. The two main parts of plants grown in some sort of media, such as soil or vermiculite, are often referred to as the "above-ground" part, also often referred to as the "shoots", and the "below-ground" part, also often referred to as the "roots".

[73] The term "plant tissue" refers to any part of a plant, such as a plant organ. Examples of plant organs include, but are not limited to the leaf, stem, root, tuber, seed, branch, pubescence, nodule, leaf axil, flower, pollen, stamen, pistil, petal, peduncle, stalk, stigma, style, bract, fruit, trunk, carpel, sepal, anther, ovule, pedicel, needle, cone, rhizome, stolon, shoot, pericarp, endosperm, placenta, berry, stamen, and leaf sheath.

[74] The term "seed" is meant to encompass the whole seed and/or all seed components, including, for example, the coleoptile and leaves, radicle and coleorhiza, scutellum, starchy endosperm, aleurone layer, pericarp and/or testa, either during seed maturation and seed germination.

[75] "Microorganism" and "microbe" mean any microscopic unicellular organism and can include bacteria, algae, yeast, or fungi.

[76] The term “transgenic” means an organism that has been transformed with one or more exogenous nucleic acids from another species. “Transformation” refers to a process by which a nucleic acid is introduced into a cell, either transiently or stably. Transformation may rely on any known method for the insertion of nucleic acid sequences into a prokaryotic or eukaryotic host cell, including *Agrobacterium*-mediated transformation protocols, viral infection, whiskers, electroporation, heat shock, lipofection, polyethylene glycol treatment, micro-injection, and particle bombardment.

[77] “Stably integrated” refers to the permanent, or non-transient retention and/or expression of a polynucleotide in and by a cell genome. Thus, a stably integrated polynucleotide is one that is a fixture within a transformed cell genome and can be replicated and propagated through successive progeny of the cell or resultant transformed plant. Transformation may occur under natural or artificial conditions using various methods well known in the art. Transformation may rely on any known method for the insertion of nucleic acid sequences into a prokaryotic or eukaryotic host cell, including *Agrobacterium*-mediated transformation protocols, viral infection, whiskers, electroporation, heat shock, lipofection, polyethylene glycol treatment, micro-injection, and particle bombardment.

[78] As used herein, the terms “stably expressed” or “stable expression” refer to expression and accumulation of a protein in a plant cell. In some embodiments, a protein may accumulate because it is not degraded by endogenous plant proteases. In some embodiments, a protein is considered to be stably expressed in a plant if it is present in the plant in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant.

[79] As used herein, the term “fusion protein” refers to a protein comprising at least two constituent proteins (or fragments or variants thereof) that are encoded by separate genes, and that have been joined so that they are transcribed and translated as a single polypeptide. In some embodiments, a fusion protein may be separated into its constituent proteins, for example by cleavage with a protease.

[80] The term “recombinant” refers to nucleic acids or proteins formed by laboratory methods of genetic recombination (e.g., molecular cloning) to bring together genetic material from multiple sources, creating sequences that would not otherwise be found in the genome. A recombinant

fusion protein is a protein created by combining sequences encoding two or more constituent proteins, such that they are expressed as a single polypeptide. Recombinant fusion proteins may be expressed *in vivo* in various types of host cells, including plant cells, bacterial cells, fungal cells, mammalian cells, etc. Recombinant fusion proteins may also be generated *in vitro*.

[81] The term "promoter" or a "transcription regulatory region" refers to nucleic acid sequences that influence and/or promote initiation of transcription. Promoters are typically considered to include regulatory regions, such as enhancer or inducer elements. The promoter will generally be appropriate to the host cell in which the target gene is being expressed. The promoter, together with other transcriptional and translational regulatory nucleic acid sequences (also termed "control sequences"), is necessary to express any given gene. In general, the transcriptional and translational regulatory sequences include, but are not limited to, promoter sequences, ribosomal binding sites, transcriptional start and stop sequences, translational start and stop sequences, and enhancer or activator sequences.

[82] The term signal peptide – also known as “signal sequence”, “targeting signal”, “localization signal”, “localization sequence”, “transit peptide”, “leader sequence”, or “leader peptide”, is used herein to refer to an N-terminal peptide which directs a newly synthesized protein to a specific cellular location or pathway. Signal peptides are often cleaved from a protein during translation or transport, and are therefore not typically present in a mature protein.

[83] The term "proteolysis" or "proteolytic" or "proteolyze" means the breakdown of proteins into smaller polypeptides or amino acids. Uncatalyzed hydrolysis of peptide bonds is extremely slow. Proteolysis is typically catalyzed by cellular enzymes called proteases, but may also occur by intra-molecular digestion. Low pH or high temperatures can also cause proteolysis non-enzymatically. Limited proteolysis of a polypeptide during or after translation in protein synthesis often occurs for many proteins. This may involve removal of the N-terminal methionine, signal peptide, and/or the conversion of an inactive or non-functional protein to an active one.

[84] The term "2A peptide", used herein, refers to nucleic acid sequence encoding a 2A peptide or the 2A peptide itself. The average length of 2A peptides is 18-22 amino acids. The designation "2A" refers to a specific region of picornavirus polyproteins and arose from a systematic nomenclature adopted by researchers. In foot-and-mouth disease virus (FMDV), a member of

Picornaviridae family, a 2A sequence appears to have the unique capability to mediate cleavage at its own C-terminus by an apparently enzyme-independent, novel type of reaction. This sequence can also mediate cleavage in a heterologous protein context in a range of eukaryotic expression systems. The 2A sequence is inserted between two genes of interest, maintaining a single open reading frame. Efficient cleavage of the polyprotein can lead to co-ordinate expression of active two proteins of interest. Self-processing polyproteins using the FMDV 2A sequence could therefore provide a system for ensuring coordinated, stable expression of multiple introduced proteins in cells including plant cells.

[85] The term "purifying" is used interchangeably with the term "isolating" and generally refers to the separation of a particular component from other components of the environment in which it was found or produced. For example, purifying a recombinant protein from plant cells in which it was produced typically means subjecting transgenic protein containing plant material to biochemical purification and/or column chromatography.

[86] When referring to expression of a protein in a specific amount per the total protein weight of the soluble protein extractable from the plant ("TSP"), it is meant an amount of a protein of interest relative to the total amount of protein that may reasonably be extracted from a plant using standard methods. Methods for extracting total protein from a plant are known in the art. For example, total protein may be extracted from seeds by bead beating seeds at about 15000 rpm for about 1 min. The resulting powder may then be resuspended in an appropriate buffer (e.g., 50 mM Carbonate-Bicarbonate pH 10.8, 1 mM DTT, 1X Protease Inhibitor Cocktail). After the resuspended powder is incubated at about 4°C for about 15 minutes, the supernatant may be collected after centrifuging (e.g., at 4000g, 20 min, 4°C). Total protein may be measured using standard assays, such as a Bradford assay. The amount of protein of interest may be measured using methods known in the art, such as an ELISA or a Western Blot.

[87] When referring to a nucleic acid sequence or protein sequence, the term "identity" is used to denote similarity between two sequences. Sequence similarity or identity may be determined using standard techniques known in the art, including, but not limited to, the local sequence identity algorithm of Smith & Waterman, *Adv. Appl. Math.* 2, 482 (1981), by the sequence identity alignment algorithm of Needleman & Wunsch, *J Mol. Biol.* 48,443 (1970), by the search for

similarity method of Pearson & Lipman, *Proc. Natl. Acad. Sci. USA* 85, 2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Drive, Madison, WI), the Best Fit sequence program described by Devereux et al., *Nucl. Acid Res.* 12, 387-395 (1984), or by inspection. Another suitable algorithm is the BLAST algorithm, described in Altschul et al., *J Mol. Biol.* 215, 403-410, (1990) and Karlin et al., *Proc. Natl. Acad. Sci. USA* 90, 5873-5787 (1993). A particularly useful BLAST program is the WU-BLAST-2 program which was obtained from Altschul et al., *Methods in Enzymology*, 266, 460-480 (1996); blast.wustl.edu/blast/README.html. WU-BLAST-2 uses several search parameters, which are optionally set to the default values. The parameters are dynamic values and are established by the program itself depending upon the composition of the particular sequence and composition of the particular database against which the sequence of interest is being searched; however, the values may be adjusted to increase sensitivity. Further, an additional useful algorithm is gapped BLAST as reported by Altschul et al, (1997) *Nucleic Acids Res.* 25, 3389-3402. Unless otherwise indicated, percent identity is determined herein using the algorithm available at the world wide web address: blast.ncbi.nlm.nih.gov/Blast.cgi.

[88] As used herein, the terms “dicot” or “dicotyledon” or “dicotyledonous” refer to a flowering plant whose embryos have two seed leaves or cotyledons. Examples of dicots include, but are not limited to, Arabidopsis, tobacco, tomato, potato, sweet potato, cassava, alfalfa, lima bean, pea, chick pea, soybean, carrot, strawberry, lettuce, oak, maple, walnut, rose, mint, squash, daisy, Quinoa, buckwheat, mung bean, cow pea, lentil, lupin, peanut, fava bean, French beans (i.e., common beans), mustard, or cactus.

[89] The terms “monocot” or “monocotyledon” or “monocotyledonous” refer to a flowering plant whose embryos have one cotyledon or seed leaf. Examples of monocots include, but are not limited to turf grass, maize (corn), rice, oat, wheat, barley, sorghum, orchid, iris, lily, onion, palm, and duckweed.

[90] As used herein, a “low lactose product” is any food composition considered by the FDA to be “lactose reduced”, “low lactose”, or “lactose free”.

[91] As used herein, a “milk protein” is any protein, or fragment or variant thereof, that is typically found in one or more mammalian milks. In some embodiments, the milk proteins described herein are casein proteins, such as kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein.

[92] As used herein, a “non-milk” protein is any protein that is not typically found in any mammalian milk composition. One non-limiting example of a non-milk protein is green fluorescent protein (GFP).

[93] As used herein, a “caseinate” is a compound derived from casein. Caseinates may be produced by adding acid to skim milk to reduce the pH to about 4.6, which causes the casein proteins to be precipitated. The resulting curd is rinsed and dried to produce acid casein. Acid casein is typically insoluble without further treatment, such as pH adjustment. Acid casein either before or after drying can be mixed with a base such as sodium hydroxide to produce sodium caseinate, or calcium hydroxide to produce calcium caseinate.

[94] As used herein, an “alternative dairy composition” is a composition that comprises an isolated, or recombinant, casein protein, and may also comprise variations of the composition, such as a low-fat alternative dairy composition.

[95] As used herein, the phrase “solid phase, protein-stabilized emulsion” refers to a homogenous and stable emulsion that is a solid at room temperature. The solid-phase, protein stabilized emulsions described herein is formed by the protein reducing the interfacial tension between the continuous aqueous phase and discontinuous lipid phase by aligning and/or unfolding at the interface. The amphiphilic nature of proteins allows them to interact with both phases and association between proteins in the aqueous phase results in decreased mobility of water in the form of increased viscosity and/or solid like behavior at different temperatures. The presence of “emulsifying salts” can enhance the emulsifying properties of the proteins.

[96] As used herein, “cheese” refers to a food that is produced by curdling animal-derived milk. The milk may be curdled using, for example, enzymes (e.g., rennet), or using acid.

[97] As used herein, “cheese composition” refers to a food that is produced by combining one or more milk proteins, optionally with other ingredients, as described herein. For example, cheese

compositions may be produced using one or more recombinant milk proteins, or one or more milk proteins isolated from bovine milk. The cheese compositions may, in some embodiments, include only one milk protein. In some embodiments, the cheese compositions may comprise 2, 3, or 4 milk proteins. In some embodiments, the cheese compositions may comprise one or more milk proteins in a ratio that does not occur in the milk produced by any mammal (i.e., a non-naturally occurring ratio).

[98] As used herein, the term “melt”, “melting”, or “meltability” refers to the liquefaction of cheese or a cheese composition by heat.

[99] As used herein, the term “viscosity” or “flow” refers to the tendency of cheese (or a cheese composition) to spread and flow when completely melted.

[100] As used herein, the term “stretch”, “stretching”, or “stretchability” refers to the formation of fibrous strands of cheese (or a cheese composition) that elongate without breaking.

[101] As used herein, the term “oiling-off” refers to the tendency of free oil separation from melted cheese or a cheese composition (also known as fat leakage).

[102] As used herein, the term “browning” or “blistering” refers to the trapped pockets of heated air and steam that may be scorched during baking with cheese (or a cheese composition).

[103] As used herein, the term “whitening” or “decolorization” refers to the bleaching of cheese (or a cheese composition).

[104] As used herein, the term “spread”, “spreading” or “spreadability” refers to the ability of cheese or a cheese composition to spread over a surface on application of slight force to form a layer, thin enough to form a coating.

[105] The term “ash” is used herein as it is well known in the art, and means one or more ions, elements, minerals and/or compounds that may be found in mammalian produced milk. Ash may comprise one or more of sodium, potassium, calcium, magnesium, phosphorus, iron, copper, zinc, chloride, manganese, selenium, iodine, phosphate, citrate, sulfate, and carbonate. In some embodiments, ash may comprise calcium carbonate and/or sodium citrate.

Milk Proteins

[106] The fusion proteins described herein may comprise one or more milk proteins. In some embodiments, the fusion proteins described herein may comprise a first protein and a second protein, wherein the first protein and/or second protein is a milk protein. In some embodiments, the first protein and the second protein are both milk proteins. As used herein the term “milk protein” refers to any protein, or fragment or variant thereof, that is typically found in one or more mammalian milks. Examples of mammalian milk include, but are not limited to, milk produced by a cow, human, goat, sheep, camel, horse, donkey, dog, cat, elephant, monkey, mouse, rat, hamster, guinea pig, whale, dolphin, seal, sheep, buffalo, water buffalo, dromedary, llama, yak, zebu, reindeer, mole, otter, weasel, wolf, raccoon, walrus, polar bear, rabbit, or giraffe. Some representative examples of milk protein species of the disclosure can be found in Table 34.

[107] The composition of milk varies depending on the mammal. For example, as shown below in Table 1, cow milk comprises β -lactoglobulin, α -S1-casein, and α -S2-casein, whereas human milk does not. However, for the purposes of this disclosure, β -lactoglobulin, α -S1-casein, and α -S2-casein are considered milk proteins.

Table 1: Protein composition of human and cow milk

Protein	Human milk (mg/mL)	Bovine (cow) milk (mg/mL)
α -lactalbumin	2.2	1.2
α -s1-casein	0	11.6
α -s2-casein	0	3.0
β -casein	2.2	9.6
κ -casein	0.4	3.6
γ -casein	0	1.6
Immunoglobulins	0.8	0.6
Lactoferrin	1.4	0.3
β -lactoglobulin	0	3.0
Lysozyme	0.5	Traces
Serum albumin	0.4	0.4
Other	0.8	0.6

[108] Illustrative milk proteins that may be used in the fusion proteins of the disclosure include, but are not limited to, α -S1 casein, α -S2 casein, β -casein, κ -casein, para- κ -casein, β -lactoglobulin,

α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, serum albumin, and immunoglobulins (e.g., IgA, IgG, IgM, IgE).

[109] Milk proteins may be further classified as structured or unstructured proteins. An “unstructured milk protein” is a milk protein that lacks a defined secondary structure, a defined tertiary structure, or a defined secondary and tertiary structure. Whether a milk protein is unstructured may be determined using a variety of biophysical and biochemical methods known in the art, such as small angle X-ray scattering, Raman optical activity, circular dichroism, nuclear magnetic resonance (NMR) and protease sensitivity. In some embodiments, a milk protein is considered to be unstructured if it is unable to be crystallized using standard techniques.

[110] Illustrative unstructured milk proteins that may be used in the fusion proteins of the disclosure includes members of the casein family of proteins, such as α -S1 casein, α -S2 casein, β -casein, and κ -casein. The caseins are phosphoproteins, and make up approximately 80% of the protein content in bovine milk and about 20-45% of the protein in human milk. Caseins form a multi-molecular, granular structure called a casein micelle in which some enzymes, water, and salts, such as calcium and phosphorous, are present. The micellar structure of casein in milk is significant in terms of a mode of digestion of milk in the stomach and intestine and a basis for separating some proteins and other components from cow milk. In practice, casein proteins in bovine milk can be separated from whey proteins by acid precipitation of caseins, by breaking the micellar structure by partial hydrolysis of the protein molecules with proteolytic enzymes, or microfiltration to separate the smaller soluble whey proteins from the larger casein micelle. Caseins are relatively hydrophobic, making them poorly soluble in water.

[111] In some embodiments, the casein proteins described herein (e.g., α -S1 casein, α -S2 casein, β -casein, and/or κ -casein) are isolated or derived from cow (*bos taurus*), goat (*capra hircus*), sheep (*ovis aries*), water buffalo (*bubalus bubalis*), dromedary camel (*camelus dromedaries*), bactrian camel (*camelus bactrianus*), wild yak (*bos mutus*), horse (*equus caballus*), donkey (*equus asinus*), reindeer (*rangifer tarandus*), eurasian elk (*alces alces*), alpaca (*vicugna pacos*), zebu (*bos indicus*), llama (*lama glama*), or human (*homo sapiens*). In some embodiments, a casein protein (e.g., α -S1 casein, α -S2 casein, β -casein, or κ -casein) has at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with a casein protein from one or more of cow (*bos taurus*), goat (*capra hircus*), sheep (*ovis aries*), water buffalo (*bubalus*

bubalis), dromedary camel (*camelus dromedaries*), bactrian camel (*camelus bactrianus*), wild yak (*bos mutus*), horse (*equus caballus*), donkey (*equus asinus*), reindeer (*rangifer tarandus*), eurasian elk (*alces alces*), alpaca (*vicugna pacos*), zebu (*bos indicus*), llama (*lama glama*), or human (*homo sapiens*).

[112] As used herein, the term “ α -S1 casein” refers to not only the α -S1 casein protein, but also fragments or variants thereof. α -S1 casein is found in the milk of numerous different mammalian species, including cow, goat, and sheep. The sequence, structure and physical/chemical properties of α -S1 casein derived from various species is highly variable. An illustrative sequence for bovine α -S1 casein can be found at Uniprot Accession No. P02662, and an illustrative sequence for goat α -S1 casein can be found at GenBank Accession No. X59836.1. The terms “ α -S1 casein” and “alpha-S1-casein” (and similar terms) are used interchangeably herein.

[113] As used herein, the term “ α -S2 casein” refers to not only the α -S2 casein protein, but also fragments or variants thereof. α -S2 is known as epsilon-casein in mouse, Gamma-casein in rat, and casein-A in guinea pig. The sequence, structure and physical/chemical properties of α -S2 casein derived from various species is highly variable. An illustrative sequence for bovine α -S2 casein can be found at Uniprot Accession No. P02663, and an illustrative sequence for goat α -S2 casein can be found at Uniprot Accession No. P33049. The terms “ α -S2 casein” and “alpha-S2-casein” (and similar terms) are used interchangeably herein.

[114] As used herein, the term “ β -casein” refers to not only the β -casein protein, but also fragments or variants thereof. For example, A1 and A2 β -casein are genetic variants of the β -casein milk protein that differ by one amino acid (at amino acid 67, A2 β -casein has a proline, whereas A1 has a histidine). Other genetic variants of β -casein include the A3, B, C, D, E, F, H1, H2, I and G genetic variants. The sequence, structure and physical/chemical properties of β -casein derived from various species is highly variable. Exemplary sequences for bovine β -casein can be found at Uniprot Accession No. P02666 and GenBank Accession No. M15132.1. The terms “ β -casein”, “beta-casein” and “B-casein” (and similar terms) are used interchangeably herein.

[115] As used herein, the term “ κ -casein” refers to not only the κ -casein protein, but also fragments or variants thereof. κ -casein is cleaved by rennet, which releases a macropeptide from the C-terminal region. The remaining product with the N-terminus and approximately two-thirds

of the original peptide chain is referred to as para- κ -casein. The sequence, structure and physical/chemical properties of κ -casein derived from various species is highly variable. Illustrative sequences for bovine κ -casein can be found at Uniprot Accession No. P02668 and GenBank Accession No. CAA25231. The terms “ κ -casein”, “k-casein” and “kappa-casein” (and similar terms) are used interchangeably herein.

[116] In some embodiments, the milk protein is a casein protein, for example, α -S1 casein, α -S2 casein, β -casein, and or κ -casein. In some embodiments, the milk protein is κ -casein and comprises the sequence of SEQ ID NO: 4, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the milk protein is para- κ -casein and comprises the sequence of SEQ ID NO: 2, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the milk protein is β -casein and comprises the sequence of SEQ ID NO: 6, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the milk protein is α -S1 casein and comprises the sequence SEQ ID NO: 8, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, milk protein is α -S2 casein and comprises the sequence SEQ ID NO: 84, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto.

[117] In some embodiments, the milk protein comprises a sequence that is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 4. In some embodiments, the milk protein comprises a sequence that is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 2. In some embodiments, the milk protein comprises a sequence that is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 6. In some embodiments, the milk protein comprises a sequence that is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at

least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 8. In some embodiments, the milk protein comprises a sequence that is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 84.

[118] In some embodiments, α -S1 casein is encoded by the sequence of SEQ ID NO: 7, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, α -S2 casein is encoded by the sequence of SEQ ID NO: 83, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, β -casein is encoded by the sequence of SEQ ID NO: 5, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, κ -casein is encoded by the sequence of SEQ ID NO: 3, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, para- κ -casein is encoded by the sequence of SEQ ID NO: 1, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto.

[119] In some embodiments, the milk protein is encoded by a sequence that is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 7. In some embodiments, the milk protein is encoded by a sequence that is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 83. In some embodiments, the milk protein is encoded by a sequence that is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 3. In some embodiments, the milk protein is encoded by a sequence that is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least

93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 1. In some embodiments, the milk protein is encoded by a sequence that is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 5.

[120] In some embodiments, the milk protein is a casein protein, and comprises a sequence that is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NO: 85-133, or 148-563. In some embodiments, the milk protein is a casein protein and comprises the sequence of any one of SEQ ID NO: 85-133 or 148-563.

[121] In some embodiments, the milk protein comprises a sequence that is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NO: 85-98 or 148-340. In some embodiments, the milk protein comprises the sequence of any one of SEQ ID NO: 85-98 or 148-340.

[122] In some embodiments, the milk protein comprises a sequence that is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NO: 99-109 or 341-440. In some embodiments, the milk protein comprises the sequence of any one of SEQ ID NO: 99-109 or 341-440.

[123] In some embodiments, the milk protein comprises a sequence that is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NO: 110-120 or 441-494. In some embodiments, the milk protein comprises the sequence of any one of SEQ ID NO: 110-120 or 441-494.

[124] In some embodiments, the milk protein comprises a sequence that is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NO: 121-133 or 495-563. In some embodiments, the milk protein comprises the sequence of any one of SEQ ID NO: 121-133 or 495-563 or 495-563.

[125] In some embodiments, the milk protein is a structured protein. Examples of structured milk proteins include, for example, β -lactoglobulin, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, serum albumin, or an immunoglobulin.

[126] In some embodiments, the milk protein is β -lactoglobulin and comprises the sequence of SEQ ID NO: 10, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the milk protein is β -lactoglobulin and is encoded by the sequence of any one of SEQ ID NO: 9, 11, 12, or 13, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NO: 9, 11, 12, or 13. In some embodiments, the milk protein comprises a sequence that is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NO: 9-13 or 564-614. In some embodiments, the milk protein comprises the sequence of any one of SEQ ID NO: 10 or 564-614.

Fusion Partners

[127] The fusion proteins described herein comprise a first protein and a second protein, wherein at least one of the first protein and the second protein is a milk protein. Accordingly, in addition to the milk protein, the fusion proteins described herein comprise a “fusion partner” (i.e., the second protein) – a protein that is fused the milk protein in a fusion protein.

[128] In some embodiments, fusion partner is a protein with a molecular weight of about 5 to about 100 kDa. For example, the fusion partner may have a molecular weight of at least 5 kDa, at least 10 kDa, at least 15 kDa, about 20 kDa, about 25 kDa, about 30 kDa, about 35 kDa, about 40 kDa, about 45 kDa, about 50 kDa, about 55 kDa, about 60 kDa, about 65 kDa, about 70 kDa, about

75 kDa, about 80 kDa, about 85 kDa, about 90 kDa, about 95 kDa, or about 100 kDa. In some embodiments, the fusion partner is a protein with a molecular weight of about 15 kDa, or more.

[129] In some embodiments, fusion partner is a protein with about 10% to about 90% hydrophobic amino acids, e.g., about 10% to about 20%, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, or about 80% to about 90%. In some embodiments, the fusion partner may comprise at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, or at least 90% hydrophobic amino acids. In some embodiments, the fusion partner is a protein with about 25% or more hydrophobic amino acids. In some embodiments, the fusion partner is a protein with about 30% or more hydrophobic amino acids. In some embodiments, the fusion partner is a protein with about 35% or more hydrophobic amino acids. In some embodiments, the fusion partner is a protein with about 40% or more hydrophobic amino acids. A hydrophobic amino acid is an amino acid with a hydrophobic side chain, such as alanine (A), valine (V), isoleucine (I), leucine (L), methionine (M), phenylalanine (F), tryptophan (W), tyrosine (Y), or proline.

[130] In some embodiments, the fusion partner is a flexible protein. In general, proteins with fewer disulfide bonds are more flexible. In some embodiments, the fusion partner comprises less than about 5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises less than about 4.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises less than about 4.0 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises less than about 3.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises less than about 3.0 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises less than about 2.0 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises less than about 1.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises less than about 1 disulfide bond per 10 kDa molecular weight. Number of disulfide bonds may be predicted using one or more computer algorithms known to those of skill in the art. For example, the software SnapGene® or

the Prot Pi tool (available on the Internet by placing [https://](https://www.protpi.ch/Calculator) in front of www.protpi.ch/Calculator) may be useful for making such predictions. Notably, as understood by those of skill in the art, the number of cysteines in a protein, on its own, is not necessarily predictive of the number of disulfide bonds in that protein. The secondary and tertiary structure of the protein must also be considered, to determine whether a given cysteine is in appropriate proximity to another cysteine in order to form a bond.

[131] In some embodiments, the fusion partner comprises at least one of the following characteristics: (i) a molecular weight of 15 kDa or higher, (ii) at least 30% hydrophobic amino acids, (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises at least two of the following characteristics: (i) a molecular weight of 15 kDa or higher, (ii) at least 30% hydrophobic amino acids, (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises all three of the following characteristics: (i) a molecular weight of 15 kDa or higher, (ii) at least 30% hydrophobic amino acids, and (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight.

[132] In some embodiments, the fusion partner comprises at least one of the following characteristics: (i) a molecular weight of 10 kDa or higher, (ii) at least 30% hydrophobic amino acids, (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises at least one of the following characteristics: (i) a molecular weight of 11 kDa or higher, (ii) at least 30% hydrophobic amino acids, (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises at least one of the following characteristics: (i) a molecular weight of 12 kDa or higher, (ii) at least 30% hydrophobic amino acids, (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises at least one of the following characteristics: (i) a molecular weight of 13 kDa or higher, (ii) at least 30% hydrophobic amino acids, (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises at least one of the following characteristics: (i) a molecular weight of 14 kDa or higher, (ii) at least 30% hydrophobic amino acids, (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises at least one of the following characteristics: (i) a molecular weight of 15 kDa or higher, (ii) at least 30% hydrophobic

amino acids, (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises at least one of the following characteristics: (i) a molecular weight of 16 kDa or higher, (ii) at least 30% hydrophobic amino acids, (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises at least one of the following characteristics: (i) a molecular weight of 17 kDa or higher, (ii) at least 30% hydrophobic amino acids, (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises at least one of the following characteristics: (i) a molecular weight of 18 kDa or higher, (ii) at least 30% hydrophobic amino acids, (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises at least one of the following characteristics: (i) a molecular weight of 19 kDa or higher, (ii) at least 30% hydrophobic amino acids, (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises at least one of the following characteristics: (i) a molecular weight of 20 kDa or higher, (ii) at least 30% hydrophobic amino acids, (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises at least one of the following characteristics: (i) a molecular weight of 21 kDa or higher, (ii) at least 30% hydrophobic amino acids, (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises at least one of the following characteristics: (i) a molecular weight of 22 kDa or higher, (ii) at least 30% hydrophobic amino acids, (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises at least one of the following characteristics: (i) a molecular weight of 23 kDa or higher, (ii) at least 30% hydrophobic amino acids, (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises at least one of the following characteristics: (i) a molecular weight of 24 kDa or higher, (ii) at least 30% hydrophobic amino acids, (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises at least one of the following characteristics: (i) a molecular weight of 25 kDa or higher, (ii) at least 30% hydrophobic amino acids, (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight.

[133] In some embodiments, the fusion partner comprises a molecular weight of 15 kDa or higher and at least 30% hydrophobic amino acids. In some embodiments, the fusion partner comprises a

molecular weight of 15 kDa or higher and less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the fusion partner comprises at least 30% hydrophobic amino acids and less than about 2.5 disulfide bonds per 10 kDa molecular weight.

[134] In some embodiments, the fusion partner is kappa-casein. In some embodiments, the fusion partner is beta-casein. In some embodiments, the fusion partner is alpha-casein. In some embodiments, the fusion partner is beta-lactoglobulin. In some embodiments, the fusion partner is green fluorescent protein. In some embodiments the fusion partner is lysozyme. In some embodiments, fusion partner is 2S globulin. In some embodiments, the fusion partner is oleosin A. In some embodiments, the fusion partner is oleosin B. In some embodiments, the fusion partner is the Kunitz-Trypsin inhibitor. In some embodiments the fusion partner is the Bowman-Birk inhibitor. In some embodiments, the fusion partner is Hydrophobin II.

Non-Milk Proteins

[135] In some embodiments, the fusion partner is a non-milk protein. Accordingly, in some embodiments, the fusion proteins described herein may comprise one or more non-milk proteins, including any fragment or variant thereof. As used herein, the term “non-milk protein” refers to any protein that is not typically present in any mammalian milk composition. In some embodiments, the fusion proteins described herein may comprise a first protein and a second protein, wherein the first protein is a milk protein and the second protein (i.e., the fusion partner) is a non-milk protein. The non-milk protein may be, for example, an animal protein or a plant protein. In some embodiments, the animal protein is a mammalian protein. In some embodiments, the animal protein is an avian protein. The non-milk proteins described herein may be classified as structured or unstructured. In some embodiments, the non-milk protein is a structured protein. In some embodiment, the non-milk protein is an unstructured protein.

[136] Whether a protein is structured may be determined using a variety of biophysical and biochemical methods known in the art, such as small angle X-ray scattering, Raman optical activity, circular dichroism, and protease sensitivity. In some embodiments, a protein is considered to be structured if it has been crystallized or if it may be crystallized using standard techniques.

[137] In some embodiments, the non-milk protein is a protein that is typically used as a marker. As used herein, the term “marker” refers to a protein that produces a visual or other signal and is

used to detect successful delivery of a vector (e.g., a DNA sequence) into a cell. Proteins typically used as a marker may include, for example, fluorescent proteins (e.g., green fluorescent protein (GFP)). Other examples include yellow fluorescent protein (YFP), orange fluorescent protein, blue fluorescent protein (BFP), cyan fluorescent protein (CFP), or red fluorescent protein (RFP). Non-limiting examples of proteins within these color classes are shown below in Table 2 (*See also*, Schaner, N. et al., A guide to choosing fluorescent proteins, 2005, *Nature*, 2:12, 905-909).

Table 2: Examples of fluorescent proteins

Color class	Protein
Far-red	mPlum
Red	mCherry
	tdTomato
	mStrawberry
	J-Red
	DsRed-monomer
Orange	mOrange
	mKO
Yellow-green	mCitrine
	Venus
	YPet
	EYFP
Green	Emerald
	EGFP
	GFP
Cyan	CyPet
	mCFPm
	Cerulean
UV-excitable green	T-Sapphire

[138] Other examples of marker proteins include, but are not limited to, bacterial or other enzymes (e.g., β -glucuronidase (GUS), β -galactosidase, luciferase, chloramphenicol acetyltransferase).

[139] Additional non-limiting examples of non-milk proteins that may be used in the fusion proteins described herein are provided in Table 3. In some embodiments, a fragment or variant of any one of the proteins listed in Table 3 may be used.

Table 3: Non-milk proteins for use as fusion partners

Categories	Protein or Protein family	Native Species	Exemplary Uniprot Accession No.
Mammalian	Collagen family	Human (<i>Homo sapiens</i>)	Q02388, P02452, P08123, P02458
	Hemoglobin	Bovine (<i>Bos taurus</i>)	P02070
Avian proteins	Ovalbumin	Chicken (<i>Gallus gallus</i>)	P01012
	Ovotransferrin	Chicken (<i>Gallus gallus</i>)	P02789
	Ovoglobulin	Chicken (<i>Gallus gallus</i>)	I0J170
	Lysozyme	Chicken (<i>Gallus gallus</i>)	P00698
Plant Proteins	Oleosins	Soybean (<i>Glycine max</i>)	P29530, P29531
	Leghemoglobin	Soybean (<i>Glycine max</i>)	Q41219
	Extensin-like protein family	Soybean (<i>Glycine soja</i>)	A0A445JU93
	Prolamin	Rice (<i>Oryza sativa</i>)	Q0DJ45
	Glutenin	Wheat (<i>Sorghum bicolor</i>)	P10388
	Gamma-kafirin	Wheat (<i>Sorghum bicolor</i>)	Q41506

Categories	Protein or Protein family	Native Species	Exemplary Uniprot Accession No.
	preprotein		
	Alpha globulin	Rice (<i>Oryza sativa</i>)	P29835
	Basic 7S globulin precursor	Soybean (<i>Glycine max</i>)	P13917
	2S albumin	Soybean (<i>Glycine max</i>)	P19594
	Beta-conglycinins	Soybean (<i>Glycine max</i>)	P0DO16, P0DO15, P0DO15
	Glycinins	Soybean (<i>Glycine max</i>)	P04347, P04776, P04405
	Canein	Sugar cane (<i>Saccharum officinarum</i>)	ABP64791.1
	Zein	Corn (<i>Zea Mays</i>)	ABP64791.1
	Patatin	Tomato (<i>Solanum lycopersicum</i>)	P07745
	Kunitz-Trypsin inhibitor	Soybean (<i>Glycine max</i>)	Q39898
	Bowman-Birk inhibitor	Soybean (<i>Glycine max</i>)	I1MQD2
	Cystatine	Tomato (<i>Solanum lycopersicum</i>)	Q9SE07
Fungal proteins	Hydrophobin I	Fungus (<i>Trichoderma reesei</i>)	P52754
	Hydrophobin II	Fungus (<i>Trichoderma reesei</i>)	P79073

[140] In some embodiments, the non-milk protein may be an animal protein. For example, in some embodiments, the non-milk protein may be a mammalian protein. The mammalian protein may be, for example, hemoglobin or collagen. In some embodiments, the non-milk protein is an avian protein, such as ovalbumin, ovotransferrin, lysozyme or ovoglobulin.

[141] In some embodiments, the non-milk protein is a plant protein. In some embodiments, the non-milk protein is a protein that is typically expressed in a seed. In some embodiments, the plant protein is a protein that is not typically expressed in a seed. In some embodiments, the plant protein is a storage protein, e.g., a protein that acts as a storage reserve for nitrogen, carbon, and/or sulfur. In some embodiments, the plant protein may inhibit one or more proteases. In some embodiments, the non-milk protein is a plant protein selected from: oleosins, leghemoglobin, extension-like protein family, prolamin, glutenin, gamma-kafirin preprotein, α -globulin, basic 7S globulin precursor, 2S albumin, β -conglycinins, glycinins, canelin, zein, patatin, kunitz-trypsin inhibitor, Bowman-Birk inhibitor, and cystatine. Illustrative plant proteins that may be used to inhibit one or more proteases are shown below in Table 4. In some embodiments, the non-milk protein comprises the sequence of any one of SEQ ID NO: 840, 842, 844, 846, 848 or 850. In some embodiments, the non-milk protein comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to any one of SEQ ID NO: 840, 842, 844, 846, 848 or 850. In some embodiments, the non-milk protein comprises a sequence having the sequence of any one of SEQ ID NO: 840, 842, 844, 846, 848 or 850 plus at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, or more amino acid substitutions.

Table 4: Proteins capable of inhibiting plant proteases

Protein Name	Short Name	Accession No. (Uniprot)	DNA Sequence	Protein Sequence
Bowman-Birk serine protease inhibitor D-II	GmBBID-II	Glyma16g33400	839	840

Bowman-Birk serine protease inhibitor A1	GMBBI-A1	Glyma14g26410	841	842
Kunitz-type trypsin inhibitor gene 1	GmKTi1	Glyma01g10900	843	844
Kunitz-type trypsin inhibitor gene 2	GmKTi2	AAB23483	845	846
Kunitz-type trypsin inhibitor gene 3	GmKTi3	Glyma08g45531	847	848
Cystatine proteinase inhibitor (Cystatin)	SICYS8		849	850

[142] In some embodiments, the structured protein is a fungal protein. For example, the fungal protein may be selected from hydrophobin I and hydrophobin II.

Fusion Proteins

[143] Described herein are fusion proteins comprising at least first protein and a second protein. In some embodiments, at least one of the first protein and the second protein is a milk protein. In some embodiments, a fusion protein comprises at least two proteins, such as three, four, five, six, seven, eight, nine, or ten proteins, or more. In some embodiments, the proteins in the fusion proteins are linked via a linker. In some embodiments, the fusion proteins comprise one or more protease cleavage sites, such as one or more chymosin cleavage sites. Various illustrative embodiments of the fusion proteins of the disclosure are described in further detail below.

Fusion protein comprising a milk protein and a non-milk protein

[144] In some embodiments, a fusion protein comprises at least first protein and a second protein, wherein at least one of the first protein and the second protein is a milk protein, and at least one of the first protein and the second protein is a non-milk protein. In some embodiments, a fusion

protein comprises at least two proteins, such as three, four, five, six, seven, eight, nine, or ten proteins, or more.

[145] In some embodiments, the first protein is a milk protein and the second protein is a non-milk protein. In some embodiments, the non-milk protein is an avian protein. For example, the non-milk protein may be an avian protein selected from: ovalbumin, ovotransferrin, and ovoglobulin. In some embodiments, the non-milk protein is a protein capable of inhibiting one or more proteases, such as the proteins shown above in Table 4, or variants thereof.

[146] In some embodiments, the fusion protein comprises α -S1 casein, or fragment thereof; and ovalbumin. In some embodiments, the fusion protein comprises α -S2 casein, or fragment thereof; and ovalbumin. In some embodiments, the fusion protein comprises β -casein, or fragment thereof; and ovalbumin. In some embodiments, the fusion protein comprises κ -casein, or fragment thereof; and ovalbumin. In some embodiments, the recombinant fusion protein comprises para- κ -casein, or fragment thereof; and ovalbumin.

[147] In some embodiments, the fusion protein comprises α -S1 casein, or fragment thereof; and ovotransferrin. In some embodiments, the fusion protein comprises α -S2 casein, or fragment thereof; and ovotransferrin. In some embodiments, the fusion protein comprises β -casein, or fragment thereof; and ovotransferrin. In some embodiments, the fusion protein comprises κ -casein, or fragment thereof; and ovotransferrin. In some embodiments, the fusion protein comprises para- κ -casein, or fragment thereof; and ovotransferrin.

[148] In some embodiments, the fusion protein comprises α -S1 casein, or fragment thereof; and ovoglobulin. In some embodiments, the fusion protein comprises α -S2 casein, or fragment thereof; and ovoglobulin. In some embodiments, the fusion protein comprises β -casein, or fragment thereof; and ovoglobulin. In some embodiments, the fusion protein comprises κ -casein, or fragment thereof; and ovoglobulin. In some embodiments, the fusion protein comprises para- κ -casein, or fragment thereof; and ovoglobulin.

[149] In some embodiments, the fusion protein comprises a non-milk protein that functions as a marker, such as green fluorescent protein (GFP). In some embodiments, the fusion protein comprises α -S1-casein, or fragment thereof; and GFP. In some embodiments, the fusion protein comprises α -S2-casein, or fragment thereof; and GFP. In some embodiments, the fusion protein comprises β -casein, or fragment thereof; and GFP. In some embodiments, the fusion protein

comprises κ -casein, or fragment thereof; and GFP. In some embodiments, the fusion protein comprises para- κ -casein, or fragment thereof; and GFP.

[150] In some embodiments, the fusion protein comprises a non-milk protein that is a plant protein. In some embodiments, the fusion protein comprises α -S1 casein, or fragment thereof; and a plant protein selected from the group consisting of hydrophobin I, hydrophobin II, oleosins, leghemoglobin, extension-like protein family, prolamin, glutenin, gamma-kafirin preprotein, α -globulin, basic 7S globulin precursor, 2S albumin, β -conglycinins, glycinins, canein, zein, patatin, kunitz-trypsin inhibitor, Bowman-Birk inhibitor, and cystatine.

[151] In some embodiments, the fusion protein comprises α -S2-casein, or fragment thereof; and a plant protein selected from the group consisting of hydrophobin I, hydrophobin II, oleosins, leghemoglobin, extension-like protein family, prolamin, glutenin, gamma-kafirin preprotein, α -globulin, basic 7S globulin precursor, 2S albumin, β -conglycinins, glycinins, canein, zein, patatin, kunitz-trypsin inhibitor, Bowman-Birk inhibitor, and cystatine.

[152] In some embodiments, the fusion protein comprises β -casein, or fragment thereof; and a plant protein selected from the group consisting of hydrophobin I, hydrophobin II, oleosins, leghemoglobin, extension-like protein family, prolamin, glutenin, gamma-kafirin preprotein, α -globulin, basic 7S globulin precursor, 2S albumin, β -conglycinins, glycinins, canein, zein, patatin, kunitz-trypsin inhibitor, Bowman-Birk inhibitor, and cystatine.

[153] In some embodiments, the fusion protein comprises κ -casein, or fragment thereof; and a plant protein selected from the group consisting of hydrophobin I, hydrophobin II, oleosins, leghemoglobin, extension-like protein family, prolamin, glutenin, gamma-kafirin preprotein, α -globulin, basic 7S globulin precursor, 2S albumin, β -conglycinins, glycinins, canein, zein, patatin, kunitz-trypsin inhibitor, Bowman-Birk inhibitor, and cystatine.

[154] In some embodiments, the fusion protein comprises para- κ -casein, or fragment thereof; and a plant protein selected from the group consisting of hydrophobin I, hydrophobin II, oleosins, leghemoglobin, extension-like protein family, prolamin, glutenin, gamma-kafirin preprotein, α -globulin, basic 7S globulin precursor, 2S albumin, β -conglycinins, glycinins, canein, zein, patatin, kunitz-trypsin inhibitor, Bowman-Birk inhibitor, and cystatine.

Fusion proteins comprising a milk protein and an animal (e.g., mammalian) protein

[155] In some embodiments, the fusion proteins described herein comprise (i) a milk protein (which may be unstructured or structured), and (ii) an animal protein. In some embodiments, the fusion proteins described herein comprise (i) an unstructured milk protein, and (ii) a mammalian protein. In some embodiments, the fusion proteins described herein comprise (i) an unstructured milk protein, and (ii) an avian protein. In some embodiments, the fusion proteins described herein comprise (i) an unstructured milk protein, and (ii) a fungal protein.

[156] In some embodiments, the fusion proteins comprise a milk protein, such as a casein protein. In some embodiments, the fusion protein comprises a milk protein selected from α -S1 casein, α -S2 casein, β -casein, and κ -casein. In some embodiments, the fusion protein comprises a milk protein isolated or derived from cow (*bos taurus*), goat (*capra hircus*), sheep (*ovis aries*), water buffalo (*bubalus bubalis*), dromedary camel (*camelus dromedaries*), bactrian camel (*camelus bactrianus*), wild yak (*bos mutus*), horse (*equus caballus*), donkey (*equus asinus*), reindeer (*rangifer tarandus*), eurasian elk (*alces alces*), alpaca (*vicugna pacos*), zebu (*bos indicus*), llama (*lama glama*), or human (*homo sapiens*). In some embodiments, the fusion protein comprises a casein protein (e.g., α -S1 casein, α -S2 casein, β -casein, para- κ -casein or κ -casein) from cow (*bos taurus*), goat (*capra hircus*), sheep (*ovis aries*), water buffalo (*bubalus bubalis*), dromedary camel (*camelus dromedaries*), bactrian camel (*camelus bactrianus*), wild yak (*bos mutus*), horse (*equus caballus*), donkey (*equus asinus*), reindeer (*rangifer tarandus*), eurasian elk (*alces alces*), alpaca (*vicugna pacos*), zebu (*bos indicus*), llama (*lama glama*), or human (*homo sapiens*).

[157] In some embodiments, the fusion protein comprises a milk protein found in Table 34. In some embodiments, the fusion protein comprises a milk protein that is a variant of a protein found in Table 34. In some embodiments, the fusion protein comprises a casein protein as found in Table 34 and/or a variant thereof. In some embodiments, the fusion protein comprises a beta-lactoglobulin as found in Table 34 and/or a variant thereof. One of skill in the art would be able to utilize the numerous milk proteins taught in Table 34, along with their associated SEQ ID NO and/or accession number and find such other milk proteins as encompassed by the disclosure.

[158] In some embodiments, the fusion protein comprises a milk protein that shares at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or at least 100% sequence identity to a protein in Table 34 and/or a variant

thereof. In some embodiments, the fusion protein comprises a milk protein that shares at least from about 70% to about 100% sequence identity to a protein in Table 34 and/or a variant thereof. In some embodiments, the fusion protein comprises a milk protein that shares at least from about 80% to about 100% sequence identity to a protein in Table 34 and/or a variant thereof. In some embodiments, the fusion protein comprises a milk protein that shares at least from about 90% to about 100% sequence identity to a protein in Table 34 and/or a variant thereof. In some embodiments, the fusion protein comprises a milk protein that shares at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100% sequence identity with any one of SEQ ID NO: 148-614. In some embodiments, the fusion protein comprises a milk protein that comprises a sequence of any one of SEQ ID NO: 148-614.

[159] In some embodiments, the fusion protein is α -S1 casein. In some embodiments, the α -S1 casein comprises the sequence SEQ ID NO: 8, or a sequence at least 70%, 80%, 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the α -S1 casein comprises the sequence of any one of SEQ ID NO: 99-109, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto.

[160] In some embodiments, the fusion protein comprises α -S2 casein. In some embodiments, the α -S2 casein comprises the sequence SEQ ID NO: 84, or a sequence at least 70%, 80%, 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the α -S2 casein comprises the sequence of any one of SEQ ID NO: 110-120, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto.

[161] In some embodiments, the fusion protein comprises β -casein. In some embodiments, the β -casein comprises the sequence of SEQ ID NO: 6, or a sequence at least 70%, 80%, 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the β -casein comprises the sequence of any one of SEQ ID NO: 121-133, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto.

[162] In some embodiments, the fusion protein comprises κ -casein. In some embodiments, the κ -casein comprises the sequence of SEQ ID NO: 4, or a sequence at least 70%, 80%, 90%, at least

95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the κ -casein comprises the sequence of any one of SEQ ID NO: 85-98, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto.

[163] In some embodiments, the fusion protein comprises para- κ -casein. In some embodiments, the para- κ -casein comprises the sequence of SEQ ID NO: 2, or a sequence at least 70%, 80%, 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto.

[164] In some embodiments, the fusion protein comprises β -lactoglobulin, α -lactalbumin, albumin, lysozyme, lactoferrin, lactoperoxidase, or an immunoglobulin (e.g., IgA, IgG, IgM, or IgE).

[165] In some embodiments, the fusion protein comprises β -lactoglobulin. In some embodiments, the β -lactoglobulin comprises the sequence of SEQ ID NO: 10, or a sequence at least 70%, 80%, 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto.

[166] In some embodiments, the fusion protein comprises a mammalian protein selected from hemoglobin and collagen. In some embodiments, the fusion protein comprises an avian protein selected from ovalbumin, ovotransferrin, lysozyme and ovoglobulin.

[167] In some embodiments, a fusion protein comprises a casein protein (e.g., κ -casein, para- κ -casein, β -casein, or α -S1 casein) and β -lactoglobulin. In some embodiments, a fusion protein comprises κ -casein and β -lactoglobulin (see, e.g., **FIG. 4**, **FIG. 9**, **FIG. 12A-12B**). In some embodiments, a fusion protein comprises para- κ -casein and β -lactoglobulin (see, e.g., **FIG. 7**, **FIG. 8**, **FIG. 12A-12B**). In some embodiments, a fusion protein comprises β -casein and β -lactoglobulin. In some embodiments, a fusion protein comprises α -S1 casein and β -lactoglobulin.

[168] In some embodiments, a plant-expressed recombinant fusion protein comprises κ -casein, or fragment thereof; and β -lactoglobulin, or fragment thereof. In some embodiments, the fusion protein comprises, in order from N-terminus to C-terminus, the κ -casein and the β -lactoglobulin.

[169] In some embodiments, a plant-expressed recombinant fusion protein comprises β -casein, or fragment thereof; and β -lactoglobulin, or fragment thereof. In some embodiments, the fusion protein comprises, in order from N-terminus to C-terminus, the β -casein and the β -lactoglobulin.

Fusion proteins comprising a milk protein and a plant protein

[170] In some embodiments, the fusion proteins described herein comprise (i) a milk protein (which may be unstructured or structured), and (ii) a plant protein. In some embodiments, the milk protein is a casein protein, such as α -S1 casein, α -S2 casein, β -casein, or κ -casein. In some embodiments, the milk protein is β -lactoglobulin, α -lactalbumin, albumin, lysozyme, lactoferrin, lactoperoxidase, or an immunoglobulin (e.g., IgA, IgG, IgM, or IgE). In some embodiments, the plant protein is selected from the group consisting of: hydrophobin I, hydrophobin II, oleosins, leghemoglobin, extension-like protein family, prolamin, glutenin, gamma-kafirin preprotein, α -globulin, basic 7S globulin precursor, 2S albumin, β -conglycinins, glycinins, canelin, zein, patatin, kunitz-trypsin inhibitor, Bowman-Birk inhibitor, and cystatine. In some embodiments, the plant protein is a protein that is capable of forming a protein body (PB), such as a prolamin. In some embodiments, the protein that is capable of forming a protein body comprises one or more repeat sequences, such as a repeat sequence selected from PPPPVHL (SEQ ID NO: 828); PPPPVXS, wherein X=S, Y, Q, or F (SEQ ID NO: 829); PPPV (SEQ ID NO: 830); PPVHX, wherein X=S or F (SEQ ID NO: 831); PPPVHS (SEQ ID NO: 832); PPPPVXS, wherein X=Y, H, or F (SEQ ID NO: 833); PPPVXL, wherein X=H, or D (SEQ ID NO: 834); PPPVHL (SEQ ID NO: 835); PPPPPVYS (SEQ ID NO: 836); PPPPVHS (SEQ ID NO: 837); and PPPVHL (SEQ ID NO: 838). In some embodiments, the repeat sequence repeats at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9 or at least 10 times.

Fusion protein comprising a milk protein and prolamin

[171] In some embodiments, the fusion protein comprises a prolamin protein, or a fragment or derivative thereof. Prolamins are a group of plant storage proteins having a high proline and glutamine amino acid content, and have poor solubility in water. They are found in plants, mainly in the seeds of cereal grains such as wheat (e.g., the gliadin class of proteins), barley (e.g., the hordein class of proteins), rye (e.g., the secalin class of proteins), corn (e.g., the zein class of proteins), sorghum (e.g., the kafirin class of proteins), and oats (e.g., the avenin class of proteins).

[172] In some embodiments, a fusion protein comprises a canein, such as a gamma canein. For example, the canein may be a 27 kD gamma canein (gCan27), or a fragment or derivative thereof. gCan27 is a zein-like protein, known to be resident in the endoplasmic reticulum. An illustrative sequence for gCAN27 from sugar cane (*Saccharum officinarum*) can be found at Uniprot Ref. No. ABP64791.1 (SEQ ID NO: 800).

[173] In some embodiments, the fusion protein comprises a canein, wherein the canein has the sequence of SEQ ID NO: 800, or a sequence at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the fusion protein comprises a canein, wherein the canein has the sequence of SEQ ID NO: 800 with 1-5, 5-10, 10-20, 20-30, or 30-50 amino acid substitutions relative thereto. In some embodiments, the fusion protein comprises a canein, wherein the canein has a sequence corresponding to amino acids 42-237 of SEQ ID NO: 800, or a sequence at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the fusion protein comprises a canein, wherein the canein has a sequence corresponding to amino acids 42-237 of SEQ ID NO: 800 with 1-5, 5-10, 10-20, 20-30, or 30-50 amino acid substitutions relative thereto. In some embodiments, the fusion protein comprises a canein, wherein the canein has the sequence of SEQ ID NO: 805, or a sequence at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the fusion protein comprises a canein, wherein the canein has the sequence of SEQ ID NO: 805 with 1-5, 5-10, 10-20, 20-30, or 30-50 amino acid substitutions relative thereto. In some embodiments, the canein is encoded by the DNA sequence of SEQ ID NO: 804.

[174] In some embodiments, the fusion protein comprises a milk protein and canein, or a fragment thereof. In some embodiments, the fusion protein comprises a casein protein and canein, or a fragment thereof. In some embodiments, the fusion protein comprises α -S1 casein and canein. In some embodiments, the fusion protein comprises α -S2-casein and canein. In some embodiments, the fusion protein comprises β -casein and canein. In some embodiments, the fusion protein comprises κ -casein and canein. In some embodiments, the fusion protein comprises para- κ -casein and canein. In some embodiments, the fusion protein comprises β -lactoglobulin and

canein. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 803, or a sequence at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 803, or a sequence with 1-5, 5-10, 10-20, 20-30, or 30-50 amino acid substitutions relative thereto. In some embodiments, the fusion protein is encoded by the DNA sequence of SEQ ID NO: 802.

[175] In some embodiments, the fusion protein comprises a zein, such as gamma zein (γ Zein or glutenin 2). Zein is a storage protein of the prolamin class. It is found in the seeds of cereal plants and is able to accumulate within the endoplasmic reticulum (ER). In maize, for example, there are four classes of zeins (α , β , δ , γ). During endosperm development, γ - and β -zeins are synthesized first, forming a polymer termed protein bodies (PBs) where α - and δ -zein will later accumulate (Mainieri et al, 2018). Proteins in the ER lumen usually have a tetrapeptide at the C terminus (KDEL or variations), which is necessary and sufficient for ER localization; however, zeins do not have this signal. The interactions that retain zeins in the ER are not well understood, but γ -zein is able to form ER-located PBs when expressed in storage (Coleman et al., 1996) or vegetative (Geli et al., 1994, Torrent et al., 2009, Marques et al 2020) tissues of transgenic plants in the absence of its partner zein subunits, indicating that no tissue-specific helper factors are required.

[176] The γ -zein sequence (including the 27kDa form of the protein) contains a signal peptide for translocation to the ER (co-translationally removed) followed by a region containing eight repeats of the hexapeptide PPPVHL (SEQ ID NO: 812), the prox domain and seven Cys residues involved in inter-chain bonds that make the protein insoluble in non-reducing conditions, and finally a second region (C-term) homologous to 2S albumins, which are vacuolar storage proteins present in various amounts in all land plants.

[177] An illustrative sequence for γ -zein from corn (*zea mays*) can be found at Uniprot Ref. No. P04706 (SEQ ID NO: 801). In some embodiments, the fusion protein comprises γ -zein, wherein the γ -zein has the sequence of SEQ ID NO: 801, or a sequence at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the fusion protein comprises a γ -zein, wherein the γ -zein has the sequence of SEQ ID NO: 801 with 1-5, 5-10, 10-20, 20-30, or 30-50 amino acid substitutions relative thereto. In some embodiments, the fusion protein

comprises γ -zein, wherein the γ -zein has a sequence corresponding to amino acids 17-112 of SEQ ID NO: 801, or a sequence at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the fusion protein comprises γ -zein, wherein the γ -zein has a sequence corresponding to amino acids 17-112 of SEQ ID NO: 801 with 1-5, 5-10, 10-20, 20-30, or 30-50 amino acid substitutions relative thereto. In some embodiments, the fusion protein comprises a γ -zein, wherein the γ -zein has a sequence corresponding to amino acids 20-223 of SEQ ID NO: 801, or a sequence at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the fusion protein comprises a γ -zein, wherein the γ -zein has a sequence corresponding to amino acids 20-223 of SEQ ID NO: 801 with 1-5, 5-10, 10-20, 20-30, or 30-50 amino acid substitutions relative thereto. In some embodiments, the fusion protein comprises a γ -zein, wherein the γ -zein has the sequence of SEQ ID NO: 809, or a sequence at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the fusion protein comprises a γ -zein, wherein the γ -zein has the sequence of SEQ ID NO: 809 with 1-5, 5-10, 10-20, 20-30, or 30-50 amino acid substitutions relative thereto. In some embodiments, the γ -zein is encoded by the DNA sequence of SEQ ID NO: 808. In some embodiments, the fusion protein comprises a γ -zein, wherein the γ -zein has the sequence of SEQ ID NO: 811, or a sequence at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the fusion protein comprises a γ -zein, wherein the γ -zein has the sequence of SEQ ID NO: 811 with 1-5, 5-10, 10-20, 20-30, or 30-50 amino acid substitutions relative thereto. In some embodiments, the γ -zein is encoded by the DNA sequence of SEQ ID NO: 810.

[178] In some embodiments, the fusion protein comprises a milk protein and γ -zein, or a fragment thereof. In some embodiments, the fusion protein comprises a casein protein and γ -zein, or a fragment thereof. In some embodiments, the fusion protein comprises α -S1 casein and γ -zein. In some embodiments, the fusion protein comprises α -S2-casein and γ -zein. In some embodiments, the fusion protein comprises β -casein and γ -zein. In some embodiments, the fusion protein

comprises κ -casein and γ -zein. In some embodiments, the fusion protein comprises para- κ -casein and γ -zein. In some embodiments, the fusion protein comprises β -lactoglobulin and γ -zein. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 807, or a sequence at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 807, or a sequence with 1-5, 5-10, 10-20, 20-30, or 30-50 amino acid substitutions relative thereto. In some embodiments, the fusion protein is encoded by the DNA sequence of SEQ ID NO: 806.

Fusion protein comprising two or more milk proteins

[179] In some embodiments, the fusion proteins described herein comprise at least first protein and a second protein, wherein the first protein and/or second protein is a milk protein. In some embodiments, the first protein and the second protein are milk proteins. In some embodiments, each of the first protein and the second protein are independently selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, para- κ -casein, β -lactoglobulin, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, serum albumin, and immunoglobulins.

[180] In some embodiments, the recombinant fusion protein comprises α -S1 casein, or fragment thereof; and β -lactoglobulin. In some embodiments, the recombinant fusion protein comprises α -S2 casein, or fragment thereof; and β -lactoglobulin. In some embodiments, the recombinant fusion protein comprises β -casein, or fragment thereof; and β -lactoglobulin. In some embodiments, the recombinant fusion protein comprises κ -casein, or fragment thereof; and β -lactoglobulin. In some embodiments, the recombinant fusion protein comprises para- κ -casein, or fragment thereof; and β -lactoglobulin.

[181] In some embodiments, the recombinant fusion protein comprises α -S1 casein, or fragment thereof; and α -lactalbumin. In some embodiments, the recombinant fusion protein comprises α -S2 casein, or fragment thereof; and α -lactalbumin. In some embodiments, the recombinant fusion protein comprises β -casein, or fragment thereof; and α -lactalbumin. In some embodiments, the recombinant fusion protein comprises κ -casein, or fragment thereof; and α -lactalbumin. In some

embodiments, the recombinant fusion protein comprises para- κ -casein, or fragment thereof; and α -lactalbumin.

[182] In some embodiments, the recombinant fusion protein comprises α -S1 casein, or fragment thereof; and lysozyme. In some embodiments, the recombinant fusion protein comprises α -S2 casein, or fragment thereof; and lysozyme. In some embodiments, the recombinant fusion protein comprises β -casein, or fragment thereof; and lysozyme. In some embodiments, the recombinant fusion protein comprises κ -casein, or fragment thereof; and lysozyme. In some embodiments, the recombinant fusion protein comprises para- κ -casein, or fragment thereof; and lysozyme.

[183] In some embodiments, the recombinant fusion protein comprises α -S1 casein, or fragment thereof; and lactoferrin. In some embodiments, the recombinant fusion protein comprises α -S2 casein, or fragment thereof; and lactoferrin. In some embodiments, the recombinant fusion protein comprises β -casein, or fragment thereof; and lactoferrin. In some embodiments, the recombinant fusion protein comprises κ -casein, or fragment thereof; and lactoferrin. In some embodiments, the recombinant fusion protein comprises para- κ -casein, or fragment thereof; and lactoferrin.

[184] In some embodiments, the recombinant fusion protein comprises α -S1 casein, or fragment thereof; and lactoperoxidase. In some embodiments, the recombinant fusion protein comprises α -S2 casein, or fragment thereof; and lactoperoxidase. In some embodiments, the recombinant fusion protein comprises β -casein, or fragment thereof; and lactoperoxidase. In some embodiments, the recombinant fusion protein comprises κ -casein, or fragment thereof; and lactoperoxidase. In some embodiments, the recombinant fusion protein comprises para- κ -casein, or fragment thereof; and lactoperoxidase.

[185] In some embodiments, the recombinant fusion protein comprises α -S1 casein, or fragment thereof; and an immunoglobulin. In some embodiments, the recombinant fusion protein comprises α -S2 casein, or fragment thereof; and an immunoglobulin. In some embodiments, the recombinant fusion protein comprises β -casein, or fragment thereof; and an immunoglobulin. In some embodiments, the recombinant fusion protein comprises κ -casein, or fragment thereof; and an immunoglobulin. In some embodiments, the recombinant fusion protein comprises para- κ -casein, or fragment thereof; and an immunoglobulin.

[186] In some embodiments, the first protein and the second protein are casein proteins. In some embodiments, the fusion protein comprises κ -casein and para- κ -casein. In some embodiments, the fusion protein comprises κ -casein and β -casein. In some embodiments, the fusion protein comprises κ -casein and α -S1-casein. In some embodiments, the fusion protein comprises κ -casein and α -S2-casein. In some embodiments, the fusion protein comprises para- κ -casein and β -casein. In some embodiments, the fusion protein comprises para- κ -casein and α -S1-casein. In some embodiments, the fusion protein comprises para- κ -casein and α -S2-casein. In some embodiments, the fusion protein comprises β -casein and α -S1-casein. In some embodiments, the fusion protein comprises β -casein and α -S2-casein. In some embodiments, the fusion protein comprises α -S1-casein and α -S2-casein.

[187] In some embodiments, the fusion protein comprises two of the same casein proteins. In some embodiments, the fusion protein comprises a first protein and a second protein, wherein each of the first and second proteins are κ -casein. In some embodiments, the fusion protein comprises a first protein and a second protein, wherein each of the first and second proteins are β -casein. In some embodiments, the fusion protein comprises a first protein and a second protein, wherein each of the first and second proteins are para- κ -casein. In some embodiments, the fusion protein comprises a first protein and a second protein, wherein each of the first and second proteins are α -S1-casein. In some embodiments, the fusion protein comprises a first protein and a second protein wherein each of the first and second proteins are α -S2-casein.

[188] In some embodiments, the fusion protein comprises, from N-terminus to C-terminus, a para-kappa-casein and a beta-lactoglobulin. In some embodiments, the fusion protein comprises, from N-terminus to C-terminus, a beta-lactoglobulin and a para-kappa-casein. In some embodiments, the fusion protein comprises, from N-terminus to C-terminus, an alpha-S1-casein and a beta-lactoglobulin. In some embodiments, the fusion protein comprises, from N-terminus to C-terminus, a beta-lactoglobulin and an alpha-S1-casein. In some embodiments, the fusion protein comprises, from N-terminus to C-terminus, a beta-casein and a beta-lactoglobulin. In some embodiments, the fusion protein comprises from N-terminus to C-terminus, a beta-lactoglobulin and a beta-casein.

Fusion proteins comprising a milk protein and a fusion partner

[189] In some embodiments, a fusion protein comprises a milk protein and a fusion partner having one or more desirable characteristics. For example, in some embodiments, a fusion protein comprises a first protein and a second protein, wherein the first protein is a milk protein, and the second protein comprises at least one of the following characteristics: (i) a molecular weight of 15 kDa or higher; (ii) at least 30% hydrophobic amino acids; and/or (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, the second protein comprises at least two of the characteristics (i), (ii) and (iii). In some embodiments, the second protein comprises all three of the characteristics (i), (ii) and (iii).

[190] In some embodiments, a fusion protein comprises a milk protein and a fusion partner, wherein the fusion partner has a molecular weight of 15 kDa or higher. In some embodiments, a fusion protein comprises a milk protein and a fusion partner, wherein the fusion partner has at least 30% hydrophobic amino acids. In some embodiments, a fusion protein comprises a milk protein and a fusion partner, wherein the fusion partner has less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, a fusion protein comprises a milk protein and a fusion partner, wherein the fusion partner has a molecular weight of 15 kDa or higher, and at least 30% hydrophobic amino acids. In some embodiments, a fusion protein comprises a milk protein and a fusion partner, wherein the fusion partner has at least 30% hydrophobic amino acids, and less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, a fusion protein comprises a milk protein and a fusion partner, wherein the fusion partner has a molecular weight of 15 kDa or higher, and less than about 2.5 disulfide bonds per 10 kDa molecular weight. In some embodiments, a fusion protein comprises a milk protein and a fusion partner, wherein the fusion partner has a molecular weight of 15 kDa or higher, at least 30% hydrophobic amino acids, and less than about 2.5 disulfide bonds per 10 kDa molecular weight.

[191] In some embodiments, the fusion protein comprises a protease cleavage site located between the first protein and the second protein. In some embodiments, the protease cleavage site is a chymosin cleavage site. In some embodiments, cleavage of the fusion protein with a protease separates the first protein from the second protein. In some embodiments, after being separated from one another, the first protein and/or the second protein optionally comprise at their N-terminus or C-terminus one or more amino acids that do not occur in the native form of the first protein or the second protein and that are derived from the protease cleavage site.

Fusion proteins comprising more than two proteins

[192] Fusion proteins may also be created that comprise more than two proteins, such as at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, or at least 10, or more proteins. In some embodiments, a fusion protein comprising more than two proteins may comprise at least one milk protein. In some embodiments, a fusion protein comprising more than two proteins may comprise at least one casein protein. In some embodiments, each of the proteins in a fusion protein comprising more than two proteins may be a milk protein. In some embodiments, each of the proteins in a fusion protein comprising more than two proteins may be a casein protein.

[193] In some embodiments, a fusion protein comprising more than two proteins may comprise at least one structured protein and at least one structured protein. In some embodiments, a fusion protein comprising more than two proteins may comprise at least one milk protein (e.g., a casein) and at least one non-milk protein. In some embodiments, a fusion protein comprising more than two proteins may comprise at least one milk protein (e.g., a casein) and at least one plant protein. In some embodiments, a fusion protein comprising more than two proteins may comprise at least one milk protein (e.g., a casein) and at least one animal (e.g., mammalian) protein.

[194] In some embodiments, a fusion protein comprises three proteins, wherein each protein is individually selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, para- κ -casein, β -lactoglobulin, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, serum albumin, and an immunoglobulin. In some embodiments, a fusion protein comprises four proteins, wherein each protein is individually selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, para- κ -casein, β -lactoglobulin, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, serum albumin, and an immunoglobulin. In some embodiments, a fusion protein comprises five proteins, wherein each protein is individually selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, para- κ -casein, β -lactoglobulin, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, serum albumin, and an immunoglobulin. In some embodiments, a fusion protein comprises six proteins, wherein each protein is individually selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, para- κ -casein, β -lactoglobulin, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, serum albumin, and an immunoglobulin. In some embodiments, a fusion protein comprises seven proteins, wherein each protein is individually selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, para- κ -casein,

β -lactoglobulin, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, serum albumin, and an immunoglobulin. In some embodiments, a fusion protein comprises eight proteins, wherein each protein is individually selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, para- κ -casein, β -lactoglobulin, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, serum albumin, and an immunoglobulin. In some embodiments, a fusion protein comprises nine proteins, wherein each protein is individually selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, para- κ -casein, β -lactoglobulin, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, serum albumin, and an immunoglobulin. In some embodiments, a fusion protein comprises ten proteins, wherein each protein is individually selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, para- κ -casein, β -lactoglobulin, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, serum albumin, and an immunoglobulin.

[195] In some embodiments, a fusion protein comprises three proteins, wherein each protein is individually selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, and para- κ -casein. In some embodiments, a fusion protein comprises four proteins, wherein each protein is individually selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, and para- κ -casein. In some embodiments, a fusion protein comprises five proteins, wherein each protein is individually selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, and para- κ -casein. In some embodiments, a fusion protein comprises six proteins, wherein each protein is individually selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, and para- κ -casein. In some embodiments, a fusion protein comprises seven proteins, wherein each protein is individually selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, and para- κ -casein. In some embodiments, a fusion protein comprises eight proteins, wherein each protein is individually selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, and para- κ -casein. In some embodiments, a fusion protein comprises nine proteins, wherein each protein is individually selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, and para- κ -casein. In some embodiments, a fusion protein comprises ten proteins, wherein each protein is individually selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, and para- κ -casein.

[196] In some embodiments, a fusion protein comprises between 3 and 10 proteins, wherein each protein is different. In some embodiments, a fusion protein comprises between 3 and 10 proteins, wherein each protein is the same. In some embodiments, a fusion protein comprises between 3 and

10 proteins, wherein each protein is a milk protein. In some embodiments, a fusion protein comprises between 3 and 10 proteins, wherein each protein is a casein.

[197] In some embodiments, a fusion protein comprises a first, a second, and a third protein, wherein the first protein is beta casein, the second protein is kappa casein, and the third protein is beta-lactoglobulin. See, e.g., SEQ ID NO: 652.

[198] In some embodiments, a fusion protein comprises a first, second, a third, and a fourth protein, wherein the first protein is kappa casein, the second protein is beta casein, the third protein is alpha-S1-casein, and the fourth protein is beta-lactoglobulin. In some embodiments, a fusion protein comprises a first, second, and third protein, wherein the first protein is kappa casein, the second protein is beta casein, the third protein is beta-lactoglobulin. In some embodiments, a fusion protein comprises a first, second, and third protein, wherein the first protein is kappa casein, the second protein is alpha-S1-casein, the third protein is beta-lactoglobulin. In some embodiments, a fusion protein comprises a first, second, and third protein, wherein the first protein is beta-casein, the second protein is alpha-S1-casein, the third protein is beta-lactoglobulin. In some embodiments, a fusion protein comprises a first, second, and third protein, wherein the first protein is kappa-casein, the second protein is beta-casein, the third protein is alpha-S1-casein.

[199] In some embodiments, a fusion protein comprising a first, second, third, and fourth protein, wherein the third protein is kappa-casein. In some embodiments, a fusion protein comprising a first, second, third, and fourth protein, wherein the third protein is kappa-casein and the fourth protein is beta-lactoglobulin. In some embodiments, the kappa-casein comprises a chymosin cleavage site. In some embodiments, cleavage of the fusion protein with chymosin produces the following polypeptides: (a) a first polypeptide comprising the first protein, the second protein, and para-kappa-casein; (b) a second polypeptide comprising a kappa-casein macropeptide and the fourth protein.

[200] In some embodiments, a fusion protein comprises a first, second, third, and fourth protein, wherein the first protein is beta-casein, the second protein is beta-casein, the third protein is kappa-casein, and the fourth protein is beta-lactoglobulin. See, e.g., SEQ ID NO: 652.

[201] In some embodiments, a fusion protein comprises a first, second, third, fourth, and fifth protein wherein the first protein is beta-casein, the second protein is beta-casein, the third protein

is beta-casein, the fourth protein is kappa-casein, and the fifth protein is beta-lactoglobulin. See, e.g., SEQ ID NO: 654.

[202] In some embodiments, a fusion protein comprises a first, second, third, fourth, fifth, and sixth protein wherein the first protein is beta-casein, the second protein is beta-casein, the third protein is beta-casein, the fourth protein is beta-casein, the fifth protein is kappa-casein, and the sixth protein is beta-lactoglobulin. See, e.g., SEQ ID NO: 656.

[203] In some embodiments, a fusion protein comprises a first, second, third, fourth, and fifth protein wherein the first protein is beta-casein, the second protein is beta-casein, the third protein is beta-casein, the fourth protein is beta-casein, and the fifth protein is beta-lactoglobulin. See, e.g., SEQ ID NO: 658 and 662.

[204] In some embodiments, a fusion protein comprises a first, second, third, and fourth protein, wherein the first protein is beta-casein, the second protein is beta-casein, the third protein is beta-casein, and the fourth protein is beta-lactoglobulin. See, e.g., SEQ ID NO: 660.

[205] In some embodiments, a fusion protein comprises a first, second, third, and fourth protein, wherein the first protein is beta-casein, the second protein is beta-casein, the third protein is beta-casein, and the fourth protein is beta-casein. See, e.g., SEQ ID NO: 664.

[206] Table 5 lists illustrative fusion proteins contemplated by the instant disclosure. The fusion proteins comprise the listed constituent proteins in order from N-terminus to C-terminus. As will be understood by those of skill in the art, in some embodiments, a fusion protein may comprise the constituent proteins in order from C-terminus to N-terminus. In some embodiments, one or more of the fusion proteins may comprise a protease cleavage site, such as a protease cleavage site located between two of the constituent proteins.

Table 5: Illustrative Fusion Proteins

Fusion Protein No.	First Protein	Second Protein	Third Protein	Fourth Protein	Fifth Protein	Sixth Protein
1	BC	LG				

2	BC	BC	LG			
3	BC	BC	KCN	LG		
4	BC	BC	BC	KCN	LG	
5	BC	BC	BC	BC	BC	
6	BC	aS1	aS1	BC		
7	BC	aS1	aS1	BC	LG	
8	BC	aS1	BC			
9	ZN	BC				
10	ZN27	BC				
11	BC	BC				
12	BC	BC	BC			
13	BC	BC	BC	LG		
14	BC	BC	BC	BC	LG	
15	KCN	BC	BC	BC		
16	KCN	BC	BC			
17	KCN	BC	aS1	LG		
18	BC	BC	aS1	aS1	BC	BC
19	paraKCN	paraKCN	paraKCN	BC	BC	
20	BC	aS1	KCN			
21	aS1	LG				
22	KCN	LG				
23	paraKCN	LG				
24	aS1	aS1	aS1	aS1		
25	KCN	KCN	KCN	KCN		
26	aS1	aS1	aS1	aS1	LG	
27	KCN	KCN	KCN	KCN	LG	
28	paraKCN	paraKCN	paraKCN	paraKCN	LG	
29	paraKCN	paraKCN	paraKCN	paraKCN		
30	KCN	BC	aS1	LG		
31*	KCN	BC	aS1			

32*	KCN	BC				
33*	BC	BC	BC	BC		

BC = beta-casein, LG=beta-lactoglobulin, KCN=kappa-casein; paraKCN=para-kappa-casein, aSI=alpha-S1-casein, ZN=truncated zein, ZN27=full-length zein

* indicates that the vector used to express the listed fusion protein also comprises a sequence encoding a Fam kinase, wherein the Fam kinase is expressed under the control of a different promoter.

Fusion protein structure

[207] The fusion proteins described herein may have various different structures, in order to increase expression and/or accumulation in a plant or other host organism or cell. The designation of “first protein”, “second protein”, “third protein”, and/or “fourth protein” is not intended to imply any order.

[208] In some embodiments, the fusion protein may comprise, from N-terminus to C-terminus, the first protein and the second protein. In some embodiments, the fusion protein may comprise, from N-terminus to C-terminus, the second protein and the first protein. In some embodiments, a fusion protein comprises, in order from N-terminus to C-terminus, a first protein and a second protein, wherein the first protein and/or the second protein is a milk protein. In some embodiments, a fusion protein comprises, in order from N-terminus to C-terminus, a second protein and a first protein, wherein the first protein and/or the second protein is a milk protein. For example, in some embodiments, a fusion protein comprises, in order from N-terminus to C-terminus, κ -casein and β -lactoglobulin. In some embodiments, a fusion protein comprises, in order from N-terminus to C-terminus, β -lactoglobulin and κ -casein. In some embodiments, a fusion protein comprises, in order from N-terminus to C-terminus, para- κ -casein and β -lactoglobulin. In some embodiments, a fusion protein comprises, in order from N-terminus to C-terminus, β -lactoglobulin and para- κ -casein. In some embodiments, a fusion protein comprises, in order from N-terminus to C-terminus, β -casein and β -lactoglobulin. In some embodiments, a fusion protein comprises, in order from N-terminus to C-terminus, β -lactoglobulin and β -casein. In some embodiments, a fusion protein comprises, in order from N-terminus to C-terminus, α -S1 casein and β -lactoglobulin. In some embodiments, a fusion protein comprises, in order from N-terminus to C-terminus, β -lactoglobulin and α -S1 casein.

[209] In some embodiments, a fusion protein comprises, in order from N-terminus to C-terminus, a milk protein and a plant protein. In some embodiments, a fusion protein comprises, in order from N-terminus to C-terminus, a plant protein and a milk protein. In some embodiments, a fusion protein comprises, in order from N-terminus to C-terminus, a casein protein and a plant protein. In some embodiments, a fusion protein comprises, in order from N-terminus to C-terminus, a plant protein and a casein protein.

Cleavable fusion proteins

[210] In some embodiments, it may be desirable to cleave the fusion protein to separate its constituent proteins. For example, it may be desirable to cleave the fusion protein to separate its constituent proteins so that the proteins may individually be used in one or more food compositions.

[211] In some embodiments, a fusion protein comprises a protease cleavage site. For example, in some embodiments, the fusion protein comprises an endoprotease, endopeptidase, and/or endoproteinase cleavage site. In some embodiments, the fusion protein comprises a rennet cleavage site. In some embodiments, the fusion protein comprises a chymosin cleavage site. In some embodiments, the fusion protein comprises a trypsin cleavage site.

[212] The protease cleavage site may be located between the first protein and the second protein. In some embodiments, the protease cleavage site may be located between a milk protein and the non-milk protein. For example, the protease cleavage site may be located between the milk protein and the animal (e.g., mammalian or avian) protein, or between the milk protein and the plant protein, such that cleavage of the protein at the protease cleavage site will separate the two proteins. In some embodiments, the protease cleavage site may be located between a first milk protein and a second milk protein. In some embodiments, the protease cleavage site may be located between a first casein protein and a second casein protein.

[213] In some embodiments, the protease cleavage site may be contained within the sequence of the first protein or the second protein. In some embodiments, the protease cleavage site may be located in either the milk protein or the non-milk protein, for example, the animal (e.g., mammalian or animal) or plant protein. In some embodiments, the protease cleavage site may be added separately, for example, between the two proteins.

[214] In some embodiments, a fusion protein comprises a chymosin cleavage site. In some embodiments, a fusion protein comprises a chymosin cleavage site selected from any one of the sequences shown in Table 6, below. In some embodiments, a fusion protein comprises a chymosin cleavage site that is not shown in Table 6, below. In some embodiments, a fusion protein comprises a chymosin cleavage site having at least 1, at least 2, at least 3 or at least 4 amino acid substitutions relative to any one of the sequences shown in Table 6. In some embodiments, a fusion protein comprises a chymosin cleavage site with a sequence of any one of SEQ ID NO: 665-668, or a sequence having 1, 2, 3, 4, or more amino acid substitutions relative thereto. In the sequences of Table 6, cleavage typically occurs after the underlined residue.

Table 6: Chymosin cleavage sites

Chymosin Cleavage Site	SEQ ID NO:
RHPHPHLS <u>F</u> MAIPPKK	665
HHPHPHLS <u>F</u> MAIPPKK	666
RHPHPHLS <u>F</u> M	667
EDFLQKQQYGISSK <u>E</u> R	668
RHPHPHLS <u>F</u> MAIPPKK	669
HHPHPHLS <u>F</u> MAIPPKK	670
RHPHPRL <u>S</u> FMAIPPKK	671
RRPRPHLS <u>F</u> MAIPPKK	672
HQTFQHASE <u>F</u> IATPPQK	673
RRPNLHPS <u>F</u> IAIPPKK	674
PYAIPNP <u>S</u> ELAMPTNE	675
PHPIPNP <u>S</u> FLAIPTNE	676
RHPCPHPS <u>F</u> IAIPPKK	677
ARRPPHAS <u>F</u> IAIPPKK	678
VGRHSHPF <u>F</u> MAILPNK	679
RRPRPRPS <u>F</u> IAIPPKK	680
RHPRPHPS <u>F</u> IAIPPKX	681
RHPYRRPS <u>F</u> IAIPPKK	682
RHPHLPAS <u>F</u> IVIPPKK	683
CRRRPHPS <u>F</u> LAIPPXK	684
HRPNLHPS <u>F</u> IAIPPKK	685

HRPQLHPSFIAIPPKK	686
HRPHIHPSFIAIPPKK	687
HRPHLHPSFIAIPPKK	688
HRPHLHPSFIAIPAKK	689
HHPHPCPSFLAIPPKK	690
HRPHLHPSFTAIPAKK	691
HHPHPRPSFTAIPPKK	692
HHPHPRPSFLAIPPKK	693
HRPHLHPSFIAIPTKK	694
HHKYLKPSFIVIPPTK	695
RHPRPHPSFIAIPPKK	696
YHQAKHPSFMAILSKK	697
PHTYLKPPFIVIPPKK	698
HRPKLHPSFIAVPPKK	699
RRPHPRLSFMAIPPKK	700
KPAEFFRL	701
KPAEFKRL	702
KPAEFERL	703
KPAEFTRL	704
KPAEFGRL	705
KPAEFARL	706
KPAEFVRL	707
KPAEFLRL	708
KPAEFIRL	709
HPHLSFMAI	710
HPHLSFEAI	711
YGIFLRF	712
YGIFKRF	713
YGAFLRF	714
KYSSWYVAL	715
KYSSWKVAL	716
KYSSWEVAL	717
KYSSWLVAL	718
RPKPQQFFGLM	719
RPKPQQFKGLM	720

AFPLEFKREL	721
AFPLEFKREL	722
AFPLEFEREL	723
AFPLEFEREL	724
AFPLEFIREL	725
AFPLEFFREL	726
KIPYILKRQL	727
KIPYILRRQL	728
KIPYILERQL	729
KIPYILSRQL	730
KIPYILARQL	731
KIPYILIRQL	732
KIPYILFRQL	733
KIPYILFRQL	734
KIPYILWRQL	735
EDFLKQQQYGISSKYSGE	736
EDFLKQQQYGISSKFM	737
EDFLKQQQYGISSKFA	738
EDFLKQQQYGISSKFC	739
EDFLKQQQYGISSKFF	740
EDFLKQQQYGISSKFH	741
EDFLKQQQYGISSKFI	742
EDFLKQQQYGISSKFK	743
EDFLKQQQYGISSKFL	744
EDFLKQQQYGISSKEN	745
EDFLKQQQYGISSKFR	746
EDFLKQQQYGISSKFT	747
EDFLKQQQYGISSKFEV	748
EDFLKQQQYGISSKFW	749
EDFLKQQQYGISSKYSGFV	750
EDFLKQQQYGISSKYSGFV	751
EDFLKQQQYGISSKYSGF	752
EDFLKQQQYGISSKYSGF	753
EDFLKQQQYGISSKYSGF	754
EDFLKQQQYGISSKSSGFV	755

EDFLQKQQY <u>G</u> ISSKSSGFV	756
EDFLQKQQYGI <u>S</u> SKSSGFV	757
EDFLQKQQYGISSK <u>S</u> GFV	758
EDFLQKQQYGISSKFS	759

[215] In some embodiments, a fusion protein comprises a cleavage site recognized by an endoprotease. For example, in some embodiments, a fusion protein comprises a cleavage site selected from any one of the sequences shown in Table 7, below. In some embodiments, a fusion protein comprises a cleavage site having at least 1, at least 2, at least 3 or at least 4 amino acid substitutions relative to any one of the sequences shown in Table 7. In the sequences of Table 7, cleavage typically occurs after the underlined residue.

Table 7: Endoprotease Cleavage Sites

Cleavage Site	SEQ ID NO:	Endoprotease
DDD <u>D</u> K	760	Enterokinase
HPHLS <u>F</u> MAI	761	Pepsin A
HPHLS <u>F</u> EAI	762	Pepsin A
LVPR <u>G</u>	763	Thrombin
ELSLSRL <u>R</u> DSA	764	Thrombin
ELSLSRLR	765	Thrombin
DNYTRLR <u>K</u>	766	Thrombin
YTRLRKQ <u>M</u>	767	Thrombin
APSGR <u>V</u> SM	768	Thrombin
VSMIK <u>N</u> LQ	769	Thrombin
RIRPKLK <u>W</u>	770	Thrombin
AMAPR <u>E</u> RK	771	Thrombin
NFFWK <u>T</u> FT	772	Thrombin
KMYPR <u>G</u> NH	773	Thrombin
QTYPR <u>T</u> NT	774	Thrombin
IQGR	775	Factor Xa
IEGR	776	Factor Xa
ENLYFQ(G/S) (G/S = G or S)	777	TEV protease
EXXYXQ(G/S) (x= any amino acid, G/S = G or S)	778	TEV protease

Cleavage Site	SEQ ID NO:	Endoprotease
VDVAD <u>X</u> (x= any amino acid)	779	Caspase 2
RXXR <u>R</u> (x= any amino acid)	780	Furin
XX(T/A/S/V) <u>X</u> XX (x= any amino acid)	781	Alpha-lytic protease

[216] In some embodiments, a fusion protein comprises a cleavage site that is sensitive to cleavage by one or more chemical agents, such as nickel, formic acid, or hydroxylamine. For example, in some embodiments, a fusion protein comprises a chemical cleavage site selected from any one of the sequences shown in Table 8, below. In the sequences of Table 8, cleavage typically occurs after the underlined residue.

Table 8: Chemical Cleavage Sites

Chemical Cleavage Site	SEQ ID NO:	Chemical agent
GSHHW	782	Nickel
DP	--	Formic Acid
NG	--	Hydroxylamine

[217] In some embodiments, the fusion protein comprises a protease cleavage site that comprises the amino acids residues F and M (phenylalanine and methionine). Without being bound by any theory, it is believed that one or more enzymes (e.g., chymosin) and cleave between the F and the M. When a protease, such as chymosin, is used to cleave a fusion protein comprising an FM cleavage site, the first protein comprises the F at its C terminus and the second protein comprises a M at its N terminus when liberated from the fusion protein. For example, a protein separated from a fusion protein by cleavage of an FM site may comprise the sequence of any one of SEQ ID NO: 782-791. Thus, in some embodiments, a protein derived from (i.e., separated from) a fusion protein may comprise at least one non-native amino acid. In some embodiments, the non-native amino acid is derived from a protease cleavage site.

[218] In some embodiments, a fusion protein comprises a linker between the first protein and the second protein. In some embodiments, the linker is between the milk protein and the animal (e.g., mammalian or avian) protein, or between the milk protein and the plant protein. In some embodiments, the linker is between a first milk protein and a second milk protein. In some

embodiments, the linker is between a first casein protein and a second casein protein. In some embodiments, the linker may comprise a peptide sequence recognizable by an endoprotease. In some embodiments, the linker may comprise a protease cleavage site. In some embodiments, the linker may comprise a self-cleaving peptide, such as a 2A peptide.

[219] In some embodiments, a fusion protein may comprise a signal peptide. The signal peptide may be cleaved from the fusion protein, for example, during processing or transport of the protein within the cell. In some embodiments, the signal peptide is located at the N-terminus of the fusion protein. In some embodiments, the signal peptide is located at the C-terminus of the fusion protein.

[220] In some embodiments, the signal peptide is selected from the group consisting of GmSCB1, StPat21, 2Sss, Sig2, Sig12, Sig8, Sig10, Sig11, and Coixss. In some embodiments, the signal peptide is Sig10 and comprises SEQ ID NO: 15, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the signal peptide is Sig2 and comprises SEQ ID NO: 17, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto.

[221] In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 71. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 73. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 75. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 77. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 79. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 81. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 135. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 137. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 616. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 618. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 620. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 622. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 624. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 626. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 628. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 630. In some

embodiments, the fusion protein comprises the sequence of SEQ ID NO: 632. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 634. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 636. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 638. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 640. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 642. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 644. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 646. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 648. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 650. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 652. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 654. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 656. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 658. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 660. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 662. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 664. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 793. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 795. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 797. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 799.

[222] In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 71, with 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acid substitutions. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 73, with 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acid substitutions. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 75, with 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acid substitutions. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 77, with 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acid substitutions. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 79, with 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acid substitutions. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 81, with 1, 2, 3, 4, 5, 6,

[illegible]

amino acid substitutions. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 650, with 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acid substitutions. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 652, with 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acid substitutions. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 654, with 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acid substitutions. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 656, with 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acid substitutions. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 658, with 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acid substitutions. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 660, with 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acid substitutions. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 662, with 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acid substitutions. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 664, with 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acid substitutions. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 793, with 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acid substitutions. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 795, with 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acid substitutions. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 797, with 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acid substitutions. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 799, with 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acid substitutions.

[223] In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 71, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 73, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 75, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 77, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 79, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least

[illegible]

[illegible]

sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 795, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 797, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the fusion protein comprises the sequence of SEQ ID NO: 799, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto.

[224] In some embodiments, the fusion proteins have a molecular weight in the range of about 1 kDa to about 500 kDa, about 1 kDa to about 250 kDa, about 1 to about 100 kDa, about 10 to about 50 kDa, about 1 to about 10 kDa, about 10 to about 200 kDa, about 30 to about 150 kDa, about 30 kDa to about 50 kDa, or about 20 to about 80 kDa.

Nucleic Acids Encoding Fusion Proteins and Vectors Comprising the Same

[225] Also provided herein are nucleic acids encoding the fusion proteins of the disclosure. In some embodiments, the nucleic acids are DNAs. In some embodiments, the nucleic acids are RNAs.

[226] Also provided herein are examples of expression cassettes for the expression of casein proteins in non-mammalian systems, such as plants and microorganisms, to produce recombinant casein proteins. The expression cassette may comprise, for example, a promoter, a 5' untranslated region (UTR), a sequence encoding one or more casein proteins, and a terminator. The expression cassette may further comprise a selectable marker and retention signal.

[227] In some embodiments, a nucleic acid comprises a sequence encoding a fusion protein. In some embodiments, a nucleic acid comprises a sequence encoding a fusion protein, which is operably linked to a promoter. In some embodiments, a nucleic acid comprises, in order from 5' to 3', a promoter, a 5' untranslated region (UTR), a sequence encoding a fusion protein, and a terminator.

[228] The promoter may be a plant promoter. A "plant promoter" is a promoter capable of initiating transcription in plant cells. Examples of promoters under developmental control include

promoters that preferentially initiate transcription in certain organs, such as leaves, roots, flowers, seeds and tissues such as fibers, xylem vessels, tracheids, or sclerenchyma. Such promoters are referred to as "tissue-preferred." Promoters which initiate transcription only in certain tissue are referred to as "tissue-specific." A "cell-type" specific promoter primarily drives expression in certain cell types in one or more organs, for example, vascular cells in leaves, roots, flowers, or seeds. An "inducible" promoter is a promoter which is under environmental control. Examples of environmental conditions that may affect transcription by inducible promoters include anaerobic conditions or the presence of light. Tissue-specific, tissue-preferred, cell-type specific, and inducible promoters constitute the class of "non-constitutive" promoters. A "constitutive" promoter is a promoter which is active under most environmental conditions.

[229] In some embodiments, the promoter is a plant promoter derived from, for example soybean, lima bean, Arabidopsis, tobacco, rice, maize, barley, sorghum, wheat, pea, and/or oat. In some embodiments, the promoter is a constitutive or an inducible promoter. Exemplary constitutive promoters include, but are not limited to, the promoters from plant viruses such as the 35S promoter from CaMV and the promoters from such genes as rice actin; ubiquitin; pEMU; MAS and maize H3 histone. In some embodiments, the constitutive promoter is the ALS promoter, XbaI/NcoI fragment 5' to the Brassica napus ALS3 structural gene (or a nucleotide sequence similarity to said XbaI/NcoI fragment).

[230] In some embodiments, the promoter is a plant tissue-specific or tissue-preferential promoter. In some embodiments, the promoter is isolated or derived from a soybean gene. Illustrative soybean tissue-specific promoters include AR-Pro1, AR-Pro2, AR-Pro3, AR-Pro4, AR-Pro5, AR-Pro6, AR-Pro7, AR-Pro8, and AR-Pro9.

[231] In some embodiments, the plant is a seed-specific promoter. In some embodiments, the seed-specific promoter is selected from the group consisting of PvPhas, BnNap, AtOle1, GmSeed2, GmSeed3, GmSeed5, GmSeed6, GmSeed7, GmSeed8, GmSeed10, GmSeed11, GmSeed12, pBCON, GmCEP1-L, GmTHIC, GmBg7S1, GmGRD, GmOLEA, GmOLER, Gm2S-1, and GmBBld-II. In some embodiments, the seed-specific promoter is PvPhas and comprises the sequence of SEQ ID NO: 18, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the seed-specific

promoter is GmSeed2 and comprises the sequence of SEQ ID NO: 19, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the promoter is a Cauliflower Mosaic Virus (CaMV) 35S promoter.

[232] In some embodiments, the promoter is a soybean polyubiquitin (Gmubi) promoter, a soybean heat shock protein 90-like (GmHSP90L) promoter, a soybean Ethylene Response Factor (GmERF) promoter. In some embodiments, the promoter is a constitutive soybean promoter derived from GmScreamM1, GmScreamM4, GmScreamM8 genes or GmubiXL genes.

[233] In some embodiments, the 5' UTR is selected from the group consisting of *Arc5'UTR* and *glnBIUTR*. In some embodiments, the 5' untranslated region is *Arc5'UTR* and comprises the sequence of SEQ ID NO: 20, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto.

[234] In some embodiments, the terminator sequence is isolated or derived from a gene encoding Nopaline synthase, *Arc5-1*, an Extensin, Rb7 matrix attachment region, a Heat shock protein, Ubiquitin 10, Ubiquitin 3, and M6 matrix attachment region. In some embodiments, the terminator sequence is isolated or derived from a *Nopaline synthase* gene and comprises the sequence of SEQ ID NO: 22, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto.

[235] In some embodiments, the nucleic acid comprises a first terminator sequence and a second terminator sequence (i.e., a dual terminator). In some embodiments, the dual terminator is EU:Rb7. In some embodiments, the dual terminator is AtHSP:AtUbi10. In some embodiments, the dual terminator is EU:StUbi3. In some embodiments, the dual terminator is EU:TM6.

[236] In some embodiments, the dual terminator is EU:Rb7 and comprises the sequence of SEQ ID NO: 138, or a sequence at least 90% at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto.

[237] In some embodiments, the dual terminator is AtHSP:AtUbi10 and comprises the sequence of SEQ ID NO: 141, or a sequence at least 90% at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto.

[238] In some embodiments, the dual terminator is EU:StUbi3 and comprises the sequence of SEQ ID NO: 144, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto.

[239] In some embodiments, the dual terminator is EU:TM6 and comprises the sequence of SEQ ID NO: 146, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto.

[240] In some embodiments, the nucleic acid comprises a 3' UTR. For example, the 3' untranslated region may be *Arc5-1* and comprise SEQ ID NO: 21, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto.

[241] In some embodiments the nucleic acid comprises a gene encoding a selectable marker. One illustrative selectable marker gene for plant transformation is the neomycin phosphotransferase II (nptII) gene, isolated from transposon Tn5, which, when placed under the control of plant regulatory signals, confers resistance to kanamycin. Another exemplary marker gene is the hygromycin phosphotransferase gene which confers resistance to the antibiotic hygromycin. In some embodiments, the selectable marker is of bacterial origin and confers resistance to antibiotics such as gentamycin acetyl transferase, streptomycin phosphotransferase, and aminoglycoside-3'-adenyl transferase, the bleomycin resistance determinant. In some embodiments, the selectable marker genes confer resistance to herbicides such as glyphosate, glufosinate or bromoxynil. In some embodiments, the selectable marker is mouse dihydrofolate reductase, plant 5-enolpyruvylshikimate-3-phosphate synthase and plant acetolactate synthase. In some embodiments, the selectable marker is acetolactate synthase (e.g., AtCsr1.2).

[242] In some embodiments, a nucleic acid comprises an endoplasmic reticulum retention signal. For example, in some embodiments, a nucleic acid comprises a KDEL sequence (SEQ ID NO: 23). In some embodiments, the nucleic acid may comprise an endoplasmic reticulum retention signal selected from any one of SEQ ID NO: 23-70.

[243] Shown in Table 9 are exemplary promoters, 5' UTRs, signal peptides, and terminators that may be used in the nucleic acids of the disclosure.

Table 9: Promoters, 5' UTRs, signal peptides and terminators

Type	Name	Description	Native Species	Illustrative Accession No. (Glyma, GenBank)
Promoter	PvPhas	Phaseolin-1 (aka β -phaseolin)	Common bean (<i>Phaseolus vulgaris</i>)	J01263.1
	BnNap	Napin-1	Rapeseed (<i>Brassica napus</i>)	J02798.1
	AtOle1	Oleosin-1 (Ole1)	<i>Arabidopsis</i> (<i>Arabidopsis thaliana</i>)	X62353.1, AT4G25140
	GmSeed2	Gy1 (Glycinin 1)	Soybean (Glycine max)	Glyma.03G163500
	GmSeed3	cysteine protease	Soybean (Glycine max)	Glyma.08G116300
	GmSeed5	Gy5 (Glycinin 5)	Soybean (Glycine max)	Glyma.13G123500
	GmSeed6	Gy4 (Glycinin 4)	Soybean (Glycine max)	Glyma.10G037100
	GmSeed7	Kunitz trypsin protease inhibitor	Soybean (Glycine max)	Glyma.01G095000
	GmSeed8	Kunitz trypsin protease inhibitor	Soybean (Glycine max)	Glyma.08G341500
	GmSeed10	Legume Lectin Domain	Soybean (Glycine max)	Glyma.02G012600
	GmSeed11	β -conglycinin a subunit	Soybean (Glycine max)	Glyma.20G148400
	GmSeed12	β -conglycinin a' subunit	Soybean (Glycine max)	Glyma.10G246300
	pBCON	β -conglycinin β subunit	Soybean (Glycine max)	Glyma.20G148200
	GmCEP1-L	KDEL-tailed cysteine endopeptidase CEP1-like	Soybean (Glycine max)	Glyma06g42780

Type	Name	Description	Native Species	Illustrative Accession No. (Glyma, GenBank)
	GmTHIC	phosphomethylpyrimidine synthase	Soybean (Glycine max)	Glyma11g26470
	GmBg7S1	Basic 7S globulin precursor	Soybean (Glycine max)	Glyma03g39940
	GmGRD	glucose and ribitol dehydrogenase-like	Soybean (Glycine max)	Glyma07g38790
	GmOLEA	Oleosin isoform A	Soybean (Glycine max)	Glyma.19g063400
	GmOLEB	Oleosin isoform B	Soybean (Glycine max)	Glyma.16g071800
	Gm2S-1	2S albumin	Soybean (Glycine max)	Glyma13g36400
	GmBBId-II	Bowman-Birk protease inhibitor	Soybean (Glycine max)	Glyma16g33400
5'UTR	Arc5'UTR	arc5-1 gene	Phaseolus vulgaris	J01263.1
	glnB1UTR	65 bp of native glutamine synthase	Soybean (Glycine max)	AF301590.1
Signal peptide	GmSCB1	Seed coat BURP domain protein	Soybean (Glycine max)	Glyma07g28940.1
	StPat21	Patatin	<i>Tomato (Solanum lycopersicum)</i>	CAA27588
	2Sss	2S albumin	Soybean (Glycine max)	Glyma13g36400
	Sig2	Glycinin G1 N-terminal peptide	Soybean (Glycine max)	Glyma.03G163500
	Sig12	Beta-conglycinin alpha prime subunit N-terminal peptide	Soybean (Glycine max)	Glyma.10G246300
	Sig8	Kunitz trypsin inhibitor N-terminal peptide	Soybean (Glycine max)	Glyma.08G341500

Type	Name	Description	Native Species	Illustrative Accession No. (Glyma, GenBank)
	Sig10	Lectin N-terminal peptide from Glycine max	Soybean (Glycine max)	Glyma.02G012600
	Sig11	Beta-conglycinin alpha subunit N-terminal peptide	Soybean (Glycine max)	Glyma.20G148400
	Coixss	Alpha-coixin N-terminal peptide from Coix lacryma-job	Coix lacryma-job	
	KDEL	C-terminal amino acids of sulfhydryl endopeptidase	Phaseolus vulgaris	
Terminator	NOS	Nopaline synthase gene termination sequence	Agrobacterium tumefaciens	
	ARC	arc5-1 gene termination sequence	Phaseolus vulgaris	J01263.1
	EU	Extensin termination sequence	Nicotiana tabacum	
	Rb7	Rb7 matrix attachment region termination sequence	Nicotiana tabacum	
	HSP or AtHSP	Heat shock termination sequence	Arabidopsis thaliana	
	AtUbi10	Ubiquitin 10 termination sequence	Arabidopsis thaliana	
	Stub3	Ubiquitin 3 termination	Solanum tuberosum	
	TM6	M6 matrix attachment region termination sequence	Nicotiana tabacum	
Dual terminators	EU:Rb7	Extensin termination sequence: Rb7 matrix attachment region termination sequence	Nicotiana tabacum	

Type	Name	Description	Native Species	Illustrative Accession No. (Glyma, GenBank)
	AtHSP:AtUbi10	Heat shock termination sequence: Ubiquitin 10 termination sequence	Arabidopsis thaliana	
	EU:StUbi3	Rb7 matrix attachment region termination sequence: Ubiquitin 3 termination	Nicotiana tabacum, Solanum tuberosum	
	EU:TM6	Rb7 matrix attachment region termination sequence: M6 matrix attachment region termination sequence	Nicotiana tabacum	

[244] Illustrative nucleic acids of the disclosure are provided in **FIG. 1A-FIG. 1P** and **FIG. 2A-FIG. 2P**. In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding an unstructured milk protein, a sequence encoding a structured mammalian protein, an endoplasmic reticulum retention signal, and a terminator (See, e.g., **FIG. 1A**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding an unstructured milk protein, a sequence encoding a linker, a sequence encoding a structured mammalian protein, an endoplasmic reticulum retention signal, and a terminator (See, e.g., **FIG. 1B**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding an unstructured milk protein, a sequence encoding a linker, a sequence encoding a structured mammalian protein, and a terminator (See, e.g., **FIG. 1C**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding an unstructured milk protein, a sequence encoding a structured mammalian protein, and a terminator (See, e.g., **FIG. 1D**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a structured mammalian protein, a sequence encoding an unstructured milk protein, an endoplasmic reticulum retention signal, and a terminator (See, e.g., **FIG. 1E**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a structured mammalian protein, a sequence encoding a linker, a sequence encoding an unstructured milk protein, an endoplasmic reticulum retention signal, and a terminator

(See, e.g., **FIG. 1F**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a structured mammalian protein, a sequence encoding a linker, a sequence encoding an unstructured milk protein, and a terminator (See, e.g., **FIG. 1G**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a structured mammalian protein, a sequence encoding an unstructured milk protein, and a terminator (See, e.g., **FIG. 1H**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a signal peptide, a sequence encoding an unstructured milk protein, a sequence encoding a structured mammalian protein, an endoplasmic reticulum retention signal, and a terminator (See, e.g., **FIG. 1I**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a signal peptide, a sequence encoding an unstructured milk protein, a sequence encoding a linker, a sequence encoding a structured mammalian protein, an endoplasmic reticulum retention signal, and a terminator (See, e.g., **FIG. 1J**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a signal peptide, a sequence encoding an unstructured milk protein, a sequence encoding a linker, a sequence encoding a structured mammalian protein, and a terminator (See, e.g., **FIG. 1K**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a signal peptide, a sequence encoding an unstructured milk protein, a sequence encoding a structured mammalian protein, and a terminator (See, e.g., **FIG. 1L**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a signal peptide, a sequence encoding a structured mammalian protein, a sequence encoding an unstructured milk protein, an endoplasmic reticulum retention signal, and a terminator (See, e.g., **FIG. 1M**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a signal peptide, a sequence encoding a structured mammalian protein, a sequence encoding a linker, a sequence encoding an unstructured milk protein, an endoplasmic reticulum retention signal, and a terminator (See, e.g., **FIG. 1N**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a signal peptide, a sequence encoding a structured mammalian protein, a sequence encoding a linker, a sequence encoding an unstructured milk protein, and a terminator (See, e.g., **FIG. 1O**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence

encoding a signal peptide, a sequence encoding a structured mammalian protein, a sequence encoding an unstructured milk protein, and a terminator (See, e.g., **FIG. 1P**).

[245] In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a signal peptide, a sequence encoding first protein, a sequence encoding a second protein, an endoplasmic reticulum retention signal, and a terminator (See, e.g., **FIG. 2A**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a signal peptide, a sequence encoding first protein, a sequence encoding a linker, a sequence encoding a second protein, an endoplasmic reticulum retention signal, and a terminator (See, e.g., **FIG. 2B**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a signal peptide, a sequence encoding a first protein, a sequence encoding a linker, a sequence encoding a second protein, and a terminator (See, e.g., **FIG. 2C**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a signal peptide, a sequence encoding a first protein, a sequence encoding a second protein, and a terminator (See, e.g., **FIG. 2D**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a signal peptide, a sequence encoding a second protein, a sequence encoding a first protein, an endoplasmic reticulum retention signal, and a terminator (See, e.g., **FIG. 2E**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a signal peptide, a sequence encoding a second protein, a sequence encoding a linker, a sequence encoding a first protein, an endoplasmic reticulum retention signal, and a terminator (See, e.g., **FIG. 2F**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a signal peptide, a sequence encoding a second protein, a sequence encoding a linker, a sequence encoding a first protein, and a terminator (See, e.g., **FIG. 2G**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a signal peptide, a sequence encoding a second protein, a sequence encoding first protein, and a terminator (See, e.g., **FIG. 2H**).

[246] In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a first protein, a sequence encoding a second protein, an endoplasmic reticulum retention signal, and a terminator (See, e.g., **FIG. 2I**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a first protein, a sequence encoding a linker, a sequence encoding a second protein, an endoplasmic reticulum retention

signal, and a terminator (See, e.g., **FIG. 2J**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a first protein, a sequence encoding a linker, a sequence encoding a second protein, and a terminator (See, e.g., **FIG. 2K**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a first protein, a sequence encoding a second protein, and a terminator (See, e.g., **FIG. 2L**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a second protein, a sequence encoding a first protein, an endoplasmic reticulum retention signal, and a terminator (See, e.g., **FIG. 2M**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a second protein, a sequence encoding a linker, a sequence encoding a first protein, an endoplasmic reticulum retention signal, and a terminator (See, e.g., **FIG. 2N**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a second protein, a sequence encoding a linker, a sequence encoding a first protein, and a terminator (See, e.g., **FIG. 2O**). In some embodiments a nucleic acid comprises, from 5' to 3', a promoter, a 5'UTR, a sequence encoding a second protein, a sequence encoding a first protein, and a terminator (See, e.g., **FIG. 2P**).

[247] In some embodiments, the nucleic acid comprises an expression cassette comprising a OKC1-T:OLG1 (Optimized Kappa Casein version 1:beta-lactoglobulin version 1) fusion driven by PvPhas promoter fused with arc5'UTR:sig10, followed by the ER retention signal (KDEL) and the 3'UTR of the arc5-1 gene, "arc-terminator" (See, e.g., **FIG. 4**). In some embodiments, the nucleic acid comprises SEQ ID NO: 72.

[248] In some embodiments, the nucleic acid comprises an expression cassette comprising a OBC-T2:FM:OLG1 (Optimized Beta Casein Truncated version 2:Chymosin cleavage site:beta-lactoglobulin version 1) fusion driven by PvPhas promoter fused with arc5'UTR:sig10, followed by the 3'UTR of the arc5-1 gene, "arc-terminator" (See, e.g., **FIG. 5**). In some embodiments, the nucleic acid comprises SEQ ID NO: 74. The Beta Casein is "truncated" in that the bovine secretion signal is removed and replaced with a plant targeting signal.

[249] In some embodiments, the nucleic acid comprises an expression cassette comprising a OaS1-T:FM:OLG1 (Optimized Alpha S1 Casein Truncated version 1:Chymosin cleavage site:beta-lactoglobulin version 1) fusion driven by PvPhas promoter fused with arc5'UTR:sig10,

followed by the 3'UTR of the arc5-1 gene, "arc-terminator" (See, e.g., **FIG. 6**). In some embodiments, the nucleic acid comprises SEQ ID NO: 76. The Alpha S1 is "truncated" in that the bovine secretion signal is removed and replaced with a plant targeting signal.

[250] In some embodiments, the nucleic acid comprises an expression cassette comprising a para-OKC1-T:FM:OLG1:KDEL (Optimized paraKappa Casein version 1:Chymosin cleavage site:beta-lactoglobulin version 1) fusion driven by PvPhas promoter fused with arc5'UTR:sig 10, followed by the ER retention signal (KDEL) and the 3'UTR of the arc5-1 gene, "arc-terminator" (See, e.g., **FIG. 7**). In some embodiments, the nucleic acid comprises SEQ ID NO: 78.

[251] In some embodiments, the nucleic acid comprises an expression cassette comprising a para-OKC1-T:FM:OLG1 (Optimized paraKappa Casein version 1:Chymosin cleavage site:beta-lactoglobulin version 1) fusion driven by PvPhas promoter fused with arc5'UTR:sig 10, followed by the 3'UTR of the arc5-1 gene, "arc-terminator" (See, e.g., **FIG. 8**). In some embodiments, the nucleic acid comprises SEQ ID NO: 80.

[252] In some embodiments, the nucleic acid comprises an expression cassette comprising a OKC1-T-OLG1 (Optimized Kappa Casein version 1:beta-lactoglobulin version 1) fusion that is driven by the promoter and signal peptide of glycinin 1 (GmSeed2:sig2) followed by the ER retention signal (KDEL) and the nopaline synthase gene termination sequence (nos term) (See, e.g., **FIG. 9**). In some embodiments, the nucleic acid comprises SEQ ID NO: 82.

[253] In some embodiments, a nucleic acid encoding a fusion protein comprises the sequence of any one of SEQ ID NO: 72, 74, 76, 78, 80, 82, 134, or 136. In some embodiments, a nucleic acid encoding a fusion protein comprises the sequence of any one of SEQ ID NO: 615, 617, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 792, 794, 796, or 798.

[254] In some embodiments, the nucleic acids are codon optimized for expression in a host cell. Codon optimization is a process used to improve gene expression and increase the translational efficiency of a gene of interest by accommodating codon bias of the host organism (i.e., the organism in which the gene is expressed). Codon-optimized mRNA sequences that are produced using different programs or approaches can vary because different codon optimization strategies differ in how they quantify codon usage and implement codon changes. Some approaches use the

most optimal (frequently used) codon for all instances of an amino acid, or a variation of this approach. Other approaches adjust codon usage so that it is proportional to the natural distribution of the host organism. These approaches include codon harmonization, which endeavors to identify and maintain regions of slow translation thought to be important for protein folding. Alternative approaches involve using codons thought to correspond to abundant tRNAs, using codons according to their cognate tRNA concentrations, selectively replacing rare codons, or avoiding occurrences of codon-pairs that are known to translate slowly. In addition to approaches that vary in the extent to which codon usage is considered as a parameter, there are hypothesis-free approaches that do not consider this parameter. Algorithms for performing codon optimization are known to those of skill in the art and are widely available on the Internet.

[255] In some embodiments the nucleic acids are codon optimized for expression in a plant species. The plant species may be, for example, a monocot or a dicot. In some embodiments, the plant species is a dicot species selected from soybean, lima bean, Arabidopsis, tobacco, rice, maize, barley, sorghum, wheat and/or oat. In some embodiments, the plant species is soybean.

[256] In some embodiments, the nucleic acids are codon optimized for expression in a eukaryotic microorganism. The species may be, for example, *Saccharomyces* spp., *Kluyveromyces* spp., *Pichia* spp., *Aspergillus* spp., *Tetrahymena* spp., *Yarrowia* spp., *Hansenula* spp., *Blastobotrys* spp., *Candida* spp., *Zygosaccharomyces* spp., *Debaryomyces* spp., *Fusarium* spp., and *Trichoderma* spp.

[257] In some embodiments, the nucleic acids are codon optimized for expression in a bacterial cell. The bacterial species may be, for example, *Escherichia coli*, *Caulobacter crescentus*, *Rodhobacter sphaeroides*, *Pseudoalteromonas haloplanktis*, *Shewanella* sp., *Pseudomonas putida*, *P. aeruginosa*, *P. fluorescens*, *Halomonas elongate*, *Chromohalobacter salexigens*, *Streptomyces lividans*, *S. griseus*, *Nocardia lactamdurans*, *Mycobacterium smegmatis*, *Corynebacterium glutamicum*, *C. ammoniagenes*, *Brevibacterium lactofermentum*, *Bacillus subtilis*, *B. brevis*, *B. megaterium*, *B. licheniformis*, *B. amyloliquefaciens*, *Lactococcus lactis*, *L. plantarum*, *L. casei*, *L. reuteri*, *L. gasseri*.

[258] In some embodiments, a nucleic acid may encode more than one fusion protein. For example, in some embodiments, a nucleic acid may encode two, three, four, five, six, seven, eight,

nine, or ten fusion proteins, or more. Expression of each fusion protein from the nucleic acid may be driven by a separate promoter. For example, in some embodiments, a nucleic acid comprises a first promoter configured to drive expression of a sequence encoding a first fusion protein, and a second promoter configured to drive expression of a sequence encoding a second fusion protein. In some embodiments, a nucleic acid comprises a first promoter operably linked to a sequence encoding a first fusion protein, and a second promoter operably linked to a sequence encoding a second fusion protein.

[259] The nucleic acids of the disclosure may be contained within a vector. The vector may be, for example, a viral vector or a non-viral vector. In some embodiments, the non-viral vector is a plasmid, such as an *Agrobacterium* Ti plasmid. In some embodiments, the non-viral vector is a lipid nanoparticle.

[260] In some embodiments, the vector comprises a nucleic acid encoding multiple fusion proteins. For example, in some embodiments, a vector comprises a nucleic acid comprising a sequence encoding a first fusion protein and a sequence encoding a second fusion protein. A first promoter may drive expression of the first fusion protein, and a second promoter may drive expression of the second fusion protein. In some embodiments, the first promoter and the second promoter are the same. In some embodiments, the first promoter and the second promoter are different.

[261] In some embodiments, a vector comprises a nucleic acid comprising a sequence encoding a first fusion protein, a sequence encoding a second fusion protein, and a sequence encoding a third fusion protein. A first promoter may drive expression of the first fusion protein, a second promoter may drive expression of the second fusion protein, and a third promoter may drive expression of the third fusion protein. In some embodiments, each of the first, second, and third promoter are different. In some embodiments, at least two of the first, second, and third promoter are different. In some embodiments, the first, second, and third promoter are the same.

[262] In some embodiments, a vector comprises a nucleic acid encoding a recombinant fusion protein, wherein the recombinant fusion protein comprises: (i) an unstructured milk protein, and (ii) a structured animal (e.g., mammalian or avian) protein. In some embodiments, the vector is an *Agrobacterium* Ti plasmid. In some embodiments, a vector comprises a nucleic acid encoding a

recombinant fusion protein, wherein the recombinant fusion protein comprises: (1) a milk protein, and (2) a second protein. In some embodiments, the second protein is also a milk protein. In some embodiments, the second protein is beta-lactoglobulin. In some embodiments, the second protein is a mammalian or avian protein. In some embodiments, the vector is an *Agrobacterium* Ti plasmid. In some embodiments, the vector is a vector for use with an *Agrobacterium* binary vector transformation system. In some embodiments, the fusion protein is cleaved to liberate the milk protein and the second protein before either one is used to prepare a composition as described herein (See, e.g., **FIG. 13**). The fusion protein may be cleaved, for example, with one or more proteases.

[263] In some embodiments, a method for expressing a casein protein (including fusion proteins comprising a casein protein) in a plant comprises contacting the plant with a vector of the disclosure. In some embodiments, a method for expression of a casein protein in a plant comprises contacting the plant with an *Agrobacterium* cell comprising a vector of the disclosure. In some embodiments, the method comprises maintaining the plant or part thereof under conditions in which the fusion protein is expressed.

[264] In some embodiments, a method for expressing a fusion protein in a plant comprises contacting the plant with a vector of the disclosure. In some embodiments, the method comprises maintaining the plant or part thereof under conditions in which the fusion protein is expressed.

Plants Expressing Fusion Proteins

[265] Also provided herein are transgenic plants expressing one or more fusion proteins of the disclosure. In some embodiments, the transgenic plants stably express the fusion protein. In some embodiments, the transgenic plants transiently express the fusion protein. In some embodiments, the transgenic plants stably express the fusion protein in the plant in an amount of at least 1% per the total protein weight of the soluble protein extractable from the plant. For example, the transgenic plants may stably express the fusion protein in an amount of at least 1%, at least 1.5%, at least 2%, at least 2.5%, at least 3%, at least 3.5%, at least 4%, at least 4.5%, at least 5%, at least 5.5%, at least 6%, at least 6.5%, at least 7%, at least 7.5%, at least 8%, at least 8.5%, at least 9%, at least 9.5%, at least 10%, at least 10.5%, at least 11%, at least 11.5%, at least 12%, at least 12.5%, at least 13%, at least 13.5%, at least 14%, at least 14.5%, at least 15%, at least 15.5%, at least 16%,

at least 16.5%, at least 17%, at least 17.5%, at least 18%, at least 18.5%, at least 19%, at least 19.5%, at least 20%, or more of total protein weight of soluble protein extractable from the plant.

[266] In some embodiments, the transgenic plants stably express the fusion protein in an amount of less than about 1% of the total protein weight of soluble protein extractable from the plant. In some embodiments, the transgenic plants stably express the fusion protein in the range of about 1% to about 2%, about 3% to about 4%, about 4% to about 5%, about 5% to about 6%, about 6% to about 7%, about 7% to about 8%, about 8% to about 9%, about 9% to about 10%, about 10% to about 11%, about 11% to about 12%, about 12% to about 13%, about 13% to about 14%, about 14% to about 15%, about 15% to about 16%, about 16% to about 17%, about 17%, to about 18%, about 18% to about 19%, about 19% to about 20%, or more than about 20% of the total protein weight of soluble protein extractable from the plant.

[267] In some embodiments, the transgenic plant stably expresses the fusion protein in an amount in the range of about 0.5% to about 3%, about 1% to about 4%, about 1% to about 5%, about 2% to about 5%, about 1% to about 10%, about 2% to about 10%, about 3% to about 10%, about 5 to about 12%, about 4% to about 10%, or about 5% to about 10%, about 4% to about 8%, about 5% to about 15%, about 5% to about 18%, about 10% to about 20%, or about 1% to about 20% of the total protein weight of soluble protein extractable from the plant.

[268] In some embodiments, the fusion protein is expressed at a level at least 2-fold higher than a milk protein expressed individually (i.e., expressed alone, not as part of a fusion protein) in a plant. For example, in some embodiments, the fusion protein is expressed at a level at least 2-fold, at least 2.5-fold, at least 3-fold, at least 3.5-fold, at least 4-fold, at least 4.5-fold, at least 5-fold, at least 5.5-fold, at least 6-fold, at least 7-fold, at least 7.5-fold, at least 8-fold, at least 8.5-fold, at least 9-fold, at least 9.5-fold, at least 10-fold, at least 25-fold, at least 50-fold, or at least 100-fold higher than a milk protein expressed individually in a plant.

[269] In some embodiments, the fusion protein allows for accumulation of a casein protein in the plant at least 2-fold higher than a casein protein expressed individually (i.e., expressed alone, not as a part of a fusion protein) in a plant. For example, in some embodiments, the casein protein expressed in a fusion protein accumulates in the plant at least 2-fold, at least 2.5-fold, at least 3-fold, at least 3.5-fold, at least 4-fold, at least 4.5-fold, at least 5-fold, at least 5.5-fold, at least 6-

fold, at least 7-fold, at least 7.5-fold, at least 8-fold, at least 8.5-fold, at least 9-fold, at least 9.5-fold, at least 10-fold, at least 25-fold, at least 50-fold, or at least 100-fold higher than a casein protein expressed individually.

[270] In some embodiments, the fusion protein is stably expressed in the plant in an amount of 1% or higher per the total protein weight of the soluble protein extractable from the plant. In some embodiments, the fusion protein is stably expressed in the plant in an amount of 2% or higher per the total protein weight of the soluble protein extractable from the plant. In some embodiments, the fusion protein is stably expressed in the plant in an amount of 3% or higher per the total protein weight of the soluble protein extractable from the plant. In some embodiments, the fusion protein is stably expressed in the plant in an amount of 4% or higher per the total protein weight of the soluble protein extractable from the plant. In some embodiments, the fusion protein is stably expressed in the plant in an amount of 5% or higher per the total protein weight of the soluble protein extractable from the plant. In some embodiments, the fusion protein is stably expressed in the plant in an amount of 6% or higher per the total protein weight of the soluble protein extractable from the plant. In some embodiments, the fusion protein is stably expressed in the plant in an amount of 7% or higher per the total protein weight of the soluble protein extractable from the plant. In some embodiments, the fusion protein is stably expressed in the plant in an amount of 8% or higher per the total protein weight of the soluble protein extractable from the plant. In some embodiments, the fusion protein is stably expressed in the plant in an amount of 9% or higher per the total protein weight of the soluble protein extractable from the plant. In some embodiments, the fusion protein is stably expressed in the plant in an amount of 10% or higher per the total protein weight of the soluble protein extractable from the plant. In some embodiments, the fusion protein is stably expressed in the plant in an amount of 11% or higher per the total protein weight of the soluble protein extractable from the plant. In some embodiments, the fusion protein is stably expressed in the plant in an amount of 12% or higher per the total protein weight of the soluble protein extractable from the plant. In some embodiments, the fusion protein is stably expressed in the plant in an amount of 13% or higher per the total protein weight of the soluble protein extractable from the plant. In some embodiments, the fusion protein is stably expressed in the plant in an amount of 14% or higher per the total protein weight of the soluble protein extractable from the plant. In some embodiments, the fusion protein is stably expressed in the plant in an amount

of 15% or higher per the total protein weight of the soluble protein extractable from the plant. In some embodiments, the fusion protein is stably expressed in the plant in an amount of 16% or higher per the total protein weight of the soluble protein extractable from the plant. In some embodiments, the fusion protein is stably expressed in the plant in an amount of 17% or higher per the total protein weight of the soluble protein extractable from the plant. In some embodiments, the fusion protein is stably expressed in the plant in an amount of 18% or higher per the total protein weight of the soluble protein extractable from the plant. In some embodiments, the fusion protein is stably expressed in the plant in an amount of 19% or higher per the total protein weight of the soluble protein extractable from the plant. In some embodiments, the fusion protein is stably expressed in the plant in an amount of 20% or higher per the total protein weight of the soluble protein extractable from the plant.

[271] In some embodiments, a transformed plant comprises in its genome: a recombinant DNA construct encoding a first protein and a second protein, wherein the first protein and/or the second protein is a milk protein. In some embodiments, a transformed plant comprises in its genome: a recombinant DNA construct encoding a first protein and a second protein, wherein the first protein is a milk protein and the second protein is a non-milk protein. In some embodiments, a transformed plant comprises in its genome a recombinant DNA construct encoding a fusion protein, wherein the fusion protein comprises a first protein and a second protein, wherein the first protein and the second protein are milk proteins. In some embodiments, a transformed plant comprises in its genome a recombinant DNA construct encoding a fusion protein, wherein the fusion protein comprises from N-terminus to C-terminus, the first protein and the second protein. In some embodiments, the fusion protein comprises, from N-terminus to C-terminus, the second protein and the first protein.

[272] In some embodiments, a transformed plant comprises in its genome: a recombinant DNA construct encoding a fusion protein, wherein the fusion protein comprises (i) a milk protein, and (ii) an animal (e.g., mammalian or avian) protein. In some embodiments, a transformed plant comprises in its genome a recombinant DNA construct encoding a fusion protein, wherein the fusion protein comprises from N-terminus to C-terminus, the milk protein and the animal (e.g., mammalian or avian) protein. In some embodiments, the fusion protein comprises, from N-terminus to C-terminus, the animal (e.g., mammalian or avian) protein and the milk protein.

[273] In some embodiments, a transformed plant comprises in its genome: a recombinant DNA construct encoding a fusion protein, wherein the fusion protein comprises a milk protein such as a casein protein. In some embodiments, a transformed plant comprises in its genome: a recombinant DNA construct encoding a fusion protein, wherein the fusion protein comprises a milk protein selected from α -S1 casein, α -S2 casein, β -casein, and κ -casein. In some embodiments, the milk protein is α -S1 casein. In some embodiments, the milk protein is α -S1 casein and comprises the sequence SEQ ID NO: 8, or a sequence at least 90% identical thereto. In some embodiments, the milk protein is α -S2 casein. In some embodiments, the milk protein is α -S2 casein and comprises the sequence SEQ ID NO: 84, or a sequence at least 90% identical thereto. In some embodiments, the milk protein is β -casein. In some embodiments, the milk protein is β -casein and comprises the sequence of SEQ ID NO: 6, or a sequence at least 90% identical thereto. In some embodiments, the milk protein is κ -casein. In some embodiments, the milk protein is κ -casein and comprises the sequence of SEQ ID NO: 4, or a sequence at least 90% identical thereto. In some embodiments, the milk protein is para- κ -casein. In some embodiments, the milk protein is para- κ -casein and comprises the sequence of SEQ ID NO: 2, or a sequence at least 90% identical thereto. In some embodiments, the milk protein is β -lactoglobulin. In some embodiments, the milk protein is β -lactoglobulin and comprises the sequence of SEQ ID NO: 10, or a sequence at least 90% identical thereto. In some embodiments, the milk protein is α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, or an immunoglobulin (e.g., IgA, IgG, IgM, or IgE).

[274] In some embodiments, a transformed plant comprises in its genome: a recombinant DNA construct encoding a fusion protein, wherein the fusion protein comprises a mammalian protein selected from hemoglobin, or collagen, IgM, or IgE. In some embodiments, a transformed plant comprises in its genome: a recombinant DNA construct encoding a fusion protein, wherein the fusion protein comprises an avian protein selected from lysozyme, ovalbumin, ovotransferrin, and ovoglobulin.

[275] In some embodiments, a transformed plant comprises in its genome: a recombinant DNA construct encoding a fusion protein, wherein the fusion protein comprises a casein protein and β -lactoglobulin. In some embodiments, a transformed plant comprises in its genome: a recombinant DNA construct encoding a fusion protein, wherein the fusion protein comprises κ -casein and β -lactoglobulin. In some embodiments, the fusion protein comprises para- κ -casein and β -

lactoglobulin. In some embodiments, the fusion protein comprises β -casein and β -lactoglobulin. In some embodiments, the fusion protein comprises α -S1 casein and β -lactoglobulin. In some embodiments, the fusion protein comprises two, three, four, five, or six β -caseins.

[276] In some embodiments, a transformed plant comprises in its genome: a recombinant DNA construct encoding a fusion protein; wherein the fusion protein comprises (i) κ -casein, and (ii) β -lactoglobulin. In some embodiments the fusion protein is expressed in the plant in an amount of 1% or higher per the total protein weight of the soluble protein extractable from the plant.

[277] In some embodiments, a transformed plant comprises in its genome: a recombinant DNA construct encoding a fusion protein, wherein the fusion protein comprises a first protein and a second protein, wherein the first protein and the second protein are each casein proteins. In some embodiments, the recombinant fusion protein comprises κ -casein and para- κ -casein. In some embodiments, the recombinant fusion protein comprises κ -casein and β -casein. In some embodiments, the recombinant fusion protein comprises κ -casein and α -S1-casein. In some embodiments, the recombinant fusion protein comprises κ -casein and α -S2-casein. In some embodiments, the recombinant fusion protein comprises para- κ -casein and β -casein. In some embodiments, the recombinant fusion protein comprises para- κ -casein and α -S1-casein. In some embodiments, the recombinant fusion protein comprises para- κ -casein and α -S2-casein. In some embodiments, the recombinant fusion protein comprises β -casein and α -S1-casein. In some embodiments, the recombinant fusion protein comprises β -casein and α -S2-casein. In some embodiments, the recombinant fusion protein comprises α -S1-casein and α -S2-casein.

[278] In some embodiments, the recombinant fusion protein comprises two or more of the same casein proteins. In some embodiments, the recombinant fusion protein comprises κ -casein and κ -casein. In some embodiments, the recombinant fusion protein comprises β -casein and β -casein. In some embodiments, the recombinant fusion protein comprises para- κ -casein and para- κ -casein. In some embodiments, the recombinant fusion protein comprises α -S1-casein and α -S1-casein. In some embodiments, the recombinant fusion protein comprises α -S2-casein and α -S2-casein.

[279] In some embodiments, the transformed plant is a monocot. For example, in some embodiments, the plant may be a monocot selected from turf grass, maize (corn), rice, oat, wheat, barley, sorghum, orchid, iris, lily, onion, palm, and duckweed.

[280] In some embodiments, the transformed plant is a dicot. For example, in some embodiments, the plant may be a dicot selected from Arabidopsis, tobacco, tomato, potato, sweet potato, cassava, alfalfa, lima bean, pea, chick pea, soybean, carrot, strawberry, lettuce, oak, maple, walnut, rose, mint, squash, daisy, Quinoa, buckwheat, mung bean, cow pea, lentil, lupin, peanut, fava bean, French beans (i.e., common beans), mustard, or cactus. In some embodiments, the plant is a soybean (*Glycine max*).

[281] In some embodiments, the plant is a non-vascular plant selected from moss, liverwort, hornwort or algae. In some embodiments, the plant is a vascular plant reproducing from spores (e.g., a fern).

[282] In some embodiments, the recombinant DNA construct is codon-optimized for expression in the plant. For example, in some embodiments, the recombinant DNA construct is codon-optimized for expression in a soybean plant.

[283] The transgenic plants described herein may be generated by various methods known in the art. For example, a nucleic acid encoding a fusion protein may be contacted with a plant, or a part thereof, and the plant may then be maintained under conditions wherein the fusion protein is expressed. In some embodiments, the nucleic acid is introduced into the plant, or part thereof, using one or more methods for plant transformation known in the art, such as *Agrobacterium*-mediated transformation, particle bombardment-mediated transformation, electroporation, and microinjection.

[284] In some embodiments, a method for stably expressing a recombinant fusion protein in a plant comprises (i) transforming a plant with a plant transformation vector comprising an expression cassette comprising: a sequence encoding a fusion protein, wherein the fusion protein comprises a milk protein, and an animal (e.g., mammalian or avian) protein; and (ii) growing the transformed plant under conditions wherein the recombinant fusion protein is expressed. In some embodiments, the milk protein is κ -casein. In some embodiments, the animal protein is β -lactoglobulin. In some embodiments, the milk protein is κ -casein and the animal protein is β -lactoglobulin. In some embodiments, the recombinant fusion protein is expressed in an amount of 1% or higher per the total protein weight of the soluble protein extractable from the plant.

Casein Accumulation in Plants

[285] As described herein, fusion proteins comprising one or more milk proteins (e.g., casein proteins) accumulate to a greater extent in plant cells than the milk proteins expressed individually (not as fusion proteins). Caseins aggregate and bind to calcium-phosphate to form micelles. Without being bound by any theory, it is believed that native plant proteases are capable of degrading caseins by cleavage at various protease recognition sites (**FIG. 11A**). Thus, when caseins are expressed alone (i.e., not as a fusion protein), they are degraded quickly and do not accumulate in the cells. When caseins are fused to a second protein (**FIG. 11B**, **FIG. 11C**), the second protein may partially or fully limit protease access to the cleavage site on the caseins and may reduce degradation thereof. The extent of protection may vary depending on the properties of the second protein. For example, fusion proteins comprising two caseins (e.g., homodimers or heterodimers, **FIG. 11C**) may be able adopt a conformation that partially or fully prevents access to one or more protease cleavage sites. Some non-casein proteins, such as beta-lactoglobulin, GFP, or lysozyme, may also partially or fully block protease access, allowing casein accumulation at high levels in the cell (**FIG. 11B**). Without being bound by any theory, it is believed that fusion of a casein to a second protein comprising one, two or all three of the following characteristics is able to prevent access to one or more protease cleavage sites on the casein: (i) a molecular weight of 15 kDa or higher; (ii) at least 30% hydrophobic amino acids; and/or (iii) less than about 2.5 disulfide bonds per 10 kDa molecular weight.

[286] Protease access to cleavage sites on a casein protein may also be blocked, for example, by the addition of one or more post-translational modifications to the casein, such as phosphorylation, glycosylation (**FIG. 11D**) or lipidation (**FIG. 11E**). Thus, in some embodiments, a recombinant casein protein described herein comprises one or more post-translational modifications. The post-translational modifications may, in some embodiments, prevent proteolysis by endogenous plant proteases. For example, the presence of one or more post-translational modifications on a recombinant casein may reduce proteolysis of the casein in a plant cell by at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 100%, at least 200% or more, relative to the proteolysis of a casein that does not have the one or more post-translational modifications. In some embodiments, the presence of one or more post-translational modifications on a recombinant casein may lead to an increase in expression of

at least 2-fold, at least 3-fold, at least 4-fold, at least 5-fold, at least 10-fold, at least 20-fold, at least 30-fold, at least 40-fold, at least 50-fold or more, relative to the expression of a casein that does not have the one or more post-translational modifications. The recombinant casein proteins comprising post-translational modifications described herein may be expressed alone or may be expressed in a fusion protein (e.g., a casein protein homo- or hetero-multimer).

[287] In some embodiments, the post-translational modifications may be non-mammalian post-translational modifications. For example, the post-translational modifications may be plant post-translational modifications. In some embodiments, the post-translational modifications may not typically occur in a casein protein when expressed in a plant or an animal cell. A non-limiting list of post-translational modifications that may be used to prevent proteolysis by endogenous plant proteases includes glycosylation (e.g., O-glycans, N-glycans, or glycosaminoglycans such as heparin, heparan sulfate, chondroitin sulfate, keratan sulfate or dermatan sulfate), phosphorylation, lipidation, ubiquitylation, nitrosylation, methylation, acetylation, amidation, prenylation, alkylation, gamma-carboxylation, biotinylation, oxidation, or sulfation. In some embodiments, the post-translational modification is phosphorylation.

[288] In some embodiments, a recombinant milk protein (e.g., a casein protein) comprises a site for post-translational modification that is not present in the native form of the protein. In some embodiments, a recombinant milk protein (e.g., a casein protein) comprises at least one, at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, or more sites for post-translational modifications that are not present in the native form of the protein. In some embodiments, a recombinant milk protein (e.g., a casein protein) comprises at least one, at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, or more post-translational modifications at sites that are not present in the native form of the protein.

[289] In some embodiments, a recombinant milk protein (e.g., a casein protein) comprises an amino acid sequence that is modified to promote addition of one or more post-translational modifications in a plant cell. In some embodiments, the one or more post-translational modifications are selected from glycosylation, phosphorylation, lipidation, ubiquitylation, nitrosylation, methylation, acetylation, amidation, prenylation, alkylation, gamma-carboxylation,

biotinylation, oxidation, and sulfation. In some embodiments, the amino acid sequence of a recombinant casein protein may be modified to introduce one or more glycosylation or phosphorylation sites.

[290] In some embodiments, a milk protein is expressed in a plant, wherein the milk protein comprises an amino acid sequence that is modified to promote addition of one or more post-translational modifications, and wherein the milk protein comprises one or more post-translational modifications that are not present in a non-modified milk protein expressed in the same type of plant. In some embodiments, the milk protein is expressed in a plant in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant. In some embodiments, the milk protein is a casein protein selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, and para- κ -casein.

[291] In some embodiments, a fusion protein comprises (i) a recombinant milk protein that comprises an amino acid sequence that is modified to promote addition of one or more post-translational modifications in a plant cell, and (ii) at least one additional protein. In some embodiments, the at least one additional protein is a milk protein. In some embodiments, the at least one additional protein is a casein protein selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, and para- κ -casein. In some embodiments, the at least one additional protein is β -lactoglobulin. In some embodiments, the recombinant milk protein is κ -casein or para- κ -casein and the at least one additional protein is β -lactoglobulin. In some embodiments, the recombinant milk protein is β -casein and the at least one additional protein is β -lactoglobulin. In some embodiments, the recombinant milk protein is α -S1 casein or α -S2 casein and the at least one additional protein is β -lactoglobulin. In some embodiments, the fusion protein is expressed in a plant in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant. In some embodiments, the plant is soybean.

[292] In some embodiments, a transgenic plant expresses a milk protein comprising an amino acid sequence that is modified to promote addition of one or more post-translational modifications, or a fusion protein comprising the same.

[293] Proteolysis of recombinant caseins in plant cells may also be prevented by modifying the plant cell itself. Without being bound by any theory, it is believed that in wildtype seeds, proteases

present in one or more cellular compartments may bind to and cleave casein expressed therein. Thus, casein does not accumulate at high levels in the seeds (See **FIG. 18**, top panel). In contrast, when expression of one or more proteases is knocked-down or knocked-out in the seed (indicated by “X” in the bottom panel of **FIG. 18**), degradation of the casein is substantially prevented. Accordingly, the casein can accumulate in the seed. This strategy may be used to increase expression in the seed of casein monomers (i.e., caseins expressed alone, not as a fusion protein), or fusion proteins comprising one or more caseins.

[294] In some embodiments, expression of one or more endogenous plant proteases may be knocked down or knocked out in a plant cell (e.g., a seed). The one or more proteases may be, for example, one or more proteases endogenously expressed in a plant (e.g., a soybean), such as cysteine proteases, serine proteases, threonine proteases, or aspartic proteases, glutamic proteases, metalloproteases, or asparagine peptide lyases. A non-limiting list of genes encoding proteases that may be knocked down or knocked out in a plant cell is provided below in Table 10. Additional proteases that may be knocked down or knocked out in a soybean cell are described in Shamimuzzaman M., Vodkin L (2018) Ribosome profiling reveals changes in translational status of soybean transcripts during immature cotyledon development. *PLoS ONE* 13(3): e0194596.

[295] In some embodiments, expression of at least one, at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten or more proteases may be knocked down or knocked out in a plant cell.

Table 10: Genes encoding proteases that are transcriptionally active in soybeans

Soybean Gene ID		
Glyma.02g213000	Glyma.09g187200	Glyma.14g085800
Glyma.03g125400	Glyma.09g226700	Glyma.14g216300
Glyma.03g239700	Glyma.09g249500	Glyma.15g177800
Glyma.04g022500	Glyma.10g207100	Glyma.15g234300
Glyma.04g027600	Glyma.12G010100	Glyma.16G018900

Soybean Gene ID		
Glyma.04g091800	Glyma.13g027600	Glyma.17g164100
Glyma.06g022600	Glyma.13g196200	Glyma.17g239000
Glyma.06g027700	Glyma.13g208200	Glyma.17g254900
Glyma.06G272700	Glyma.13g255900	Glyma.18G242900
Glyma.06g275300	Glyma.13g321700	Glyma.18g250100
Glyma.08g116300	Glyma.14g048000	Glyma.19G236600
Glyma.08G116400	Glyma.14g064600	

Table 11: Proteases that may be knocked down or knocked out in a plant cell

Protein Name	Accession No. (Uniprot)	DNA Sequence	Protein Sequence
Peptidase A1 domain-containing protein	Glyma.04g091800	851	852
Cysteine proteinase	Glyma.10g207100	853	854
34kDa maturing seed protein	Glyma.08g116300	855	856
Uncharacterized protein (cysteine protease family C1-related)	Glyma.06g275300	857	858
Uncharacterized protein (Subsilin-like serine peptidase)	Glyma.17g164100	859	560

[296] In some embodiments, a plant cell for expressing recombinant milk proteins is provided, wherein expression of one or more proteases is reduced (e.g., knocked down or knocked out) in the cell. The expression of the one or more proteases may be reduced (e.g., knocked down or knocked out), for example, using a gene editing technology (e.g., CRISPR, TALENs, Zn Finger Nuclease, etc.) or base editing technology (e.g., using a cytidine deaminase or an adenosine deaminase). In some embodiments, expression of the one or more proteases may be reduced using RNA interference (e.g., microRNAs or siRNAs). In some embodiments the one or more proteases that is knocked down or knocked out is a cysteine protease, a serine protease, or an aspartyl

protease. In some embodiments, the one or more proteases that is knocked down or knocked out is any one of the proteases listed in Table 10 or Table 11. In some embodiments, the one or more proteases that is knocked down or knocked out comprises the sequence of any one of SEQ ID NO: 852, 584, 856, 858, or 860. In some embodiments, the one or more proteases that is knocked down or knocked out comprises a sequence with at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity with any one of SEQ ID NO: 852, 584, 856, 858, or 860. In some embodiments, the one or more proteases that is knocked down or knocked out comprises a sequence of any one of SEQ ID NO: 852, 584, 856, 858, or 860 plus at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, or more amino acid substitutions. The expression or activity of endogenous plant proteases may also be reduced using small molecule inhibitors thereof (i.e., protease inhibitors).

[297] Also provided is a transgenic plant comprising a plant cell for expressing recombinant milk proteins, wherein expression of one or more proteases is reduced (e.g., knocked down or knocked out) in the plant.

[298] In some embodiments, a method for stably expressing a recombinant milk protein in a plant comprises: (i) reducing expression of one or more proteases in the plant, (ii) transforming the plant with a plant transformation vector comprising an expression cassette encoding a recombinant milk protein or a fusion protein comprising the same, (iii) growing the transformed plant under conditions wherein the recombinant milk protein is expressed in an amount of 1% or higher per total weight of soluble protein extractable from the plant.

[299] In some embodiments, a recombinant casein protein that comprises one or more post-translational modifications is produced in a plant cell by expressing or over-expressing one or more enzymes in the plant cell, such as an enzyme known to perform post-translational modifications (e.g., a kinase, a phosphatase, or glycosyltransferase). In some embodiments, a recombinant casein protein that comprises one or more post-translational modifications is produced in a plant cell by knocking out or knocking down one or more enzymes the plant cell known to remove or prevent addition of post-translational modifications (e.g., a phosphatase or an endoglycosidase). In some embodiments, a recombinant casein protein that comprises one or more

post-translational modifications is produced in a plant cell by contacting the cell with one or more precursors of the post-translational modification (e.g., a nucleotide sugar precursor).

[300] In some embodiments, a recombinant casein protein comprises one or more glycoprotein tags. For example, in some embodiments, a recombinant casein protein may comprise a glycoprotein tag derived from a hydroxyproline (Hyp)-rich glycoprotein (GRGP). In some embodiments, the glycoprotein tag comprises SP repeats. For example, the glycoprotein tag may be derived from a glycoprotein comprising 11 tandem SP repeats (See Glyma.02g204500, annotated as early nodulin-like protein 10 in soy). In some embodiments, the fusion protein comprises the M domain of CD45 (receptor-type tyrosine-protein phosphatase C), or a fragment or derivative thereof. For example, in some embodiments, the fusion protein comprises amino acids Ala231 to Asp 290 of Uniprot Accession No. P08575. In some embodiments, the glycoprotein tag comprises the sequence of SEQ ID NO: 824. In some embodiments, the glycoprotein tag is encoded by the sequence SEQ ID NO: 825. In some embodiments, the glycoprotein tag comprises the sequence of SEQ ID NO: 827. In some embodiments, the glycoprotein tag is encoded by the sequence of SEQ ID NO: 826. The glycoprotein tag may be fused, in some embodiments, to the N-terminus or the C-terminus of the casein protein.

[301] Illustrative expression cassettes for expressing a gene of interest (GOI; e.g., a casein) fused to a glycoprotein tag are provided in **FIG. 25A-25F**. In some embodiments, an expression cassette comprises a promoter, a signal peptide, a glycoprotein tag, a GOI (e.g., a casein) and a terminator (See **FIG. 25A**). In some embodiments, an expression cassette comprises a promoter, a signal peptide, a GOI, a glycoprotein tag, and a terminator. (See **FIG. 25B**). In some embodiments, an expression cassette comprises the GmSeed 2 promoter (SEQ ID NO: 813), the pat21ss signal peptide (SEQ ID NO: 823), a (SP)11 glycoprotein tag (SEQ ID NO: 825), a GOI (e.g., a casein) and the AtHSP/AtUBi10 Terminator (SEQ ID NO: 815, 816) (See **FIG. 25C**). In some embodiments, an expression cassette comprises the GmSeed 2 promoter (SEQ ID NO: 813), the pat21ss signal peptide (SEQ ID NO: 823), a GOI (e.g., a casein), a (SP)11 glycoprotein tag (SEQ ID NO: 825), and the AtHSP/AtUBi10 Terminator (SEQ ID NO: 815,816) (See **FIG. 25D**). In some embodiments, an expression cassette comprises the GmSeed 2 promoter (SEQ ID NO: 813), the sig2 signal peptide (SEQ ID NO: 814), a CD45 tag (SEQ ID NO: 827), a GOI (e.g., a casein), a KDEL sequence, and the AtHSP/AtUBi10 Terminator (SEQ ID NO: 815, 816) (See **FIG. 25E**).

In some embodiments, an expression cassette comprises the GmSeed 2 promoter (SEQ ID NO: 813), the sig2 signal peptide (SEQ ID NO: 814), a GOI (e.g., a casein), a CD45 tag (SEQ ID NO: 827), a KDEL sequence, and the AtHSP/AtUBi10 Terminator (SEQ ID NO: 815, 816) (See **FIG. 25F**).

[302] Following protein synthesis, many eukaryotic proteins undergo post-translational modification (PTM). These modifications may be for example, the covalent addition of a function group, and contributes to protein diversity and function. Examples of PTMs include, but are not limited to, phosphorylation, glycosylation, ubiquitination, nitrosylation, methylation, acetylation, and lipidation. The proteins within milk also undergo PTM (Greenberg et al., "Human beta-casein. Amino acid sequence and identification of phosphorylation sites," *J. Biol. Chem.*, 1984, 259(8):5132-5138, Imafidon et al., "Isolation, purification, and alteration of some functional groups of major milk proteins: a review," *Crit. Rev. Food. Sci. Nutr.* 37(7):663-689, 1997). For example, alpha and beta caseins are phosphorylated, and kappa casein is glycosylated. It has been reported that caseins assemble in a colloidal complex with calcium phosphate and other minerals.

[303] In some embodiments, a casein protein expressed in a plant cell comprises different post-translational modifications relative to the same casein protein expressed by a mammalian cell. In some embodiments, a casein protein expressed in a plant cell does not comprise any post-translational modifications. In some embodiments, a casein protein expressed in a plant cell has reduced phosphorylation compared to the same casein protein expressed in a mammalian cell. In some embodiments, a casein protein expressed in a plant cell has increased phosphorylation compared to the same casein protein expressed in a mammalian cell.

[304] In some embodiments, the compositions and methods described herein can be used to produce a casein protein that does not comprise any post-translational modifications. In some embodiments, the compositions and methods described herein can be used to produce a casein protein that is substantially free of phosphorylation. In some embodiments, the compositions and methods described herein can be used to produce a casein protein in a plant cell that comprises substantially the same level of post-translational modifications relative to the same casein protein expressed in a mammalian cell. In some embodiments, the compositions and methods described herein can be used to produce a casein protein that comprises substantially the same level of phosphorylation relative to the same casein protein expressed in a mammalian cell. For example,

in some embodiments, a casein protein expressed in a plant cell may comprise at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95% of the number of phosphates relative to the same casein protein expressed in a mammalian cell.

Methods for Producing Recombinant Milk Proteins, including Casein proteins

[305] The recombinant milk proteins (e.g., casein proteins) described herein may be produced in a number of non-mammalian species, including for example, plants and microorganisms such as yeast and bacteria.

[306] The recombinant casein proteins may be expressed in one or more non-mammalian cells using genetic sequences (e.g., DNA or RNA sequences) isolated or derived from cow (*bos taurus*), goat (*capra hircus*), sheep (*ovis aries*), water buffalo (*bubalus bubalis*), dromedary camel (*camelus dromedaries*), bactrian camel (*camelus bactrianus*), wild yak (*bos mutus*), horse (*equus caballus*), donkey (*equus asinus*), reindeer (*rangifer tarandus*), Eurasian elk (*alces alces*), alpaca (*vicugna pacos*), zebu (*bos indicus*), llama (*lama glama*), or human (*homo sapiens*). In some embodiments, a genetic sequence used to encode the recombinant casein has at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with the genetic sequence used to encode a casein protein in one or more of cow (*bos taurus*), goat (*capra hircus*), sheep (*ovis aries*), water buffalo (*bubalus bubalis*), dromedary camel (*camelus dromedaries*), bactrian camel (*camelus bactrianus*), wild yak (*bos mutus*), horse (*equus caballus*), donkey (*equus asinus*), reindeer (*rangifer tarandus*), eurasian elk (*alces alces*), alpaca (*vicugna pacos*), zebu (*bos indicus*), llama (*lama glama*), or human (*homo sapiens*). In some embodiments, the recombinant casein protein expressed in a non-mammalian cell has at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with a casein protein from one or more of cow (*bos taurus*), goat (*capra hircus*), sheep (*ovis aries*), water buffalo (*bubalus bubalis*), dromedary camel (*camelus dromedaries*), bactrian camel (*camelus bactrianus*), wild yak (*bos mutus*), horse (*equus caballus*), donkey (*equus asinus*), reindeer (*rangifer tarandus*), eurasian elk (*alces alces*), alpaca (*vicugna pacos*), zebu (*bos indicus*), llama (*lama glama*), or human (*homo sapiens*).

[307] When expressed in a plant, the recombinant casein proteins may be extracted using standard methods known in the art. For example, the casein proteins may be extracted using solvent or aqueous extraction or using phenol extraction. Once extracted, the casein proteins may be maintained in a buffered environment (e.g., Tris, MOPS, HEPES), in order to avoid sudden changes in the pH. The casein proteins may also be maintained at a particular temperature, such as 4°C. One or more additives may be used to aid the extraction process (e.g., salts, protease/peptidase inhibitors, osmolytes, reducing agents, etc.)

Protein Co-Expression in Plants

[308] Another way to increase accumulation of one or more recombinant proteins, such as milk proteins, in a plant cell is to co-express the protein with a second protein, such as a protein capable of forming a protein body (e.g., a prolamin). Without being bound by any theory, it is believed that co-expressing a milk protein and a prolamin protein in a plant cell will cause protein body formation in the plant cell, wherein the milk protein gets sequestered into and/or associated with the protein body. This protects the milk protein from degradation by one or more proteases and increases accumulation thereof in the plant cell.

[309] In some embodiments, two or more recombinant proteins may be co-expressed in a plant cell. In some embodiments, one of the two or more recombinant proteins is a milk protein (e.g., casein protein). In some embodiments, the milk protein is selected from the group consisting of: α -S1 casein, α -S2 casein, β -casein, κ -casein, para- κ -casein, β -lactoglobulin, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, and an immunoglobulin. In some embodiments, the milk protein is β -casein or β -lactoglobulin.

[310] In some embodiments, one of the two or more proteins is a protein capable of forming a protein body. For example, in some embodiments, one of the two or more proteins is a prolamin (e.g., zein and/or canein). In some embodiments, the prolamin is selected from the group consisting of: gliadin, a hordein, a secalin, a zein, a kafirin, and an avenin. In some embodiments, the protein capable of forming a protein body is a hydrophobin or an elastin-like protein. In some embodiments, at least two proteins are co-expressed in a plant cell (e.g., a casein protein and a prolamin). In some embodiments, the at least two proteins are casein and zein (e.g., gamma-zein). In some embodiments, the at least two proteins are casein and canein.

[311] In some embodiments, a method for expressing a first recombinant protein in a cell comprises: (i) contacting the cell with a vector encoding a first recombinant protein, and (ii) contacting the cell with a vector encoding a second recombinant protein, wherein the second recombinant protein is capable of forming a protein body (e.g., a prolamin.) In some embodiments, the first recombinant protein is a casein protein, such as a milk protein.

[312] A milk protein (e.g., a casein protein) may, in some embodiments, be co-expressed with a protein capable of forming a protein body (e.g., a prolamin) in a transgenic plant. In some embodiments, co-expressing a milk protein (e.g., a casein protein) with a protein capable of forming a protein body (e.g., a prolamin) in a transgenic plant leads to accumulation of the milk protein in an amount of at least 1%, at least 1.5%, at least 2%, at least 2.5%, at least 3%, at least 3.5%, at least 4%, at least 4.5%, at least 5%, at least 5.5%, at least 6%, at least 6.5%, at least 7%, at least 7.5%, at least 8%, at least 8.5%, at least 9%, at least 9.5%, at least 10%, at least 10.5%, at least 11%, at least 11.5%, at least 12%, at least 12.5%, at least 13%, at least 13.5%, at least 14%, at least 14.5%, at least 15%, at least 15.5%, at least 16%, at least 16.5%, at least 17%, at least 17.5%, at least 18%, at least 18.5%, at least 19%, at least 19.5%, at least 20%, or more of total protein weight of soluble protein extractable from the plant.

[313] Illustrative constructs for co-expressing a milk protein (e.g., a casein protein) and a protein capable of inducing formation of a protein body in a plant cell are provided in **FIG. 26A-26G**. In some embodiments, a construct comprises (i) a first expression cassette comprising a promoter, a signal peptide, a Gene of Interest (e.g., a casein protein) and a terminator, and (ii) a second expression cassette comprising a promoter, a signal peptide, a protein that induces protein body formation, and a terminator (See **FIG. 26A**). In some embodiments, a construct comprises (i) a first expression cassette comprising a promoter, a signal peptide, a Gene of Interest (e.g., a casein protein) and a terminator, and (ii) a second expression cassette comprising a promoter, a signal peptide, a prolamin, and a terminator (See **FIG. 26B**). In some embodiments, a construct comprises (i) a first expression cassette comprising a promoter, a signal peptide, a Gene of Interest (e.g., a casein protein) and a terminator, and (ii) a second expression cassette comprising a promoter, a signal peptide, a zein, and a terminator (See **FIG. 26C**). In some embodiments, a construct comprises (i) a first expression cassette comprising a promoter, a signal peptide, a Gene of Interest (e.g., a casein protein) and a terminator, and (ii) a second expression cassette comprising a

promoter, a signal peptide, a casein, and a terminator (See **FIG. 26D**). In some embodiments, a construct comprises (i) a first expression cassette comprising a promoter, a signal peptide, a Gene of Interest (e.g., a casein protein) and a terminator, and (ii) a second expression cassette comprising a promoter, a signal peptide, a hydrophobin, and a terminator (See **FIG. 26E**). In some embodiments, a construct comprises (i) a first expression cassette comprising a promoter, a signal peptide, a Gene of Interest (e.g., a casein protein) and a terminator, and (ii) a second expression cassette comprising a promoter, a signal peptide, an elastin-like protein, and a terminator (See **FIG. 26F**). In some embodiments, a construct comprises (i) a first expression cassette comprising a GmSeed2 promoter, a Sig2 signal peptide, a Gene of Interest (e.g., a casein protein) and a AtHSP/AtUbi10 terminator, and (ii) a second expression cassette comprising a GmSeed 12 promoter, a Coixss signal peptide, a protein that induces protein body formation, and a EU Term/Tm6 terminator (See **FIG. 26G**). An illustrative binary vector for use in coexpressing a casein and a protein that can induce protein body formation is provided in **FIG. 27**.

[314] In some embodiments, a milk protein (e.g., a casein protein) can be co-expressed with one or more proteins capable of adding or removing a post-translational modification to/from a milk protein. For example, in some embodiments, the milk protein may be co-expressed with one or more of a kinase, a phosphatase, or a glycosyltransferase. In some embodiments, the milk protein is co-expressed with a kinase. The kinase may be for example, a kinase that phosphorylates Ser-X-Glu/pSer motifs. In some embodiments, the kinase may be a kinase in the family 20C, such as the Fam20C kinase. In some embodiments, the kinase may be a fragment or derivative of the Fam20C kinase, such as a truncated Fam20C comprising amino acids 94-586 of the native protein. In some embodiments, the kinase comprises amino acids 94-586 of SEQ ID NO: 821, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto. In some embodiments, the kinase is encoded by the sequence of SEQ ID NO: 820, or a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical thereto.

[315] Illustrative expression cassettes that may be used to co-express a milk protein (e.g., a casein protein) with a kinase (or other enzyme capable of adding/removing a PTM) are shown in **FIG. 24A-24E**. In some embodiments, a construct for co-expression of a milk protein in a cell comprises: (i) a first expression cassette comprising a promoter, a signal peptide, a Gene of Interest (GOI, e.g., a casein protein) and a terminator, and (ii) a second expression cassette comprising a

promoter, a 5'UTR, a signal peptide, a Gene of Interest (GOI, e.g., a kinase), and a terminator (See **FIG. 24B**). In some embodiments, a construct for co-expression of a milk protein in a cell comprises: (i) a first expression cassette comprising a promoter, a signal peptide, a Gene of Interest (GOI, e.g., a casein protein) and a terminator, and (ii) a second expression cassette comprising a promoter, a 5'UTR, a signal peptide, a Gene of Interest (GOI, e.g., a kinase in the 20C family), and a terminator (See **FIG. 24A**). In some embodiments, a construct for co-expression of a milk protein in a cell comprises: (i) a first expression cassette comprising a promoter, a signal peptide, a Gene of Interest (GOI, e.g., a casein protein) and a terminator, and (ii) a second expression cassette comprising a promoter, a 5'UTR, a signal peptide, a Gene of Interest (GOI, e.g., a Fam20C kinase), and a terminator (See **FIG. 24C**). In some embodiments, a construct for co-expression of a milk protein in a cell comprises: (i) a first expression cassette comprising a promoter, a signal peptide, a Gene of Interest (GOI, e.g., a casein protein) and a terminator, and (ii) a second expression cassette comprising a promoter, a 5'UTR, a signal peptide, a Gene of Interest (GOI, e.g., a truncated Fam20C kinase), and a terminator (See **FIG. 24D**). In some embodiments, the promoter may be the GmSeed2 promoter (SEQ ID NO: 813) or the PvPhas promoter (SEQ ID NO: 817). In some embodiments, the promoter may be the Sig2 signal peptide (SEQ ID NO: 814) or the sig10 signal peptide (SEQ ID NO: 819). In some embodiments, the terminator may be the AtHSP/AtUbi10 Terminator (SEQ ID NO: 815, 816) or the 3arc Terminator (SEQ ID NO: 822). In some embodiments, the 5'UTR may be the Arc 5'UTR (SEQ ID NO: 818). In some embodiments, the construct for co-expression of a milk protein in a cell comprises the construct of **FIG. 24E**. An illustrative binary vector is provided in **FIG. 23**.

[316] In some embodiments, a milk protein (e.g., a casein protein) can be co-expressed with one or more proteins capable of inhibiting a protease. Illustrative plant proteins that may be used to inhibit one or more proteases are shown above in Table 4. In some embodiments, a milk protein may be co-expressed with any one of the proteins shown in Table 4. In some embodiments, a milk protein is co-expressed with a protein that comprises the sequence of any one of SEQ ID NO: 840, 842, 844, 846, 848 or 850. In some embodiments, a milk protein may be co-expressed with a protein having a sequence with at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to any one of SEQ ID NO: 840, 842, 844, 846, 848 or 850. In some embodiments, the milk protein may

be co-expressed with a protein having the sequence of any one of SEQ ID NO: 840, 842, 844, 846, 848 or 850 plus at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, or more amino acid substitutions.

[317] In some embodiments, protein co-expression can be utilized to reduce or prevent degradation of the one or more proteins in the plant cell, such as protease-mediated degradation in the plant cell. In some embodiments, the protein-co-expression is useful to reduce or prevent degradation of one or more milk proteins by proteases in a plant cell. In some embodiments co-expressing one or more milk proteins (e.g., casein protein) and a prolamin (e.g., a canein or a zein) may lead to the formation of a protein body in a seed of a plant. In some embodiments, the one or more milk proteins can be sequestered in and/or associated with the protein body, which in turn partially or fully shields the one or more milk proteins from degradation by plant cell proteases thereby allowing for accumulation of the one or more milk proteins. In some embodiments, the one or more milk proteins can be sequestered in the protein body, which in turn may protect a plant cell from potential toxic effects of recombinant proteins, such as any toxic effects of the one or more proteins.

[318] In some embodiments, protein co-expression is effective in increasing at least one of concentration, stability, or expression of one or more proteins in a plant cell. In some embodiments, protein co-expression is effective in increasing concentration of one or more proteins in a plant cell as determined by detecting the amount of the one or more protein in the plant cell. In some embodiments, protein co-expression is effective in increasing stability of one or more proteins in a plant cell. Increased stability can be determined by detecting persistence of the one or more proteins in the plant cell over time or detecting a level of degradation. In some embodiments, protein co-expression is effective in increasing expression of one or more proteins in a plant cell. Increased expression can be determined by measuring protein level and/or accumulation in the plant cell. In some embodiments, protein co-expression is effective in increasing at least one of: concentration, stability, or expression of one or more proteins by at least about 1-fold, 10-fold, 19-fold, 28-fold, 37-fold, 46-fold, 55-fold, 64-fold, 73-fold, 82-fold, 91-fold, 100-fold, 109-fold, 118-fold, 127-fold, 136-fold, 145-fold, 154-fold, 163-fold, 172-fold, 181-fold, 190-fold, 199-fold, 208-fold, 217-fold, 226-fold, 235-fold, 244-fold, 253-fold, 262-fold, 271-fold, 280-fold, 289-fold, 298-

fold, or up to about 300-fold as compared to an otherwise comparable method lacking the protein co-expression. In some embodiments, protein co-expression is effective in increasing at least one of concentration, stability, or expression of one or more proteins in a plant cell by at least about 1-fold to 10-fold, 5-fold to 30-fold, 20-fold to 50-fold, 40-fold to 100-fold, or 100-fold to 200-fold as compared to an otherwise comparable method lacking the protein co-expression.

[319] In some embodiments, protein co-expression is effective in reducing toxicity of recombinant expression of the one or more proteins in a plant cell. In some embodiments, protein co-expression is effective in reducing toxicity of recombinant expression of one or more proteins in a plant cell by at least about 1-fold, 10-fold, 19-fold, 28-fold, 37-fold, 46-fold, 55-fold, 64-fold, 73-fold, 82-fold, 91-fold, 100-fold, 109-fold, 118-fold, 127-fold, 136-fold, 145-fold, 154-fold, 163-fold, 172-fold, 181-fold, 190-fold, 199-fold, 208-fold, 217-fold, 226-fold, 235-fold, 244-fold, 253-fold, 262-fold, 271-fold, 280-fold, 289-fold, 298-fold, or up to about 300-fold as compared to an otherwise comparable method lacking the protein co-expression. In some embodiments, protein co-expression is effective in reducing toxicity associated with recombinant expression of one or more proteins in a plant cell by at least about 1-fold to 10-fold, 5-fold to 30-fold, 20-fold to 50-fold, 40-fold to 100-fold, or 100-fold to 200-fold as compared to an otherwise comparable method lacking the protein co-expression.

[320] In some embodiments, protein co-expression may be achieved via transformation of a composition comprising one or more vectors encoding the one or more proteins into a plant cell. In some embodiments, one or more vectors are binary *agrobacterium* vectors. In some embodiments, one or more vectors encodes for one or more protein sequences. In some embodiments, a single vector encodes for two or more protein sequences. In some embodiments, two or more vectors are used to introduced two or more sequences into a plant cell. In some embodiments, a vector encodes for a milk protein (e.g., casein protein) and a prolamin (e.g., a canein or a zein). In some embodiments, a vector encodes for a milk protein and a protein capable of forming a protein body. In some embodiments a first vector encodes for a milk protein and a second vector encodes for a prolamin. In some embodiments, a first vector encodes for a milk protein and a second vector encodes for a prolamin. Also provided are compositions that comprise one or more vectors described herein.

Food Compositions Comprising a Fusion Protein or a Protein Derived Therefrom

[321] The fusion proteins, recombinant proteins, and transgenic plants described herein may be used to prepare food compositions. The fusion protein may be used directly to prepare the food composition (i.e., used in the form of a fusion protein), or the fusion protein may first be separated into its constituent proteins. For example, in some embodiments, a food composition may comprise (i) a fusion protein, (ii) a milk protein (structured or unstructured) or (iii) a non-milk protein, such as a structured mammalian, avian, or plant protein.

[322] More specifically, the present disclosure provides alternative dairy compositions, solid phase protein-stabilized emulsions (including cheese compositions), and colloidal suspensions, each comprising one or more casein proteins. The casein proteins may be isolated or recombinant and may be selected from the group consisting of kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein and alpha-S2-casein. The compositions, emulsions, or suspensions described herein may be used to produce food compositions (e.g., cheese, yogurt, ice cream, etc.) that have organoleptic properties similar to traditional animal-derived dairy compositions. For example, the food compositions described herein may have one or more characteristics of a traditional animal-derived dairy composition, such as taste, aroma, appearance, handling, mouthfeel, density, structure, texture, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess and emulsification. The food compositions described herein offer a sustainable, environmentally-friendly, cruelty-free alternative to traditional animal-derived dairy compositions.

[323] In some embodiments, the alternative dairy compositions, solid phase, protein-stabilized emulsions, and colloidal suspensions comprising recombinant casein proteins have non-mammalian PTMs. In some embodiments, the recombinant casein proteins are not phosphorylated or glycosylated. In some embodiments, the recombinant casein proteins have an alternative PTM pattern, as compared to naturally occurring casein proteins.

[324] PTMs have been reported to be important for the casein micelle structure, which determines the physical properties of milk. Unexpectedly, the recombinant proteins described herein are still able to confer to the compositions described herein one or more organoleptic properties similar to animal-derived dairy compositions, such as taste, appearance, mouthfeel,

structure, texture, density, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess, and emulsification.

[325] Food compositions, including alternative dairy compositions, solid phase protein-stabilized emulsions, and colloidal suspensions, are described in more detail below.

Solid Phase Protein-Stabilized Emulsions

[326] Provided herein are solid phase, protein-stabilized emulsions comprising at least one milk protein. For example, in some embodiments, a solid phase, protein-stabilized emulsion comprises at least one casein protein. In some embodiments, a protein-stabilized emulsion comprises at least one recombinant casein protein. In some embodiments, a protein-stabilized emulsion comprises at least one plant-expressed casein protein. In some embodiments, a protein-stabilized emulsion comprises at least one casein protein isolated from milk (e.g., bovine milk). In some embodiments, the protein-stabilized emulsion is a cheese composition.

[327] In some embodiments, a solid-phase protein stabilized protein emulsion comprises only one casein protein. In some embodiments, the one casein protein is recombinant beta-casein protein.

[328] In some embodiments, a solid-phase protein stabilized protein emulsion comprises only two casein proteins. In some embodiments, the two casein proteins are recombinant beta-casein protein and kappa-casein protein. In some embodiments, the two casein proteins are recombinant beta-casein protein and para-kappa-casein protein. In some embodiments, the two casein proteins are recombinant beta-casein protein and alpha-S1-casein protein. In some embodiments, the two casein proteins are recombinant beta-casein protein and alpha-S2-casein protein.

[329] In some embodiments, a solid-phase, protein stabilized emulsion comprises only three casein proteins. In some embodiments, the three casein proteins are recombinant beta-casein, kappa-casein, and para-kappa-casein. In some embodiments, the three casein proteins are recombinant beta-casein, kappa-casein, and alpha-S1-casein. In some embodiments, the three casein proteins are recombinant beta-casein, kappa-casein, and alpha-S2-casein. In some embodiments, the three casein proteins are recombinant beta-casein, para-kappa-casein, and alpha-

S1-casein. In some embodiments, the three casein proteins are recombinant beta-casein, para-kappa-casein, and alpha-S2-casein.

[330] In some embodiments, a solid-phase, protein stabilized emulsion comprises only four casein proteins. In some embodiments, one of the four casein proteins is recombinant beta-casein.

[331] The casein proteins used in the solid-phase, protein-stabilized emulsions described herein may be selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein and alpha-S2-casein. In some embodiments, the solid-phase protein stabilized emulsions may comprise, in addition to the casein protein(s), one or more additional milk proteins. In some embodiments, the solid-phase protein stabilized emulsions may comprise, in addition to the casein protein(s), one or more plant proteins.

[332] In some embodiments, the emulsion has a firmness of at least 150 grams. In some embodiments, the emulsion has a melting point of about 35°C to about 100°C. In some embodiments, the emulsion has an ability to stretch to at least 3 cm in length without breaking. In some embodiments, the emulsion has a firmness of at least 150 grams and a melting point of about 35°C to about 100°C. In some embodiments, the emulsion has a firmness of at least 150 grams and an ability to stretch to at least 3 cm in length without breaking. In some embodiments, the emulsion has a melting point of about 35°C to about 100°C and an ability to stretch to at least 3 cm in length without breaking. In some embodiments, the emulsion has a firmness of at least 150 grams, a melting point of about 35°C to about 100°C, and an ability to stretch to at least 3 cm in length without breaking.

[333] Firmness, also referred to herein as hardness, may be measured by a number of methods known in the art, such as by compression, or using an instrument such as the Instron Testing Machine (A.H. Chen et al., Textural analysis of cheese, 1979, *J. Dairy Sci.* 62:901-907). For example, a cylindrical-shaped sample of a solid-phase, protein stabilized emulsion may be compressed from 50% to 100% relative to its original height and/or width. The cylindrical shaped-sample may have a height in the range of about 1 to about 10 cm, or more, and a diameter in the range of about 1 to about 10 cm, or more. The compression may occur at a predetermined temperature, such as a temperature in the range of about 0°C to about 5°C, about 5°C to about 10°C, about 10°C to about 20°C, about 15°C to about 25°C, about 20°C to about 25°C, about 25°C

to about 25°C. In some embodiments, firmness may be determined by compressing a cylindrical-shaped sample having a height of about 3 cm, and a diameter of about 3 cm may be compressed to a height of 1.5 cm at 5°C. The compositions described herein may have a firmness in the range of about 50 to 100 grams, about 100 to about 150 grams, about 150 grams to about 200 grams, about 200 to about 300 grams, about 300 grams to about 400 grams, about 400 grams to about 500 grams, about 500 grams to about 600 grams, about 600 grams to about 700 grams, about 700 grams to about 800 grams, about 800 grams to about 900 grams, about 900 grams to 1 kilogram, or more.

[334] Stretch ability may be analyzed by standard assays known in the art. For example, stretch ability may be determined by heating a 100 gram mass of an emulsion at a temperature of 225°C for 4 minutes, cooling to about 90°C, and then pulling with a fork placed beneath the mass. Other methods to test stretch ability are well known in the art. See for example, Fife R.L et al, Test for measuring the stretch ability of melted cheese, 2002, *J. Dairy Sci.* 85(12):3539-3545.

[335] In some embodiments, the recombinant casein protein may be expressed by a plant (i.e., it is a “plant-expressed” protein). In some embodiments, the recombinant protein may be expressed in a monocot, such as turf grass, maize (corn), rice, oat, wheat, barley, sorghum, orchid, iris, lily, onion, palm, or duckweed. In some embodiments, the recombinant casein protein may be expressed in a dicot, such as Arabidopsis, tobacco, tomato, potato, sweet potato, cassava, alfalfa, lima bean, pea, chick pea, soybean, carrot, strawberry, lettuce, oak, maple, walnut, rose, mint, squash, daisy, Quinoa, buckwheat, mung bean, cow pea, lentil, lupin, peanut, fava bean, French beans (i.e., common beans), mustard, or cactus. In some embodiments, the recombinant casein protein may be expressed in a non-vascular plant selected from moss, liverwort, hornwort or algae. In some embodiments, the recombinant casein protein may be expressed in a vascular plant reproducing from spores (e.g., a fern). In some embodiments, the recombinant casein protein is expressed in a soybean plant.

[336] In some embodiments, the recombinant casein protein is expressed in a microorganism. Microorganisms used for recombinant protein production are well known in the art (see for example, Ferrer-Miralles et al., Bacterial cell factories for recombinant protein production; expanding the catalogue, 2013, *Microb Cell Fact.* 2013; 12:113). In some embodiments, the recombinant casein protein is expressed in a yeast or a bacterium (i.e., it is “yeast-expressed” or

“bacterial-expressed”). For example, the recombinant casein protein may be expressed in bacteria such as *Escherichia coli*, *Caulobacter crescentus*, *Rodhobacter sphaeroides*, *Pseudoalteromonas haloplanktis*, *Shewanella* sp., *Pseudomonas putida*, *P. aeruginosa*, *P. fluorescens*, *Halomonas elongate*, *Chromohalobacter salexigens*, *Streptomyces lividans*, *S. griseus*, *Nocardia lactamdurans*, *Mycobacterium smegmatis*, *Corynebacterium glutamicum*, *C. ammoniagenes*, *Brevibacterium lactofermentum*, *Bacillus subtilis*, *B. brevis*, *B. megaterium*, *B. licheniformis*, *B. amyloliquefaciens*, *Lactococcus lactis*, *L. plantarum*, *L. casei*, *L. reuteri*, or *L. gasseri*.

[337] In some embodiments, the recombinant casein protein is expressed in a eukaryotic microorganism, such as *Saccharomyces* spp., *Kluyveromyces* spp., *Pichia* spp., *Aspergillus* spp., *Tetrahymena* spp., *Yarrowia* spp., *Hansenula* spp., *Blastobotrys* spp., *Candida* spp., *Zygosaccharomyces* spp., *Debrayomyces* spp., *Fusarium* spp., and *Trichoderma* spp.

[338] In some embodiments, the solid-phase, protein stabilized emulsions comprise ash. In some embodiments, the solid-phase, protein stabilized emulsions comprise at least one lipid and at least one salt. “Lipid” means any of a class of molecules that are soluble in nonpolar solvents (such as ether and hexane) and relatively or completely insoluble in water. Lipid molecules are typically composed of long hydrocarbon tails that are hydrophobic in nature. Examples of lipids include fatty acids (saturated and unsaturated); glycerides or glycerolipids (such as monoglycerides, diglycerides, triglycerides or neutral fats, and phosphoglycerides or glycerophospholipids); and nonglycerides (sphingolipids, tocopherols, tocotrienols, sterol lipids including cholesterol and steroid hormones, prenol lipids including terpenoids, fatty alcohols, waxes, and polyketides).

[339] Examples of lipids that may be included in the solid-phase, protein stabilized emulsion include, for example, dairy fats or vegetable oils such as palm oil or palm kernel oil, butter oil, anhydrous milkfat, soybean oil, corn oil, rapeseed oil, canola oil, sunflower oil, safflower oil, coconut oil, rice bran oil, olive oil, sesame oil, flaxseed oil, hemp oil, cottonseed oil, peanut oil, almond oil, beech nut oil, brazil nut oil, cashew oil, hazelnut oil, macadamia oil, mongongo nut oil, pecan oil, pine nut oil, pistachio oil, walnut oil, pumpkin seed oil, grapefruit seed oil, lemon oil, apricot oil, apple seed oil, argan oil, avocado oil, or orange oil. In some embodiments, the solid-phase, protein stabilized emulsion comprises butter or margarine.

[340] Examples of salts that may be included in the emulsion include, but are not limited to, magnesium chloride, sodium chloride, calcium chloride, sodium phosphates and trisodium citrate.

[341] In some embodiments, the emulsion comprises at least two plant-expressed casein proteins each selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein and alpha-S2-casein. In some embodiments, the emulsion comprises at least three plant-expressed casein proteins each selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein and alpha-S2-casein. In some embodiments, the emulsion comprises at least four plant-expressed casein proteins each selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein and alpha-S2-casein. In some embodiments, the emulsion comprises at least one additional mammalian or plant protein that is not a casein protein.

[342] Examples of combinations of casein, mammalian, and/or plant proteins that may be used in the solid phase, protein stabilized emulsions are shown below in Table 12. The casein or casein protein combination shown in Column 1 may be combined with one or more of the mammalian proteins listed in Column 2, and/or one or more of the plant proteins listed in Column three. In some embodiments, the solid-phase protein stabilized emulsions described herein comprise proteins from Column 1, and do not include any proteins from Column 2 or Column 3.

Table 12: Example combinations of casein, mammalian, and/or plant proteins

Casein proteins (Column 1)	Mammalian proteins (Column 2)	Plant proteins (Column 3)
κ -casein	Alpha-lactalbumin	Oleosins
Para- κ -casein	Beta-lactoglobulin	Leghemoglobin
β -casein	Albumin	Extensin-like protein family
α -S1-casein	Lysozyme	Prolamine
α -S2-casein	Collagen family	Glutenin
κ -casein & para- κ -casein	Hemoglobin	Gamma-kafirin
κ -casein & β -casein		preprotein
κ -casein & α -S1-casein		Alpha globulin
κ -casein & α -S2-casein		Basic 7S globulin precursor
Para- κ -casein & β -casein		2S albumin
Para- κ -casein & α -S1-casein		Beta-conglycinins
Para- κ -casein & α -S2-casein		Glycinins
β -casein & α -S1-casein		
β -casein & α -S2-casein		

Casein proteins (Column 1)	Mammalian proteins (Column 2)	Plant proteins (Column 3)
α -S1-casein & α -S2-casein		Canein
κ -casein, para- κ -casein, & β -casein		Zein
κ -casein, para- κ -casein, & α -S1-casein		Patatin
κ -casein, para- κ -casein, & α -S2-casein		Kunitz-Trypsin inhibitor
Para- κ -casein, β -casein, & α -S1-casein		Bowman-Birk inhibitor
Para- κ -casein, β -casein, & α -S2-casein		Cystatine
β -casein, α -S1-casein, & α -S2-casein		
κ -casein, β -casein, & α -S1-casein		
κ -casein, β -casein, & α -S2-casein		
κ -casein, α -S1-casein & α -S2-casein		
para- κ -casein, α -S1-casein & α -S2-casein		
κ -casein, para- κ -casein, β -casein, α -S1-casein		
κ -casein, para- κ -casein, β -casein, & α -S2-casein		
Para- κ -casein, β -casein, α -S1-casein, & α -S2-casein		
κ -casein, β -casein, α -S1-casein, & α -S2-casein		
κ -casein, para- κ -casein, α -S1-casein & α -S2-casein		

[343] In some embodiments, the emulsion further comprises plant protein. For example, in some embodiments, the emulsion comprises protein from a legume, such as, for example, soybeans, chickpeas, kidney beans, black beans, pinto beans, green peas, and lentils. In some embodiments, the emulsion comprises protein from a grain, such as, for example, wheat, millet, barley, oats, rice, spelt, teff, amaranth, and quinoa. In some embodiments, the emulsion comprises protein from nuts, hempseed, chia seed, nutritional yeast, or spirulina. In some embodiment, the emulsion further comprises protein from potato. In some embodiments, the emulsion further comprises protein from a plant of the family *Fabaceae*.

[344] In some embodiments, the emulsion has a pH of about 5.0 to about 6.7. In some embodiments, the emulsion has a pH of about 5.2 to about 5.9. In some embodiments, the emulsion has a pH of about 5.0, about 5.1, about 5.2, about 5.3, about 5.4, about 5.5, about 5.6, about 5.7, about 5.8, about 5.9, about 6.0, about 6.1, about 6.2, about 6.3, about 6.4, about 6.5, about 6.6, about 6.7, about 6.8, or about 6.9.

[345] In some embodiments, the emulsion may further comprise one or more additional agents, such as an edible gum, starch, and/or gelling agent. Examples of edible gums include, but are not limited to, curdian, locust bean gum, carrageenan, gellan gum, xanthan gum, guar gum, agar agar, gelatin, sodium alginate, or combinations thereof. Examples of starch include, but are not limited to, potato starch, corn starch, rice flour, pea flour, modified starch, and combinations thereof. Examples of gelling agents include, but are not limited to, pectin, alginate, vegetable gums, gelatin, agar, methyl cellulose, enzymes (transglutaminase) and hydroxypropylmethyl cellulose. In some embodiments, the emulsion may further comprise an acid or a base, such as lemon juice, lactic acid, acetic acid, citric acid, sodium citrate, sodium orthophosphates, sodium pyrophosphates, sodium polyphosphates, potassium citrate, potassium orthophosphates, potassium pyrophosphates, sorbic acid, potassium sorbate, tartaric acid, and sodium aluminum phosphate.

[346] In some embodiments, the emulsion does not contain an organoleptically functional amount of beta-lactoglobulin. In some embodiments, the emulsion may comprise beta-lactoglobulin in the amount of about 0.01% (w/v) to about 0.1% (w/v), about 0.1% (w/v) to about 0.5% (w/v), about 0.5% (w/v) to about 1.0% (w/v), about 1.0% (w/v) to about 2% (w/v), about 2% (w/v) to about 3% (w/v), about 3% (w/v) to about 5% (w/v), about 5% (w/v) to about 10% (w/v), about 10% (w/v) to about 20% (w/v), about 20% (w/v) to about 40% (w/v), or more, of the emulsion.

[347] As used herein, an “organoleptically functional amount of beta-lactoglobulin” refers to an amount of beta-lactoglobulin that significantly impacts one or more organoleptic properties of the composition. An organoleptic property is “significantly impacted” if it represents a change that can be detected by a human, using one or more of the senses taste, sight, smell, and/or touch. In some embodiments, a solid-phase, protein stabilized emulsion that does not comprise an organoleptically functional amount of beta-lactoglobulin may comprise only trace amounts of beta-lactoglobulin. In some embodiments, the emulsion may comprise beta-lactoglobulin in the range of about 0.01% (w/v) to about 0.1% (w/v), about 0.1% (w/v) to about 0.5% (w/v), about 0.5% (w/v) to about 1.0% (w/v), about 1.0% (w/v) to about 2% (w/v), about 2% (w/v) to about 3% (w/v), about 3% (w/v) to about 5% (w/v), about 5% (w/v) to about 10% (w/v), about 10% (w/v) to about 20% (w/v), about 20% (w/v) to about 40% (w/v), or more, of the emulsion.

[348] In some embodiments, a solid phase, protein-stabilized emulsion comprises one plant-expressed casein protein selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein; wherein the emulsion does not contain any additional casein proteins; and wherein the emulsion has at least one of the following characteristics: i) a firmness of at least 150 grams; ii) a melting point of about 35°C to about 100°C; or iii) ability to stretch to at least 3 cm in length without breaking. In some embodiments, the emulsion further comprises at least one lipid and at least one salt. In some embodiments, the plant-expressed casein protein is expressed in a soybean plant. In some embodiments, the emulsion has a pH of about 5.2 to about 5.9. In some embodiments, the emulsion does not contain an organoleptically functional amount of beta-lactoglobulin. In some embodiments, the emulsion may comprise beta-lactoglobulin in the amount of about 0.01% (w/v) to about 0.1% (w/v), about 0.1% (w/v) to about 0.5% (w/v), about 0.5% (w/v) to about 1.0% (w/v), about 1.0% (w/v) to about 2% (w/v), about 2% (w/v) to about 3% (w/v), about 3% (w/v) to about 5% (w/v), about 5% (w/v) to about 10% (w/v), about 10% (w/v) to about 20% (w/v), about 20% (w/v) to about 40% (w/v), or more.

[349] In some embodiments, a solid phase, protein-stabilized emulsion comprises: a plant-expressed casein protein selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein; and further comprises plant-expressed beta-lactoglobulin; wherein the ratio of the casein protein to the beta-lactoglobulin is about 8:1 to about 1:2. In some embodiments, the emulsion has at least one of the following characteristics: i) a firmness of at least 150 grams; ii) a melting point of about 35°C to about 100°C; or iii) ability to stretch to at least 3 cm in length without breaking. In some embodiments, the emulsion comprises at least at least one additional mammalian or plant protein that is not a casein protein. In some embodiments, the ratio of the casein protein to the beta-lactoglobulin is 1:2. In some embodiments, the ratio of the casein protein to the beta-lactoglobulin is about 2:1. In some embodiments, the emulsion has a pH of about 5.2 to about 5.9.

[350] In some embodiments, a solid-phase protein-stabilized emulsion comprises about 8% (w/v) to about 25% (w/v) total protein, such as about 8% to about 10%, about 10% to about 15%, about 15% to about 20%, or about 20 to about 25% total protein. In some embodiments, a solid-phase protein stabilized emulsion comprises about 1% to about 10% (w/v) total protein. In some embodiments, a solid-phase protein stabilized emulsion comprises about 25% to about 35%, about

35% to about 45%, about 45% to about 55%, about 55% to about 65%, about 65% to about 75% (w/v), or more total protein.

[351] In some embodiments, about 1% to about 5% of the total protein in the solid-phase protein stabilized emulsion is casein protein. In some embodiments, about 5% to about 10% of the total protein in the solid-phase protein stabilized emulsion is casein protein. In some embodiments, about 10% to about 20% of the total protein in the solid-phase protein stabilized emulsion is casein protein. In some embodiments, about 20% to about 30% of the total protein in the solid-phase protein stabilized emulsion is casein protein. In some embodiments, about 30% to about 40% of the total protein in the solid-phase protein stabilized emulsion is casein protein. In some embodiments, about 40% to about 50% of the total protein in the solid-phase protein stabilized emulsion is casein protein. In some embodiments, about 50% to about 60% of the total protein in the solid-phase protein stabilized emulsion is casein protein. In some embodiments, about 60% to about 70% of the total protein in the solid-phase protein stabilized emulsion is casein protein. In some embodiments, about 70% to about 80% of the total protein in the solid-phase protein stabilized emulsion is casein protein. In some embodiments, about 80% to about 90% of the total protein in the solid-phase protein stabilized emulsion is casein protein. In some embodiments, about 90% to about 100% of the total protein in the solid-phase protein stabilized emulsion is casein protein.

[352] In some embodiments, at least 1%, at least 2%, at least 3%, at least 4%, at least 5%, at least 6%, at least 7%, at least 8%, at least 9%, at least 10%, at least 11%, at least 12%, at least 13%, at least 14%, at least 15%, at least 16%, at least 17%, at least 18%, at least 19%, at least 20%, or more of the total protein in the solid-phase protein stabilized emulsion is casein protein.

[353] In some embodiments, about 20% to about 100% of the casein protein in the solid-phase protein-stabilized emulsion is kappa casein. For example, the emulsion may comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% kappa casein. In some embodiments, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the solid-phase protein-stabilized emulsion is kappa casein.

[354] In some embodiments, about 20% to about 100% of the casein protein in the solid-phase protein-stabilized emulsion is para-kappa casein. For example, the emulsion may comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% para-kappa casein. In some embodiments, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the solid-phase protein-stabilized emulsion is para-kappa casein.

[355] In some embodiments, about 20% to about 100% of the casein protein in the solid-phase protein-stabilized emulsion is beta casein. In some embodiments, about 50% to about 100% of the casein protein in the solid-phase protein-stabilized emulsion is beta casein. For example, the emulsion may comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% beta casein. In some embodiments, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the solid-phase protein-stabilized emulsion is beta casein.

[356] In some embodiments, about 20% to about 100% of the casein protein in the solid-phase protein-stabilized emulsion is alpha-S1-casein. In some embodiments, about 50% to about 100% of the casein protein in the solid-phase protein-stabilized emulsion is alpha-S1-casein. For example, the emulsion may comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% alpha-S1-casein. In some embodiments, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the solid-phase protein-stabilized emulsion is alpha-S1-casein.

[357] In some embodiments, about 20% to about 100% of the casein protein in the solid-phase protein-stabilized emulsion is alpha-S2-casein. In some embodiments, about 50% to about 100% of the casein protein in the solid-phase protein-stabilized emulsion is alpha-S2-casein. For

example, the emulsion may comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% alpha-S2-casein. In some embodiments, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the solid-phase protein-stabilized emulsion is alpha-S2-casein.

[358] In some embodiments, a solid-phase protein-stabilized emulsion comprises about 8% (w/v) to about 25% (w/v) total protein, one or more lipids, and one or more salts; wherein at least 4% of the total protein comprises casein proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein; wherein the emulsion has at least one of the following characteristics: i) a firmness of at least 150 grams; ii) a melting point of about 35°C to about 100°C; or iii) ability to stretch to at least 3 cm in length without breaking. In some embodiments, at least 20% to 100% of the casein protein is kappa casein. In some embodiments, at least 20% to 100% of the casein protein is para-kappa casein. In some embodiments, at least 50% to 100% of the casein protein is beta-casein. In some embodiments, at least 50% to 100% of the casein protein is alpha-S1-casein. In some embodiments, at least 20% to 100% of the casein protein is alpha-S2-casein. In some embodiments, casein protein is expressed in a plant. In some embodiments, the emulsion has a pH of about 5.2 to about 5.9. In some embodiments, the composition comprises only one, only two, only three, or only four casein proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein. In some embodiments, the emulsion does not contain an organoleptically functional amount of beta-lactoglobulin. In some embodiments, the emulsion may comprise beta-lactoglobulin in the amount of about 0.01% (w/v) to about 0.1% (w/v), about 0.1% (w/v) to about 0.5% (w/v), about 0.5% (w/v) to about 1.0% (w/v), about 1.0% (w/v) to about 2% (w/v), about 2% (w/v) to about 3% (w/v), about 3% (w/v) to about 5% (w/v), about 5% (w/v) to about 10% (w/v), about 10% (w/v) to about 20% (w/v), about 20% (w/v) to about 40% (w/v), or more.

Alternative Dairy Compositions Comprising One or More Isolated or Recombinant Casein Proteins

[359] The milk or casein proteins described herein may also be used to prepare alternative dairy compositions. For example, in some embodiments, an alternative dairy composition comprises one or more casein proteins, such as recombinant casein proteins. In some embodiments, the casein proteins are selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein and alpha-S2-casein. In some embodiments, the alternative dairy composition comprises only one casein protein. In some embodiments, the alternative dairy composition comprises two, three, or four casein proteins.

[360] In some embodiments, the disclosure relates to an alternative dairy composition comprising a casein protein selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein; and a beta-lactoglobulin. In some embodiments the casein protein is recombinant. In some embodiments, the beta-lactoglobulin is recombinant. In some embodiments, both the casein protein and the beta-lactoglobulin are recombinant. In some embodiments, the ratio of the casein protein to the beta-lactoglobulin is about 8:1 to about 1:2. In some embodiments, the ratio of the casein protein to the beta-lactoglobulin is about 8:1 to about 2:1.

[361] In some embodiments, an alternative dairy composition comprises about 8% (w/v) to about 25% (w/v) total protein, such as about 8% to about 10%, about 10% to about 15%, about 15% to about 20%, or about 20 to about 25% total protein. In some embodiments, an alternative dairy composition comprises about 1% to about 10% (w/v) total protein. In some embodiments, an alternative dairy composition comprises about 25% to about 35%, about 35% to about 45%, about 45% to about 55%, about 55% to about 65%, about 65% to about 75% (w/v), or more total protein.

[362] In some embodiments, about 1% to about 5% of the total protein in the alternative dairy composition is casein protein. In some embodiments, about 5% to about 10% of the total protein in the alternative dairy composition is casein protein. In some embodiments, about 10% to about 20% of the total protein in the alternative dairy composition is casein protein. In some embodiments, about 20% to about 30% of the total protein in the alternative dairy composition is casein protein. In some embodiments, about 30% to about 40% of the total protein in the alternative dairy composition is casein protein. In some embodiments, about 40% to about 50% of the total protein in the alternative dairy composition is casein protein. In some embodiments, about 50% to about 60% of the total protein in the alternative dairy composition is casein protein. In some embodiments, about 60% to about 70% of the total protein in the alternative dairy composition is

casein protein. In some embodiments, about 70% to about 80% of the total protein in the alternative dairy composition is casein protein. In some embodiments, about 80% to about 90% of the total protein in the alternative dairy composition is casein protein. In some embodiments, about 90% to about 100% of the total protein in the alternative dairy composition is casein protein.

[363] In some embodiments, at least 1%, at least 2%, at least 3%, at least 4%, at least 5%, at least 6%, at least 7%, at least 8%, at least 9%, at least 10%, at least 11%, at least 12%, at least 13%, at least 14%, at least 15%, at least 16%, at least 17%, at least 18%, at least 19%, at least 20%, or more of the total protein in the alternative dairy composition is casein protein.

[364] In some embodiments, about 20% to about 100% of the casein protein in the alternative dairy composition is kappa casein. For example, the alternative dairy composition may comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% kappa casein. In some embodiments, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the alternative dairy composition is kappa casein.

[365] In some embodiments, about 20% to about 100% of the casein protein in the alternative dairy composition is para-kappa casein. For example, the alternative dairy composition may comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% para-kappa casein. In some embodiments, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the alternative dairy composition is para-kappa casein.

[366] In some embodiments, about 20% to about 100% of the casein protein in the alternative dairy composition is beta casein. In some embodiments, about 50% to about 100% of the casein protein in the alternative dairy composition is beta casein. For example, the alternative dairy composition may comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% beta casein. In some embodiments, about 20% to about

30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the alternative dairy composition is beta casein.

[367] In some embodiments, about 20% to about 100% of the casein protein in the alternative dairy composition is alpha-S1-casein. In some embodiments, about 50% to about 100% of the casein protein in the alternative dairy composition is alpha-S1-casein. For example, the alternative dairy composition may comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% alpha-S1-casein. In some embodiments, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the alternative dairy composition is alpha-S1-casein.

[368] In some embodiments, about 20% to about 100% of the casein protein in the alternative dairy composition is alpha-S2-casein. In some embodiments, about 50% to about 100% of the casein protein in the alternative dairy composition is alpha-S2-casein. For example, the alternative dairy composition may comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% alpha-S2-casein. In some embodiments, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the alternative dairy composition is alpha-S2-casein.

[369] In some embodiments, an alternative dairy composition comprises kappa casein and essentially no para-kappa casein. For example, in some embodiments, the alternative dairy composition comprises less than about 1%, less than about 0.9%, less than about 0.8%, less than about 0.7%, less than about 0.6%, less than about 0.5%, less than about 0.4%, less than about 0.3%, less than about 0.2%, or less than about 0.1%, para-kappa casein. In some embodiments, the alternative dairy composition comprises about 0.01% to about 1%, about 0.01% to about 0.9%, about 0.01% to about 0.8%, about 0.01% to about 0.7%, about 0.01% to about 0.6%, about 0.1% to about 0.5%, about 0.1% to about 0.4%, about 0.1% to about 0.3%, about 0.1% to about 0.2%, or about 0.01% to about 0.1% para-kappa casein. In some embodiments, the kappa casein is

recombinant. In some embodiments, the kappa casein is expressed in a plant. In some embodiments, the kappa casein is expressed in a soybean plant.

[370] In some embodiments, an alternative dairy composition comprises one to four recombinant milk proteins, each selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein. In some embodiments, an alternative dairy composition comprises 1, 2, 3, or 4 casein proteins. In some embodiments, an alternative dairy composition comprises only one casein protein.

[371] In some embodiments, an alternative dairy composition comprises recombinant beta-casein and at least one lipid and does not comprise an organoleptically functional amount of beta-lactoglobulin. In some embodiments, the composition does not comprise any additional casein proteins. In some embodiments, the composition comprises at least one additional casein protein. In some embodiments, the at least one additional casein protein is selected from kappa-casein, para-kappa-casein, alpha-S1-casein and alpha-S2-casein. In some embodiments, the at least one additional casein is kappa-casein or para-kappa-casein. In some embodiments, at least 50%, at least 75%, or at least 90% by weight of the total casein protein in an alternative dairy composition is beta-casein. In some embodiments, the beta-casein is expressed in a plant. In some embodiments, the beta-casein is expressed in a soybean plant. In some embodiments, all caseins in the composition are plant expressed. In some embodiments, the composition comprises a fusion protein comprising recombinant beta-casein.

[372] In some embodiments, the alternative dairy composition comprises two of the milk proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein. In some embodiments, the alternative dairy composition comprises three of the milk proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein. In some embodiments, the alternative dairy composition comprises four of the milk proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein. In some embodiments, the one or more milk protein(s) is(are) plant-expressed. In some embodiments, the milk protein(s) is(are) expressed in a soybean plant. In some embodiments, the milk protein(s) is(are) yeast- or bacterial-expressed. Exemplary combinations of 1, 2, 3, or 4 casein

proteins that may be used in the alternative dairy compositions described herein are shown above in Table 12.

[373] In some embodiments, the disclosure relates to an alternative dairy composition comprising one to four plant-expressed recombinant milk proteins (i.e., 2, 3, or 4 plant-expressed recombinant milk proteins), wherein the recombinant milk proteins confer one, two, three or more organoleptic properties similar to a dairy composition (i.e., a dairy composition comprising mammalian milk such as bovine milk) selected from the group consisting of taste, appearance, mouthfeel, structure, texture, density, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess, and emulsification. In some embodiments, the plant-expressed milk proteins are selected from beta lactoglobulin, kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein. In some embodiments, the recombinant beta-casein protein confers on the alternative dairy composition one, two, or more characteristics of a dairy food product selected from the group consisting of: taste, aroma, appearance, handling, mouthfeel, density, structure, texture, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess and emulsification.

[374] In some embodiments, the alternative dairy compositions described above comprise at least one additional mammalian or plant protein that is not a casein protein. Examples of combinations of casein, mammalian, and/or plant proteins are shown above in Table 12.

[375] In some embodiments, the alternative dairy compositions described herein may comprise plant protein. For example, in some embodiments, the alternative dairy compositions comprise protein from a legume, such as, for example, soybeans, chickpeas, kidney beans, black beans, pinot beans, green peas, and lentils. In some embodiments, the alternative dairy compositions comprise protein from a grain, such as, for example, wheat, millet, barley, oats, rice, spelt, teff, amaranth, and quinoa. In some embodiments, the alternative dairy compositions comprise protein from nuts, hempseed, chia seed, nutritional yeast, or spirulina. In some embodiments, the alternative dairy composition comprises protein from potato. In some embodiments, the alternative dairy composition comprises protein from a plant of the family *Fabaceae*.

[376] In some embodiments, the alternative dairy compositions described above have at least one of the following characteristics: i) a firmness of at least 150 grams; ii) a melting point of about

35°C to about 100°C; or iii) ability to stretch to at least 3 cm in length without breaking. In some embodiments, the alternative dairy compositions described above have the ability to stretch to at least 4 cm, at least 5 cm, at least 6 cm, at least 7 cm, at least 8 cm, at least 9 cm, at least 10 cm, at least 11 cm, at least 12 cm, at least 13 cm, at least 14 cm, at least 15 cm, at least 16 cm, at least 17 cm, at least 18 cm, at least 19 cm, or at least 10 cm in length without breaking. In some embodiments, the alternative dairy compositions described above have the ability to stretch to at least 5 cm in length without breaking. Testing methods and ranges firmness, melting point, and stretch are disclosed above.

[377] In some embodiments, the alternative dairy compositions comprise ash. In some embodiments, the alternative dairy compositions comprise at least one lipid and/or at least one salt. Examples of lipids include fatty acids (saturated and unsaturated); glycerides or glycerolipids (such as monoglycerides, diglycerides, triglycerides or neutral fats, and phosphoglycerides or glycerophospholipids); and nonglycerides (sphingolipids, tocopherols, tocotrienols, sterol lipids including cholesterol and steroid hormones, prenol lipids including terpenoids, fatty alcohols, waxes, and polyketides).

[378] Examples of lipids that may be included in the alternative dairy compositions include, for example, dairy fats or vegetable oils such as palm oil or palm kernel oil, soybean oil, corn oil, rapeseed oil, canola oil, sunflower oil, safflower oil, coconut oil, rice bran oil, olive oil, sesame oil, flaxseed oil, hemp oil, cottonseed oil, peanut oil, almond oil, beech nut oil, brazil nut oil, cashew oil, hazelnut oil, macadamia oil, mongongo nut oil, pecan oil, pine nut oil, pistachio oil, walnut oil, pumpkin seed oil, grapefruit seed oil, lemon oil, apricot oil, apple seed oil, argan oil, avocado oil, or orange oil. In some embodiments, the solid-phase, protein stabilized emulsion comprises butter or margarine.

[379] Examples of salts that may be included in the alternative dairy composition include, but are not limited to, magnesium chloride, sodium chloride, calcium chloride, sodium phosphate and trisodium citrate.

[380] In some embodiments, the alternative dairy compositions do not contain an organoleptically functional amount of beta-lactoglobulin. In some embodiments, the alternative dairy composition may comprise beta-lactoglobulin in the amount of about 0.01% (w/v) to about

0.1% (w/v), about 0.1% (w/v) to about 0.5% (w/v), about 0.5% (w/v) to about 1.0% (w/v), about 1.0% (w/v) to about 2% (w/v), about 2% (w/v) to about 3% (w/v), about 3% (w/v) to about 5% (w/v), about 5% (w/v) to about 10% (w/v), about 10% (w/v) to about 20% (w/v), about 30% (w/v) to about 40% (w/v), or more, of the composition.

[381] In some embodiments, the alternative dairy compositions comprise one or more recombinant casein proteins that are expressed in a microorganism. In some embodiments, the recombinant casein protein is yeast-expressed or bacterial-expressed. In some embodiments, the recombinant casein protein is expressed in a bacterium. Microorganisms used for recombinant protein production are well known in the art (see for example, Ferrer-Miralles et al., Bacterial cell factories for recombinant protein production; expanding the catalogue, 2013, *Microb Cell Fact.* 2013; 12:113). For example, the recombinant casein protein may be expressed in a bacteria such as *Escherichia coli*, *Caulobacter crescentus*, *Rodhobacter sphaeroides*, *Pseudoalteromonas haloplanktis*, *Shewanella* sp., *Pseudomonas putida*, *P. aeruginosa*, *P. fluorescens*, *Halomonas elongate*, *Chromohalobacter salexigens*, *Streptomyces lividans*, *S. griseus*, *Nocardia lactamdurans*, *Mycobacterium smegmatis*, *Corynebacterium glutamicum*, *C. ammoniagenes*, *Brevibacterium lactofermentum*, *Bacillus subtilis*, *B. brevis*, *B. megaterium*, *B. licheniformis*, *B. amyloliquefaciens*, *Lactococcus lactis*, *L. plantarum*, *L. casei*, *L. reuteri*, or *L. gasseri*.

[382] In some embodiments, the recombinant casein proteins are expressed in a microorganism that is a eukaryotic cell, such as *Saccharomyces* spp., *Kluyveromyces* spp., *Pichia* spp., *Aspergillus* spp., *Tetrahymena* spp., *Yarrowia* spp., *Hansenula* spp., *Blastobotrys* spp., *Candida* spp., *Zygosaccharomyces* spp., *Debrayomyces* spp., *Fusarium* spp., and *Trichoderma* spp.

[383] In some embodiments, the one or more recombinant casein proteins are expressed in a plant. In some embodiments, the plant may be a monocot selected from turf grass, maize (corn), rice, oat, wheat, barley, sorghum, orchid, iris, lily, onion, palm, and duckweed. In some embodiments, the plant is a dicot selected from Arabidopsis, tobacco, tomato, potato, sweet potato, cassava, alfalfa, lima bean, pea, chick pea, soybean, carrot, strawberry, lettuce, oak, maple, walnut, rose, mint, squash, daisy, Quinoa, buckwheat, mung bean, cow pea, lentil, lupin, peanut, fava bean, French beans (i.e., common beans), mustard, or cactus. In some embodiments, the plant is a non-vascular plant selected from moss, liverwort, hornwort or algae. In some embodiments, the plant

is a vascular plant reproducing from spores (e.g., a fern). In some embodiments, the recombinant casein protein is expressed in a soybean plant.

[384] In some embodiments, the alternative dairy compositions described above have a pH of about 2 to about 8. In some embodiments, the alternative dairy compositions described above have a pH of about 4 to about 8. Table 13 below shows exemplary ranges of pH for common mammalian derived dairy products.

Table 13: pH ranges of common dairy products

Dairy product	pH range
Milk	6.7-6.9
Butter	6.1-6.4
Yogurt	2.0-4.5
Brie	6.0-6.5
Cheddar	5.1-5.3
Cream cheese	4.6-5.1
Feta	4.1-4.5
Parmesan	5.2-5.3
Ricotta	6.0

[385] Examples of alternative dairy compositions that may be produced as described herein include, but are not limited to, alternative versions of milk, cream, butter, and cheese. Other example alternative dairy compositions include ice cream, frozen desserts, frozen yogurt or custard, yogurt, cottage cheese, cream cheese, curds, crème fraiche, toppings, icings, fillings, low-fat spreads, dairy-based dry mixes, geriatric nutrition compositions, coffee creamers, analog dairy products, follow-up formula, baby formula, infant formula, milk, dairy beverages, acid dairy drinks, smoothies, milk tea, margarine, butter alternatives, growing up milks, low-lactose products, buttermilk, sour cream, skyr, leben, lassi, kefir, and beverages. In some embodiments, the alternative dairy compositions may be cultured milks, such as drinkable yogurts. The alternative dairy compositions may also be powders containing a milk protein, or a low-lactose product. An illustrative method for preparing an alternative dairy composition is provided in **FIG. 13**.

[386] An alternative milk composition may be produced, for example, by mixing a liquid comprising at least one isolated or recombinant milk or casein protein, with ash, lipids, and/or a

sweetener, and optionally one or more flavor compounds and/or color agents. In some embodiments, one or more vitamins are added to the alternative milk composition, such as retinal, carotene, vitamins, vitamin D, vitamin E, vitamin B12, thiamin, or riboflavin. This milk alternative may then be used to produce, for example, butter, ice cream, frozen desserts, frozen yogurt or custard, yogurt, cottage cheese, cream cheese, curds, and crème fraiche.

[387] In some embodiments, the alternative dairy composition comprises one or more sweeteners. Examples of sweeteners include, but are not limited to, saccharides, such as glucose, mamiose, maltose, fructose, galactose, lactose, sucrose, monatin, and tagatose. In some embodiments the sweetener is selected from stevia, aspartame, cyclamate, saccharin, sucralose, mogrosides, brazzein, curculin, erythritol, glycyrrhizin, inulin, isomalt, lacticitol, mabinlin, malititol, mamiitol, miraculin, monatin, monelin, osladin, pentadin, sorbitol, thaumatin, xylitol, acesulfame, potassium, advantame, alitame, aspartame-acesulfame, sodium cyclamate, dulcin, glucin, neohesperidin, dihydochalcone, neotame, and P-4000.

[388] In some embodiments, an alternative dairy food composition comprises calcium. In some embodiments, the composition comprises calcium at a concentration of about 0% to about 2% by weight. In some embodiments, the composition comprises calcium at a concentration of about 0.001% to about 2% by weight. In some embodiments, the composition comprises calcium at a concentration of about 0.01% to about 2% by weight. In some embodiments, the composition comprises calcium at a concentration of about 0.1% to about 2% by weight. In some embodiments, the composition comprises calcium at a concentration of about 1% to about 2% by weight. In some embodiments, the composition comprises calcium at a concentration of about 0.01%, about 0.02%, about 0.03%, about 0.04%, about 0.05%, about 0.06%, about 0.07%, about 0.08%, about 0.09%, about 0.1%, about 0.2%, about 0.3%, about 0.4%, about 0.5%, about 0.6%, about 0.7%, about 0.8%, about 0.9%, about 1.0%, about 1.1%, about 1.2%, about 1.3%, about 1.4%, about 1.5%, about 1.6%, about 1.7%, about 1.8%, about 1.9%, or about 2.0% by weight.

[389] Thus, in some embodiments, the alternative dairy composition is a milk composition. In some embodiments, the alternative dairy composition is a cheese composition. In some embodiments, the alternative dairy composition is cream composition. In some embodiments, the alternative dairy composition is a yogurt composition (e.g., a frozen yogurt composition, a sugar-

free yogurt composition, a low-fat yogurt composition, a Greek yogurt composition, a drinkable yogurt composition, etc). In some embodiments, the alternative dairy composition is ice cream. In some embodiments, alternative dairy composition is a frozen custard composition. In some embodiments, the alternative dairy composition is a frozen dessert. In some embodiments, the alternative dairy composition is a crème fraiche composition. In some embodiments, the alternative dairy composition is curd composition. In some embodiments, the alternative dairy composition is a cottage cheese composition. In some embodiments, the alternative dairy composition is cream composition. In some embodiments, the alternative dairy composition is a sour cream composition.

Cheese Compositions

[390] Traditionally, cheese is made with milk, which comprises a number of proteins including various casein proteins (see Table 14 below for exemplary compositions of human and cow milk). Coagulation of the milk proteins occurs by way of an acid and/or rennet addition, which causes the milk to curdle. Rennet is a bacterial enzyme that cleaves kappa-casein, generating para-kappa-casein, which then links up with the calcium and phosphate present in milk to join casein micelles together. These solids curds are collected and/or separated from the liquid (whey) and various procedures of pressing, forming, and aging yield different cheese products.

Table 14: Illustrative Milk Protein Compositions

Protein	Human milk (mg/mL)	Bovine (cow) milk (mg/mL)
α -lactalbumin	2.2	1.2
α -s1-casein	0	11.6
α -s2-casein	0	3.0
β -casein	2.2	9.6
κ -casein	0.4	3.6
γ -casein	0	1.6
Immunoglobulins	0.8	0.6
Lactoferrin	1.4	0.3
β -lactoglobulin	0	3.0
Lysozyme	0.5	Traces
Serum albumin	0.4	0.4
Other	0.8	0.6

[391] Described herein are cheese compositions comprising a different protein composition compared to that of any mammalian milk (i.e., a non-naturally occurring protein composition). For example, in some embodiments, a cheese composition can be prepared using only one milk protein. In some embodiments, a cheese composition can be prepared using only two milk proteins. In some embodiments, a cheese composition may be prepared using only three milk proteins. In some embodiments, a cheese composition may be prepared using only four milk proteins. In some embodiments, a cheese composition comprises one or more milk proteins at a ratio that is not found in any mammalian milk (e.g., a non-naturally occurring ratio).

[392] In some embodiments, a cheese composition comprises one milk protein, which may be derived from animal-produced milk, or recombinantly expressed. In some embodiments, a cheese composition comprises two, three, or four milk proteins, wherein each milk protein is derived from animal-produced milk or is recombinantly expressed. In some embodiments, the milk protein is a casein protein.

[393] In some embodiments, a cheese composition may comprise beta-casein as the only casein protein (i.e., 100% beta-casein). In some embodiments, a cheese composition comprises beta-casein and at least one additional casein protein. In some embodiments, the at least one additional casein protein is selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein and alpha-S2-casein. In some embodiments, the at least one additional casein protein is kappa-casein. In some embodiments, the at least one additional casein protein is para-kappa-casein.

[394] In some embodiments, a cheese composition comprises two or more casein proteins, wherein about 95% by weight of the casein protein in the composition is beta-casein. In some embodiments, a cheese composition comprises two or more casein proteins, wherein about 90% by weight of the casein protein in the composition is beta-casein. In some embodiments, a cheese composition comprises two or more casein proteins, wherein about 85% by weight of the casein protein in the composition is beta-casein. In some embodiments, a cheese composition comprises two or more casein proteins, wherein about 80% by weight of the casein protein in the composition is beta-casein. In some embodiments, a cheese composition comprises two or more casein proteins, wherein about 75% by weight of the casein protein in the composition is beta-casein. In some

embodiments, a cheese composition comprises two or more casein proteins, wherein about 70% by weight of the casein protein in the composition is beta-casein. In some embodiments, a cheese composition comprises two or more casein proteins, wherein about 65% by weight of the casein protein in the composition is beta-casein. In some embodiments, a cheese composition comprises two or more casein proteins, wherein about 60% by weight of the casein protein in the composition is beta-casein. In some embodiments, a cheese composition comprises two or more casein proteins, wherein about 55% by weight of the casein protein in the composition is beta-casein. In some embodiments, a cheese composition comprises two or more casein proteins, wherein about 50% by weight of the casein protein in the composition is beta-casein.

[395] In some embodiments, a cheese composition may comprise 95%, beta-casein and 5% of one or more additional casein proteins. In some embodiments, it may comprise 90%, beta-casein and 10% of one or more additional casein proteins. In some embodiments, it may comprise 85%, beta-casein and 15% of one or more additional casein proteins. In some embodiments, it may comprise 80%, beta-casein and 20% of one or more additional casein proteins. In some embodiments, it may comprise 75%, beta-casein and 25% of one or more additional casein proteins. In some embodiments, it may comprise 70%, beta-casein and 30% of one or more additional casein proteins. The other casein proteins may be kappa-casein, para-kappa-casein, alpha-S1-casein, and/or alpha-S2-casein.

[396] In some embodiments, the cheese composition comprises 75% beta-casein and 25% alpha caseins (i.e., a mixture of alpha-S1-casein and alpha-S2-casein). In some embodiments, the cheese composition comprises 75% beta-casein and 25% kappa-casein. In some embodiments, the cheese composition comprises 50% beta-casein and 50% kappa-casein. In some embodiments, the cheese composition comprises 50% beta-casein and 50% alpha caseins.

[397] In some embodiments the beta-casein is recombinant beta-casein. In some embodiments, the recombinant beta-casein protein is plant-expressed. In some embodiments, the recombinant beta-casein is expressed in a soybean. In some embodiments, all the caseins in the cheese composition are plant-expressed. In some embodiments, the recombinant casein protein is derived from a fusion protein. In some embodiments, the cheese composition does not contain an organoleptically functional amount of beta-lactoglobulin.

[398] In some embodiments, a cheese composition comprises para-kappa-casein produced without the use of any enzyme that cleaves kappa-casein to para-kappa-casein. In some embodiments, a cheese composition comprises para-kappa-casein produced without the use of any acid that cleaves kappa-casein to para-kappa-casein. In some embodiments, a cheese composition comprises para-kappa-casein produced without the use of any enzyme or acid that cleaves kappa-casein to para-kappa-casein. In some embodiments, a cheese composition comprises a recombinantly expressed para-kappa-casein. In some embodiments, a cheese composition comprises substantially no casein, such as 0.01% (w/v) to 0.1% (w/v) or 0.1% (w/v) to 0.1% (w/v) casein.

[399] In some embodiments, a cheese composition comprises about 8% (w/v) to about 25% (w/v) total protein, such as about 8% to about 10%, about 10% to about 15%, about 15% to about 20%, or about 20 to about 25% total protein. In some embodiments, a cheese composition comprises about 1% to about 10% (w/v) total protein. In some embodiments, a cheese composition comprises about 25% to about 35%, about 35% to about 45%, about 45% to about 55%, about 55% to about 65%, about 65% to about 75% (w/v), or more total protein.

[400] In some embodiments, about 1% to about 5% of the total protein in the cheese composition is casein protein. In some embodiments, about 5% to about 10% of the total protein in the cheese composition is casein protein. In some embodiments, about 10% to about 20% of the total protein in the cheese composition is casein protein. In some embodiments, about 20% to about 30% of the total protein in the cheese composition is casein protein. In some embodiments, about 30% to about 40% of the total protein in the cheese composition is casein protein. In some embodiments, about 40% to about 50% of the total protein in the cheese composition is casein protein. In some embodiments, about 50% to about 60% of the total protein in the cheese composition is casein protein. In some embodiments, about 60% to about 70% of the total protein in the cheese composition is casein protein. In some embodiments, about 70% to about 80% of the total protein in the cheese composition is casein protein. In some embodiments, about 80% to about 90% of the total protein in the cheese composition is casein protein. In some embodiments, about 90% to about 100% of the total protein in the cheese composition is casein protein.

[401] In some embodiments, at least 1%, at least 2%, at least 3%, at least 4%, at least 5%, at least 6%, at least 7%, at least 8%, at least 9%, at least 10%, at least 11%, at least 12%, at least 13%, at

least 14%, at least 15%, at least 16%, at least 17%, at least 18%, at least 19%, at least 20%, or more of the total protein in the cheese composition is casein protein.

[402] In some embodiments, about 20% to about 100% of the casein protein in the cheese composition is kappa casein. For example, the cheese composition may comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% kappa casein. In some embodiments, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the cheese composition is kappa casein.

[403] In some embodiments, about 20% to about 100% of the casein protein in the cheese composition is para-kappa casein. For example, the cheese composition may comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% para-kappa casein. In some embodiments, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the cheese composition is para-kappa casein.

[404] In some embodiments, about 20% to about 100% of the casein protein in the cheese composition is beta casein. In some embodiments, about 50% to about 100% of the casein protein in the cheese composition is beta casein. For example, the cheese composition may comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% beta casein. In some embodiments, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the cheese composition is beta casein.

[405] In some embodiments, about 20% to about 100% of the casein protein in the cheese composition is alpha-S1-casein. In some embodiments, about 50% to about 100% of the casein protein in the cheese composition is alpha-S1-casein. For example, the cheese composition may

comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% alpha-S1-casein. In some embodiments, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the cheese composition is alpha-S1-casein.

[406] In some embodiments, about 20% to about 100% of the casein protein in the cheese composition is alpha-S2-casein. In some embodiments, about 50% to about 100% of the casein protein in the cheese composition is alpha-S2-casein. For example, the cheese composition may comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% alpha-S2-casein. In some embodiments, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the cheese composition is alpha-S2-casein.

[407] In some embodiments, a cheese composition comprises a stable, protein-stabilized emulsion described herein. In some embodiments, a cheese composition comprises more than one of the stable, protein-stabilized emulsions described herein. In some embodiments, a cheese composition is made using at least one stable, protein-stabilized emulsion described herein.

[408] In some embodiments, a cheese composition comprises a colloidal suspension described herein. In some embodiments, a cheese composition comprises more than one colloidal suspension described herein. In some embodiments, a cheese composition is made using at least one of the colloidal suspensions described herein.

[409] In some embodiments, a cheese composition described herein may comprise plant protein. For example, in some embodiments, the cheese composition comprises protein from a legume, such as, for example, soybeans, chickpeas, kidney beans, black beans, pinto beans, green peas, and lentils. In some embodiments, the cheese composition comprises protein from a grain, such as, for example, wheat, millet, barley, oats, rice, spelt, teff, amaranth, and quinoa. In some embodiments, the cheese composition comprises protein from nuts, hempseed, chia seed, nutritional yeast, or spirulina. In some embodiments, the cheese composition comprises protein

from potato. In some embodiments, the cheese composition comprises protein from a plant of the family *Fabaceae*.

[410] In some embodiments, the cheese compositions described herein may be substantially transparent. As used herein, “substantially transparent” means having an opacity of about 50%, about 40%, about 30%, about 20%, about 10% or less. In some embodiments, the cheese composition has about 0% opacity. In some embodiments, the cheese compositions described herein are substantially transparent when in solid form. In some embodiments, the cheese compositions described herein are substantially transparent when melted.

[411] In some embodiments, the cheese compositions described herein may have at least one, at least two, or at least three desirable organoleptic properties. In some embodiments, the cheese compositions described herein may have at least one, at least two, or at least three organoleptic properties that is similar to that of cheese (i.e., cheese produced using mammalian milk, such as bovine milk or goat milk). For example, in some embodiments, the cheese compositions may have at least one, at least two, or at least three organoleptic properties found in the cheeses of Table 15 or Table 16.

[412] In some embodiments, the cheese compositions described herein may be used in a similar manner (e.g., for cooking, etc.) as one or more of the cheeses listed in Table 15 or Table 16. In some embodiments, the cheese compositions described herein may be used as a substitute for one or more of the cheeses listed in Table 15 or Table 16.

Table 15: Illustrative types of cheese

Category	Examples
Soft Fresh Cheeses	Cottage Cheese Cream Cheese Feta Mascarpone Neufchâtel Queso Blanco Ricotta
Soft-Ripened Cheeses	Brie (single, double and triple cream and flavored) Camembert
Semi-Soft Chesses	Brick, dry- and washed-rind Fontina

	Havarti Limburger Monterey Jack Muenster Pepper Jack
Blue-Veined Cheeses	Blue Cheese Gorgonzola, creamy style Gorgonzola, crumbly style
Gouda & Edam	Gouda Smoked Gouda Edam
Pasta Filata and Related Cheeses	Fresh Mozzarella Low-Moisture, Part-Skim Mozzarella Low-Moisture, Whole Milk Mozzarella Part-Skim Mozzarella Whole Milk Mozzarella Provolone, mild, aged and smoked String Cheese Pizza Cheese Individually Quick Frozen mozzarella (IQF)
Cheddar & Colby	Cheddar Smoked Cheddar Colby
Swiss Cheeses	Baby Swiss Swiss Gruyère
Hard Cheeses	Asiago Parmesan Romano Pepato
Process Cheeses	Pasteurized Process Cheese Pasteurized Process Cheese Food Pasteurized Process Cheese Spread Pasteurized Process Cheese Product Cold-Pack High-Melt Cheeses
Powder & Enzyme-modified Cheeses	Cheese Powders Enzyme Modified Cheeses (EMCs)
Custom & Convenience Cheese Products	Pre-blends Pre-cut Cheese Shredded Cheese Grated Cheese Cheese Sauce

	Portion Packaged Cheese
Cheeses for Special Needs	Low-fat Cheeses No-fat Cheeses Low-sodium Cheeses Kosher Cheeses Halal Cheeses Organic Cheeses

[413] Cheese may also be categorized based on moisture content. Shown below in Table 16 are example categories of cheeses and their respective moisture content (from Jana AH et al., *J. Food Sci Technol* (2017) 54(12):3776-3778).

Table 16: Moisture content of cheeses

Cheese type	Moisture content (%)	Examples
Soft cheese	50-80	Cottage, Quark, Baker's, Mozzarella, Camembert, Feta
Semi-soft cheese	39-50	Blue, Limburger, Provolone, Tilsiter
Hard cheese	Max. 39	Cheddar, Colby, Edam, Swiss, Gouda
Very hard cheese	Max. 34	Parmesan, Romano, Sardo, Grana

[414] In some embodiments, a cheese composition described herein has a moisture content of between about 30% and about 80%. In some embodiments, a cheese composition described herein has between about 45% to 60% moisture content.

[415] Cheese and cheese compositions have functional properties such as moisture content, firmness, stretchability, melting, viscosity/flow, oiling off, browning/blistering, whitening/decolorization, spreadability, grating, slicing, dicing, shredding/mincing, mouthfeel, flavor, aroma, freezing ability, and overall appearance. These properties can be determined by any number of means well known in the art.

[416] Firmness and stretch may be analyzed as described above. Moisture content may be measured for example, as described in Bradley, R. L., Jr., and M. A. Vanderwarn. 2001, Determination of moisture in cheese and cheese products, *J. AOAC* 84:570–592. Texture may be analyzed as described in Kapoor et al., 2005, Small-scale manufacture of process cheese using a rapid visco analyzer, *J. Dairy Sci.* 88:3382-3391, using a TA.XT2 Texture Analyzer (see also

Drake et al., 1999 Relationship between instrumental and sensory measurements of cheese texture, *J. Texture Stud.* 30:451–476) or for example by Breene 1975, Application of texture profile analysis to instrumental food texture evaluation, *J. Texture Stud.* 6:53–82.

[417] In some embodiments, a cheese composition has the ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the composition at a temperature of 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass. In some embodiments, a cheese composition has the ability to stretch to at least 4 cm in length without breaking, as determined by heating a 100 gram mass of the composition at a temperature of 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass. In some embodiments, a cheese composition has the ability to stretch to at least 5 cm in length without breaking, as determined by heating a 100 gram mass of the composition at a temperature of 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass. In some embodiments, a cheese composition has the ability to stretch to at least 6 cm in length without breaking, as determined by heating a 100 gram mass of the composition at a temperature of 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass. In some embodiments, a cheese composition has the ability to stretch to at least 9 cm in length without breaking, as determined by heating a 100 gram mass of the composition at a temperature of 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass. In some embodiments, a cheese composition has the ability to stretch to at least 12 cm in length without breaking, as determined by heating a 100 gram mass of the composition at a temperature of 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass. In some embodiments, a cheese composition has the ability to stretch to at least 15 cm in length without breaking, as determined by heating a 100 gram mass of the composition at a temperature of 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass. In some embodiments, a cheese composition has the ability to stretch to at least 18 cm in length without breaking, as determined by heating a 100 gram mass of the composition at a temperature of 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[418] In some embodiments, a cheese composition described herein has a firmness of at least 150 grams, as determined by compressing a cylindrical-shaped sample of the cheese composition

having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C. In some embodiments, a cheese composition described herein has a firmness of at least 300 grams, as determined by compressing a cylindrical-shaped sample of the cheese composition having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C. In some embodiments, a cheese composition described herein has a firmness of at least 600 grams, as determined by compressing a cylindrical-shaped sample of the cheese composition having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C. In some embodiments, a cheese composition described herein has a firmness of at least 1000 grams, as determined by compressing a cylindrical-shaped sample of the cheese composition having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C. In some embodiments, a cheese composition described herein has a firmness of at least 2000 grams, as determined by compressing a cylindrical-shaped sample of the cheese composition having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C. In some embodiments, a cheese composition described herein has a firmness in the range of about 600 to about 3000 grams, for example about 650 to about 1000 grams, about 1000 grams to about 1500 grams, about 1500 grams to about 2000 grams, about 2500 grams to about 3000 grams, as determined by compressing a cylindrical-shaped sample of the cheese composition having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C.

[419] As will be understood by those of skill in the art, melting properties can be influenced by a number of factors, including water content, fat content, protein content, and other the presence of other ingredients such as salt, acid, and stabilizers. Meltability may be measured with a rapid visco analyzer (RVA) (Metzger et al., 2002, RVA: Process cheese manufacture, *Aust. J. Dairy Technol.* 57:136; Kapoor et al., 2004, Comparison of pilot scale and rapid visco analyzer process cheese manufacture, *J. Dairy Sci.* 87:2813–2821; Prow et al., 2005, Melt analysis of process cheese spread or product using a rapid visco analyzer, *J. Dairy Sci.* 88:1277–1287). Meltability may also be measured by the Schreiber melt test (1977), wherein a 0.5 cm high plug of cheese is placed in a glass petri dish and heated in an oven at 450°F for 5 minutes. Other melting tests include the Arnott test (1957), the tube test (1958), the melt analysis/UW meltmeter (1997), and the Dynamic Stress Rheometry (DSR) (1998). Shown in Table 17 are some examples of cheeses and their melting temperatures. In some embodiments, the cheese compositions described herein have a melting temperature similar to one or more of the cheeses in Table 17. In some

embodiments, the cheese compositions described herein have a melting temperature in the range of 100°F to 200°F, such as about 120°F, 130°F, 150°F, or 180°F.

Table 17: Melting ranges for cheese

Cheese type	Melt temperature	Examples
Process cheese	120°F/49°C	Pasteurized Process Cheese
Soft or semi-soft cheese	130°F/54°C	Mozzarella
Hard cheese	150°F/66°C	Cheddar, Colby, Edam, Swiss, Gouda
Very hard cheese	180°F/82°C	Parmesan, Romano, Sardo, Grana

[420] In some embodiments, the cheese composition has a melting point of about 35°C to about 100°C. In some embodiments, the cheese composition has a melting point of about 40°C to about 50°C. In some embodiments, the cheese composition has a melting point of about 50°C to about 60°C. In some embodiments, the cheese composition has a melting point of about 60°C to about 70°C. In some embodiments, the cheese composition has a melting point of about 70°C to about 90°C.

[421] As mentioned above, the properties of cheese can be influenced by a number of factors, such as lipids, salts, and/or calcium. Lipids that may be added to the cheese compositions disclosed herein include, for example, dairy fats or vegetable oils such as palm oil or palm kernel oil, butter oil, anhydrous milkfat, soybean oil, corn oil, rapeseed oil, canola oil, sunflower oil, safflower oil, coconut oil, rice bran oil, olive oil, sesame oil, flaxseed oil, hemp oil, cottonseed oil, peanut oil, almond oil, beech nut oil, brazil nut oil, cashew oil, hazelnut oil, macadamia oil, mongongo nut oil, pecan oil, pine nut oil, pistachio oil, walnut oil, pumpkin seed oil, grapefruit seed oil, lemon oil, apricot oil, apple seed oil, argan oil, avocado oil, or orange oil.

[422] Examples of salts that may be included in a cheese composition include, but are not limited to, magnesium chloride, sodium chloride, calcium chloride, sodium phosphates and trisodium citrate. In some embodiments, a cheese composition comprises at least one lipid and at least one salt. In some embodiments, a cheese composition comprises calcium. In some embodiments, a cheese composition comprises calcium at a concentration of about 0 to about 2% by weight. In some embodiments, a cheese composition comprises calcium at a concentration of about 0.001 to

about 2% by weight. In some embodiments, a cheese composition comprises calcium at a concentration of about 0.01 to about 2% by weight. In some embodiments, a cheese composition comprises calcium at a concentration of about 0.1 to about 2% by weight. In some embodiments, a cheese composition comprises calcium at a concentration of about 1 to about 2% by weight. In some embodiments, a cheese composition has a pH of about 5.2 to about 5.9. In some embodiments, a cheese composition comprises at least one organoleptic property similar to cheese (i.e., cheese produced using mammalian milk, such as bovine milk or goat milk) selected from the group consisting of taste, appearance, mouthfeel, structure, texture, density, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess, and emulsification. In some embodiments, the cheese composition comprises at least two organoleptic properties similar to cheese (i.e., cheese produced using mammalian milk, such as bovine milk or goat milk) selected from the group consisting of taste, appearance, mouthfeel, structure, texture, density, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess, and emulsification. In some embodiments, the cheese composition comprises at least three organoleptic properties similar to cheese (i.e., cheese produced using mammalian milk, such as bovine milk or goat milk) selected from the group consisting of taste, appearance, mouthfeel, structure, texture, density, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess, and emulsification.

[423] In some embodiments, a cheese composition comprises one or more vitamins, such as retinal, carotene, vitamins, vitamin D, vitamin E, vitamin B12, thiamin, or riboflavin.

Colloidal Suspensions Comprising One or More Isolated or Recombinant Casein Proteins

[424] A colloidal suspension is a mixture having particles suspended in a continuous phase with another component. The particles may be, for example, proteins. The other component may be, for example water. Many different kinds of foods may be colloidal suspensions, including beverages and other foods such as jam, ice cream, mayonnaise, etc. One example of a colloidal suspension is milk.

[425] The colloidal suspensions described herein may be a Newtonian fluid or a non-Newtonian fluid. Newtonian fluids are characterized by a viscosity that is independent of shear rate; they

follow Newton's law of viscosity. Apparent viscosity is the shear stress applied to a fluid divided by the shear rate (expressed in Pascal-second or centipoise units). For a Newtonian fluid, the apparent viscosity is constant. Water is an example of a Newtonian fluid. Non-Newtonian fluids do not follow Newton's law of viscosity; their viscosity can change (for example, become more liquid or more solid) when under force. Ketchup is an example of a non-Newtonian fluid.

[426] In some embodiments, a colloidal suspension comprises: 1-4 milk proteins (i.e., 1, 2, 3, or 4 recombinant milk proteins). The milk proteins may be recombinant or may be isolated from a mammalian milk. In some embodiments, the milk proteins may be plant-expressed.

[427] In some embodiments, a colloidal suspension comprises recombinant beta-casein and at least one lipid and does not contain an organoleptically functional amount of beta-lactoglobulin. In some embodiments, the colloidal suspension does not comprise any additional casein proteins. In some embodiments, the colloidal suspension comprises at least one additional casein protein. In some embodiments, the at least one additional casein protein is selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein and alpha-S2-casein. In some embodiments, the at least one additional casein protein is kappa-casein or para-kappa-casein. In some embodiments, the colloidal suspension is a non-Newtonian fluid.

[428] In some embodiments, at least 80%, at least 90%, or at least 95% by weight of the total casein protein in a colloidal suspension is beta-casein. In some embodiments, the beta-casein is expressed in a plant. In some embodiments, the beta-casein is expressed in a soybean plant. In some embodiments, all caseins in the composition are plant expressed. In some embodiments, the composition comprises a fusion protein comprising recombinant beta-casein.

[429] In some embodiments, a colloidal suspension is a non-Newtonian fluid. In some embodiments, a colloidal suspension is characterized as a shear thinning fluid with an apparent viscosity greater than 10 centipoise, at a shear rate of 1 sec^{-1} . In some embodiments, the suspension is an aqueous suspension.

[430] In some embodiments, the milk proteins comprise between 0.5% (w/v) to 15 % (w/v) of the composition, such as about 0.5% (w/v); about 1.0% (w/v), about 1.5% (w/v), about 2.0% (w/v), about 2.5% (w/v), about 3.0% (w/v), about 3.5% (w/v), about 4.0% (w/v), about 4.5% (w/v), about 5.0% (w/v), about 5.5% (w/v), about 6.0% (w/v), about 6.5% (w/v), about 7.0% (w/v), about 7.5%

(w/v), about 8.0% (w/v), about 8.5% (w/v), about 9.0% (w/v), about 9.5% (w/v), about 10.1% (w/v), about 10.5% (w/v), about 11.0% (w/v), about 11.5% (w/v), about 12.0% (w/v), about 12.5% (w/v), about 13.0% (w/v), about 13.5% (w/v), about 14.0% (w/v), about 14.5% (w/v), or about 15.0% (w/v). In some embodiments, the colloidal suspension may comprise one or more additional components, such as ash. In some embodiments, the colloidal suspension may comprise one or more vitamins such as retinal, carotene, vitamins, vitamin D, vitamin E, vitamin B12, thiamin, or riboflavin.

[431] In some embodiments, a colloidal suspension comprises about 8% (w/v) to about 25% (w/v) total protein, such as about 8% to about 10%, about 10% to about 15%, about 15% to about 20%, or about 20 to about 25% total protein. In some embodiments, a colloidal suspension comprises about 1% to about 10% (w/v) total protein. In some embodiments, a colloidal suspension comprises about 25% to about 35%, about 35% to about 45%, about 45% to about 55%, about 55% to about 65%, about 65% to about 75% (w/v), or more total protein.

[432] In some embodiments, about 1% to about 5% of the total protein in the colloidal suspension is casein protein. In some embodiments, about 5% to about 10% of the total protein in the colloidal suspension is casein protein. In some embodiments, about 10% to about 20% of the total protein in the colloidal suspension is casein protein. In some embodiments, about 20% to about 30% of the total protein in the colloidal suspension is casein protein. In some embodiments, about 30% to about 40% of the total protein in the colloidal suspension is casein protein. In some embodiments, about 40% to about 50% of the total protein in the colloidal suspension is casein protein. In some embodiments, about 50% to about 60% of the total protein in the colloidal suspension is casein protein. In some embodiments, about 60% to about 70% of the total protein in the colloidal suspension is casein protein. In some embodiments, about 70% to about 80% of the total protein in the colloidal suspension is casein protein. In some embodiments, about 80% to about 90% of the total protein in the colloidal suspension is casein protein. In some embodiments, about 90% to about 100% of the total protein in the colloidal suspension is casein protein.

[433] In some embodiments, at least 1%, at least 2%, at least 3%, at least 4%, at least 5%, at least 6%, at least 7%, at least 8%, at least 9%, at least 10%, at least 11%, at least 12%, at least 13%, at least 14%, at least 15%, at least 16%, at least 17%, at least 18%, at least 19%, at least 20%, or more of the total protein in the colloidal suspension is casein protein.

[434] In some embodiments, about 20% to about 100% of the casein protein in the colloidal suspension is kappa casein. For example, the colloidal suspension may comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% kappa casein. In some embodiments, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the colloidal suspension is kappa casein.

[435] In some embodiments, about 20% to about 100% of the casein protein in the colloidal suspension is para-kappa casein. For example, the colloidal suspension may comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% para-kappa casein. In some embodiments, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the colloidal suspension is para-kappa casein.

[436] In some embodiments, about 20% to about 100% of the casein protein in the colloidal suspension is beta casein. In some embodiments, about 50% to about 100% of the casein protein in the colloidal suspension is beta casein. For example, the colloidal suspension may comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% beta casein. In some embodiments, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the colloidal suspension is beta casein.

[437] In some embodiments, about 20% to about 100% of the casein protein in the colloidal suspension is alpha-S1-casein. In some embodiments, about 50% to about 100% of the casein protein in the colloidal suspension is alpha-S1-casein. For example, the colloidal suspension may comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or

about 90% to about 100% alpha-S1-casein. In some embodiments, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the colloidal suspension is alpha-S1-casein.

[438] In some embodiments, about 20% to about 100% of the casein protein in the colloidal suspension is alpha-S2-casein. In some embodiments, about 50% to about 100% of the casein protein in the colloidal suspension is alpha-S2-casein. For example, the colloidal suspension may comprise about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% alpha-S2-casein. In some embodiments, about 20% to about 30%, about 30% to about 40%, about 40% to about 50%, about 50% to about 60%, about 60% to about 70%, about 70% to about 80%, about 80% to about 90%, or about 90% to about 100% of the casein protein in the colloidal suspension is alpha-S2-casein.

[439] In some embodiments, colloidal suspension has at least one organoleptic property that is substantially similar to bovine milk. In some embodiments, the organoleptic property is selected from the group consisting of taste, appearance, mouthfeel, structure, texture, density, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess, and emulsification. In some embodiments, colloidal suspension has at least two, at least three, at least four, at least five, or more organoleptic properties that are substantially similar to bovine milk. In some embodiments, the plant-expressed milk proteins are recombinant, and are selected from beta lactoglobulin, kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein.

[440] In some embodiments, the colloidal suspensions described herein may be used to produce one or more food compositions such as butter, ice cream, frozen yogurt or custard, yogurt, frozen desserts, cottage cheese, cream cheese, curds, and crème fraîche.

Methods for Making the Food Compositions Described Herein

[441] Also provided herein are methods for making solid phase, protein-stabilized emulsions, colloidal suspensions, dairy alternatives and food compositions described herein (collectively

referred to in this section as “compositions”). In some embodiments, a method for making a composition comprises isolating one or more casein proteins from a mammalian milk. In some embodiments, a method for making a composition comprises expressing a casein protein in a cell (e.g., in a plant, or microorganism), extracting the recombinant protein, and preparing a composition comprising recombinant casein protein (See, e.g., **FIG. 13**).

[442] Initially, all ingredients for the composition are provided. For example, in some embodiments, the one or more milk proteins are provided. The milk proteins may be isolated from a mammalian milk, or may be produced recombinantly (e.g., by expression in a plant). An illustrative process for preparing a recombinant protein for use in making a composition as described herein is illustrated in **FIG. 13** and is also described below. In some embodiments, one or more lipids, salts, acids, etc. are also provided. In some embodiments, ash is provided. In some embodiments, one or more vitamins is provided, such as retinal, carotene, vitamins, vitamin D, vitamin E, vitamin B12, thiamin, or riboflavin.

[443] The ingredients are then combined and mixed. In some embodiments, the mixing is performed at a pre-determined temperature, for example a temperature in the range of about 0 °C to about 10 °C, about 10 °C to about 20 °C, about 20 °C to about 40 °C, about 40 °C to about 50 °C, about 50 °C to about 60 °C, about 60 °C to about 70 °C, about 70 °C to about 80 °C, about 80 °C to about 90 °C, about 90 °C to about 100 °C or higher. In some embodiments, the mixing is performed at a temperature of about 40 °C. In some embodiments, the mixing is performed at a temperature of about 85 °C. In some embodiments, the mixing is performed at a temperature of about 90 °C. In some embodiments, the mixing is performed at a temperature of about 95 °C. In some embodiments, the mixing is performed at a speed that will not negatively affect the properties of the composition, such as a speed of about 100 RPM, 200 RPM, 300 RPM, 400 RPM, 500 RPM, 600 RPM, 700 RPM, 800 RPM, 900 RPM, 1000 RPM, or more. In some embodiments the mixing lasts for about 1 minute, about 2 minutes, about 3 minutes, about 4 minutes, about 5 minutes, about 6 minutes, about 7 minutes, about 8 minutes, about 9 minutes, about 10 minutes, or more.

[444] In some embodiments, the composition is mixed only once. In some embodiments, the composition is mixed more than once, such as twice, three times, four times, five times, or more. In some embodiments, the temperature is changed between each mix. For example, in some

embodiments, the composition is mixed a first time at a first temperature, and a second time at a second temperature. In some embodiments, the composition is mixed a first time at a first temperature, a second time at a second temperature, and a third time at a third temperature. In some embodiments, the composition is mixed a first time at a first temperature, a second time at a second temperature, a third time at a third temperature, and a fourth time at a fourth temperature. In some embodiments, the composition is mixed a first time at 40 °C, a second time at 95 °C, a third time at 90 °C, and a fourth time at 85 °C. After mixing and/or between different mixings the composition may be allowed to rest.

[445] The compositions are then poured into molds. The molds may be of any shape, such as cube-shaped, cylindrical-shaped, triangular prism-shaped, spherical-shaped, cone-shaped, or rectangular prism-shaped. The compositions may then be covered, cooled and stored. In some embodiments, the compositions may be stored for at least 1 day, at least 3 days, at least 5 days, at least 7 days, at least 30 days, at least 180 days, or at least 360 days.

[446] The pH of the composition may be monitored during production thereof. In some embodiments, the pH may be adjusted to a target pH, such as a pH in the range of about 5.5 to about 5.7. As will be understood by those of ordinary skill in the art, the pH may be adjusted up or down using acids or bases. Exemplary acids that may be used to adjust the pH include lactic acid, citric acid, or sodium citrate.

[447] An illustrative method for preparing a food composition of the disclosure is provided in **FIG. 13**. The first step in this method is production of a seed expressing a fusion protein. In this process, an expression construct is designed. The construct is then transformed into a plant. The plant is grown under conditions that allow for expression of the fusion protein. Subsequently, seeds may be collected from the plant for further processing.

[448] The next step in the method for preparing a food composition illustrated in **FIG. 13** is seed processing, to prepare one or more ingredients for use in a food composition. First, the seeds are hulled and ground. Protein (including the fusion protein and other seed proteins) is extracted from the seed. The protein fraction may then be enriched. Specifically, the protein fraction may be enriched for fusion protein. Optionally, the fusion protein may then be concentrated.

[449] The plant protein, including fusion proteins, may be extracted from a plant using standard methods known in the art. For example, the proteins may be extracted using solvent or aqueous extraction. In some embodiments, the oil may be separated from the proteins using hexane or ethanol extraction to produce a white flake. The proteins may be extracted from the white flake using controlled temperature in an aqueous buffered environment (e.g., carbonate, citrate), in order to control the pH. The fusion proteins can be separated from the plant proteins using selective precipitation of one or more of the proteins with centrifugation or filtration methods. In some embodiments, one or more additives may be used to aid the extraction processes (e.g., salts, protease/peptidase inhibitors, osmolytes, solvents, reducing agents, etc.) The following step is processing the fusion protein into a food product. In some embodiments, constituent proteins of the fusion protein may be separated from one another before they are used to formulate a product. In some embodiments, only one of the constituent proteins of the fusion protein is used in the product. In some embodiments, more than one of the constituent proteins of the fusion protein is used in the product. In some embodiments, all of the constituent proteins of the fusion protein may be used in the product. In some embodiments, the fusion protein may be used itself in the food product. The product is then formulated as desired.

[450] FIG. 17 also illustrates a method for preparing a food composition. In this method, after seeds are collected, hulled and ground, and protein has been extracted, the fusion protein is separated from other seed protein. In some embodiments, this separation is not 100% efficient, meaning that the “other seed protein” fraction may still contain some residual fusion protein. For example, in some embodiments, the other seed protein fraction may comprise about 0.1%, about 0.3%, about 0.5%, about 0.7%, about 1%, about 2%, about 3%, about 4%, about 5%, about 6%, about 7%, about 8%, about 9%, about 10%, about 20%, about 30%, or about 50% fusion protein by weight. The other seed protein fraction may then be used directly in a food composition. Alternatively, the other seed protein fraction may be combined with concentrated fusion protein. In some embodiments, the other seed protein fraction is combined with one or more of the constituent proteins from the fusion protein. In some embodiments, the other seed protein fraction is combined with all of the constituent proteins from the fusion protein.

[451] It may be advantageous to use a seed processing composition comprising plant protein and a fusion protein (e.g., about 0.1%, about 0.3%, about 0.5%, about 0.7%, about 1%, about 2%, about

3%, about 4%, about 5%, about 6%, about 7%, about 8%, about 9%, about 10%, about 20%, about 30%, or about 50% fusion protein by weight) as an ingredient in a food composition. Using both (i) a fusion protein produced by a seed and (ii) other protein extracted from the seed allows for efficient use of resources and reduces waste. Such processes may simplify food manufacturing processes, and reduce the unit cost to manufacture each product. Thus, provided herein is a method of making a food composition, the method comprising: (i) expressing a fusion protein in a transformed plant; and (ii) preparing a food composition comprising the fusion protein and plant protein from the same transformed plant in which the fusion protein was produced. In some embodiments, the transformed plant is a soybean. In some embodiments, the transformed plant is pea.

[452] Without being bound by any theory, it is believed that having a casein protein (i.e., as a monomer or as part of a fusion protein) in a plant protein composition may improve the properties of the plant protein composition. **FIG. 19** illustrates various properties that may be improved due to the presence of one or more caseins in a plant protein composition, including. In some embodiments, a plant protein composition comprising one or more casein proteins has improved nutritional properties compared to a plant protein composition that does not contain a casein protein. In some embodiments, a plant protein composition comprising one or more casein proteins has improved organoleptic properties, such as taste, compared to a plant protein composition that does not contain a casein protein. In some embodiments, a plant protein composition comprising one or more casein proteins has improved water holding capacity compared to a plant protein composition that does not contain a casein protein. In some embodiments, a plant protein composition comprising one or more casein proteins has improved emulsification compared to a plant protein composition that does not contain a casein protein. In some embodiments, a plant protein composition comprising one or more casein proteins has improved gelation compared to a plant protein composition that does not contain a casein protein. In some embodiments, a plant protein composition comprising one or more casein proteins has improved viscosity and/or adhesiveness compared to a plant protein composition that does not contain a casein protein. In some embodiments, a plant protein composition comprising one or more casein proteins has improved aeration and/or foaming compared to a plant protein composition that does not contain a casein protein. In some embodiments, a plant protein composition comprising one or more casein

proteins has improved solubility compared to a plant protein composition that does not contain a casein protein. Illustrative improvements in each one of these properties are described in further detail below.

[453] Nutrition: The presence of a casein protein (alone, or expressed as a fusion protein) in the plant protein composition may enhance the nutritional properties of the plant protein composition and/or any food compositions comprising the plant protein composition. For example, the presence of the casein protein may, in some embodiments, improve the balance of essential amino acids. Pea protein has a PDCAAS (protein digestibility corrected amino acid score) of about 0.82. Nutritionally complete proteins have a score of about 1.0. By expressing a casein protein (fused to, for example, ovalbumin and/or beta-lactoglobulin) at sufficient levels in a pea plant, the PDCAAS of the protein extracted from the pea plant may reach 1.0, provided that the limiting amino acids (e.g., methionine) are raised. In some embodiments, a plant protein composition comprising a casein protein comprises a PDCAAS of about 0.90, about 0.95, about 1.0, or about 1.05.

[454] Gelation: In some embodiments, the casein protein present in the plant protein composition may enhance gelation of the plant protein composition and/or any food compositions comprising the plant protein composition. Many of the proteins used as fusion partners in the fusion proteins described herein, including whey proteins (e.g., beta-lactoglobulin) and egg proteins are often added to a number of food products such as meats and bakery products, because the proteins gel after heating and cooling. Seed proteins are generally insoluble under the processing conditions used to prepare many foods, such as meats and bakery products. Methylcellulose is often added to plant-based meats to impart gelling, and egg white has historically been used in some vegetarian products. However, eggs are not considered vegan and do not meet the standard of “plant-based” for many individuals. Thus, by using a plant composition comprising one or more casein proteins (fused to, for example, an egg protein and/or a whey protein), enhanced gelation may be achieved without using animal products.

[455] Solubility: In some embodiments, the casein protein present in the plant protein composition may enhance solubility of the plant protein composition and/or any food compositions comprising the plant protein composition. Seed proteins typically have poor solubility at acidic

and neutral pH. Beverage formulations are suspensions utilizing hydrocolloids such as gellan gum to keep the proteins from settling out. Conversely, casein proteins are soluble at neutral pH, and whey proteins are soluble at acidic pH. Both caseins and whey are soluble at neutral pH. In some embodiments, beverages made with seed protein enhanced by the expression of casein proteins (expressed alone or fused to, for example, a whey protein) exhibit a smoother and/or less chalky mouthfeel.

[456] Emulsification: In some embodiments, the casein protein present in the plant protein composition may enhance emulsification of the plant protein composition and/or any food compositions comprising the plant protein composition. Caseinates are effective at emulsifying lipids with a low viscosity, and this property is used in spray drying to produce powdered coffee creamers and powdered sauces with lipids used in convenience foods. Seed proteins do not have these attributes, and additives such as starches chemically modified with octenyl succinic anhydride are often used as additives in plant protein compositions. Food compositions made with plant protein compositions comprising casein proteins will have improved emulsification properties for a number of different applications.

[457] Water holding capacity: In some embodiments, the casein protein may enhance the water holding capacity of the plant protein. During the processing of the plant protein, pH and heat conditions can be modified to denature the casein protein to enhance this property.

[458] Aeration/Foaming: Aeration and foaming properties of the plant protein can be improved by the addition of the casein proteins. Caseins have excellent foaming properties, as evidenced by their incorporation in frozen whipped toppings. Egg proteins and beta-lactoglobulin also demonstrate good foaming properties. The surface-active properties of these proteins are beneficial in food compositions.

[459] Viscosity/Adhesiveness: Unstructured casein proteins can unfold to interact with other components of a food composition to impart viscosity and adhesiveness. Granola bars can utilize casein proteins at specific concentrations to form a viscous solution that holds the particulates together.

[460] Flavor: The casein proteins can also improve the flavor of plant proteins. In addition to acting as binders for off flavors, casein proteins can impart desirable flavors to food

compositions. Hydrolyzed protein caseins impart a savory umami flavor similar to those from autolyzed yeast extract. Some of the expressed casein are hydrolyzed by plant enzymes in the seed, and the resultant peptides can provide savory flavors.

[461] In some embodiments, a plant protein composition comprising a fusion protein is used to produce a food composition. The food composition may be, for example, a meat analog, a nutritional bar, a bakery product, a beverage, mashed potatoes, or candy. In some embodiments, the food composition is for a human. For example, the food composition may be infant formula. In some embodiments, the food composition is for a companion animal (e.g., a dog, cat, rabbit, hamster, guinea pig, horse, etc.) For example, the food composition may be pet food.

[462] Also provided herein are various compositions prepared during a method of making a food composition. For example, in some embodiments, a seed processing composition is provided. In some embodiments, a seed processing composition comprises (a) a fusion protein comprising i) a full-length κ -casein or para- κ -casein component; and ii) a β -lactoglobulin component; and (b) plant seed tissue. In some embodiments, a seed processing composition comprises (a) a fusion protein comprising i) a beta-casein component; and ii) a β -lactoglobulin component; and (b) plant seed tissue. In some embodiments, a seed processing composition comprises (a) a fusion protein comprising i) a milk protein (e.g., a casein protein); and ii) a second protein (i.e., a fusion partner); and (b) plant seed tissue. In some embodiments, the plant seed tissue is ground. In some embodiments, the plant seed tissue is from soybean. In some embodiments, the seed processing composition comprises at least one member selected from the group consisting of: enzyme (e.g., chymosin), protease, extractant, solvent (e.g., ethanol, or hexane), buffer, additive, salt, protease inhibitor, peptidase inhibitor, osmolyte, and reducing agent.

[463] In some embodiments, a protein concentrate composition is provided. In some embodiments, the protein concentrate composition comprises: a fusion protein, comprising i) a full-length κ -casein or para- κ -casein component; and ii) a β -lactoglobulin component. In some embodiments, the protein concentrate composition comprises: a fusion protein, comprising i) a beta-casein component; and ii) a β -lactoglobulin component. In some embodiments, the protein concentrate composition comprises: a fusion protein, comprising i) a milk protein (e.g., a casein protein); and ii) a second protein (i.e., a fusion partner). In some embodiments, the fusion protein

is present in an enriched amount, relative to other components present in the composition. In some embodiments, there is substantially no plant seed tissue present in the protein concentrate composition. In some embodiments, the protein concentrate composition further comprises at least one member selected from the group consisting of: enzyme (e.g., chymosin), protease, extractant, solvent (e.g., ethanol, or hexane), buffer, additive, salt, protease inhibitor, peptidase inhibitor, osmolyte, and reducing agent.

[464] In some embodiments, a food composition comprises a fusion protein comprising a first protein and a second protein. In some embodiments, a food composition comprises a first protein, wherein the first protein is derived from (i.e., separated from) a fusion protein comprising at least the first protein and a second protein. In some embodiments, a food composition comprises (i) a fusion protein comprising a first protein and a second protein and (ii) at least one of the first protein and the second protein, wherein the first protein and/or the second protein has been separated from the fusion protein. The first protein and/or second protein which have been separated from the fusion protein may comprise, in some embodiments, at least at least one non-native amino acid from an introduced protease cleavage site (e.g., a chymosin cleavage site).

[465] In some embodiments, the food composition is a solid. In some embodiments, the food composition is a liquid. In some embodiments, the food composition is a powder.

[466] In some embodiments, the food composition is a solid phase, protein-stabilized emulsion. In some embodiments, the food composition is a colloidal suspension.

[467] In some embodiments, the fusion proteins and transgenic plants described herein may be used to prepare a food composition such as cheese or processed cheese products. In some embodiments, the food composition is an alternative dairy composition selected such as milk, cream, or butter. The alternative milk composition may be used to prepare alternative dairy compositions such as yogurt and fermented dairy products, directly acidified counterparts of fermented dairy products, cottage cheese, dressing, curds, crème fraîche, toppings, icings, fillings, low-fat spreads, dairy-based dry mixes, frozen dairy products, frozen desserts, desserts, baked goods, soups, sauces, salad dressing, geriatric nutrition, creams and creamers, analog dairy products, follow-up formula, baby formula, infant formula, milk, dairy beverages, acid dairy drinks, smoothies, milk tea, butter, margarine, butter alternatives, growing up milks, low-lactose

products and beverages, medical and clinical nutrition products, protein/nutrition bar applications, sports beverages, confections, meat products, analog meat products, meal replacement beverages, and weight management food and beverages.

[468] In some embodiments the fusion proteins and transgenic plants described herein may be used to prepare a dairy product. In some embodiments, the dairy product is a fermented dairy product. An illustrative list of fermented dairy products includes cultured buttermilk, sour cream, yogurt, skyr, leben, lassi, or kefir. In some embodiments the fusion proteins and transgenic plants described herein may be used to prepare cheese products.

[469] In some embodiments the fusion proteins and transgenic plants described herein may be used to prepare a powder containing a milk protein. In some embodiments, the fusion proteins and transgenic plants described herein may be used to prepare a low-lactose product.

[470] In some embodiments, a method for making a food composition comprises, expressing a recombinant fusion protein of the disclosure in a plant, extracting the recombinant fusion protein from the plant, optionally separating the milk protein from the mammalian or plant protein, and creating a food composition using the fusion protein and/or the milk protein.

[471] In some embodiments, a method of expressing, extracting, and making a food composition from a fusion protein, comprises: expressing a fusion protein in a host cell, the fusion protein comprising a first protein and a second protein; extracting the fusion protein from the host cell; and processing the fusion protein into a food composition. The food composition may be, for example, cheese, processed cheese product, yogurt, fermented dairy product, directly acidified counterpart of fermented dairy product, cottage cheese dressing, frozen dairy product, frozen dessert, dessert, baked good, topping, icing, filling, low-fat spread, dairy-based dry mix, soup, sauce, salad dressing, geriatric nutrition, cream, creamer, analog dairy product, follow-up formula, baby formula, infant formula, milk, dairy beverage, acid dairy drink, smoothie, milk tea, butter, margarine, butter alternative, growing up milk, low-lactose product, low-lactose beverage, medical and clinical nutrition product, protein bar, nutrition bar, sport beverage, confection, meat product, analog meat product, meal replacement beverage, weight management food and beverage, dairy product, cultured buttermilk, sour cream, yogurt, skyr, leben, lassi, kefir, powder containing a milk

protein, and low-lactose product. In some embodiments, the food composition is a dairy product. In some embodiments, the food composition is a cheese.

[472] In some embodiments, a method for making a food composition comprises, expressing a recombinant fusion protein of the disclosure in a plant, extracting one or both of the proteins, and creating a food composition using the milk protein. In some embodiments, the first protein and the second protein are separated from one another in the plant cell, prior to extraction. In some embodiments, the first protein is separated from the second protein after extraction, for example by contacting the fusion protein with an enzyme that cleaves the fusion protein. The enzyme may be, for example, chymosin. In some embodiments, the fusion protein is cleaved using rennet.

[473] All references, articles, publications, patents, patent publications, and patent applications cited herein are incorporated by reference in their entireties for all purposes. However, mention of any reference, article, publication, patent, patent publication, and patent application cited herein is not, and should not be taken as an acknowledgment or any form of suggestion that they constitute valid prior art or form part of the common general knowledge in any country in the world, or that they disclose essential matter.

EXAMPLES

[474] The following experiments demonstrate different recombinant fusion constructs comprising a milk protein (e.g., a casein) and at least one other protein, as well as methods of producing and testing the fusion proteins. While the examples below describe expression in soybean, it will be understood by those skilled in the art that the constructs and methods disclosed herein may be tailored for expression in any organism.

[475] The following examples also demonstrate the production of various cheese compositions and characterization of their properties. Traditionally cheese is made from milk, which comprises a mixture of casein proteins. To test whether a cheese composition having acceptable organoleptic and physical properties could be made using only one casein protein, or different combinations/ratios of casein proteins as compared to that found in any mammalian milk, various experiments described below were performed. While the examples below utilize isolated caseins isolated from bovine milk, it will be understood by those skilled in the art that the recipes and

methods disclosed herein may be tailored for use with other isolated caseins and recombinant caseins, including caseins expressed in a plant.

Example 1: Construction of expression vectors for plant transformation for stable expression of recombinant fusion proteins

Binary Vector Design

[476] While a number of vectors may be utilized for expression of the fusion proteins disclosed herein, the example constructs described below were built in the binary pCAMBIA3300 (Creative Biogene, VET1372) vector, which was customized for soybean transformation and selection. In order to modify the vector, pCAMBIA3300 was digested with *Hind*III and *Ase*I allowing the release of the vector backbone (LB T-DNA repeat_KanR_pBR322 ori_pBR322 bom_pVS1 oriV_pVs1 repA_pVS1 StaA_RB T-DNA repeat). The 6598 bp vector backbone was gel extracted and a synthesized multiple cloning site (MCS) was ligated via In-Fusion cloning (In-Fusion® HD Cloning System CE, available on the world wide web at clontech.com) to allow modular vector modifications. A cassette containing the *Arabidopsis thaliana* Csr1.2 gene for acetolactate synthase was added to the vector backbone to be used as a marker for herbicide selection of transgenic plants. In order to build this cassette, the regulatory sequences from *Solanum tuberosum* ubiquitin/ribosomal fusion protein promoter (StUbi3 prom; -1 to -922 bp) and terminator (StUbi3 term; 414 bp) (GenBank accession no. L22576.1) were fused to the mutant (S653N) acetolactate synthase gene (Csr1.2; GenBank accession no. X51514.1) (Sathasivan et al, 1990; Ding et al, 2006) to generate imazapyr-resistant traits in soybean plants. The selectable marker cassette was introduced into the digested (EcoRI) modified vector backbone via In-Fusion cloning to form vector pAR15-00 (FIG. 2).

[477] Recombinant DNA constructs were designed to express milk proteins in transgenic plants. The coding regions of the expression cassettes outlined below contain a fusion of codon-optimized nucleic acid sequences encoding bovine milk proteins, or a functional fragment thereof. To enhance protein expression in soybean, the nucleic acid sequences encoding β -lactoglobulin (GenBank accession no. X14712.1), κ -casein (GenBank accession no. CAA25231), β -casein (GenBank accession no. M15132.1), and α S1-casein (GenBank accession no. X59836.1) were codon optimized using *Glycine max* codon bias and synthesized (available on the world wide web

at [idtdna.com/CodonOpt](https://www.idtdna.com/CodonOpt)). The signal sequences were removed (i.e., making the constructs “truncated”) and the new versions of the genes were renamed as OLG1 (β -lactoglobulin version 1, SEQ ID NO: 9), OLG2 (β -lactoglobulin version 2, SEQ ID NO: 11), OLG3 (β -lactoglobulin version 3, SEQ ID NO: 12), OLG4 (β -lactoglobulin version 4, SEQ ID NO: 13), OKC1-T (Optimized κ -casein Truncated version 1, SEQ ID NO: 3), paraOKC1-T (only the para- κ portion of OKC1-T, SEQ ID NO: 1), OBC-T2 (Optimized β -casein Truncated version 2, SEQ ID NO: 5), and OaS1-T (Optimized α S1-casein Truncated version 1, SEQ ID NO: 7). As will be understood by those skilled in the art, codon optimized nucleic acid sequences can present from about 60% to about 100% identity to the native version of the nucleic acid sequence.

[478] All the expression cassettes described below and shown in **FIG. 4-FIG. 9** contained codon-optimized nucleic acid sequences encoding bovine milk proteins, or a functional fragment thereof, a seed specific promoter, a 5'UTR, a signal sequence (Sig) that directs foreign proteins to the protein storage vacuoles, and a termination sequence. In some versions of the constructs a linker such as a linker comprising a chymosin cleavage site (FM), was placed between the two proteins and/or a C-terminal KDEL sequence for ER retention was included. Expression cassettes were inserted in the pAR15-00 vector described above utilizing a *KpnI* restriction site with the MCS (**FIG. 3**). Coding regions and regulatory sequences are indicated as blocks (not to scale) in **FIG. 4-FIG. 9**.

κ -casein- β -lactoglobulin fusion with KDEL

[479] Shown in **FIG. 4** is an example expression cassette comprising κ -casein (OKC1-T, SEQ ID NO: 3) and β -lactoglobulin (OLG1, SEQ ID NO: 9). The regulatory sequences that were used in order to produce the heterologous milk proteins in soybean seeds include the promoter of the beta-phaseolin storage protein gene (PvPhas prom; -1 to -1543; GenBank accession no. J01263.1, SEQ ID NO: 18); the 5'UTR of the arc5-1 gene (arc5'UTR; -1 to -13; GenBank accession no. Z50202, SEQ ID NO: 20) (De Jaeger et al, 2002); the signal peptide of Lectin 1 gene 1 (sig10; +1 to +93; GenBank accession no. Glyma.02G012600, SEQ ID NO: 14) (Darnowski et al, 20020); and, the 3'UTR of the arc5-1 gene, (arc term 1197 bp; GenBank accession no. Z50202.1, SEQ ID NO: 21)(De Jaeger et al, 2002). A C-terminal KDEL (SEQ ID NO: 23) was also included for ER retention.

β -casein- β -lactoglobulin fusion with linker

[480] Shown in **FIG. 5** is an example expression cassette comprising β -casein (OBC-T2, SEQ ID NO: 5) and β -lactoglobulin (OLG1, SEQ ID NO: 9). The regulatory sequences that were used in order to produce the heterologous milk proteins in soybean seeds include the promoter of the beta-phaseolin storage protein gene (PvPhas prom; -1 to -1543; GenBank accession no. J01263.1, SEQ ID NO: 18); the 5'UTR of the arc5-1 gene (arc5'UTR; -1 to -13; GenBank accession no. Z50202, SEQ ID NO: 20) (De Jaeger et al, 2002); the signal peptide of Lectin 1 gene 1 (sig10; +1 to +93; accession no. Glyma.02G012600, SEQ ID NO: 14) (Darnowski et al, 2002); and, the 3'UTR of the arc5-1 gene, (arc term 1197 bp; accession no. Z50202.1, SEQ ID NO: 21) (De Jaeger, et al 2002). A linker comprising a chymosin cleavage site (FM) was inserted between the two proteins.

 α S1-casein- β -lactoglobulin fusion with linker

[481] Shown in **FIG. 6** is an example expression cassette comprising α S1-casein (OaS1-T, SEQ ID NO: 7) and β -lactoglobulin (OLG1, SEQ ID NO: 9). The regulatory sequences that were used in order to produce the heterologous milk proteins in soybean seeds include the promoter of the beta-phaseolin storage protein gene (PvPhas prom; -1 to -1543; GenBank accession no. J01263.1, SEQ ID NO: 18); the 5'UTR of the arc5-1 gene (arc5'UTR; -1 to -13; GenBank accession no. Z50202, SEQ ID NO: 20) (De Jaeger et al, 2002); the signal peptide of Lectin 1 gene 1 (sig10; +1 to +93; accession no. Glyma.02G012600, SEQ ID NO: 14) (Darnowski et al, 2002); and, the 3'UTR of the arc5-1 gene, (arc term 1197 bp; GenBank accession no. Z50202.1, SEQ ID NO: 21)(De Jaeger et al, 2002). A linker comprising a chymosin cleavage site (FM) was inserted between the two proteins.

Para- κ -casein- β -lactoglobulin fusion with linker and KDEL

[482] Shown in **FIG. 7** is an example expression cassette comprising para- κ -casein (paraOKC1-T, SEQ ID NO: 1) and β -lactoglobulin (OLG1, SEQ ID NO: 9). The regulatory sequences that were used in order to produce the heterologous milk proteins in soybean seeds include the promoter of the beta-phaseolin storage protein gene (PvPhas prom; -1 to -1543; GenBank accession no. J01263.1, SEQ ID NO: 18); the 5'UTR of the arc5-1 gene (arc5'UTR; -1 to -13; GenBank accession no. Z50202, SEQ ID NO: 20) (De Jaeger et al, 2002); the signal peptide of Lectin 1 gene 1 (sig10; +1 to +93; GenBank accession no. Glyma.02G012600, SEQ ID NO: 14) (Darnowski et

al, 2002); and, the 3'UTR of the arc5-1 gene, (arc term 1197 bp; GenBank accession no. Z50202.1, SEQ ID NO: 21) (De Jaeger et al 2002). A linker comprising a chymosin cleavage site (FM) was inserted between the two proteins and a C-terminal KDEL (SEQ ID NO: 23) was also included for ER retention.

Para- κ -casein- β -lactoglobulin fusion with linker

[483] Shown in **FIG. 8** is an example expression cassette comprising para- κ -casein (paraOKC1-T, SEQ ID NO: 1) and β -lactoglobulin (OLG1, SEQ ID NO: 9). The regulatory sequences that were used in order to produce the heterologous milk proteins in soybean seeds include the promoter of the beta-phaseolin storage protein gene (PvPhas prom; -1 to -1543; GenBank accession no. J01263.1, SEQ ID NO: 18); the 5'UTR of the arc5-1 gene (arc5'UTR; -1 to -13; GenBank accession no. Z50202, SEQ ID NO: 20) (De Jaeger et al, 2002); the signal peptide of Lectin 1 gene 1 (sig10; +1 to +93; GenBank accession no. Glyma.02G012600, SEQ ID NO: 14) (Darnowski et al, 2002); and, the 3'UTR of the arc5-1 gene, (arc term 1197 bp; GenBank accession no. Z50202.1, SEQ ID NO: 21) (De Jaeger et al, 2002). A linker comprising a chymosin cleavage site (FM) was inserted between the two proteins.

Fusion protein with Seed2 Promoter, Sig2 and nopaline synthase terminator

[484] Shown in **FIG. 9** is an example expression cassette comprising κ -casein (OKC1-T, SEQ ID NO: 3) and β -lactoglobulin (OLG1, SEQ ID NO: 9). The regulatory sequences that were used in order to produce the heterologous milk proteins in soybean seeds include the promoter and signal peptide of glycinin 1 (GmSeed2 (SEQ ID NO: 19):sig2 (SEQ ID NO: 16)) followed by the ER retention signal (KDEL) and the *Nopaline synthase* termination sequence (nos term, SEQ ID NO: 22).

Exemplary Protein Co-expression Vector

[485] Binary pCambia3300 vectors individually encoding for: (1) a prolamin (e.g., Canein or Zein); (2) a milk protein (e.g., Casein); or (3) both a prolamin and a milk protein are generated to co-express a milk protein and prolamin in plant cells (See **FIG. 26A-26G**, **FIG. 27**). Co-expression of a milk protein and prolamin will result in generation of a protein body in the plant cell capable of shielding the milk protein from degradation or capable of reducing toxicity, if any, associated with recombinant expression of the milk protein in the plant cell.

Example 2: Identification of transgenic events, recombinant protein extraction and detection

[486] To quantify recombinant protein expression levels, DNA constructs such as those shown in FIG. 4-FIG. 9 were transformed into soybean using transformation protocols well known in the art, for example, by bombardment or agrobacterium. Total soybean genomic DNA was isolated from the first trifoliolate leaves of transgenic events using the PureGene tissue DNA isolation kit (product #158667: QIAGEN, Valencia, CA, USA). Trifoliolates were frozen in liquid nitrogen and pulverized. Cells were lysed using the PureGene Cell Lysis Buffer, proteins were precipitated using the PureGene Protein Precipitation Buffer, and DNA was precipitated from the resulting supernatant using ethanol. The DNA pellets were washed with 70% ethanol and resuspended in water.

[487] Genomic DNA was quantified by the Quant-iT PicoGreen (product #P7589: ThermoFisher Scientific, Waltham, MA, USA) assay as described by manufacturer, and 150 ng of DNA was digested overnight with *EcoRI*, *HindIII*, *NcoI*, and/or *KpnI*, 30 ng of which was used for a BioRad ddPCR reaction, including labelled FAM or HEX probes for the transgene and Lectin1 endogenous gene respectively. Transgene copy number (CNV) was calculated by comparing the measured transgene concentration to the reference gene concentration. A CNV of greater than or equal to one was deemed acceptable.

Preparation of Total Soluble Protein Samples

[488] Total soluble soybean protein fractions were prepared from the seeds of transgenic events by bead beating seeds (seeds collected about 90 days after germination) at 15000 rpm for 1 min. The resulting powder was resuspended in 50 mM Carbonate-Bicarbonate pH 10.8, 1 mM DTT, 1X HALT Protease Inhibitor Cocktail (Product # 78438 ThermoFisher Scientific). The resuspended powder was incubated at 4°C for 15 minutes and then the supernatant collected after centrifuging twice at 4000g, 20 min, 4°C. Protein concentration was measured using a modified Bradford assay (Thermo Scientific Pierce 660 nm assay; Product #22660 ThermoFisher Scientific) using a bovine serum albumin (BSA) standard curve.

Recombinant Protein Quantification via Western Blot Densitometry

[489] SDS-PAGE was performed according to manufacturer's instructions (Product# 5678105BioRad, Hercules, CA, USA) under denaturing and reducing conditions. 5 µg of total

protein extracts were loaded per lane. For immunoblotting proteins separated by SDS-PAGE were transferred to a PVDF membrane using Trans-Blot® Turbo™ Midi PVDF Transfer Packs (Product# 1704157 BioRad) according to manufacturer's guidelines. Membranes were blocked with 3% BSA in phosphate buffered saline with 0.5% Tween-20, reacted with antigen specific antibody and subsequently reacted with fluorescent goat anti rabbit IgG (Product #60871 BioRad, CA). Membranes were scanned according to manufacturer's instructions using the ChemiDoc MP Imaging System (BioRad, CA) and analyzed using ImageLab Version 6.0.1 Standard Edition (BioRad Laboratories, Inc.) Recombinant protein from the seeds of transgenic events was quantified by densitometry from commercial reference protein spike-in standards.

[490] Shown in **FIG. 10A**, **FIG. 10B**, **FIG. 10C**, and **FIG. 10D** are Western Blots of protein extracted from transgenic soybeans expressing the κ -casein- β -lactoglobulin expression cassette shown in **FIG. 4**. **FIG. 10A** shows the fusion protein detected using a primary antibody raised against κ -casein. The first lane is a molecular weight marker. Lanes two (DCI 9.1) and three (DCI 9.2) represent individual seeds from a single transgenic line. Lane four (DCI 3.1) represents a seed from a separate transgenic line. Lane five is protein extracted from wild-type soybean plants, and lanes six – eight are protein extracted from wild-type soybean plants spiked with 0.05% commercial κ -casein (lane 6), 0.5% commercial κ -casein (lane 7), and 1.5% commercial κ -casein (lane 8). The κ -casein commercial protein is detected at an apparent molecular weight (MW) of ~26 kDa (theoretical: 19 kDa - arrow). The fusion protein is detected at an apparent MW of ~40 kDa (theoretical: 38 kDa - arrowhead).

[491] **FIG. 10B** shows the fusion protein detected using a primary antibody raised against β -lactoglobulin. The first lane is a molecular weight marker. Lanes two (DCI 9.1) and three (DCI 9.2) represent individual seeds from a single transgenic line. Lane four (DCI 3.1) represents a seed from a separate transgenic line. Lane five is protein extracted from wild-type soybean plants, and lanes six – eight are protein extracted from wild-type soybean plants spiked with 0.05% commercial β -lactoglobulin (lane 6), 1% commercial β -lactoglobulin (lane 7), and 2% commercial β -lactoglobulin (lane 8). The β -lactoglobulin commercial protein is detected at an apparent MW of ~18kDa (theoretical: 18 kDa - arrow). The fusion protein is detected at an apparent MW of ~40 kDa (theoretical: 38 kDa - arrowhead). **FIG. 10C** and **FIG. 10D** show the protein gels as control

for equal lane loading (image is taken at the end of the SDS run) for **FIG. 10A** and **FIG. 10B**, respectively.

[492] Shown in **FIG. 15A** and **FIG. 15B** are Western Blots of protein extracted from transgenic soybeans expressing a β -casein- β -lactoglobulin fusion protein. **FIG. 15A** shows the fusion protein detected using a primary antibody raised against β -casein. The first lane is a molecular weight marker. Lane two (IX2) represents individual seeds from a single transgenic line. Lanes three through seven are samples comprising protein extracted from wild-type soybean plants spiked with 3% commercial β -casein (lane 3), 1.5% commercial β -casein (lane 4), 0.75% commercial β -casein (lane 5), 0.37% commercial β -casein (lane 6), and 0% commercial β -casein (lane 7). The fusion protein was detected at an apparent MW of ~40 kDa (arrow; theoretical: 42 kDa).

[493] Other combinations of proteins were tested and evaluated for the percentage of recombinant protein. Cassettes having the same promoter (Seed2-sig), signal peptide (EUT:Rb7T), and in some instances a different terminator, were built with either α -S1-casein, β -casein, κ -casein, or the fusion of β -lactoglobulin (LG) with κ -casein (kCN) (See **FIG. 3** and **FIG. 8**). As shown below in Table 18, none of the cassettes encoding α -S1-casein, β -casein, or κ -casein alone were able to produce expression of the protein at a level that exceeded 1% total soluble protein. However, when κ -casein was fused with β -lactoglobulin, κ -casein was expressed at a level that was greater than 1% total soluble protein. Similarly, when β -casein or alpha-S1-casein were fused with β -lactoglobulin, the β -casein and the alpha-S1-casein were expressed at a level that was greater than 1% total soluble protein.

Table 18: Expression levels of milk proteins expressed alone or in a fusion protein

			Number of events ¹ accumulating the recombinant protein at the concentration:	
		Total events ¹ analyzed	0-1% TSP	Above 1% TSP
Single Proteins	κ -Casein	89	89	0
	β -Casein	12	12	0
	α S1-Casein	6	6	0
Fusion	κ -Casein -LG	23	12	11

	β -Casein-LG	25	5	20
	α S1-Casein-LG	10	4	6

¹As used in Table 18, the each “event” refers to an independent transgenic line.

[494] As will be readily understood by those of skill in the art, T-DNA insertion into the plant genome is a random process and each T-DNA lands at an unpredictable genomic position. Thus, for example, each of the 23 events generated in Table 18 for the κ -Casein-LG fusion protein have different genomic insertion loci. The genomic context greatly influences the expression levels of a gene, and each locus will be either favorable or unfavorable for the expression of the recombinant genes. The variability observed at the protein level is a reflection of that random insertion process, and explains why 12 out of 23 events present expression levels below 1%.

Example 3: Expression of Casein Multimers

[495] A casein multimer is a fusion protein comprising at least a first casein protein and a second casein protein, wherein the first and second casein proteins are the same (homo-multimer) or different (hetero-multimer). Expression vectors for producing casein multimers were created, using the methods described in Example 1. Specifically, expression vectors were created to express casein multimers comprising: (i) kappa-casein fused to kappa-casein, (ii) kappa-casein fused to beta-casein, and (iii) kappa-casein fused to alpha-S1-casein. Expression vectors were also created to express: (iv) kappa-casein fused to GFP, and (v) kappa-casein fused to beta-lactoglobulin.

[496] Illustrative casein multimers prepared during this study are shown below in Table 19. Colons (:) are used to indicate junctions between various elements of the fusion protein. KDEL indicates the use of a KDEL sequence (i.e., an endoplasmic reticulum retention signal) and FM indicates the use of a linker comprising a chymosin cleavage site.

Table 19: Illustrative Casein Multimers

Description	Abbreviated Description	DNA Sequence (SEQ ID NO)	Amino Acid Sequence (SEQ ID NO)
Optimized para kappa-casein truncated version 1 (paraOKC1-T):FM: Optimized beta-lactoglobulin version 1 (OLG1)	paraOKC1-T:FM:OLG1	615	616

Description	Abbreviated Description	DNA Sequence (SEQ ID NO)	Amino Acid Sequence (SEQ ID NO)
Optimized para kappa-casein truncated version 1 (paraOKC1-T):FM: Optimized beta-lactoglobulin version 1 (OLG1):KDEL	paraOKC1-T:FM:OLG1:KDEL	617	618
Optimized para kappa-casein truncated version 1 (paraOKC1-T): Optimized beta-lactoglobulin version 1 (OLG1)	paraOKC1-T:OLG1	619	620
Optimized para kappa-casein truncated version 1 (paraOKC1-T): Optimized beta-lactoglobulin version 1 (OLG1):KDEL	paraOKC1-T:OLG1:KDEL	621	622
Optimized beta-lactoglobulin version 1 (OLG1):FM: Optimized para kappa-casein truncated version 1 (paraOKC1-T)	OLG:FM:paraOKC1-T	623	624
Optimized beta-lactoglobulin version 1 (OLG1):FM: Optimized para kappa-casein truncated version 1 (paraOKC1-T):KDEL	OLG:FM:paraOKC1-T:KDEL	625	626
Optimized beta-lactoglobulin version 1 (OLG1):Optimized para kappa-casein truncated version 1 (paraOKC1-T)	OLG:paraOKC1-T	627	628
Optimized beta-lactoglobulin version 1 (OLG1):Optimized para kappa-casein truncated version 1 (paraOKC1-T):KDEL	OLG:paraOKC1-T:KDEL	629	630
Optimized alpha S1-casein truncated version 1 (OaS1-T): FM: Optimized beta-lactoglobulin version 1 (OLG1)	OaS1-T:FM:OLG1	631	632
Optimized alpha S1-casein truncated version 1 (OaS1-T): FM: Optimized beta-lactoglobulin version 1 (OLG1):KDEL	OaS1-T:FM:OLG1:KDEL	633	634
Optimized alpha S1-casein truncated version 1 (OaS1-T): Optimized beta-lactoglobulin version 1 (OLG1)	OaS1-T:OLG1	635	636
Optimized alpha S1-casein truncated version 1 (OaS1-T): Optimized beta-lactoglobulin version 1 (OLG1):KDEL	OaS1-T:OLG1:KDEL	637	638

Description	Abbreviated Description	DNA Sequence (SEQ ID NO)	Amino Acid Sequence (SEQ ID NO)
Optimized beta-lactoglobulin version 1 (OLG1):FM: Optimized alpha S1-casein truncated version 1 (OaS1-T)	OLG1:FM:OaS1-T	639	640
Optimized beta-lactoglobulin version 1 (OLG1):FM: Optimized alpha S1-casein truncated version 1 (OaS1-T):KDEL	OLG1:FM:OaS1-T:KDEL	641	642
Optimized beta-lactoglobulin version 1 (OLG1): Optimized alpha S1-casein truncated version 1 (OaS1-T)	OLG1:OaS1-T	643	644
Optimized beta-lactoglobulin version 1 (OLG1): Optimized alpha S1-casein truncated version 1 (OaS1-T):KDEL	OLG1:OaS1-T:KDEL	645	646
Optimized beta-lactoglobulin version 1 (OLG1): FM: Optimized beta-casein (A2 variant) truncated version 2	OLG1:FM:OBC-T2	647	648
Optimized beta-casein (A2 variant) truncated version 2: Optimized kappa-casein truncated version 1 (OKC1-T): Optimized beta-lactoglobulin version 1 (OLG1)	OBC-T2:OKC1-T:OLG1	649	650
Optimized beta-casein (A2 variant) truncated version 3: Optimized beta-casein (A2 variant) truncated version 2: Optimized kappa-casein truncated version 1 (OKC1-T): Optimized beta-lactoglobulin version 1 (OLG1)	OBC-T3:OBC-T2:OKC1-T:OLG1	651	652
Optimized beta-casein (A2 variant) truncated version 4:Optimized beta-casein (A2 variant) truncated version 3: Optimized beta-casein (A2 variant) truncated version 2: Optimized kappa-casein truncated version 1 (OKC1-T): Optimized beta- lactoglobulin version 1 (OLG1)	OBC-T4:OBC-T3:OBC-T2:OKC1-T:OLG1	653	654
Optimized beta-casein (A2 variant) truncated version 5:Optimized beta-casein (A2 variant) truncated version 5:Optimized beta-casein (A2 variant) truncated version 4:Optimized beta-casein (A2 variant) truncated version 3: Optimized beta-casein (A2 variant) truncated version 2: Optimized para	OBC-T5:OBC-T4:OBC-T3:OBC-T2:OKC1-T:OLG1	655	656

Description	Abbreviated Description	DNA Sequence (SEQ ID NO)	Amino Acid Sequence (SEQ ID NO)
kappa-casein truncated version 1 (paraOKC1-T): Optimized beta-lactoglobulin version 1 (OLG1)			
Optimized beta-casein (A2 variant) truncated version 5:Optimized beta-casein (A2 variant) truncated version 4:Optimized beta-casein (A2 variant) truncated version 3: Optimized beta-casein (A2 variant) truncated version 2: Optimized beta-lactoglobulin version 1 (OLG1)	OBC-T5:OBC-T4:OBC-T3:OBC-T2:OLG1	657	658
Optimized beta-casein (A2 variant) truncated version 5:Optimized beta-casein (A2 variant) truncated version 4:Optimized beta-casein (A2 variant) truncated version 3: Optimized beta-casein (A2 variant) truncated version 2	OBC-T5:OBC-T4:OBC-T3:OBC-T2	659	660
Optimized beta-casein (A2 variant) truncated version 5:FM:Optimized beta-casein (A2 variant) truncated version 4:FM:Optimized beta-casein (A2 variant) truncated version 3: FM:Optimized beta-casein (A2 variant) truncated version 2: FM:Optimized beta-lactoglobulin version 1 (OLG1)	OBC-T5:FM:OBC-T4:FM:OBC-T3:FM:OBC-T2:FM:OLG1	661	662
Optimized beta-casein (A2 variant) truncated version 5:FM:Optimized beta-casein (A2 variant) truncated version 4:FM:Optimized beta-casein (A2 variant) truncated version 3: FM:Optimized beta-casein (A2 variant) truncated version 2	OBC-T5:FM:OBC-T4:FM:OBC-T3:FM:OBC-T2	663	664
Optimized beta-casein (A2 variant) truncated version 4:FM:Optimized beta-casein (A2 variant) truncated version 3: FM:Optimized beta-casein (A2 variant) truncated version 2	OBC-T4:FM:OBC-T3:FM:OBC-T2	792	793
Optimized beta-casein (A2 variant) truncated version 4:FM:Optimized beta-casein (A2 variant) truncated version 3: FM:Optimized beta-casein (A2 variant) truncated version 2:FM: Optimized beta-lactoglobulin version 1 (OLG1)	OBC-T4:FM:OBC-T3:FM:OBC-T2:FM:OLG1	794	795
Optimized beta-casein (A2 variant) truncated version 4:Optimized beta-casein (A2 variant) truncated version 3: Optimized beta-casein (A2 variant) truncated version 2	OBC-T4:OBC-T3:OBC-T2	796	797

Description	Abbreviated Description	DNA Sequence (SEQ ID NO)	Amino Acid Sequence (SEQ ID NO)
Optimized beta-casein (A2 variant) truncated version 4: Optimized beta-casein (A2 variant) truncated version 3: Optimized beta-casein (A2 variant) truncated version 2: Optimized beta-lactoglobulin version 1 (OLG1)	OBC-T4:OBC-T3:OBC-T2:OLG1	798	799

[497] The expression constructs were transformed into soybean, as described in Example 2. Quantification of casein multimer expression was performed using Western Blot Densitometry. Table 20 shows expression levels of the casein proteins when expressed in the indicated multimer constructs, relative to the caseins expressed alone (i.e., not as part of a fusion protein).

Table 20: Expression levels of casein multimers relative to caseins expressed alone

Casein Multimer Fusion Protein	Fold increase in expression relative to κ -Casein alone	Fold increase in expression relative to B-Casein alone	Fold increase in expression relative to α S1-Casein alone
κ -Casein : κ -Casein	3.4	...	...
κ -Casein : β -Casein	17	2.5	...
κ -Casein : α S1-Casein	5	...	32
κ -Casein : GFP	16	...	--
α S1-Casein : GFP	--	...	77
κ -Casein : β -Lactoglobulin	68	..	--
β -Casein : β -Lactoglobulin	--	27	--

Casein Multimer Fusion Protein	Fold increase in expression relative to κ -Casein alone	Fold increase in expression relative to B-Casein alone	Fold increase in expression relative to α S1-Casein alone
α S1-Casein- β : Lactoglobulin	--	--	522
κ -Casein : α -Lactalbumin	10	--	--
β -casein : α -Lactalbumin	--	2.8	--
α S1-Casein : α -Lactalbumin	--	--	150
β -Casein : β -Casein : β -Casein	--	10.7	--
β -Casein : β -Casein : β -Casein : β -Casein	--	14.5	--

[498] As shown in Table 20, expression of the casein proteins as multimers led to significant increases in expression relative to the caseins expressed alone. Specifically, expression of kappa-casein as a casein homo-multimer led to a 3.4-fold increase in expression relative to expression of casein alone. Expression of kappa-casein as a multimer with beta-casein led to 17-fold and 2.5-fold increases in expression, respectively, relative to either protein expressed alone. Expression of kappa-casein as a multimer with alpha-S1-casein led to 5-fold and 32-fold increases in expression, respectively, relative to either protein expressed alone. Expression of kappa-casein fused to GFP led to a 16-fold increase in expression. Expression of kappa-casein fused to beta-lactoglobulin led to a 68-fold increase in expression, and expression of beta-casein fused to beta-lactoglobulin led to an 11.5-fold increase in expression. Expression of beta-casein or alpha-S-casein was also increased by fusion to alpha-lactalbumin (2.8-fold and 150-fold respectively).

[499] Expression of β -casein as a trimer or tetramer also led to significant increases in expression relative to β -casein expressed alone (18-fold and 18.5-fold, respectively).

[500] Without being bound by any theory, it is believed that fusing a first casein protein to a second protein partially or fully shields each of the proteins from degradation by host cell proteases and allows for accumulation of the casein in the cell.

Example 4: Kappa-Casein is Sensitive to Soybean Endogenous Proteolysis Activity

[501] To determine whether endogenous host cell proteases are responsible for degradation of casein proteins expressed alone, soybean total protein extracts were spiked with 100 ng of commercial kappa-casein, in the presence or absence of Halt[®] Protease Inhibitor Cocktail (Thermo Fisher Scientific[®]). All samples were incubated at 37°C for two hours. The samples were then subjected to analysis using a Western blot. The protein was detected using a primary antibody against kappa-casein.

[502] As shown in **FIG. 14A** and **FIG. 14B**, most of the kappa-casein added to the cellular extracts was degraded, and this degradation was prevented by the addition of protease inhibitors. This data confirms that kappa-casein is sensitive to soybean endogenous proteolysis activity. Inhibition of endogenous proteolysis activity may lead to increased casein accumulation in transformed cells.

Example 5: Food Compositions

[503] The transgenic plants expressing the recombinant fusion proteins described herein can produce milk proteins for the purpose of food industrial, non-food industrial, pharmaceutical, and commercial uses described in this disclosure. Illustrative methods for making a food composition are provided in **FIG. 13** and **FIG. 17**.

[504] A fusion protein comprising an unstructured milk protein (e.g. a casein such as para- κ -casein, κ -casein, β -casein, α S1-casein, or α S2-casein), and a structured mammalian protein (e.g. β -lactoglobulin) is expressed in a transgenic plant (e.g. a soybean plant). The fusion protein comprises a chymosin cleavage site between the milk protein (e.g. a casein such as para- κ -casein, κ -casein, β -casein, α S1-casein, or α S2-casein) and the β -lactoglobulin.

[505] The fusion protein is extracted from the plant. The fusion protein is then treated with chymosin, to separate the milk protein (e.g. a casein) from the β -lactoglobulin. The casein is isolated and/or purified and used to make a food composition (e.g., cheese).

Example 6: Determination of Physicochemical Parameters that Contribute to Casein Accumulation in Plants

[506] The purpose of the experiments described in this example was to determine the physicochemical parameters of proteins (i.e., fusion partners) that, when fused to a casein protein, are capable of enhancing accumulation thereof.

[507] Various proteins having distinct physicochemical properties were fused to kappa-casein. The physicochemical properties thereof are listed in Table 21. The fusion proteins were then expressed in soybean plants as described above. Protein expression levels of the fusion protein and relative increases thereof relative to casein alone (not expressed as a fusion) were measured.

[508] Results are summarized in Table 21. The term “KCN-fusion %TSP” refers to protein expression levels of the fusion protein, as a percent of total soluble protein. The term “% KCN only” refers to increases in kappa-casein expression relative to kappa casein expressed alone (not as a fusion). The %KCN only value was calculated by division the KCN-fusion %TSP value by 0.059 (i.e., the percent accumulation of kappa-casein by itself).

Table 21: Proteins fused to kappa casein and physicochemical parameters thereof

Full name	Uniprot Accession No.	MW in kDa	Percentage hydrophobic AA/Total AA (%)	Number of disulfide bonds/per 10 kDa	KCN-fusion %TSP	% KCN only
Kappa Casein	P02668	18.9	48.04	0.53	0.2	339
Beta Casein	P02668	23.5	53.11	0	1	1695
Alpha Casein	P18626	22.9	45.23	0	0.29	492
Beta Lactoglobulin	P02754	18.2	48.15	1.1	4	6780
Alpha Lactalbumin	P00711	14.1	36.59	2.2	0.34	1017

Green Fluorescent Protein	P42212	26.8	40.76	0	0.94	1593
Lysozyme	Q6B411	14.9	39.23	2.68	0.05	85
2S globulin	P19594	16.1	24.82	2.48	0.1	169
Oleosin A	P29530	23.5	51.11	0	0.1	169
Oleosin B	P29531	23.4	50.67	0	0.1	169
Kunitz-Trypsin inhibitor	Q39898	21	41.67	0.95	0.001	16.9
Bowman-Birk inhibitor	I1MQD2	9	25	3.33	0.05	85
Hydrophobin II	P79073	7.19	49.3	5.56	0.025	42

[509] An analysis of the data shown in Table 21 is provided in **FIG. 16A**, **FIG. 16B**, and **FIG. 16C**. This analysis suggests that there are several physicochemical properties of proteins that when fused to kappa-casein, may contribute to accumulation of the kappa-casein. The first is molecular weight. In general, a protein (fusion partner) with molecular weight of 15 kDa or higher tended to increase accumulation (**FIG. 16A**). The second is hydrophobicity. A protein (fusion partner) having greater than about 30% hydrophobic amino acids also tended to increase accumulation (**FIG. 16B**). The third is flexibility. A protein (fusion partner) with less than about 2.5 disulfide bonds per 10 kDa molecular weight also tended to increase accumulation (**FIG. 16B**). The disulfide bonds were predicted using a computer program. Notably, the number of cysteines in the protein, on its own, was not predictive of the protein's ability to contribute to accumulation of the kappa-casein.

[510] Notably, as evidenced by the data in Table 21 and **FIG. 16A-16C**, the fusion partner did not need to have all three of these characteristics in order to increase accumulation of kappa-casein. For example, increases in accumulation were observed in some cases where the fusion partner had only one, only two or all three of these characteristics.

Example 7: Fusion proteins comprising milk proteins and prolamin proteins

[511] To determine the impact of including a prolamin in a fusion protein on accumulation thereof in a seed, expression vectors for producing fusion proteins comprising a milk protein and a prolamin protein were created using the methods described in Example 1. Specifically,

expression vectors were created to express fusion proteins comprising: (i) canelin (gCan27) fused to β -casein, (ii) zein (γ -zein) fused to β -casein, and (iii) canelin (gCan27) fused to κ -casein.

[512] Illustrative fusion proteins used during this study are shown below in Table 22. Colons (:) are used to indicate junctions between various elements of the fusion protein. FM indicates the use of a linker comprising a chymosin cleavage site.

Table 22: Fusion Proteins Comprising a Prolamin

Description	Abbreviated Description	DNA Sequence (SEQ ID NO)	Amino Acid Sequence (SEQ ID NO)
27 kD gamma canelin (gcan27):FM:Optimized beta casein truncated version 2 (OBC-T2):FM	gCan27:FM:OBC-T2	802	803
Gamma zein (γ Zein):Optimized beta-casein truncated version 2 (OBC-T2)	yZein:OBC-T2	806	807

[513] The expression constructs were transformed into soybean, as described in Example 2. Western blots showing detection of beta-casein in transgenic seed extracts are provided in FIG. 21 and FIG. 22. Quantification of casein multimer expression was performed using Western Blot Densitometry. Table 23 shows expression levels of the beta-casein protein when expressed in the indicated fusion constructs, relative to the beta-casein expressed alone (*i.e.*, not as part of a fusion protein).

Table 23: Expression levels of beta-casein when fused to a prolamin relative to caseins expressed alone

Casein Multimer Fusion Protein	Fold increase in expression relative to κ -Casein alone	Fold increase in expression relative to B-Casein alone
gCan27 : κ -Casein	16	...
gCan27 : β -Casein	--	40
Zein : β -Casein	--	55

[514] As shown in Table 23, fusion of caseins to either casein or zein led to significant increases in expression relative to the caseins expressed alone. Specifically, expression of kappa-casein fused to gCan27 led to a 16-fold increase in expression relative to expression of kappa-casein alone. Expression of beta-casein fused to gCan27 led to a 40-fold increase in expression, relative to beta-casein expressed alone. Fusion of beta-casein to zein led to a 55-fold increase in expression, relative to beta-casein expressed alone. In each of these experiments, the casein protein accumulated in the seeds at a level well above 1% TSP.

[515] Without being bound by any theory, it is believed that fusing a casein protein to a prolamin (e.g., a casein or a zein) leads to the formation of a protein body in the seed. The casein is then sequestered in the protein body, which partially or fully shields the casein from degradation by host cell proteases, and allows for accumulation of the casein in the cell. An illustrative mechanism for protein body formation is found in **FIG. 20**.

Example 8: Phosphorylation Prevents Degradation of Caseins in a Plant Cell

[516] It was hypothesized that various post-translational modifications, such as phosphorylation, may have a “shielding” effect which prevents degradation of milk proteins, especially casein proteins, in a plant cell. Specifically, it was hypothesized that by adding one or more phosphates to a casein protein expressed in a plant cell, it may be possible to block and/or reduce the access of plant proteases to various cleavage sites on the protein. By reducing the ability of the plant proteases to degrade the milk protein, higher levels of protein accumulation may be possible.

[517] To test this hypothesis, the enzyme Fam20C was co-expressed with one or more caseins in a plant cell. Fam20C is a serine kinase and is responsible for the phosphorylation of caseins (Bauman, D. E., et al. "Major advances associated with the biosynthesis of milk." *Journal of dairy science* 89, no. 4 (2006): 1235-1243.)

[518] The expression construct used in this study is shown in **FIG. 24E**. The construct comprised (i) a first expression cassette comprising the GmSeed2 promoter (SEQ ID NO: 813), a sig2 signal peptide (SEQ ID NO: 814), a sequence encoding a fusion protein (GOI, see table below), and an AtHSP/AtUbi10 Terminator (SEQ ID NO: 815, 816), and (ii) a second expression cassette comprising the pvPhas promoter (SEQ ID NO: 817), an Arc5'UTR (SEQ ID NO: 818), a sig10

signal peptide (SEQ ID NO: 819), a sequence encoding the Fam20c kinase (SEQ ID NO: 821), and a 3 arc terminator (SEQ ID NO: 822). This expression construct was cloned into a binary *Agrobacterium* vector, as illustrated in FIG. 23. The vector was then transformed into soybean plants, and protein expression was measured in the seeds using a Western Blot. An anti-beta-casein antibody was used to detect fusion protein expression. Results are shown in Table 24.

Table 24: Expression levels of caseins when co-expressed with a kinase compared to caseins expressed alone

		Increased fold vs KCN	Increased fold vs BCN	Increased fold vs aS1
No Kinase	κ -Casein-B-Casein	17	2.5	na
Kinase	κ -Casein-B-Casein / Fam20C	254	38	na
	κ -Casein-B-Casein-aS1-Casein / Fam20C	185	25	1000

[519] Table 24 compares expression levels of the casein when expressed alone vs. as a fusion protein, with or without Fam20c co-expression. When expressed without the kinase, the kappa-casein:beta-casein fusion protein produced a 17-fold increase in kappa-casein expression relative to kappa-casein expressed alone, and a 2.5-fold increase in beta-casein expression relative to beta-casein expressed alone. When this fusion protein was co-expressed with Fam20C, the expression of kappa-casein was 254-fold greater than kappa-casein expressed alone, and 38-fold greater than beta-casein expressed alone. Notably, expression of a kappa-casein:beta-casein:alpha-S1-casein fusion with a kinase resulted in a 185-fold increase in kappa-casein relative to kappa-casein alone, 25-fold increase in expression of beta-casein relative to beta-casein alone, and 1000-fold increase in alpha-S1-casein relative to alpha-S1-casenin alone.

[520] Taken together, these data indicate that co-expression of a kinase with a fusion protein comprising one or more casein proteins in a plant cell leads to an increase in accumulation of the

casein protein in the cell. Without being bound by any theory, it is believed that the addition of one or more phosphates to the casein protein protects it from degradation by one or more plant proteases.

Example 9: Fusion to a Highly Glycosylated Peptide to Increase Accumulation of Caseins in a Plant Cell

[521] Certain genetic elements increase the secretion and stability of proteins in plant cells (Jia Li et al., Secretion of Active Recombinant Phytase from Soybean Cell-Suspension Cultures, 1997; Jianfeng Xu et al., High-Yields and Extended Serum Half-Life of Human Interferon $\alpha 2b$ Expressed in Tobacco Cells as Arabinogalactan-Protein Fusions, 2007). Many aspects of plant growth involve hydroxyproline (Hyp)-rich glycoproteins (GRGPs). Accordingly, it was hypothesized that fusion of a casein to a glycoprotein tag could be used to increase accumulation of the casein in a plant cell.

[522] In this experiment, a glycoprotein comprising 11 tandem SP repeats was identified from a native soybean protein (Glyma.02g204500), annotated as early nodulin-like protein 10. This tag, dubbed the (SP)11 tag, was codon optimized in IDT and fused to the N- or C-terminus of kappa-casein (See FIG. 25A-25C). The (SP)11-kappa-casein was then cloned into a binary *Agrobacterium* vector, and transformed into soy.

[523] Notably, the expression of (SP)11-kappa-casein in the seeds was increased 13-fold over expression of kappa-casein alone (i.e., not fused with a glycoprotein tag). This data indicates that fusion with a glycoprotein tag may be used to increase accumulation of caseins in a plant cell.

[524] In a similar experiment, the M domain of CD45 (receptor-type tyrosine-protein phosphatase C) will be fused to kappa-casein. The M domain is known to function as an ER-retention signal. Briefly, the M-domain from CD45 (Uniprot Accession No. P08575, amino acids Ala231 to Asp 290) will be codon optimized using the Glycine max codon usage bias, and fused to the N- or C-terminus of kappa-casein. In some constructs, a KDEL sequence may be added to the C-terminus of the M-domain or the GOI (see FIG. 25E-25F). It is expected that fusion of the M domain to the C-terminus will cause ER retention of the fusion protein, leading to increased accumulation thereof in the cell.

Example 10: Cheese composition made with beta-casein protein

[525] To test whether a cheese composition having acceptable organoleptic and physical properties could be made using only beta-casein protein (i.e., without any other caseins), isolated beta casein from bovine milk was the sole casein protein used in the recipe below. The beta casein was provided in the form of a powder, comprising 84% protein with >98% purity of beta casein.

100% Beta casein cheese composition	
Water	42.07%
Butter	31.25%
Beta casein powder (84% protein)	13.10 %
Modified potato starch	11.00%
Salt	1.70%
Sodium citrate	0.80%
Calcium chloride	0.08%

[526] To make the cheese composition, all ingredients were added to a rapid visco analyzer (RVA) tube. The mixture was heated to 40°C and mixed at 200 RPM for 2 minutes. The speed was increased to 500 RPM and mixed for an additional 3 minutes. The mixture was then allowed to rest for a minimum of 5 minutes before heating to 95°C and mixing at 960 RPM for 1 minute. Then, the speed and temperature were reduced to 500 RPM and 90°C for 1 minute. The temperature was reduced further to 85°C and the composition was mixed for one more minute at 500 RPM. The hot cheese composition was poured into cylindrical molds (3/4" diameter pipe, 1" in length with cap on bottom), covered with plastic wrap, and refrigerated for a minimum of 5 days. The target pH was 5.5 to 5.7, and was adjusted with lactic acid, citric acid, or sodium citrate. Meltability was analyzed as described below (see also the cheese labeled "D" in FIG. 28)

[527] A cheese composition was also made with 50% beta casein protein and reduced amounts of other casein proteins using the following recipe:

50% Beta casein cheese composition	
Water	43.7%
Butter	31.3%
Acid casein (95% protein dry basis)	10.4%
Modified potato starch	6.00%
Beta casein powder (84% protein)	2.80%
Trisodium citrate	1.70%

	Salt	1.70%
	Sodium aluminum phosphate, basic	1.70%
	Citric acid	0.70%

[528] To make the cheese composition, 60% of the water and all other ingredients were added to a RVA tube and heated to 50°C and mixed at 500 RPM for 5 minutes. The remaining water was added to the RVA tube and the mixture was heated to 95°C and mixed at 960 RPM for 1 minute, then reduced to 500 RPM and 90°C for 1 minute. The temperature was reduced to 85°C, and the composition was mixed at 500 RPM for one more minute. The hot cheese composition was poured into cylindrical molds (3/4" diameter pipe, 1" in length with cap on bottom), covered with plastic wrap, and refrigerated for 5 days. The target pH was 5.5 to 5.7, and was adjusted with lactic acid, citric acid, or sodium citrate. For cheese analysis, see samples 3 and 4 of Table 25.

Example 11: Cheese composition made with kappa casein protein

[529] To test whether a cheese composition having acceptable organoleptic and physical properties could be made using only kappa-casein protein (i.e., without any other casein proteins), isolated kappa casein from bovine was the sole casein protein used in the recipe below. The kappa-casein was provided in the form of a powder, which comprised 85% protein and greater than 70% purity of kappa casein.

100% Kappa casein cheese composition	
Water	45.0%
Butter	31.3%
Kappa casein powder	13.8%
Modified potato starch	6.0%
Salt	1.7%
Sodium citrate	0.6%
Citric acid	0.6%
Sodium aluminum phosphate (basic)	0.9%
Calcium chloride	0.1%

[530] All ingredients were added to a rapid visco analyzer (RVA) tube. The mixture was heated to 40°C and mixed at 200 RPM for 2 minutes. The speed was increased to 500 RPM, and the composition was mixed for an additional 3 minutes. The mixture was then allowed to rest for a minimum of 5 minutes before heating to 95°C, and mixing at 960 RPM for 1 minute. Then, the

speed and temperature were reduced to 500 RPM and 90°C for 1 minute. The temperature was reduced further to 85°C, and the composition was mixed for one more minute at 500 RPM. The hot cheese composition was poured into cylindrical molds (3/4" diameter pipe, 1" in length with cap on bottom), covered with plastic wrap, and refrigerated for 5 days. The target pH was 5.5 to 5.7 and was adjusted with lactic acid, citric acid, or sodium citrate. Meltability was analyzed as described below (see also the cheese labeled "B" in **FIG. 28**).

Example 12: Cheese composition made with alpha-casein protein

[531] To test whether a cheese composition having acceptable organoleptic and physical properties could be made using only alpha-casein proteins, isolated alpha casein (a mixture of alpha-S1 and alpha-S2 caseins) from bovine was the sole casein protein used in the recipe below. The alpha-casein was provided in the form of a powder, which comprised approximately 87% protein and greater than 90% purity of alpha casein.

100% alpha casein cheese composition	
Water	41.8%
Butter	31.3%
Alpha casein powder	12.6%
Modified potato starch	11.0%
Salt	1.7%
Sodium citrate	0.8%
Citric acid	0.2%
Sodium aluminum phosphate (basic)	0.5%
Calcium chloride	0.08%

[532] All ingredients were added to a rapid visco analyzer (RVA) tube. The mixture was heated to 40°C and mixed at 200 RPM for 2 minutes. The speed was increased to 500 RPM, and the composition was mixed for an additional 3 minutes. The mixture was then allowed to rest for a minimum of 5 minutes before heating to 95°C, and mixing at 960 RPM for 1 minute. Then, the speed and temperature were reduced to 500 RPM and 90°C for 1 minute. The temperature was reduced further to 85°C, and the composition was mixed for one more minute at 500 RPM. The hot cheese composition was poured into cylindrical molds (3/4" diameter pipe, 1" in length with cap on bottom), covered with plastic wrap, and refrigerated for 5 days. The target pH was 5.5 to

5.7 and was adjusted with lactic acid, citric acid, or sodium citrate. Meltability was analyzed as described below.

Example 13: Cheese composition made with alpha- and beta-casein proteins

[533] To test whether a cheese composition having acceptable organoleptic and physical properties could be made using alpha- and beta-casein, alpha-casein and beta-casein powder obtained from bovine were used to create cheese compositions comprising 50% alpha-casein and 50% beta-casein, 25% alpha-casein and 75% beta-casein, and 75% alpha-casein and 25% beta-casein, as shown in the recipes below.

	50% alpha- casein and 50% beta-casein	25% alpha- casein and 75% beta-casein	75% alpha- casein and 25% beta-casein
Water	42.0%	41.8%	42.0%
Butter	31.3%	31.3%	31.3%
Modified potato starch	10.5%	10.5%	10.5%
Alpha casein powder	6.3%	3.2%	9.5%
Beta casein powder (84% protein)	6.5%	9.8%	3.3%
Trisodium citrate	0.9%	0.9%	0.9%
Salt	1.7%	1.7%	1.7%
Sodium aluminum phosphate, basic	0.5%	0.5%	0.5%
Citric acid	0.2%	0.2%	0.2%
Calcium chloride	0.08%	0.08%	0.08%

[534] To make the cheese composition, 60% of the water and all other ingredients were added to a RVA tube and heated to 50°C. The composition was then mixed at 500 RPM for 5 minutes. The remaining water was added, and the mixture was heated to 95°C and mixed at 960 RPM for 1 minute. Then, the composition was mixed at 500 RPM at 90°C for 1 minute. The temperature was reduced to 85°C, and the composition was mixed at 500 RPM for one more minute. The hot cheese composition was poured into cylindrical molds (3/4" diameter pipe, 1" in length with cap on bottom), covered with plastic wrap, and refrigerated for 5 days. The target pH was 5.5 to 5.7, and was adjusted with lactic acid, citric acid, or sodium citrate. Meltability was analyzed as described below.

Example 14: Cheese composition made with beta- and kappa-casein proteins

[535] To test whether a cheese composition having acceptable organoleptic and physical properties could be made using beta- and kappa-casein, bovine kappa-casein and beta-casein powder were used to create two cheese compositions, one comprising 75% kappa casein and 25% beta casein, and another comprising 50% kappa casein and 50% beta casein, as shown in the recipes below.

	75% kappa casein and 25% beta casein	50% kappa casein and 50% beta casein
Water	41.9%	41.7%
Butter	31.3%	31.3%
Modified potato starch	10.5%	10.5%
Kappa casein powder	9.7%	6.5%
Beta casein powder (84% protein)	3.3%	6.6%
Trisodium citrate	0.8%	0.8%
Salt	1.7%	1.7%
Sodium aluminum phosphate, basic	0.5%	0.5%
Citric acid	0.3%	0.3%
Calcium chloride	0.04%	0.08

[536] To make the cheese composition, 60% of the water and all other ingredients were added to a RVA tube and heated to 50°C. The composition was then mixed at 500 RPM for 5 minutes. The remaining water was added, and the mixture was heated to 95°C and mixed at 960 RPM for 1 minute. Then, the composition was mixed at 500 RPM at 90°C for 1 minute. The temperature was reduced to 85°C, and the composition was mixed at 500 RPM for one more minute. The hot cheese composition was poured into cylindrical molds (3/4" diameter pipe, 1" in length with cap on bottom), covered with plastic wrap, and refrigerated for 5 days. The target pH was 5.5 to 5.7, and was adjusted with lactic acid, citric acid, or sodium citrate. Meltability was analyzed as described below (see also the cheeses labeled "A" and "C" in FIG. 28).

Example 15: Functional properties of cheese compositions made with beta- and kappa-caseins

[537] To test the organoleptic and physical properties, the cheese compositions were analyzed for various properties, including melt, stretch, firmness, and transparency. For the melting test, the cheeses were placed in a 450°F oven for 5 minutes. A score of 0 = no change from the initial

appearance; 1 = up to 25% coverage of pan; 2 = 25% to 50% pan coverage; 3 = 50% to 75% pan coverage; and 4 = greater than 75% pan coverage. Shown in **FIG. 28** are the results of the test with cheese compositions comprising (A) 75% kappa casein, 25% beta casein; (B) 100% kappa casein; (C) 50% kappa casein, 50% beta casein; and (D) 100% beta casein. Composition A had a melt score of 2; composition B was unchanged and therefore had a melt score of 0; composition C had a melt score of 3; and composition D exhibited the greatest meltability with a score of 4.

[538] Additional cheese composition samples comprising different ratios of caseins and total protein were analyzed for stretchability and meltability after aging for a minimum of 5 days (Tables 25-27). Cheese composition stretch was measured by an assay testing the ability to stretch (cm in length) without breaking, after heating a 100 gram mass of the emulsion to a temperature of about 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass. Firmness and transparency were also observed by sensory evaluation (data not shown).

[539] Shown in Table 5, below, are data collected from cheese compositions comprising between 11-11.5% protein, and in Table 26, data collected from cheese comprising between 13-13.5% protein. Compositions 1 and 2 comprise caseins at a ratio that is similar to the approximate percentages of caseins in bovine milk. Compositions 3, 4, 7, and 8 have levels of beta-casein higher than that found in milk, and compositions 5, 6, 8, 9, and 10 have levels of kappa-casein higher than that found in bovine milk.

Table 25: Stretch and melt of cheese compositions comprising between 11-11.5% protein

	Protein contribution (%)					
	Alpha-S1 + Alpha-S2	Beta	Kappa			
1	50	37.5	12.5	4	5	0
2	50	37.5	12.5	4	8.5	0.25
3	40	50	10	3	5	0
4	40	50	10	3	8.5	0.25
5	43	32	25	4	6.5	0
6	43	32	25	3	3	0.25
7	0	100	0	4	20	0.4
8	0	50	50	3	17.5	0.08
9	0	25	75	2	0	0.04
10	0	0	100	0	0	0.12

Table 26: Stretch and melt of cheese compositions comprising between 13-13.5% protein

	Protein contribution (%)			Melt	Stretch (cm)	CaCl ₂ (%) w/w)
	Alpha-S1 + Alpha-S2	Beta	Kappa			
1	50	37.5	12.5	3	5	0
2	50	37.5	12.5	2	5.5	0.25
3	34	57	9	2.5	5	0
4	34	57	9	3	5	0.25
5	39	32	29	3	4.5	0
6	39	32	29	2	4	0.25

[540] As evidenced by the data above, cheese compositions made with beta-casein exhibited good melting and stretch after aging for at least five days. Use of the tested amounts of beta-casein also softened the cheese composition compared to standard casein ratios. The beta-casein cheese composition was affected by calcium level similar to that of the control cheese composition, and it was also found to be highly soluble. The cheese composition was substantially transparent when melted.

[541] Kappa-casein imparted firmness to the cheese composition relative to standard casein ratios, but was more reactive to calcium than beta casein and control cheese compositions 1 and 2. Levels of less than 25% kappa casein did not impact stretch and stretch may improve slightly, whereas increasing levels of kappa casein restricted melt and reduced stretch after refrigeration for five days. Immediately after cooking, the 100% kappa casein cheese compositions can stretch to greater than 25 cm.

[542] The alpha-caseins, alpha-S1-casein and alpha-S2-casein, are assumed to provide cheese firmness, as cheeses with depleted levels of alpha-S1-casein and alpha-S2-casein are softer than the control cheeses.

[543] Cheese stretch was impacted by percent contribution of beta-casein. Specifically, increasing the amount of beta-casein correlated with an increase in stretch. As shown in Table 27 below, and **FIG. 29**, cheese compositions comprising, for example, 50% beta-casein had a stretch

of 8.5 cm, whereas cheese comprising 100% beta-casein had a stretch of 20 cm. All cheese compositions comprised 11.5% total protein and CaCl₂.

Table 27: Cheese stretch with increasing contribution of protein from beta-casein

% beta-casein	Stretch (cm)
37.5	8.5
50.0	8.5
66.7	7.5
83.3	11.5
100.0	20

Example 16: Functional properties of additional cheese compositions

[544] The analysis performed in Example 15 was repeated with additional compositions, as shown in Tables 28-30. The compositions comprised 100% beta-casein, 100% kappa-casein, or 100% alpha-caseins, or mixtures thereof. The actual percent of protein by weight of the compositions varied between 10-11.75%, and the CaCl₂ concentration ranged from 0% to 0.16% (by weight). Melt, stretch, and firmness were determined as described above.

Table 28: Functional properties of additional cheese compositions comprising only one casein

100% Beta-casein					
% protein	% CaCl ₂	melt	stretch	firmness	pH
11.75	0	2	1.5	very firm	5.23
10	0	3	15	soft	5.28
10	0	3	0	very soft	5.6
11	0.04	4	6.5	firm	5.63
11	0.08	4	20	firm	5.55
11	0.08	3	6.5	soft	5.59
100% Kappa-casein					
% protein	% CaCl ₂	melt	stretch	firmness	pH
11.75	0.2	2	0	firm	5.65

11.75	0.12	3	0	firm	5.44
11.75	0.06	2	1.5	firm	5.67
11	0.04	2	0	firm	5.67
11.75	0.12	1	0	firm	5.06
100% Alpha-casein					
% protein	% CaCl ₂	melt	stretch	firmness	pH
11*	0	3	0	brittle	6.01
11*	0.08	3	0	slightly soft	5.8
11*	0	3	0	firm	5.91
11*	0.08	3	0	firm	5.9
11	0	2	0	very firm	6.26

*indicates that the compositions were undercooked

Table 29: Functional properties of additional cheese compositions comprising mixtures of caseins

Beta-casein and Kappa-casein							
Sample No.	% beta	% kappa	% CaCl ₂	melt	stretch	firmness	pH
1	5.5	5.5	0.04	4	7	firm	5.5
2	2.75	8.25	0.04	2	0	very firm	5.64
3	5.5	5.5	0.08	4	17.5	firm	5.36
Alpha-caseins and beta-casein							
	% alpha	% beta	% CaCl ₂	melt	stretch	firmness	pH
1	5.5	5.5	0.16	4	0	firm	5.87
2	5.5	5.5	0.08	3	1.5	firm	6.01
3	5.5	5.5	0	3	1.5	firm	5.98
4	2.75	8.25	0.16	4	5	firm	5.94
5	2.75	8.25	0.08	3	1.5	firm	5.72
6	2.75	8.25	0	4	5.5	firm	5.68
7	8.25	2.75	0.16	3	0	firm	5.91
8	8.25	2.75	0.08	2	0	firm	5.92

9	8.25	2.75	0	2	0	firm	5.91
---	------	------	---	---	---	------	------

Table 30: Cheese melt and stretch with increasing contribution of protein from beta-casein

% Beta-casein	% Alpha-caseins	% Kappa-casein	Melt	Stretch (cm)
0	100	0	3	0
0	0	100	0	0
25	75	0	2	0
25	0	75	2	0
50	50	0	3	1.5
50	0	50	3	10.5
75	25	0	3	1.5
75	0	25	3	14.5
100	0	0	3	8.5

[545] As shown by the data in Table 28-30 above and **FIGS. 30-31**, beta-casein is the only casein that imparts both melt and stretch in a cheese composition under these conditions. Alpha caseins do not appear to impart stretch and do not significantly restrict melt. Alpha-caseins contribute to cheese firmness (data not shown). Kappa-casein also imparts firmness but negatively impacts melt. Like beta-casein it can contribute to stretch but only when combined with another protein.

Example 17: Cheese compositions with beta-lactoglobulin

[546] To determine the effect of adding beta-lactoglobulin on the functional and organoleptic properties of the compositions, various cheese compositions were generated with different amounts of beta-lactoglobulin.

Table 31: Stretch and melt of cheese composition with the addition of beta-lactoglobulin

Protein contribution (%)							
Alpha-S1 + Alpha-S2	Beta	Kappa	Beta Lactoglobulin	Soy	% Protein	Melt	Stretch (cm)
50	37.5	12.5	0	0	11.5	4	5
39.8	29.8	9.9	20.5	0	11.5	3	0
31.6	23.7	7.9	0	20.5	11.5	4	4.5
50	37.5	12.5	0	0	13.2	3	8.5
34.6	26	8.7	30.8	0	13.2	2	0
24.0	18	6	0	30.8	13.2	3	0

[547] As shown in Table 31 above, at 20% and 30% casein replacement levels, stretch was eliminated. Cheese composition melt, stretch, firmness were closer to the control with soy protein isolate compared to beta-lactoglobulin at both replacement levels tested.

[548] Similar results were achieved with compositions comprising kappa-casein and beta-lactoglobulin (Table 32), beta-casein and beta-lactoglobulin (Table 33). The addition of beta-lactoglobulin to the cheese compositions softened them, restricted melt, and imparted an opacity due to protein aggregation.

Table 32: Stretch and melt of kappa-casein cheese composition with the addition of beta-lactoglobulin

Kappa-casein + Beta-lactoglobulin						
% protein	% CaCl ₂	melt	stretch	firmness	pH	% BLG
10	0.06	2	0	firm	5.95	2.5

Table 33: Stretch and melt of beta-casein cheese composition with the addition of beta-lactoglobulin

Beta-casein + beta-lactoglobulin						
% protein	% CaCl ₂	melt	stretch	firmness	pH	% BLG
10	0	3	0	slightly soft (between soft and firm)	5.43	2.5

Example 18: Estimation of apparent viscosity

[549] An exemplary milk composition comprising beta-casein as the only casein (BC milk), a yogurt composition comprising beta-casein as the only casein (BC yogurt), and an ice cream mix composition comprising beta-casein as the only casein (BC IC mix) are described in Table 34, below.

Table 34: Food compositions for viscosity analysis

Beta casein milk	
Ingredient	%
Beta casein powder	3.0
Sodium citrate	0.2
Lactic acid (88%)	0.1
Sodium chloride	0.2
Calcium chloride	0.2
Palm oil	3.3
Soy lecithin	0.2
Glucose	4.0
Water	88.9
Beta casein yogurt (plain)	
Ingredient	%
Beta casein powder	4.0
Sodium citrate	0.2
Sodium chloride	0.2
Calcium chloride	0.2
Coconut oil	4.0
Soy lecithin	0.3
Modified tapioca starch	3.5
Glucose	4.0
Water	83.6
Beta casein ice cream mix	
Ingredient	%
Beta casein powder	4.5
Sodium citrate	0.2
Lactic acid (88%)	0.1
Sodium chloride	0.2
Tetrasodium pyrophosphate	0.2

Cocoa butter	12.0
Calcium sulfate	0.2
Mono & diglycerides	0.5
Cellulose gum	0.2
Sucrose	20.0
Vanilla extract	0.5
Water	61.5

[550] To make a BC milk composition, water is heated to 100-120°F, and lecithin and melted palm oil are added. Subsequently, the remaining ingredients (except lactic acid) are added with agitation and minimal air incorporation. The pH is adjusted to the range of 6.5-7.0 with lactic acid. The composition is then heated to 140°F and homogenized (two stage, 2000 psi total). Then, the composition is heated to 175°F and held at that temperature for 20 seconds before cooling to 40°F.

[551] To make a BC yogurt composition, water is heated to 100-120°F, and lecithin and melted coconut oil are added. Subsequently, the remaining ingredients (except lactic acid) are added with agitation and minimal air incorporation. The pH is adjusted to the range of 6.5-7.0 with lactic acid. The composition is then heated to 140°F and homogenized (two stage, 2000 psi total). Then, the composition is heated to 185°F and held at that temperature for 5 minutes. The composition is then cooled to 110°F and yogurt cultures are added. The composition is fermented at 108°F until the pH is 4.6. The composition is then stirred and cooled to 40 °F.

[552] To make a BC IC composition, water is heated to 100-120°F, and lecithin and melted cocoa butter are added. Subsequently, the remaining ingredients (except lactic acid) are added with agitation and minimal air incorporation. The pH is adjusted to the range of 6.5-7.0 with lactic acid. The composition is then heated to 140°F and homogenized (two stage, 2000 psi total). Then, the composition is heated to 175°F and held at that temperature for 20 seconds before cooling to 40°F.

[553] Theoretical approximations of the apparent viscosity for these BC milk, BC yogurt, and BC IC mix may be determined. Specifically, these approximations may be based on the rheological analysis of formulations made with bovine milk and adjustments made from observations made during the work with various cheese compositions described herein.

[554] Approximations are shown in **FIG. 32**. The BC milk compositions are estimated to have an apparent viscosity of about 2 cP over all shear rates analyzed. In contrast, the BC yogurt and

the BC IC mix compositions are estimated to have a higher apparent viscosity, which is expected to decrease at higher shear rates, characteristic of non-Newtonian compositions.

[555] Taken together, this data indicates that the BC yogurt and BC IC mix compositions are expected to be non-Newtonian compositions.

Example 19: Expression of a fusion protein comprising beta-casein in *E. Coli*

[556] To determine whether the fusion proteins of the disclosure may be detectably expressed in a bacterial system, a beta-casein tetramer (i.e., a fusion protein comprising four beta-caseins) was expressed in *E. Coli*. Specifically, the pET system (Novagen) was used for the cloning and expression of the proteins of interest in *E. coli*. A DNA sequence encoding the beta-casein tetramer was PCR amplified and cloned into the NcoI and BlnI sites of pET-28a (+) expression vector via In-Fusion (Takara) cloning. The ligated vector was transformed into StellarTM competent cells. Subsequently, the DNA of positive clones was used to transform BL21-CodonPlus strain (Agilent Technologies) which encodes a T7 RNA polymerase under control of the lacUV5 promoter for easy expression.

[557] To induce protein expression, an overnight culture grown to stationary phase was diluted (1/100) and then grown to mid-log phase (OD₆₀₀~0.4-0.6). The mid-log phase culture was pelleted, and the supernatant was removed. Protein expression was induced by the addition of IPTG (0.5 mM final concentration) to the pellet, and the cells were incubated for 3 hours at 37 °C with 160 rpm shaking. To extract the proteins of interest, the BugBuster® (Novagen) master mix was utilized following the manufacturer's instructions with the addition HALT protease inhibitor. The proteins were separated using SDS-PAGE and transferred to nitrocellulose membrane. The fusion protein was detected using a primary antibody raised against beta-casein.

[558] As shown in **FIG. 33**, the beta-casein tetramer (BC4) accumulated to detectable levels in *E.Coli*. Lanes 1-5 of **FIG. 33** show wildtype *E. coli* extracts with commercial beta-casein protein spiked in at 0, 5, 10, 20 and 40 ng per lane, as a standard. Lane 6 shows molecular weight markers. Lane 7 shows Beta-casein tetramer expressed in *E.Coli* after 3h of induction with IPTG.

Example 20: Expression of a fusion protein comprising beta-casein in tobacco leaves

[559] To determine whether the fusion proteins of the disclosure may be detectably expressed in a tobacco system, a fusion protein comprising beta-casein fused to beta-lactoglobulin was expressed in tobacco leaves.

[560] A DNA sequence encoding a fusion protein comprising, from N-terminus to C-terminus, beta-lactoglobulin and beta-casein, was inserted into the AR17 vector backbone in between a double 35S promoter and EUT:Rb7T double terminator. The plasmid was transformed into agrobacterium strain AGL1 and the positive agrobacterium colonies were cultured overnight in selective media. To prepare the infiltration solution, the agrobacterium culture was precipitated by centrifugation at 1,000 g for 10 mins and resuspended in equal volume of the infiltration medium (50 mM MES, 2 mM Na₃PO₄, 5 mg/mL D-glucose and 0.1 mM acetosyringone). This washing step was repeated a second time.

[561] The fusion protein expressing strain was co-infiltrated in tobacco leaves with the post-translational gene silencing inhibitor p19 strain and the protease inhibitor NbPR4 strain to enhance the fusion protein expression. Concentration of the fusion protein expressing-strain, and strains p19 and NbPR4, was adjusted to an optical density (OD) of 1, 0.5 and 0.5 respectively, immediately before co-infiltration into the leaves. Six to eight-week-old *Nicotiana benthamiana* plants were used for infiltrations. Four different fully expanded leaves were infiltrated as biological replicates.

[562] Protein samples were harvested three days after infiltration. Total soluble proteins were extracted with equal volume of extraction buffer (1X PBS PH7.4, 5 mM DTT, 0.1% Tween20 and 1X HALT protease inhibitors). Total protein concentrations were measured using Pierce 660 reagent. To visualize the target protein expression, 1 ug of total soluble protein were separated on SDS-PAGE, transferred to nitrocellulose membrane, and probed with a beta-casein primary antibody.

[563] Results are shown in **FIG. 34**. Lanes 1-5 show wild type tobacco protein extracts spiked with 0, 0.5, 1, 2, or 4 nanograms of commercially available beta-casein. Lane 6 shows molecular weight markers. Lanes 7-8 shows infiltration of tobacco leaves with the fusion protein. This data shows that the fusion protein accumulated in tobacco leaves at a level above 4% total soluble protein.

MILK PROTEIN SEQUENCES

[564] The following Table 35 describes various representative species of milk proteins exemplified in the disclosure.

Table 35: Milk Protein Sequences of the Disclosure

SEQ ID NO	Description	Genus/ species	Accession Number
Kappa casein sequences			
3	Optimized kappa-casein truncated version 1 (OKC1-T)	Artificial (codon optimized <i>Bos taurus</i>)	
4	Optimized kappa-casein truncated version 1 (OKC1-T)	<i>Bos taurus</i>	
85	Kappa casein	<i>Capra hircus</i>	
86	Kappa casein	<i>Ovis aries</i>	
87	Kappa casein	<i>Bubalus bubalis</i>	
88	Kappa casein	<i>Camelus dromedaries</i>	
89	Kappa casein	<i>Camelus bactrianus</i>	
90	Kappa casein	<i>Bos mutus</i>	
91	Kappa casein	<i>Equus caballus</i>	
92	Kappa casein	<i>Equus asinus</i>	
93	Kappa casein	<i>Rangifer tarandus</i>	
94	Kappa casein	<i>Alces alces</i>	
95	Kappa casein	<i>Vicugna pacos</i>	
96	Kappa casein	<i>Bos indicus</i>	
97	Kappa casein	<i>Lama glama</i>	
98	Kappa casein	<i>Homo sapiens</i>	
148	Kappa casein	<i>Bos taurus</i>	NP_776719.1
149			AAI02121.1
150			AAA30433.1
151			AAB26704.1
152			1406275A
153			AAF72097.1
154			AAD32139.1
155			XP_024848756.1
156			CAF03625.1
157			ABN42697.1
158			AAD32140.1

159			ALC76014.1
160			DAA28589.1
161			ADT82665.1
162			ADT82666.1
163			CAH56573.1
164			ADT82669.1
165	Kappa casein	<i>Capra hircus</i>	QIZ03342.1
166			AYN74373.1
167			AAM12026.1
168			AFZ92921.1
169			NP_001272516.1
170			AAM12027.1
171			AAR06605.1
172			AAL90873.1
173			AFZ92919.1
174			QIZ03345.1
175			AAR91623.1
176			AAK17010.1
177			AAL93193.1
178			AFZ92918.1
179			AAL90872.1
180			AFZ92917.1
181			AAO39432.1
182			AAL90871.1
183			AAO39431.1
184	Kappa casein	<i>Ovis aries</i>	NP_001009378.1
185			AAP69943.1
186	Kappa casein	<i>Bubalus bubalis</i>	NP_001277901.1
187			AXE74388.1
188			APQ30586.1
189			AXE74385.1
190			XP_006071184.1
191			AXE74386.1
192	Kappa casein	<i>Bos mutus</i>	XP_005897104.1
193			XP_014334109.1
194			MXQ92034.1
195	Kappa casein	<i>Bos indicus</i>	XP_019818432.1
196			ACF15188.1
197			ACF15186.1
198			ACF15190.1
199			ABY81250.1
200			ABY81251.1
201			ADT82668.1
202			ADT82663.1
203			ADT82671.1

204			ADT82670.1
205			AAQ73171.1
206	Kappa casein	<i>Jeotgalicoccus coquinae</i>	WP_188357548.1
207	(Hypothetical Protein)		WP_188357549.1
208	Kappa casein isoform X1	<i>Bison bison bison</i>	XP_010837415.1
209			XP_010837416.1
210	Kappa casein	<i>Bos grunniens</i>	AFM93768.1
211			AXE74296.1
212			AAM25910.1
213			ABU53615.1
214			AAM25909.1
215			AAF63191.1
216	Kappa casein	<i>Bos indicus x Bos taurus</i>	AAF72096.1
217			AAF72098.1
218	Kappa casein (precursor)	<i>Oreamnos americanus</i>	P50423.1
219	Kappa casein (precursor)	<i>Naemorhedus goral</i>	P50422.1
220	Kappa casein	<i>Odocoileus virginianus texanus</i>	XP_020729185.1
221	Kappa casein (precursor)	<i>Capricornis sumatraensis</i>	P50420.1
222	Kappa casein (precursor)	<i>Capricornis crispus</i>	BAA03287.1
223			P42156.1
224	Kappa casein (precursor)	<i>Capricornis swinhoei</i>	P50421.1
225	Kappa casein (precursor)	<i>Saiga tatarica</i>	P50425.1
226	Kappa casein (precursor)	<i>Rupicapra rupicapra</i>	P50424.1
227	Kappa casein (precursor)	<i>Cervus nippon</i>	P42157.1
228	Kappa casein	<i>Bos frontalis</i>	ADF58295.1
229	Kappa casein (hypothetical protein FD755_023011)	<i>Muntiacus reevesi</i>	KAB0354473.1
230	Kappa casein (hypothetical protein FD754_018150)	<i>Muntiacus muntjak</i>	KAB0341224.1
231	Kappa casein	<i>Madoqua saltiana</i>	AFY03578.1
232	Kappa casein	<i>Gazella dorcas</i>	AFY03574.1
233	Kappa casein	<i>Gazella arabica</i>	AFY03576.1
234	Kappa casein	<i>Capra ibex ibex</i>	AAP80529.1
235	Kappa casein	<i>Ovis ammon severtzovi</i>	ADB66396.1
236	Kappa casein	<i>Ovis orientalis gmelini</i>	ADB66423.1
237			ADB66420.1
238	Kappa casein (hypothetical protein G4228_004474)	<i>Cervus hanglu yarkandensis</i>	KAF4013038.1
239	Kappa casein	<i>Procapra gutturosa</i>	AFY03581.1
240			AFY03580.1

1	Optimized para -kappa-casein truncated version 1 (paraOKC1-T)	Artificial (codon optimized <i>Bos taurus</i>)	
2	Optimized para -kappa-casein truncated version 1 (paraOKC1-T)	<i>Bos taurus</i>	
241 242 243 244 245 246 247 248 249 250 251 252 253	Kappa casein isoform X1	<i>Bos taurus</i>	AAA30433.1 1406275A AAI02121.1 NP_776719.1 DAA28589.1 AAB26704.1 XP_024848756.1 ABN42697.1 AAF72097.1 721588A AAD32139.1 AAD32140.1 CAF03625.1
254 255	Kappa casein (hypothetical protein)	<i>Jeotgalicoccus coquinae</i>	WP_188357548.1 WP_188357549.1
256 257 258	Kappa casein isoform X1	<i>Bos mutus</i>	XP_005897104.1 XP_014334109.1 MXQ92034.1
259 260 261 262 263 264 265	Kappa casein	<i>Bos indicus</i>	XP_019818432.1 ACF15188.1 ABY81250.1 ABY81251.1 ACF15186.1 ACF15190.1 ADT82668.1
266 267 268 269 270	Kappa casein	<i>Bos grunniens</i>	AXE74296.1 AFM93768.1 AAM25910.1 AAM25909.1 ABU53615.1
271 272	Kappa casein isoform X1	<i>Bison bison bison</i>	XP_010837415.1 XP_010837416.1
273 274 275 276 277 278	Kappa casein (precursor)	<i>Bubalus bubalis</i>	NP_001277901.1 XP_006071184.1 AXE74388.1 AXE74385.1 APQ30586.1 AXE74386.1

279	Kappa casein (precursor)	<i>Oreamnos americanus</i>	P50423.1
280	Kappa casein (precursor)	<i>Capricornis swinhoei</i>	P50421.1
281	Kappa casein (precursor)	<i>Naemorhedus goral</i>	P50422.1
282	Kappa casein (precursor)	<i>Capricornis sumatraensis</i>	P50420.1
283	Kappa casein (precursor)	<i>Capricornis crispus</i>	BAA03287.1
284			P42156.1
285	Kappa casein (precursor)	<i>Saiga tatarica</i>	P50425.1
286	Kappa casein	<i>Bos indicus x Bos taurus</i>	AAF72096.1
287			AAF72098.1
288	Kappa casein (precursor)	<i>Capra hircus</i>	NP_001272516.1
289			AYN74373.1
290			QIZ03345.1
291			QIZ03342.1
292			AFZ92921.1
293			AAR06605.1
294			AAM12026.1
295			AAL93193.1
296			AAR91623.1
297			AFZ92917.1
298			AAM12027.1
299			AAL90873.1
300			AFZ92918.1
301			AAL90871.1
302			AAL90872.1
303			AAL31535.1
304			AAL31534.1
305			ABK59545.1
306			AAO39432.1
307			AFZ92919.1
308			AAK17010.1
309			AAO39431.1
310			AAP80475.1
311	Kappa casein	<i>Odocoileus virginianus texanus</i>	XP_020729185.1
312	Kappa casein (precursor)	<i>Rupicapra rupicapra</i>	P50424.1
313	Kappa casein (precursor)	<i>Ovis aries</i>	NP_001009378.1
314			AAP69943.1
315	Kappa casein (precursor)	<i>Cervus nippon</i>	P42157.1
316	Kappa casein	<i>Gazella arabica</i>	AFY03576.1
317	Kappa casein (hypothetical protein FD754_018150)	<i>Muntiacus muntjak</i>	KAB0341224.1

318	Kappa casein (hypothetical protein FD755_023011)	<i>Muntiacus reevesi</i>	KAB0354473.1
319	Kappa casein	<i>Gazella dorcas</i>	AFY03575.1
320	Kappa casein	<i>Procapra gutturosa</i>	AFY03581.1
321			AFY03580.1
322	Kappa casein	<i>Madoqua saltiana</i>	AFY03578.1
323	Kappa casein	<i>Ammotragus lervia</i>	QIN85723.1
324			QIN85720.1
325			QIN85721.1
326	Kappa casein	<i>Capra sibirica</i>	AAP80568.1
327	Kappa casein	<i>Ovis canadensis canadensis</i>	ADB66397.1
328			ADB66402.1
329	Kappa casein	<i>Gazella subgutturosa marica</i>	AFY03577.1
330	Kappa casein	<i>Antelope cervicapra</i>	AFY03573.1
331	Kappa casein	<i>Capra ibex ibex</i>	AAP80529.1
332	Kappa casein	<i>Ovis vignei arkal</i>	ADB66436.1
333			ADB66442.1
334	Kappa casein	<i>Ovis ammon collium</i>	ADB66395.1
335	Kappa casein	<i>Ovis vignei blanfordi</i>	ADB66445.1
336	Kappa casein	<i>Ovis orientalis gmelini</i>	ADB66423.1
337			ADB66420.1
338	Kappa casein	<i>Ovis orientalis x vignei</i>	ADB66465.1
339	Kappa casein	<i>Ovis vignei vignei</i>	ADB66456.1
340	Kappa casein	<i>Ovis ammon severtzovi</i>	ADB66396.1
Alpha S1 casein sequences			
7	Optimized alpha S1- casein truncated version 1(OaS1-T)	Artificial (codon optimized <i>Bos taurus</i>)	
8	Optimized alpha S1- casein truncated version 1(OaS1-T)	<i>Bos taurus</i>	
99	Alpha S1 casein	<i>Capra hircus</i>	
100	Alpha S1 casein	<i>Ovis aries</i>	
101	Alpha S1 casein	<i>Bubalus bubalis</i>	
102	Alpha S1 casein	<i>Camelus dromedaries</i>	
103	Alpha S1 casein	<i>Camelus bactrianus</i>	
104	Alpha S1 casein	<i>Bos mutus</i>	
105	Alpha S1 casein	<i>Equus caballus</i>	
106	Alpha S1 casein	<i>Equus asinus</i>	

107	Alpha S1 casein	<i>Bos indicus</i>	
108	Alpha S1 casein	<i>Lama glama</i>	
109	Alpha S1 casein	<i>Homo sapiens</i>	
341	Alpha S1 casein	<i>Bos taurus</i>	ABW98943.1
342			XP_024848771.1
343			ABW98940.1
344			ACG63494.1
345			XP_015327132.1
346			XP_024848772.1
347			1308122A
348			ABW98949.1
349			AAA30429.1
350			XP_015327135.1
351			XP_015327134.1
352			XP_024848773.1
353			XP_015327133.1
354			XP_024848774.1
355			XP_015327136.1
356			XP_024848775.1
357			XP_005208084.1
358			XP_024848776.1
359			XP_015327137.1
360			XP_015327138.1
361			XP_024848777.1
362			XP_024848778.1
363			XP_015327139.1
364			ABW98944.1
365			XP_015327140.1
366			XP_024848779.1
367			XP_015327141.1
368			XP_024848780.1
369			XP_015327142.1
370			ABW98945.1
371			XP_024848782.1
372			ABW98951.1
373			XP_024848784.1
374			XP_024848783.1
375			ABW98950.1
376			ABW98941.1
377			XP_005208086.1
378			ABW98942.1
379			ABW98937.1
380			ABW98952.1
381			ABW98954.1

382			ABW98953.1
383			ABW98955.1
384			ABW98957.1
385	Alpha S1 casein	<i>Capra hircus</i>	XP_017904616.1
386			QIZ03312.1
387			ALJ30147.1
388			P18626.2
389			XP_017904617.1
390			AFN44013.1
391			QIZ03319.1
392			CAA51022.1
393			NP_001272624.1
394			ALJ30148.1
395			QIZ03317.1
396			QIZ03310.1
397			QIZ03318.1
398			XP_017904618.1
399			XP_017904620.1
400			XP_017904619.1
401			XP_017904621.1
402			XP_017904622.1
403	Alpha S1 casein	<i>Ovis aries</i>	XP_012034747.1
404			P04653.3
405			AAB34797.1
406			ACJ46472.1
407			XP_027826521.1
408			XP_027826520.1
409			ACR58469.1
410			ACJ46473.1
411			AAB34798.1
412			NP_001009795.1
413	Alpha S1 casein	<i>Bubalus bubalis</i>	AAZ14098.1
414			APQ30583.1
415			O62823.2
416			XP_006071187.1
417			QCP57314.1
418			XP_025145744.1
419			QPO15022.1
420			XP_025145745.1
421			ACJ14317.1
422			XP_006071188.1
423			XP_025145747.1
424			XP_025145746.1
425			XP_025145748.1
426			XP_025145749.1

427			XP_025145750.1
428			XP_025145751.1
429			XP_025145752.1
430			XP_025145753.1
431	Alpha S1 casein	<i>Bos mutus</i>	XP_005902100.1
432	Alpha S1 casein	<i>Bos indicus</i>	XP_019818428.1
433	Alpha S1 casein	<i>Jeotgalicoccus coquinae</i>	WP_188357546.1
434	(hypothetical protein)		GGE26809.1
435	Alpha S1 casein	<i>Bison bison bison</i>	XP_010850445.1
436	Alpha S1 casein	<i>Bos grunniens</i>	AXE74293.1
437	Alpha S1 casein	<i>Jeotgalicoccus aerolatus</i>	WP_188349304.1
438	(hypothetical protein)		WP_188352531.1
439	Alpha S1 casein (hypothetical protein FD754_018154)	<i>Muntiacus muntjak</i>	KAB0341228.1
440	Alpha S1 casein (hypothetical protein FD755_023008)	<i>Muntiacus reevesi</i>	KAB0354470.1
Alpha S2 casein sequences			
83	Optimized alpha S2- casein truncated version 1(OaS2-T)	Artificial (codon optimized <i>Bos taurus</i>)	
84	Optimized alpha S2- casein truncated version 1(OaS2-T)	<i>Bos taurus</i>	
110	Alpha S2 casein	<i>Capra hircus</i>	
111	Alpha S2 casein	<i>Ovis aries</i>	
112	Alpha S2 casein	<i>Bubalus bubalis</i>	
113	Alpha S2 casein	<i>Camelus dromedaries</i>	
114	Alpha S2 casein	<i>Camelus bactrianus</i>	
115	Alpha S2 casein	<i>Bos mutus</i>	
116	Alpha S2 casein	<i>Equus caballus</i>	
117	Alpha S2 casein	<i>Equus asinus</i>	
118	Alpha S2 casein	<i>Vicugna pacos</i>	
119	Alpha S2 casein	<i>Bos indicus</i>	
120	Alpha S2 casein	<i>Lama glama</i>	
441	Alpha S2 casein	<i>Bos taurus</i>	AAI14774.1
442			XP_024848786.1
443			XP_015327143.1
444	Alpha S2 casein	<i>Capra hircus</i>	QIS93310.1
445			NP_001272514.1
446			CAB94236.1
447			QIS93322.1

448			AAB32166.1
449			QIS93306.1
450			XP_013820127.2
451			QIS93323.1
452			QIZ03322.1
453			QIS93316.1
454			CAB59920.1
455			CAC21704.2
456			QIS93307.1
457			XP_013820130.2
458			QIS93319.1
459			QIS93321.1
460			XP_013820128.2
461			QIS93304.1
462			XP_013820129.2
463			QIS93305.1
464			QIS93314.1
465			QIS93317.1
466			XP_013820132.2
467			XP_013820131.2
468	Alpha S2 casein	<i>Ovis aries</i>	ADB65931.1
469			NP_001009363.1
470			ADB65933.1
471			ADB65935.1
472			ADB65934.1
473			ADB65932.1
474	Alpha S2 casein	<i>Bubalus bubalis</i>	NP_001277794.1
475			AAZ80050.1
476			CAA06534.2
477			AFB69498.1
478			XP_006071185.2
479			AAZ57423.1
480			APQ30584.1
481			XP_025145302.1
482			XP_025145301.1
483	Alpha S2 casein	<i>Bos mutus</i>	XP_014335716.1
484			ELR51813.1
485	Alpha S2 casein	<i>Jeotgalicoccus aerolatus</i>	WP_188352530.1
486	(hypothetical protein)		GGE08804.1
487	Alpha S2 casein	<i>Jeotgalicoccus coquinae</i>	WP_188357545.1
	(hypothetical protein)		
488	Alpha S2 casein	<i>Bos grunniens</i>	AXE74294.1
489	Alpha S2 casein	<i>Bison bison bison</i>	XP_010850447.1
490	Alpha S2 casein	<i>Bos indicus x Bos taurus</i>	XP_027401112.1

491	Alpha S2 casein	<i>Odocoileus virginianus texanus</i>	XP_020729187.1
492	Alpha S2 casein (hypothetical protein FD754_018155)	<i>Muntiacus muntjak</i>	KAB0341229.1
493	Alpha S2 casein (hypothetical protein FD755_022792)	<i>Muntiacus reevesi</i>	KAB0354254.1
494	Alpha S2 casein (CSN1S2)	<i>Cervus elaphus hippelaphus</i>	OWK13818.1
Beta-casein sequences			
5	Optimized beta-casein truncated version 2 (OBC-T2)	Artificial (codon optimized <i>Bos taurus</i>)	
6	Optimized beta-casein truncated version 2 (OBC-T2)	<i>Bos taurus</i>	
121	Beta casein	<i>Capra hircus</i>	
122	Beta casein	<i>Ovis aries</i>	
123	Beta casein	<i>Bubalus bubalis</i>	
124	Beta casein	<i>Camelus dromedaries</i>	
125	Beta casein	<i>Camelus bactrianus</i>	
126	Beta casein	<i>Bos mutus</i>	
127	Beta casein	<i>Equus caballus</i>	
128	Beta casein	<i>Equus asinus</i>	
129	Beta casein	<i>Alces alces</i>	
130	Beta casein	<i>Vicugna pacos</i>	
131	Beta casein	<i>Bos indicus</i>	
132	Beta casein	<i>Lama glama</i>	
133	Beta casein	<i>Homo sapiens</i>	
495	Beta casein	<i>Bos taurus</i>	AAB29137.1
496			AAA30431.1
497			1314242A
498			AGT56763.1
499			AAI11173.1
500			XP_010804480.2
501			AAA30430.1
502			XP_015327157.2
503			ABR10906.1
504			ABL74247.1
505			QCI03091.1
506			QCI03090.1
507			CAC37028.1
508	Beta casein	<i>Capra hircus</i>	P33048.1

509			QIZ03333.1
510			CAB39200.1
511			AAK97639.1
512			XP_005681778.2
513			QLI42602.1
514			XP_013820153.1
515			QLI42606.1
516			QHN12643.1
517			ABQ52487.1
518			QHN12642.1
519			CAB39313.1
520			QHN12644.1
521			AWN06750.1
522	Beta casein	<i>Ovis aries</i>	P11839.3
523			NP_001009373.1
524	Beta casein	<i>Bubalus bubalis</i>	QHB80269.1
525			APQ30585.1
526			QHB80272.1
527			QHB80273.1
528			NP_001277808.1
529			Q9TSI0.1
530			XP_006071186.1
531			CAA06535.1
532			1004269A
533			ADD31643.1
534			ADD31644.1
535			AAT09469.1
536			ABL10285.1
537			ABA41625.1
538			ABA41623.1
539	Beta casein	<i>Bos mutus</i>	MXQ92033.1
540			XP_014335713.1
541			XP_005902099.2
542			XP_014335715.1
543			XP_014335714.1
544	Beta casein	<i>Bos indicus</i>	AQY78354.1
545			AQY78355.1
546			ABL75279.1
547			ABY27644.1
548			AWN06759.1
549			AGZ84117.1
550	Beta casein	<i>Bison bison bison</i>	XP_010850446.1
551	Beta casein (hypothetical protein)	<i>Jeotgalicoccus aerolatus</i>	WP_188352529.1

552	Beta casein (hypothetical protein)	<i>Jeotgalicoccus coquinae</i>	WP_188357544.1
553	Beta casein (precursor)	<i>Bos indicus x Bos taurus</i>	ARU83745.1
554			AWN06757.1
555			AWN06758.1
556	Beta casein	<i>Bos grunniens</i>	AXE74295.1
557			AEY63644.1
558			AEY63645.1
559			AEC13563.1
560	Beta casein	<i>Neophocaena asiaeorientalis asiaeorientalis</i>	XP_024597374.1
561	Beta casein	<i>Odocoileus virginianus texanus</i>	XP_020729180.1
562	Beta casein (hypothetical protein FD755_022863)	<i>Muntiacus reevesi</i>	KAB0354325.1
563	Beta casein (hypothetical protein FD754_022431)	<i>Muntiacus muntjak</i>	KAB0345505.1
Beta-Lactoglobulin sequences			
9	Optimized Beta Lactoglobulin 1 (OLG1)	Artificial (codon optimized <i>Bos taurus</i>)	
10	Optimized Beta Lactoglobulin 1 (OLG1)	<i>Bos taurus</i>	
11	Optimized Beta Lactoglobulin 2 (OLG2)	Artificial (codon optimized <i>Bos taurus</i>)	
12	Optimized Beta Lactoglobulin 3 (OLG3)	Artificial (codon optimized <i>Bos taurus</i>)	
13	Optimized Beta Lactoglobulin 4 (OLG4)	Artificial (codon optimized <i>Bos taurus</i>)	
564	Beta Lactoglobulin	<i>Bos taurus</i>	5K06_A
565			1B0O_A
566			NP_776354.2
567			3PH5_A
568			1BEB_A
569			6QPD_A
570			6QI7_A
571			DAA24277.1
572			5HTD_A
573			6QPE_A
574			6RWR_A
575			1BSO_A
576			6RWQ_A
577			ACG59280.1
578			5NUJ_A
579			5NUM_A

580			1UZ2_X
581			CAA32835.1
582			1CJ5_A
583			5NUK_A
584			5NUN_A
585			732164A
586			XP_024854027.1
587			AAA30411.1
588	Beta Lactoglobulin	<i>Capra hircus</i>	4OMW_A
589			NP_001272468.1
590			ABQ51182.1
591	Beta Lactoglobulin	<i>Ovis aries</i>	4NLI_A
592			NP_001009366.1
593			4CK4_A
594			4CK4_B
595	Beta Lactoglobulin	<i>Bubalus bubalis</i>	0601265A
596			P02755.2
597			NP_001277893.1
598			QOQ34530.1
599			APQ30587.1
600			ABG78270.1
601	Beta Lactoglobulin	<i>Bos mutus</i>	XP_005888577.1
602			MXQ94840.1
603	Beta Lactoglobulin	<i>Bos indicus</i>	XP_019826641.1
604	Beta Lactoglobulin (lipocalin/fatty-acid binding family protein)	<i>Jeotgalicoccus coquinae</i>	WP_188357550.1
605	Beta Lactoglobulin (lipocalin/fatty-acid binding family protei	<i>Jeotgalicoccus schoeneichii</i>	WP_188349305.1
606	Beta Lactoglobulin	<i>Bison bison bison</i>	XP_010855058.1
607	Beta Lactoglobulin	<i>Ovis sp.</i>	AAA31510.1
608	Beta Lactoglobulin	<i>Ovis aries musimon</i>	P67975.1
609	Beta Lactoglobulin	<i>Odocoileus virginianus texanus</i>	XP_020744123.1
610	Beta Lactoglobulin, Chain A	<i>Rangifer tarandus</i>	1YUP_A
611	Beta Lactoglobulin	<i>Rangifer tarandus tarandus</i>	AAZ57420.1
612	Beta Lactoglobulin (hypothetical protein FD754_009020)	<i>Muntiacus muntjak</i>	KAB0364864.1
613	Beta Lactoglobulin (hypothetical protein FD755_007442)	<i>Muntiacus reevesi</i>	KAB0379658.1

614	Beta Lactoglobulin, Chain A	<i>Equus caballus</i>	3KZA_A
-----	--------------------------------	-----------------------	--------

NUMBERED EMBODIMENTS

[565] Notwithstanding the appended claims, the following numbered embodiments also form part of the instant disclosure.

Embodiment Set 1: Stably transformed plant expressing a fusion protein comprising bovine kappa-casein and bovine beta-lactoglobulin

[566] 1. A stably transformed plant, comprising in its genome: a recombinant DNA construct encoding a fusion protein, the fusion protein comprising: a) bovine kappa-casein; and b) bovine beta-lactoglobulin, wherein the fusion protein is stably expressed in the plant.

[567] 2. The stably transformed plant of embodiment 1, wherein the fusion protein comprises, in order from N-terminus to C-terminus, the kappa-casein and the beta-lactoglobulin.

[568] 3. The stably transformed plant of embodiment 1, wherein the fusion protein comprises a protease cleavage site.

[569] 4. The stably transformed plant of embodiment 3, wherein the protease cleavage site is a chymosin cleavage site.

[570] 5. The stably transformed plant of embodiment 1, wherein the fusion protein comprises a signal peptide.

[571] 6. The stably transformed plant of embodiment 5, wherein the signal peptide is located at the N-terminus of the fusion protein.

[572] 7. The stably transformed plant of embodiment 1, wherein the plant is soybean.

[573] 8. The stably transformed plant of embodiment 1, wherein the recombinant DNA construct comprises codon-optimized nucleic acids for expression in the plant.

[574] 9. The stably transformed plant of embodiment 1, wherein the fusion protein has a molecular weight of 30 kDa to 50 kDa.

[575] 10. The stably transformed plant of embodiment 1, wherein the fusion protein is expressed at a level at least 2-fold higher than kappa-casein expressed individually in a plant.

[576] 11. The stably transformed plant of embodiment 1, wherein the fusion protein accumulates in the plant at least 2-fold higher than kappa-casein expressed without beta-lactoglobulin.

[577] 12. The stably transformed plant of embodiment 1, wherein the fusion protein is stably expressed in the plant in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant.

[578] 13. A transgenic soybean plant, comprising: a recombinant DNA construct encoding a fusion protein, the fusion protein comprising: a) bovine kappa-casein; and b) bovine beta-lactoglobulin, wherein the fusion protein is expressed in the soybean plant.

[579] 14. The transgenic soybean plant of embodiment 13, wherein the fusion protein comprises, in order from N-terminus to C-terminus, the kappa-casein and the beta-lactoglobulin.

[580] 15. The transgenic soybean plant of embodiment 13, wherein the fusion protein comprises a protease cleavage site.

[581] 16. The transgenic soybean plant of embodiment 13, wherein the fusion protein comprises a chymosin cleavage site.

[582] 17. The transgenic soybean plant of embodiment 13, wherein the fusion protein has a molecular weight of 30 kDa to 50 kDa.

[583] 18. A transgenic soybean plant, comprising: a recombinant DNA construct encoding a fusion protein, the fusion protein comprising a bovine casein and bovine beta-lactoglobulin.

[584] 19. The transgenic soybean plant of embodiment 18, wherein the fusion protein comprises, in order from N-terminus to C-terminus, the bovine casein and the beta-lactoglobulin.

[585] 20. The transgenic soybean plant of embodiment 18, wherein the fusion protein comprises a protease cleavage site.

[586] 21. The transgenic soybean plant of embodiment 18, wherein the fusion protein comprises a chymosin cleavage site.

[587] 22. The transgenic soybean plant of embodiment 18, wherein the fusion protein has a molecular weight of 30 kDa to 50 kDa.

Embodiment Set 2: Stably transformed plant expressing a fusion protein comprising kappa-casein or para-kappa-casein and beta-lactoglobulin

[588] 1. A recombinant fusion protein, comprising: a) full-length kappa-casein or para-kappa-casein; and b) beta-lactoglobulin.

[589] 2. The recombinant fusion protein of embodiment 1, wherein the fusion protein comprises, in order from N-terminus to C-terminus, the full-length kappa-casein or the para-kappa-casein and the beta-lactoglobulin.

[590] 3. The recombinant fusion protein of embodiment 1, further comprising a protease cleavage site.

[591] 4. The recombinant fusion protein of embodiment 3, wherein the protease cleavage site is a chymosin cleavage site.

[592] 5. The recombinant fusion protein of embodiment 1, further comprising a signal peptide.

[593] 6. The recombinant fusion protein of embodiment 5, wherein the signal peptide is located at the N-terminus of the fusion protein.

[594] 7. The recombinant fusion protein of embodiment 1, wherein the fusion protein comprises the full-length kappa-casein.

[595] 8. The recombinant fusion protein of embodiment 1, wherein the fusion protein comprises para-kappa-casein.

[596] 9. The recombinant fusion protein of embodiment 1, wherein the fusion protein has a molecular weight of 30 kDa to 50 kDa.

[597] 10. A plant transformed to express the recombinant fusion protein of embodiment 1, wherein the fusion protein is expressed in the plant in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant.

[598] 11. A plant transformed to express the recombinant fusion protein of embodiment 1, wherein the fusion protein is expressed in the plant at a level at least 2-fold higher than kappa-casein expressed individually in a plant.

[599] 12. A plant transformed to express the recombinant fusion protein of embodiment 1, wherein the fusion protein accumulates in the plant at least 2-fold higher than kappa-casein expressed without beta-lactoglobulin.

[600] 13. A fusion protein comprising kappa-casein and beta-lactoglobulin, wherein the kappa-casein is full-length kappa-casein comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 4 or the kappa-casein is para-kappa-casein comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2 and wherein the beta-lactoglobulin is full-length beta-lactoglobulin comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 10.

[601] 14. The fusion protein of embodiment 13, wherein the kappa-casein is full-length kappa-casein comprising an amino acid sequence SEQ ID NO: 4.

[602] 15. The fusion protein of embodiment 13, wherein the kappa-casein is para-kappa-casein comprising an amino acid sequence SEQ ID NO: 2.

[603] 16. The fusion protein of embodiment 13, wherein the beta-lactoglobulin comprises the amino acid sequence SEQ ID NO: 10.

[604] 17. The fusion protein of embodiment 13, further comprising a protease cleavage site between the kappa-casein and beta-lactoglobulin.

[605] 18. The fusion protein of embodiment 17, wherein the protease cleavage site is a chymosin cleavage site.

[606] 19. The fusion protein of embodiment 13, further comprising a signal peptide.

[607] 20. A nucleic acid molecule encoding a fusion protein comprising kappa-casein and beta-lactoglobulin, wherein the kappa-casein is full-length kappa-casein comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 4 or the kappa-casein is para-kappa-casein comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2 wherein the beta-lactoglobulin is full-length beta-lactoglobulin comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 10.

- [608] 21. The nucleic acid molecule of embodiment 20, wherein the nucleic acid sequence is codon optimized for expression in a plant.
- [609] 22. The nucleic acid molecule of embodiment 21, wherein the plant is soybean.
- [610] 23. An expression vector comprising the nucleic acid molecule of embodiment 20.
- [611] 24. A host cell comprising the expression vector of embodiment 23.
- [612] 25. The host cell of embodiment 24, wherein the host cell is selected from the group consisting of plant cells, bacterial cells, fungal cells, and mammalian cells.
- [613] 26. The host cell of embodiment 25, wherein the host cell is a plant cell.
- [614] 27. A plant stably transformed with the nucleic acid molecule of embodiment 20.
- [615] 28. The plant of embodiment 27, wherein the plant is a monocot selected from the group consisting of turf grass, maize, rice, oat, wheat, barley, sorghum, orchid, iris, lily, onion, palm, and duckweed.
- [616] 29. The plant of embodiment 27, wherein the plant is a dicot selected from the group consisting of Arabidopsis, tobacco, tomato, potato, sweet potato, cassava, alfalfa, lima bean, pea, chick pea, soybean, carrot, strawberry, lettuce, oak, maple, walnut, rose, mint, squash, daisy, quinoa, buckwheat, mung bean, cow pea, lentil, lupin, peanut, fava bean, French beans, mustard, and cactus.
- [617] 30. The plant of embodiment 29, wherein the plant is soybean.
- [618] 31. The plant of embodiment 27, wherein the plant is a non-vascular plant selected from the group consisting of moss, liverwort, hornwort, and algae.
- [619] 32. The plant of embodiment 27, wherein the plant is a vascular plant reproducing from spores.
- [620] 33. A method for stably expressing a recombinant fusion protein comprising kappa-casein and beta-lactoglobulin in a plant, wherein the kappa-casein is full-length kappa-casein comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 4 or the kappa-casein is para-kappa-casein comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2 and wherein the beta-lactoglobulin is full-length beta-lactoglobulin

comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 10, the method comprising: (i) transforming a plant with a plant transformation vector comprising an expression cassette comprising a nucleic acid molecule encoding the fusion protein; and (ii) growing the transformed plant under conditions wherein the recombinant fusion protein is expressed.

[621] 34. The method of embodiment 33, wherein the fusion protein is expressed in an amount of 1% or higher per the total protein weight of the soluble protein extractable from the plant.

[622] 35. The method of embodiment 33, wherein the fusion protein is expressed in the plant at a level at least 2-fold higher than kappa-casein expressed individually in a plant.

[623] 36. The method of embodiment 33, wherein the fusion protein accumulates in the plant at least 2-fold higher than kappa-casein is expressed without beta-lactoglobulin.

[624] 37. A food composition comprising a fusion protein comprising kappa-casein and beta-lactoglobulin, wherein the kappa-casein is full-length kappa-casein comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 4 or the kappa-casein is para-kappa-casein comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2 and wherein the beta-lactoglobulin is full-length beta-lactoglobulin comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 10.

[625] 38. The food composition of embodiment 37, wherein the food composition is selected from the group consisting of cheese and processed cheese products, yogurt and fermented dairy products, directly acidified counterparts of fermented dairy products, cottage cheese dressing, frozen dairy products, frozen desserts, desserts, baked goods, toppings, icings, fillings, low-fat spreads, dairy-based dry mixes, soups, sauces, salad dressing, geriatric nutrition, creams and creamers, analog dairy products, follow-up formula, baby formula, infant formula, milk, dairy beverages, acid dairy drinks, smoothies, milk tea, butter, margarine, butter alternatives, growing up milks, low-lactose products and beverages, medical and clinical nutrition products, protein/nutrition bar applications, sports beverages, confections, meat products, analog meat products, meal replacement beverages, weight management food and beverages, cultured buttermilk, sour cream, yogurt, skyr, leben, lassi, kefir, powder containing a milk protein, and low-lactose products.

Embodiment Set 3: Recombinant fusion protein comprising beta-casein and beta-lactoglobulin

- [626] 1. A recombinant fusion protein, comprising: a) beta-casein; and b) beta-lactoglobulin.
- [627] 2. The recombinant fusion protein of embodiment 1, further comprising a protease cleavage site.
- [628] 3. The recombinant fusion protein of embodiment 1, further comprising a chymosin cleavage site.
- [629] 4. A fusion protein, comprising: beta-casein and beta-lactoglobulin, wherein the beta-casein comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6 and wherein the beta-lactoglobulin comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 10.
- [630] 5. The fusion protein of embodiment 4, further comprising a protease cleavage site.
- [631] 6. The fusion protein of embodiment 4, further comprising a chymosin cleavage site.
- [632] 7. A nucleic acid molecule encoding a fusion protein comprising beta-casein and beta-lactoglobulin, wherein the beta-casein comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6 and wherein the beta-lactoglobulin comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 10.
- [633] 8. The nucleic acid molecule of embodiment 7, wherein the nucleic acid sequence is codon optimized for expression in a plant.
- [634] 9. The nucleic acid molecule of embodiment 8, wherein the plant is a soybean plant.
- [635] 10. An expression vector comprising the nucleic acid molecule of embodiment 7.
- [636] 11. A host cell comprising the expression vector of embodiment 10.
- [637] 12. The host cell of embodiment 11, wherein the host cell is selected from the group consisting of plant cells, bacterial cells, fungal cells, and mammalian cells.
- [638] 13. The host cell of embodiment 11, wherein the host cell is a plant cell.

- [639] 14. A plant stably transformed with the nucleic acid molecule of embodiment 7.
- [640] 15. The plant of embodiment 14, wherein the plant is a monocot selected from the group consisting of turf grass, maize, rice, oat, wheat, barley, sorghum, orchid, iris, lily, onion, palm, and duckweed.
- [641] 16. The plant of embodiment 14, wherein the plant is a dicot selected from the group consisting of Arabidopsis, tobacco, tomato, potato, sweet potato, cassava, alfalfa, lima bean, pea, chick pea, soybean, carrot, strawberry, lettuce, oak, maple, walnut, rose, mint squash, daisy, quinoa, buckwheat, mung bean, cow pea, lentil, lupin, peanut, fava bean, French beans, mustard, and cactus.
- [642] 17. The plant of embodiment 14, wherein the plant is a soybean plant.
- [643] 18. A food composition, comprising: a fusion protein comprising beta-casein and beta-lactoglobulin.
- [644] 19. The food composition of embodiment 18, wherein the food composition is a solid.
- [645] 20. The food composition of embodiment 18, wherein the food composition is a liquid.
- [646] 21. The food composition of embodiment 18, wherein the food composition is a powder.
- [647] 22. The food composition of embodiment 18, wherein the food composition is selected from the group consisting of: cheese, processed cheese product, yogurt, fermented dairy product, directly acidified counterpart of fermented dairy product, cottage cheese, dressing, frozen dairy product, frozen dessert, dessert, baked good, topping, icing, filling, low-fat spread, dairy-based dry mix, soup, sauce, salad dressing, geriatric nutrition, cream, creamer, analog dairy product, follow-up formula, baby formula, infant formula, milk, dairy beverage, acid dairy drink, smoothie, milk tea, butter, margarine, butter alternative, growing up milk, low-lactose product, low-lactose beverage, medical and clinical nutrition product, protein bar, nutrition bar, sport beverage, confection, meat product, analog meat product, meal replacement beverage, weight management food and beverage, dairy product, cultured buttermilk, sour cream, yogurt, skyr, leben, lassi, kefir, powder containing a milk protein, and low-lactose product.

[648] 23. The food composition of embodiment 18, wherein the food composition is a dairy product.

[649] 24. The food composition of embodiment 18, wherein the food composition is an analog dairy product.

[650] 25. The food composition of embodiment 18, wherein the food composition is a low lactose product.

[651] 26. The food composition of embodiment 18, wherein the food composition is a milk.

[652] 27. The food composition of embodiment 18, wherein the food composition is a cheese.

[653] 28. The food composition of embodiment 18, wherein the food composition is fermented.

Embodiment Set 4: Seed processing composition

[654] 1. A seed processing composition, comprising: a) a fusion protein, comprising i) a full-length kappa-casein or para-kappa-casein component; and ii) a beta-lactoglobulin component; and b) plant seed tissue.

[655] 2. The seed processing composition of embodiment 1, wherein the plant seed tissue is ground.

[656] 3. The seed processing composition of embodiment 1, wherein the plant seed tissue is soybean.

[657] 4. The seed processing composition of embodiment 1, further comprising at least one member selected from the group consisting of: enzyme, protease, chymosin, extractant, solvent, phenol, buffer, additive, salt, protease inhibitor, peptidase inhibitor, osmolyte, and reducing agent.

[658] 5. A food composition comprising the seed processing composition of embodiment 1.

[659] 6. A protein concentrate composition, comprising a protein concentrate of a fusion protein comprising i) a full-length kappa-casein or para-kappa-casein component, and ii) a beta-lactoglobulin component.

- [660] 7. The protein concentrate composition of embodiment 6, wherein there is no plant seed tissue present.
- [661] 8. The protein concentrate composition of embodiment 6, further comprising at least one member selected from the group consisting of: enzyme, protease, chymosin, extractant, solvent, phenol, buffer, additive, salt, protease inhibitor, peptidase inhibitor, osmolyte, and reducing agent.
- [662] 9. The protein concentrate composition of embodiment 6, further comprising chymosin.
- [663] 10. A food composition comprising the protein concentrate composition of embodiment 6.
- [664] 11. A food composition, comprising: a fusion protein, comprising i) a full-length kappa-casein or para-kappa-casein component, and ii) a beta-lactoglobulin component.
- [665] 12. The food composition of embodiment 11, wherein the food composition comprises the fusion protein comprising the full-length kappa-casein component and a beta-lactoglobulin component.
- [666] 13. The food composition of embodiment 11, wherein the food composition comprises the fusion protein comprising the para-kappa-casein component and a beta-lactoglobulin component.
- [667] 14. The food composition of embodiment 11, wherein the food composition is a solid.
- [668] 15. The food composition of embodiment 11, wherein the food composition is a liquid.
- [669] 16. The food composition of embodiment 11, wherein the food composition is a powder.
- [670] 17. The food composition of embodiment 11, wherein the food composition is selected from the group consisting of: cheese, processed cheese product, fermented dairy product, directly acidified counterpart of fermented dairy product, cottage cheese, dressing, frozen dairy product, frozen dessert, dessert, baked good, topping, icing, filling, low-fat spread, dairy-based dry mix, soup, sauce, salad dressing, geriatric nutrition, cream, creamer, analog dairy product, follow-up formula, baby formula, infant formula, milk, dairy beverage, acid dairy drink, smoothie, milk tea, butter, margarine, butter alternative, growing up milk, low-lactose product, low-lactose beverage, medical and clinical nutrition product, protein bar, nutrition bar, sport beverage, confection, meat product, analog meat product, meal replacement beverage, weight management food and beverage,

dairy product, cultured buttermilk, sour cream, yogurt, skyr, leben, lassi, kefir, powder containing a milk protein, and low-lactose product.

[671] 18. The food composition of embodiment 11, wherein the food composition is a dairy product.

[672] 19. The food composition of embodiment 11, wherein the food composition is an analog dairy product.

[673] 20. The food composition of embodiment 11, wherein the food composition is a low lactose product.

[674] 21. The food composition of embodiment 11, wherein the food composition is a milk.

[675] 22. The food composition of embodiment 11, wherein the food composition is a cheese.

[676] 23. The food composition of embodiment 11, wherein the food composition is fermented.

[677] 24. A method of making a food composition, comprising: combining a fusion protein, comprising i) a full-length kappa-casein or para-kappa-casein component, and ii) a beta-lactoglobulin component, into a food composition.

[678] 25. The method of embodiment 24, wherein the food composition is selected from the group consisting of: cheese, processed cheese product, yogurt, fermented dairy product, directly acidified counterpart of fermented dairy product, cottage cheese, dressing, frozen dairy product, frozen dessert, dessert, baked good, topping, icing, filling, low-fat spread, dairy-based dry mix, soup, sauce, salad dressing, geriatric nutrition, cream, creamer, analog dairy product, follow-up formula, baby formula, infant formula, milk, dairy beverage, acid dairy drink, smoothie, milk tea, butter, margarine, butter alternative, growing up milk, low-lactose product, low-lactose beverage, medical and clinical nutrition product, protein bar, nutrition bar, sport beverage, confection, meat product, analog meat product, meal replacement beverage, weight management food and beverage, dairy product, cultured buttermilk, sour cream, yogurt, skyr, leben, lassi, kefir, powder containing a milk protein, and low-lactose product.

[679] 26. The method of embodiment 24, wherein the food composition is a dairy product.

[680] 27. The method of embodiment 24, wherein the food composition is a cheese.

- [681] 28. A food composition made by the method of embodiment 24.
- [682] 29. A method for making a fusion protein, comprising: (a) transforming a host cell with a vector comprising an expression cassette comprising a nucleic acid molecule encoding the fusion protein, wherein the fusion protein comprises i) a full-length kappa-casein or para-kappa-casein component, and ii) a beta-lactoglobulin component, and (b) growing the transformed host cell under conditions wherein the fusion protein is expressed.
- [683] 30. The method of embodiment 29, wherein the host cell is selected from the group consisting of plant cells, bacterial cells, fungal cells, and mammalian cells.
- [684] 31. The method of embodiment 29, wherein the host cell is a plant cell.
- [685] 32. A fusion protein made by the method of embodiment 29.

Embodiment Set 5: Transgenic plant comprising a recombinant DNA encoding a fusion protein comprising bovine casein and bovine beta-lactoglobulin

- [686] 1. A transgenic plant, comprising: a recombinant DNA construct encoding a fusion protein, the fusion protein comprising a bovine casein and bovine beta-lactoglobulin.
- [687] 2. The transgenic plant of embodiment 1, wherein the fusion protein comprises a protease cleavage site.
- [688] 3. The transgenic plant of embodiment 1, wherein the fusion protein comprises a chymosin cleavage site.
- [689] 4. The transgenic plant of embodiment 1, wherein the fusion protein has a molecular weight of 30 kDa to 50 kDa.
- [690] 5. A method of making a food composition, comprising: a) extracting the bovine casein and bovine beta-lactoglobulin fusion protein from the transgenic plant of embodiment 1; b) optionally separating the bovine casein from the bovine beta-lactoglobulin; and c) combining the fusion protein or the bovine casein or the bovine beta-lactoglobulin into a food composition.
- [691] 6. The method of embodiment 5, wherein the food composition is selected from the group consisting of: cheese, processed cheese product, fermented dairy product, directly acidified

counterpart of fermented dairy product, cottage cheese, dressing, frozen dairy product, frozen dessert, dessert, baked good, topping, icing, filling, low-fat spread, dairy-based dry mix, soup, sauce, salad dressing, geriatric nutrition, cream, creamer, analog dairy product, follow-up formula, baby formula, infant formula, milk, dairy beverage, acid dairy drink, smoothie, milk tea, butter, margarine, butter alternative, growing up milk, low-lactose product, low-lactose beverage, medical and clinical nutrition product, protein bar, nutrition bar, sport beverage, confection, meat product, analog meat product, meal replacement beverage, weight management food and beverage, dairy product, cultured buttermilk, sour cream, yogurt, skyr, leben, lassi, kefir, powder containing a milk protein, and low-lactose product.

[692] 7. The method of embodiment 5, wherein the bovine casein is not separated from the bovine beta-lactoglobulin and the food composition comprises the fusion protein.

[693] 8. The method of embodiment 5, wherein the bovine casein is separated from the bovine beta-lactoglobulin and the food composition comprises the bovine casein.

[694] 9. The method of embodiment 5, wherein the bovine casein is separated from the bovine beta-lactoglobulin and the food composition comprises the bovine beta-lactoglobulin.

[695] 10. The method of embodiment 5, wherein the food composition is a solid.

[696] 11. The method of embodiment 5, wherein the food composition is a liquid.

[697] 12. The method of embodiment 5, wherein the food composition is a powder.

[698] 13. The method of embodiment 5, wherein the food composition is a dairy product.

[699] 14. The method of embodiment 5, wherein the food composition is an analog dairy product.

[700] 15. The method of embodiment 5, wherein the food composition is a low lactose product.

[701] 16. The method of embodiment 5, wherein the food composition is a milk.

[702] 17. The method of embodiment 5, wherein the food composition is a cheese.

Embodiment Set 6: Recombinant fusion protein comprising casein and beta-lactoglobulin

[703] 1. A recombinant fusion protein, comprising: a) casein; and b) beta-lactoglobulin.

- [704] 2. The recombinant fusion protein of embodiment 1, further comprising a protease cleavage site.
- [705] 3. The recombinant fusion protein of embodiment 1, further comprising a chymosin cleavage site.
- [706] 4. The recombinant fusion protein of embodiment 1, wherein the casein is bovine.
- [707] 5. The recombinant fusion protein of embodiment 1, wherein the β -lactoglobulin is bovine.
- [708] 6. The recombinant fusion protein of embodiment 1, wherein the casein and β -lactoglobulin are bovine.
- [709] 7. A nucleic acid molecule encoding the recombinant fusion protein of embodiment 1.
- [710] 8. The nucleic acid molecule of embodiment 7, wherein the nucleic acid sequence is codon optimized for expression in a plant.
- [711] 9. The nucleic acid molecule of embodiment 8, wherein the plant is a soybean plant.
- [712] 10. An expression vector comprising the nucleic acid molecule of embodiment 7.
- [713] 11. A host cell comprising the expression vector of embodiment 10.
- [714] 12. The host cell of embodiment 11, wherein the host cell is selected from the group consisting of plant cells, bacterial cells, fungal cells, and mammalian cells.
- [715] 13. The host cell of embodiment 11, wherein the host cell is a plant cell.
- [716] 14. A plant stably transformed with the nucleic acid molecule of embodiment 7.
- [717] 15. The plant of embodiment 14, wherein the plant is a monocot selected from the group consisting of turf grass, maize, rice, oat, wheat, barley, sorghum, orchid, iris, lily, onion, palm, and duckweed.
- [718] 16. The plant of embodiment 14, wherein the plant is a dicot selected from the group consisting of Arabidopsis, tobacco, tomato, potato, sweet potato, cassava, alfalfa, lima bean, pea, chick pea, soybean, carrot, strawberry, lettuce, oak, maple, walnut, rose, mint, squash, daisy, quinoa, buckwheat, mung bean, cow pea, lentil, lupin, peanut, fava bean, French beans, mustard, and cactus.

- [719] 17. The plant of embodiment 14, wherein the plant is a soybean plant.
- [720] 18. A food composition, comprising: a fusion protein comprising casein and β -lactoglobulin.
- [721] 19. The food composition of embodiment 18, wherein the food composition is a solid.
- [722] 20. The food composition of embodiment 18, wherein the food composition is a liquid.
- [723] 21. The food composition of embodiment 18, wherein the food composition is a powder.
- [724] 22. The food composition of embodiment 18, wherein the food composition is selected from the group consisting of: cheese, processed cheese product, yogurt, fermented dairy product, directly acidified counterpart of fermented dairy product, cottage cheese, dressing, frozen dairy product, frozen dessert, dessert, baked good, topping, icing, filling, low-fat spread, dairy-based dry mix, soup, sauce, salad dressing, geriatric nutrition, cream, creamer, analog dairy product, follow-up formula, baby formula, infant formula, milk, dairy beverage, acid dairy drink, smoothie, milk tea, butter, margarine, butter alternative, growing up milk, low-lactose product, low-lactose beverage, medical and clinical nutrition product, protein bar, nutrition bar, sport beverage, confection, meat product, analog meat product, meal replacement beverage, weight management food and beverage, dairy product, cultured buttermilk, sour cream, skyr, leben, lassi, kefir, powder containing a milk protein, and low-lactose product.
- [725] 23. The food composition of embodiment 18, wherein the food composition is a dairy product.
- [726] 24. The food composition of embodiment 18, wherein the food composition is an analog dairy product.
- [727] 25. The food composition of embodiment 18, wherein the food composition is a low lactose product.
- [728] 26. The food composition of embodiment 18, wherein the food composition is a milk.
- [729] 27. The food composition of embodiment 18, wherein the food composition is a cheese.
- [730] 28. The food composition of embodiment 18, wherein the food composition is fermented.

Embodiment Set 7: Food composition comprising at least one component of a fusion protein

[731] 1. A food composition, comprising: at least one component of a fusion protein, the fusion protein comprising i) a bovine casein component and ii) a bovine β -lactoglobulin component, wherein the component has been separated from the fusion protein.

[732] 2. The food composition of embodiment 1, wherein the food composition comprises the bovine casein component.

[733] 3. The food composition of embodiment 1, wherein the food composition comprises the bovine β -lactoglobulin component.

[734] 4. The food composition of embodiment 1, wherein the food composition is selected from the group consisting of: cheese, processed cheese product, fermented dairy product, directly acidified counterpart of fermented dairy product, cottage cheese, dressing, frozen dairy product, frozen dessert, dessert, baked good, topping, icing, filling, low-fat spread, dairy-based dry mix, soup, sauce, salad dressing, geriatric nutrition, cream, creamer, analog dairy product, follow-up formula, baby formula, infant formula, milk, dairy beverage, acid dairy drink, smoothie, milk tea, butter, margarine, butter alternative, growing up milk, low-lactose product, low-lactose beverage, medical and clinical nutrition product, protein bar, nutrition bar, sport beverage, confection, meat product, analog meat product, meal replacement beverage, weight management food and beverage, dairy product, cultured buttermilk, sour cream, yogurt, skyr, leben, lassi, kefir, powder containing a milk protein, and low-lactose product.

[735] 5. The food composition of embodiment 1, wherein the food composition is a solid.

[736] 6. The food composition of embodiment 1, wherein the food composition is a liquid.

[737] 7. The food composition of embodiment 1, wherein the food composition is a powder.

[738] 8. The food composition of embodiment 1, wherein the food composition is a dairy product.

[739] 9. The food composition of embodiment 1, wherein the food composition is an analog dairy product.

[740] 10. The food composition of embodiment 1, wherein the food composition is a low lactose product.

- [741] 11. The food composition of embodiment 1, wherein the food composition is a milk.
- [742] 12. The food composition of embodiment 1, wherein the food composition is a cheese.
- [743] 13. The food composition of embodiment 1, wherein the food composition is fermented.
- [744] 14. A food composition, comprising: a fusion protein comprising bovine casein and bovine β -lactoglobulin.
- [745] 15. The food composition of embodiment 14, wherein the fusion protein comprises a protease cleavage site.
- [746] 16. The food composition of embodiment 14, wherein the fusion protein comprises a chymosin cleavage site.
- [747] 17. The food composition of embodiment 14, wherein the fusion protein has a molecular weight of 30 kDa to 50 kDa.
- [748] 18. The food composition of embodiment 14, wherein the food composition is selected from the group consisting of: cheese, processed cheese product, fermented dairy product, directly acidified counterpart of fermented dairy product, cottage cheese, dressing, frozen dairy product, frozen dessert, dessert, baked good, topping, icing, filling, low-fat spread, dairy-based dry mix, soup, sauce, salad dressing, geriatric nutrition, cream, creamer, analog dairy product, follow-up formula, baby formula, infant formula, milk, dairy beverage, acid dairy drink, smoothie, milk tea, butter, margarine, butter alternative, growing up milk, low-lactose product, low-lactose beverage, medical and clinical nutrition product, protein bar, nutrition bar, sport beverage, confection, meat product, analog meat product, meal replacement beverage, weight management food and beverage, dairy product, cultured buttermilk, sour cream, yogurt, skyr, leben, lassi, kefir, powder containing a milk protein, and low-lactose product.
- [749] 19. The food composition of embodiment 14, wherein the food composition is a solid.
- [750] 20. The food composition of embodiment 14, wherein the food composition is a liquid.
- [751] 21. The food composition of embodiment 14, wherein the food composition is a powder.
- [752] 22. The food composition of embodiment 14, wherein the food composition is a dairy product.

[753] 23. The food composition of embodiment 14, wherein the food composition is an analog dairy product.

[754] 24. The food composition of embodiment 14, wherein the food composition is a low lactose product.

[755] 25. The food composition of embodiment 14, wherein the food composition is a milk.

[756] 26. The food composition of embodiment 14, wherein the food composition is a cheese.

[757] 27. The food composition of embodiment 14, wherein the food composition is fermented.

[758] 28. The food composition of embodiment 14, wherein the fusion protein is a plant expressed fusion protein.

[759] 29. The food composition of embodiment 14, wherein the fusion protein is a soybean expressed fusion protein.

Embodiment Set 8: Alternative dairy food composition

[760] 1. An alternative dairy food composition comprising: i) a recombinant beta-casein protein, and ii) at least one lipid, wherein the alternative dairy food composition does not comprise any other milk proteins; and wherein the recombinant beta-casein protein confers on the alternative dairy food composition one or more characteristics of a dairy food product selected from the group consisting of: taste, aroma, appearance, handling, mouthfeel, density, structure, texture, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess and emulsification.

[761] 2. The alternative dairy food composition of embodiment 1, wherein the recombinant beta-casein is plant-expressed.

[762] 3. The alternative dairy food composition of embodiment 2, wherein the recombinant beta-casein is expressed in a soybean plant.

[763] 4. The alternative dairy food composition of embodiment 1, wherein the composition comprises a fusion protein comprising the recombinant beta-casein.

[764] 5. The alternative dairy food composition of embodiment 1, wherein the composition is a milk composition, a cream composition, a yogurt composition, an ice cream composition, a frozen custard composition, a frozen dessert composition, a crème fraiche composition, a curd composition, a cottage cheese composition, or a cream cheese composition.

[765] 6. The alternative dairy food composition of embodiment 1, wherein the composition comprises at least one salt.

[766] 7. The alternative dairy food composition of embodiment 1, wherein the composition comprises calcium.

[767] 8. The alternative dairy food composition of embodiment 1, wherein the composition comprises calcium at a concentration of about 0.01% to about 2% by weight.

[768] 9. The alternative dairy food composition of embodiment 1, wherein the composition has a pH of about 4 to about 8.

[769] 10. The alternative dairy food composition of embodiment 1, wherein the composition comprises a fusion protein comprising the recombinant beta-casein.

Embodiment Set 9: Alternative dairy food composition

[770] 1. A cheese composition comprising recombinant casein protein; wherein about 32% to 100% by weight of the total casein protein in the cheese composition is beta-casein; and wherein the cheese composition has the ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the composition at a temperature of 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[771] 2. The cheese composition of embodiment 1, wherein the composition does not comprise any casein proteins other than beta-casein.

[772] 3. The cheese composition of embodiment 1, wherein the composition comprises at least one additional casein protein.

[773] 4. The cheese composition of embodiment 3, wherein at least 80% by weight of the total casein protein in the composition is beta-casein.

[774] 5. The cheese composition of embodiment 3, wherein at least 90% by weight of the total casein protein in the composition is beta-casein.

[775] 6. The cheese composition of embodiment 3, wherein at least 95% by weight of the total casein protein in the composition is beta-casein.

[776] 7. The cheese composition of embodiment 3, wherein the at least one additional casein protein is selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein and alpha-S2-casein.

[777] 8. The cheese composition of embodiment 3, wherein the at least one additional casein protein is kappa-casein.

[778] 9. The cheese composition of embodiment 3, wherein the at least one additional casein protein is para-kappa casein.

[779] 10. The cheese composition of embodiment 1, wherein the recombinant beta-casein is plant-expressed.

[780] 11. The cheese composition of embodiment 10, wherein the recombinant beta-casein is expressed in a soybean plant.

[781] 12. The cheese composition of embodiment 3, wherein all caseins in the composition are plant-expressed.

[782] 13. The cheese composition of embodiment 1, wherein the recombinant beta-casein protein is derived from a fusion protein.

[783] 14. The cheese composition of embodiment 1, wherein the composition does not contain an organoleptically functional amount of beta-lactoglobulin.

[784] 15. The cheese composition of embodiment 1, wherein the composition has the ability to stretch to at least 5 cm in length without breaking, as determined by heating a 100 gram mass of the composition at a temperature of 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[785] 16. The cheese composition of embodiment 1, wherein the composition has the ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of

the composition at a temperature of 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass; and a firmness of at least 150 grams, as determined by compressing a cylindrical-shaped sample of the cheese composition having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C.

[786] 17. The cheese composition of embodiment 1, wherein the composition comprises at least one lipid and at least one salt.

[787] 18. The cheese composition of embodiment 1, wherein the composition comprises calcium.

[788] 19. The cheese composition of embodiment 18, wherein the composition comprises calcium at a concentration of about 0.01% to about 2% by weight.

[789] 20. The cheese composition of embodiment 1, wherein the composition has a pH of about 5.2 to about 5.9.

[790] 21. The cheese composition of embodiment 1, wherein the composition comprises at least one organoleptic property similar to cheese produced from mammalian milk selected from the group consisting of taste, appearance, mouthfeel, structure, texture, density, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess, and emulsification.

[791] 22. A method of making the cheese composition of embodiment 1, the method comprising expressing the recombinant beta-casein protein in a plant, extracting the beta-casein from the plant, and combining the beta-casein with at least one lipid and/or salt.

[792] 23. A cheese composition comprising a recombinant casein protein; wherein about 32% to 100% by weight of the total casein protein in the cheese composition is beta-casein; and wherein the cheese composition has ability to stretch to at least 5 cm in length without breaking, as determined by heating a 100 gram mass of the composition at a temperature of 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[793] 24. The cheese composition of embodiment 23, wherein the composition does not comprise any additional casein proteins.

[794] 25. The cheese composition of embodiment 23, wherein the composition comprises at least one additional casein protein, and wherein at least 80% by weight of the total casein protein in the composition is beta-casein.

[795] 26. The cheese composition of embodiment 25, wherein the at least one additional casein protein is kappa-casein or para-kappa casein.

[796] 27. The cheese composition of embodiment 23, wherein the recombinant beta-casein is plant-expressed.

[797] 28. The cheese composition of embodiment 23, wherein the recombinant beta-casein protein is derived from a fusion protein.

[798] 29. The cheese composition of embodiment 23, wherein the composition has at least one of the following characteristics: i) a firmness of at least 150 grams, as determined by compressing a cylindrical-shaped sample of the cheese composition having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C; or ii) a melting point of about 35°C to about 100°C.

[799] 30. A method of making the cheese composition of embodiment 23, the method comprising expressing the recombinant beta-casein protein in a plant, extracting the beta-casein from the plant, and combining the beta-casein with at least one lipid and/or salt.

Embodiment Set 10: Fusion protein comprising first and second milk proteins, and transformed plants expressing the same

[800] 1. A transformed plant comprising in its genome: a recombinant DNA construct encoding a fusion protein, the fusion protein comprising a first protein and a second protein, wherein the first protein and/or second protein is a milk protein, and wherein the fusion protein is expressed in the plant in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant.

[801] 2. The transformed plant of embodiment 1, wherein the fusion protein comprises, from N-terminus to C-terminus, the first protein and the second protein.

[802] 3. The transformed plant of embodiment 1, wherein the fusion protein comprises, from N-terminus to C-terminus, the second protein and the first protein.

[803] 4. The transformed plant of any one of embodiments 1-3, wherein the milk protein is α -S1 casein, α -S2 casein, β -casein, κ -casein, para- κ -casein, β -lactoglobulin, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, or an immunoglobulin.

[804] 5. The transformed plant of any one of embodiments 1-3, wherein the milk protein is selected from the group consisting of: SEQ ID NO: 4, or a sequence at least 90% identical thereto; SEQ ID NO: 2, or a sequence at least 90% identical thereto; SEQ ID NO: 6, or a sequence at least 90% identical thereto; SEQ ID NO: 8, or a sequence at least 90% identical thereto; SEQ ID NO: 84, or a sequence at least 90% identical thereto; and SEQ ID NO: 10, or a sequence at least 90% identical thereto.

[805] 6. The transformed plant of any one of embodiments 1-5, wherein each of the first protein and the second protein are milk proteins.

[806] 7. The transformed plant of any one of embodiments 1-5, wherein the first protein is a milk protein and the second protein is a non-milk protein.

[807] 8. The transformed plant of embodiment 7, wherein the non-milk protein is albumin, hemoglobin, collagen, ovalbumin, ovotransferrin, GFP, or ovoglobulin.

[808] 9. The transformed plant of embodiment 6, wherein the first protein and the second protein are each casein proteins.

[809] 10. The transformed plant of any one of embodiments 1-9, wherein the plant is a dicot.

[810] 11. The transformed plant of embodiment 10, wherein the dicot is Arabidopsis, tobacco, tomato, potato, sweet potato, cassava, alfalfa, lima bean, pea, chick pea, soybean, carrot, strawberry, lettuce, oak, maple, walnut, rose, mint, squash, daisy, or cactus.

[811] 12. The transformed plant of any one of embodiments 1-9, wherein the plant is soybean.

[812] 13. The transformed plant of any one of embodiments 1-12, wherein the fusion protein is stably expressed.

[813] 14. The transformed plant of any one of embodiments 1-12, wherein the fusion protein is transiently expressed.

[814] 15. The transformed plant of any one of embodiments 1-14, wherein the recombinant DNA construct is codon-optimized for expression in the plant.

[815] 16. The transformed plant of any one of embodiments 1-15, wherein the fusion protein comprises a protease cleavage site.

[816] 17. The transformed plant of embodiment 16, wherein the protease cleavage site is a chymosin cleavage site.

[817] 18. The transformed plant of any one of embodiments 1-17, wherein the fusion protein is expressed at a level at least 2-fold higher than a casein protein expressed individually in a plant.

[818] 19. A recombinant fusion protein comprising a first protein and a second protein, wherein at least one of the first protein and the second protein is a milk protein.

[819] 20. The recombinant fusion protein of embodiment 19, wherein the fusion protein comprises, from N-terminus to C-terminus, the first protein and the second protein.

[820] 21. The recombinant fusion protein of embodiment 19, wherein the fusion protein comprises, from N-terminus to C-terminus, the second protein and the first protein.

[821] 22. The recombinant fusion protein of any one of embodiments 19-21, wherein the milk protein is α -S1 casein, α -S2 casein, β -casein, κ -casein, para- κ -casein, β -lactoglobulin, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, or an immunoglobulin.

[822] 23. The recombinant fusion protein of any one of embodiments 19-21, wherein the milk protein is selected from the group consisting of: SEQ ID NO: 4, or a sequence at least 90% identical thereto; SEQ ID NO: 2, or a sequence at least 90% identical thereto; SEQ ID NO: 6, or a sequence at least 90% identical thereto; SEQ ID NO: 8, or a sequence at least 90% identical thereto; SEQ ID NO: 84, or a sequence at least 90% identical thereto; and SEQ ID NO: 10, or a sequence at least 90% identical thereto.

[823] 24. The recombinant fusion protein of any one of embodiments 19-23, wherein the first protein and the second protein are milk proteins.

- [824] 25. The recombinant fusion protein of any one of embodiments 19-23, wherein the first protein is a milk protein and the second protein is a non-milk protein.
- [825] 26. The recombinant fusion protein of embodiment 25, wherein the non-milk protein is albumin, hemoglobin, collagen, ovalbumin, ovotransferrin, GFP, or ovoglobulin.
- [826] 27. The recombinant fusion protein of embodiment 24, wherein the first protein and the second protein are each casein proteins.
- [827] 28. The recombinant fusion protein of embodiment 27, wherein the first protein and the second protein are the same casein protein.
- [828] 29. The recombinant fusion protein of embodiment 27, wherein the first protein and the second protein are both α -S1 casein, α -S2 casein, β -casein, κ -casein, or para- κ -casein.
- [829] 30. The recombinant fusion protein of embodiment 24, wherein the first protein and the second protein are each casein proteins and are different from one another.
- [830] 31. The recombinant fusion protein of embodiment 30, wherein the first protein and the second protein are each independently selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, and para- κ -casein.
- [831] 32. A recombinant fusion protein comprising a casein protein and lysozyme, wherein the casein protein is selected from the group consisting of α -S1 casein, α -S2 casein, β -casein, κ -casein, and para- κ -casein.
- [832] 33. A recombinant fusion protein comprising a casein protein and β -lactoglobulin, wherein the casein protein is selected from the group consisting of α -S1 casein, α -S2 casein, β -casein, κ -casein, and para- κ -casein.
- [833] 34. The recombinant fusion protein of any one of embodiments 19-33, wherein the fusion protein comprises a protease cleavage site.
- [834] 35. The recombinant fusion protein of embodiment 34, wherein the protease cleavage site is a chymosin cleavage site.
- [835] 36. A nucleic acid encoding the recombinant fusion protein of any one of embodiments 19-35.

[836] 37. The nucleic acid of embodiment 36, wherein the nucleic acid is codon optimized for expression in a plant species.

[837] 38. The nucleic acid of embodiment 36, wherein the nucleic acid is codon optimized for expression in soybean.

[838] 39. A vector comprising a nucleic acid encoding a recombinant fusion protein, wherein the recombinant fusion protein comprises a first protein and a second protein, wherein at least one of the first protein and the second protein is a milk protein.

[839] 40. The vector of embodiment 39, wherein the vector is a plasmid.

[840] 41. The vector of embodiment 40, wherein the vector is an *Agrobacterium* Ti plasmid.

[841] 42. The vector of any one of embodiments 39-41, wherein the nucleic acid comprises, in order from 5' to 3': a promoter; a 5' untranslated region; a sequence encoding the fusion protein of any one of embodiments 19-35; and a terminator.

[842] 43. The vector of embodiment 42, wherein the promoter is a seed-specific promoter.

[843] 44. The vector of embodiment 43, wherein the seed-specific promoter is selected from the group consisting of PvPhas, BnNap, AtOle1, GmSeed2, GmSeed3, GmSeed5, GmSeed6, GmSeed7, GmSeed8, GmSeed10, GmSeed11, GmSeed12, pBCON, GmCEP1-L, GmTHIC, GmBg7S1, GmGRD, GmOLEA, GmOLER, Gm2S-1, and GmBBId-II.

[844] 45. The vector of embodiment 43, wherein the seed-specific promoter is PvPhas and comprises the sequence of SEQ ID NO: 18, or a sequence at least 90% identical thereto.

[845] 46. The vector of embodiment 43, wherein the seed-specific promoter is GmSeed2 and comprises the sequence of SEQ ID NO: 19, or a sequence at least 90% identical thereto.

[846] 47. The vector of embodiment 42, wherein the 5' untranslated region is selected from the group consisting of Arc5'UTR and glnB1UTR.

[847] 48. The vector of embodiment 47, wherein the 5' untranslated region is Arc5'UTR and comprises the sequence of SEQ ID NO: 20, or a sequence at least 90% identical thereto.

[848] 49. The vector of embodiment 42, wherein the expression cassette comprises a 3' untranslated region.

[849] 50. The vector of embodiment 49, wherein the 3' untranslated region is Arc5-1 and comprises SEQ ID NO: 21, or a sequence at least 90% identical thereto.

[850] 51. The vector of embodiment 42, wherein the terminator sequence is a terminator isolated or derived from a gene encoding Nopaline synthase, Arc5-1, an Extensin, Rb7 matrix attachment region, a Heat shock protein, Ubiquitin 10, Ubiquitin 3, and M6 matrix attachment region.

[851] 52. The vector of embodiment 42, wherein the terminator sequence is isolated or derived from a Nopaline synthase gene and comprises the sequence of SEQ ID NO: 22, or a sequence at least 90% identical thereto.

[852] 53. The vector of embodiment 42, wherein the terminator sequence is a dual terminator and is selected from the group consisting of: SEQ ID NO: 138, or a sequence at least 90% identical thereto; SEQ ID NO: 141, or a sequence at least 90% identical thereto; SEQ ID NO: 144, or a sequence at least 90% identical thereto; and SEQ ID NO: 146, or a sequence at least 90% identical thereto.

[853] 54. A plant-expressed recombinant fusion protein, comprising: κ -casein and β -lactoglobulin.

[854] 55. The plant-expressed recombinant fusion protein of embodiment 54, wherein the fusion protein comprises, in order from N-terminus to C-terminus, the κ -casein and the β -lactoglobulin.

[855] 56. The plant-expressed recombinant fusion protein of embodiment 54 or 55, wherein the fusion protein comprises a protease cleavage site.

[856] 57. The plant-expressed recombinant fusion protein of embodiment 56, wherein the protease cleavage site is a chymosin cleavage site.

[857] 58. The plant-expressed recombinant fusion protein of any one of embodiments 55-57, wherein the fusion protein comprises a signal peptide.

[858] 59. The plant-expressed recombinant fusion protein of embodiment 58, wherein the signal peptide is located at the N-terminus of the fusion protein.

[859] 60. The plant-expressed recombinant fusion protein of any one of embodiments 55-59, wherein the fusion protein is encoded by a nucleic acid that is codon optimized for expression in a plant.

[860] 61. The plant-expressed recombinant fusion protein of any one of embodiments 55-60, wherein the fusion protein is expressed in a soybean.

[861] 62. The plant-expressed recombinant fusion protein of any one of embodiments 55-61, wherein the fusion protein has a molecular weight of 30 kDa to 50 kDa.

[862] 63. The plant-expressed recombinant fusion protein of any one of embodiments 55-62, wherein the fusion protein is expressed in a plant in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant.

[863] 64. The plant-expressed recombinant fusion protein of any one of embodiments 55-62, wherein the fusion protein is expressed in the plant at a level at least 2-fold higher than κ -casein expressed individually in a plant.

[864] 65. The plant-expressed recombinant fusion protein of any one of embodiments 55-62, wherein the fusion protein accumulates in the plant at least 2-fold higher than κ -casein expressed without β -lactoglobulin.

[865] 66. A stably transformed plant, comprising in its genome: a recombinant DNA construct encoding a fusion protein, the fusion protein comprising: κ -casein and β -lactoglobulin; wherein the fusion protein is stably expressed in the plant in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant.

[866] 67. The stably transformed plant of embodiment 66, wherein the fusion protein comprises, in order from N-terminus to C-terminus, the κ -casein and the β -lactoglobulin.

[867] 68. The stably transformed plant of embodiment 66 or 67, wherein the fusion protein comprises a protease cleavage site.

[868] 69. The stably transformed plant of embodiment 68, wherein the protease cleavage site is a chymosin cleavage site.

[869] 70. The stably transformed plant of any one of embodiments 66-69, wherein the fusion protein comprises a signal peptide.

[870] 71. The stably transformed plant of embodiment 70, wherein the signal peptide is located at the N-terminus of the fusion protein.

[871] 72. The stably transformed plant of any one of embodiments 66-71, wherein the plant is soybean.

[872] 73. The stably transformed plant of any one of embodiments 66-72, wherein the recombinant DNA construct comprises codon-optimized nucleic acids for expression in the plant.

[873] 74. The stably transformed plant of any one of embodiments 66-73, wherein the fusion protein has a molecular weight of 30 kDa to 50 kDa.

[874] 75. The stably transformed plant of any one of embodiments 66-74, wherein the fusion protein is expressed at a level at least 2-fold higher than κ -casein expressed individually in a plant.

[875] 76. The stably transformed plant of any one of embodiments 66-74, wherein the fusion protein accumulates in the plant at least 2-fold higher than κ -casein expressed without β -lactoglobulin.

[876] 77. A plant-expressed recombinant fusion protein comprising: a casein protein and β -lactoglobulin.

[877] 78. The plant-expressed recombinant fusion protein of embodiment 77, wherein the casein protein is α -S1 casein, α -S2 casein, β -casein, or κ -casein.

[878] 79. A stably transformed plant, comprising in its genome: a recombinant DNA construct encoding a fusion protein, the fusion protein comprising: a casein protein and β -lactoglobulin; wherein the fusion protein is stably expressed in the plant in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant.

[879] 80. The stably transformed plant of embodiment 79, wherein the casein protein is α -S1 casein, α -S2 casein, β -casein, or κ -casein.

[880] 81. A method for stably expressing a recombinant fusion protein in a plant, the method comprising: (a) transforming a plant with a plant transformation vector comprising an expression

cassette comprising: a sequence encoding a fusion protein, wherein the fusion protein comprises a first protein and a second protein, wherein at least one of the first protein and the second protein is a milk protein; and (b) growing the transformed plant under conditions wherein the recombinant fusion protein is expressed in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant.

[881] 82. The method of embodiment 81, wherein the wherein the milk protein is α -S1 casein, α -S2 casein, β -casein, κ -casein, para- κ -casein, β -lactoglobulin, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, or an immunoglobulin.

Embodiment Set 11: Casein multimers

[882] 1. A fusion protein comprising a first, second, third, and fourth protein, wherein the third protein is kappa-casein.

[883] 2. The fusion protein of embodiment 1 wherein: the first protein beta-casein; the second protein is beta-casein; and the fourth protein is beta-lactoglobulin.

[884] 3. The fusion protein of embodiment 1 or 2, wherein the kappa-casein comprises a chymosin cleavage site.

[885] 4. The fusion protein of embodiment 1, wherein cleavage of the fusion protein with chymosin produces the following polypeptides: a first polypeptide comprising the first protein, the second protein, and para-kappa-casein; and a second polypeptide comprising a kappa-casein macropeptide and the fourth protein.

[886] 5. A nucleic acid encoding the fusion protein of any one of embodiments 1-4.

[887] 6. A transformed plant comprising the fusion protein of any one of embodiments 1-4 or the nucleic acid of embodiment 5.

[888] 7. A food composition comprising the fusion protein of any one of embodiments 1-6.

[889] 8. A food composition comprising a first, second, third or fourth protein, wherein the first, second, third, our fourth protein is derived from the fusion protein of any one of embodiments 1-7.

[890] 9. A method of making a food composition, the method comprising: (i) expressing a fusion protein in a transformed plant; (ii) preparing a food composition comprising the fusion protein and plant protein from the same transformed plant in which the fusion protein was produced.

[891] 10. The method of embodiment 9, wherein the transformed plant is soybean.

[892] 11. A food composition produced using the method of any one of embodiments 9-10.

Embodiment Set No. 12: Fusion protein comprising milk protein and a fusion partner

[893] 1. A fusion protein comprising a first protein and a second protein, wherein the first protein is a milk protein, and the second protein comprises at least one of the following characteristics: a molecular weight of 15 kDa or higher; at least 30% hydrophobic amino acids; and/or less than about 2.5 disulfide bonds per 10 kDa molecular weight.

[894] 2. The fusion protein of embodiment 1, wherein the second protein comprises at least two of the characteristics (i), (ii) and (iii).

[895] 3. The fusion protein of embodiment 1, wherein the second protein comprises all three of the characteristics (i), (ii) and (iii).

[896] 4. The fusion protein of any one of embodiments 1-3, wherein the fusion protein comprises a protease cleavage site located between the first protein and the second protein.

[897] 5. The fusion protein of embodiment 4, wherein the protease cleavage site is a chymosin cleavage site.

[898] 6. The fusion protein of embodiment 4 or 5, wherein cleavage of the fusion protein with a protease separates the first protein from the second protein.

[899] 7. The fusion protein of embodiment 6, wherein after being separated from one another, the first protein and/or the second protein optionally comprise at their N-terminus or C-terminus one or more amino acids that do not occur in the native form of the first protein or the second protein and that are derived from the protease cleavage site.

[900] 8. A nucleic acid encoding the fusion protein of any one of embodiments 1-7.

[901] 9. A transformed plant comprising the fusion protein of any one of embodiments 1-7 or the nucleic acid of embodiment 8.

[902] 10. A food composition comprising the fusion protein of any one of embodiments 1-7.

Embodiment Set 13: Co-Expression of a Milk Protein and a Protein Capable of Forming a Protein Body

[903] 1. A composition comprising a first vector and a second vector, wherein the first vector comprises a sequence that encodes a milk protein, and the second vector comprises a sequence that encodes a prolamin.

[904] 2. A composition comprising a vector, wherein the vector comprises: a first sequence that encodes a milk protein; and a second sequence that encodes a prolamin.

[905] 3. The composition of any one of embodiments 1-2, wherein the milk protein is selected from the group consisting of: α -S1 casein, α -S2 casein, β -casein, κ -casein, para- κ -casein, β -lactoglobulin, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, and an immunoglobulin.

[906] 4. The composition of embodiment 3, wherein the milk protein is the β -casein.

[907] 5. The composition of embodiment 3, wherein the milk protein is the β -lactoglobulin.

[908] 6. The composition of any one of embodiments 1-5, wherein the prolamin is selected from the group consisting of: gliadin, a hordein, a secalin, a zein, a kafirin, and an avenin.

[909] 7. The composition of embodiment 6, wherein the prolamin is a zein.

[910] 8. A plant comprising the composition of any one of embodiments 1-7.

[911] 9. A method for stably expressing one or more recombinant proteins in a plant, the method comprising transforming a plant with the composition of any one of embodiments 1-7, thereby stably expressing one or more recombinant proteins in the plant.

[912] 10. The method of embodiment 9, wherein the method is effective in: (a) increasing expression of the one or more recombinant proteins in the plant, relative to expression of the milk protein alone, without co-expression of the prolamin; (b) increasing accumulation of the milk

protein in the plant, relative to expression of the milk protein alone, without co-expression of the prolamin; or (c) (a) and (b).

[913] 11. The method of embodiment 9 or 10, comprising (a), wherein the method is effective in increasing expression of the milk protein by at least about 1-fold, 5-fold, 50-fold, or 100-fold.

[914] 12. The method of embodiment 9 or 10, comprising (b), wherein the method is effective in increasing accumulation of the milk protein in the plant by at least about 1-fold, 5-fold, 10-fold, or 50-fold.

[915] 13. A food composition that comprises a recombinant protein isolated from the plant of any one of embodiments 9-12.

Embodiment Set 14: Fusion Protein Comprising a Milk Protein and a Protein Capable of Forming a Protein Body

[916] 1. A recombinant fusion protein comprising a prolamin protein and a milk protein.

[917] 2. The recombinant fusion protein of embodiment 1, wherein the milk protein is a casein protein.

[918] 3. The recombinant fusion protein of embodiment 2, wherein the casein protein is α -S1 casein, α -S2 casein, β -casein, κ -casein, or para- κ -casein.

[919] 4. The recombinant fusion protein of embodiment 1, wherein the milk protein is β -lactoglobulin, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, or an immunoglobulin.

[920] 5. The recombinant fusion protein of any one of embodiments 1-4, wherein the prolamin protein is a gliadin, a hordein, a secalin, a zein, a kafirin, or an avenin.

[921] 6. The recombinant fusion protein of embodiment 5, wherein the prolamin protein is a zein.

[922] 7. The recombinant fusion protein of embodiment 6, wherein the zein has the sequence of any one of SEQ ID NO: 800, 809 or 811, or a sequence at least 90% identical thereto.

[923] 8. The recombinant fusion protein of embodiment 1, wherein the prolamin protein is a canein.

- [924] 9. The recombinant fusion protein of embodiment 8, wherein the canein has the sequence of any one of SEQ ID NO: 800, 809 or 811, or a sequence at least 90% identical thereto.
- [925] 10. The recombinant fusion protein of embodiment 1, wherein the fusion protein has a sequence of SEQ ID NO: 803 or 807, or a sequence at least 90% identical thereto.
- [926] 11. A nucleic acid encoding the recombinant fusion protein of any one of embodiments 1-10.
- [927] 12. A transgenic plant comprising the recombinant fusion protein of any one of embodiments 1-10 or the nucleic acid of embodiment 11.
- [928] 13. The transgenic plant of embodiment 12, wherein the plant is a dicot.
- [929] 14. The transgenic plant of embodiment 13, wherein the dicot is arabidopsis, tobacco, tomato, potato, sweet potato, cassava, alfalfa, lima bean, pea, chick pea, soybean, carrot, strawberry, lettuce, oak, maple, walnut, rose, mint, squash, daisy, or cactus.
- [930] 15. The transgenic plant of embodiment 13, wherein dicot is a soybean.
- [931] 16. A food composition comprising the recombinant fusion protein of any one of embodiments 1-10, or a prolamin protein or milk protein derived therefrom.
- [932] 17. A protein body comprising a recombinant fusion protein of any one of embodiments 1-10.
- [933] 18. The protein body of embodiment 17, wherein a transgenic plant comprises the protein body.
- [934] 19. The protein body of embodiment 18, wherein the transgenic plant is a dicot.
- [935] 20. The protein body of embodiment 19, wherein the dicot is a soybean.

Embodiment Set 15: Fusion Protein Comprising an Unstructured Milk Protein and a Structured Protein; Transgenic Plants Expressing the Same

- [936] 1. A stably transformed plant comprising in its genome: a recombinant DNA construct encoding a fusion protein, the fusion protein comprising: (i) an unstructured milk protein, and (ii)

a structured animal protein; wherein the fusion protein is stably expressed in the plant in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant.

[937] 2. The stably transformed plant of embodiment 1, wherein the fusion protein comprises, from N-terminus to C-terminus, the unstructured milk protein and the animal protein.

[938] 3. The stably transformed plant of any one of embodiments 1-2, wherein the unstructured milk protein is α -S1 casein, α -S2 casein, β -casein, or κ -casein.

[939] 4. The stably transformed plant of embodiment 1, wherein the unstructured milk protein is κ -casein and comprises the sequence of SEQ ID NO: 4, or a sequence at least 90% identical thereto.

[940] 5. The stably transformed plant of embodiment 1, wherein the unstructured milk protein is para- κ -casein and comprises the sequence of SEQ ID NO: 2, or a sequence at least 90% identical thereto.

[941] 6. The stably transformed plant of embodiment 1, wherein the unstructured milk protein is β -casein and comprises the sequence of SEQ ID NO: 6, or a sequence at least 90% identical thereto.

[942] 7. The stably transformed plant of embodiment 1, wherein the unstructured milk protein is α -S1 casein and comprises the sequence SEQ ID NO: 8, or a sequence at least 90% identical thereto.

[943] 8. The stably transformed plant of embodiment 1, wherein the unstructured milk protein is α -S2 casein and comprises the sequence SEQ ID NO: 84, or a sequence at least 90% identical thereto.

[944] 9. The stably transformed plant of any one of embodiments 1-8, wherein the structured animal protein is a structured mammalian protein.

[945] 10. The stably transformed plant of embodiment 9, wherein the structured mammalian protein is β -lactoglobulin, α -lactalbumin, albumin, lysozyme, lactoferrin, lactoperoxidase, hemoglobin, collagen, or an immunoglobulin.

[946] 11. The stably transformed plant of embodiment 9, wherein the structured mammalian protein is β -lactoglobulin and comprises the sequence of SEQ ID NO: 10, or a sequence at least 90% identical thereto.

- [947] 12. The stably transformed plant of any one of embodiments 1-8, wherein the structured animal protein is a structured avian protein.
- [948] 13. The stably transformed plant embodiment 12, wherein the structured avian protein is ovalbumin, ovotransferrin, lysozyme or ovoglobulin.
- [949] 14. The stably transformed plant of embodiment 9, wherein the milk protein is κ -casein and the structured mammalian protein is β -lactoglobulin.
- [950] 15. The stably transformed plant of embodiment 9, wherein the milk protein is para- κ -casein and the structured mammalian protein is β -lactoglobulin.
- [951] 16. The stably transformed plant of embodiment 9, wherein the milk protein is β -casein and the structured mammalian protein is β -lactoglobulin.
- [952] 17. The stably transformed plant of embodiment 9, wherein the milk protein is α -S1 casein or α -S2 casein and the structured mammalian protein is β -lactoglobulin.
- [953] 18. The stably transformed plant of any one of embodiments 1-17, wherein the plant is a dicot.
- [954] 19. The stably transformed plant of embodiment 18, wherein the dicot is Arabidopsis, tobacco, tomato, potato, sweet potato, cassava, alfalfa, lima bean, pea, chick pea, soybean, carrot, strawberry, lettuce, oak, maple, walnut, rose, mint, squash, daisy, Quinoa, buckwheat, mung bean, cow pea, lentil, lupin, peanut, fava bean, French beans (i.e., common beans), mustard, or cactus.
- [955] 20. The stably transformed plant of any one of embodiments 1-19, wherein the plant is soybean.
- [956] 21. The stably transformed plant of any one of embodiments 1-20, wherein the recombinant DNA construct is codon-optimized for expression in the plant.
- [957] 22. The stably transformed plant of any one of embodiments 1-21, wherein the fusion protein comprises a protease cleavage site.
- [958] 23. The stably transformed plant of embodiment 22, wherein the protease cleavage site is a chymosin cleavage site.

[959] 24. The stably transformed plant of any one of embodiments 1-23, wherein the fusion protein is expressed at a level at least 2-fold higher than an unstructured milk protein expressed individually in a plant.

[960] 25. The stably transformed plant of any one of embodiments 1-24, wherein the fusion protein accumulates in the plant at least 2-fold higher than an unstructured milk protein expressed without the structured animal protein.

[961] 26. A recombinant fusion protein comprising: (i) an unstructured milk protein, and (ii) a structured animal protein.

[962] 27. The recombinant fusion protein of embodiment 26, wherein the fusion protein is expressed in a plant.

[963] 28. The recombinant fusion protein of embodiment 26 or 27, wherein the unstructured milk protein is α -S1 casein, α -S2 casein, β -casein, or κ -casein.

[964] 29. The recombinant fusion protein of embodiment 28, wherein the milk protein is κ -casein and comprises the sequence of SEQ ID NO: 4, or a sequence at least 90% identical thereto.

[965] 30. The recombinant fusion protein of embodiment 28, wherein the milk protein is para- κ -casein and comprises the sequence of SEQ ID NO: 2, or a sequence at least 90% identical thereto.

[966] 31. The recombinant fusion protein of embodiment 28, wherein the milk protein is β -casein and comprises the sequence of SEQ ID NO: 6, or a sequence at least 90% identical thereto.

[967] 32. The recombinant fusion protein of embodiment 28, wherein the milk protein is α -S1 casein and comprises the sequence SEQ ID NO: 8, or a sequence at least 90% identical thereto.

[968] 33. The recombinant fusion protein of embodiment 28, wherein the milk protein is α -S2 casein and comprises the sequence SEQ ID NO: 84, or a sequence at least 90% identical thereto.

[969] 34. The recombinant fusion protein of any one of embodiments 26-33, wherein the structured animal protein is a structured mammalian protein.

[970] 35. The recombinant fusion protein of embodiment 34, wherein the structured mammalian protein is β -lactoglobulin, α -lactalbumin, albumin, lysozyme, lactoferrin, lactoperoxidase, hemoglobin, collagen, or an immunoglobulin.

[971] 36. The recombinant fusion protein of embodiment 34, wherein the structured mammalian protein is β -lactoglobulin and comprises the sequence of SEQ ID NO: 10, or a sequence at least 90% identical thereto.

[972] 37. The recombinant fusion protein of any one of embodiments 26-33, wherein the structured animal protein is a structured avian protein.

[973] 38. The recombinant fusion protein of embodiment 37, wherein the structured avian protein is ovalbumin, ovotransferrin, lysozyme or ovoglobulin.

[974] 39. The recombinant fusion protein embodiment 34, wherein the milk protein is κ -casein and the structured mammalian protein is β -lactoglobulin.

[975] 40. The recombinant fusion protein of embodiment 34, wherein the milk protein is para- κ -casein and the structured mammalian protein is β -lactoglobulin.

[976] 41. The recombinant fusion protein of embodiment 34, wherein the milk protein is β -casein and the structured mammalian protein is β -lactoglobulin.

[977] 42. The recombinant fusion protein of embodiment 34, wherein the milk protein is α -S1 casein or α -S2 casein and the structured mammalian protein is β -lactoglobulin.

[978] 43. The recombinant fusion protein of embodiment 34, wherein the fusion protein comprises a protease cleavage site.

[979] 44. The recombinant fusion protein of embodiment 34, wherein the protease cleavage site is a chymosin cleavage site.

[980] 45. A nucleic acid encoding the recombinant fusion protein of any one of embodiments 26 to 44.

[981] 46. The nucleic acid of embodiment 45, wherein the nucleic acid is codon optimized for expression in a plant species.

[982] 47. The nucleic acid of embodiment 45 or 46, wherein the nucleic acid is codon optimized for expression in soybean.

- [983] 48. A vector comprising a nucleic acid encoding a recombinant fusion protein, wherein the recombinant fusion protein comprises: (i) an unstructured milk protein, and (ii) a structured animal protein.
- [984] 49. The vector of embodiment 48, wherein the vector is a plasmid.
- [985] 50. The vector of embodiment 49, wherein the vector is an *Agrobacterium* Ti plasmid.
- [986] 51. The vector of any one of embodiments 48-50, wherein the nucleic acid comprises, in order from 5' to 3': a promoter; a 5' untranslated region; a sequence encoding the fusion protein; and a terminator.
- [987] 52. The vector of embodiment 51, wherein the promoter is a seed-specific promoter.
- [988] 53. The vector of embodiment 52, wherein the seed-specific promoter is selected from the group consisting of *PvPhas*, *BnNap*, *AtOle1*, *GmSeed2*, *GmSeed3*, *GmSeed5*, *GmSeed6*, *GmSeed7*, *GmSeed8*, *GmSeed10*, *GmSeed11*, *GmSeed12*, *pBCON*, *GmCEP1-L*, *GmTHIC*, *GmBg7S1*, *GmGRD*, *GmOLEA*, *GmOLER*, *Gm2S-1*, and *GmBBld-II*.
- [989] 54. The vector of embodiment 53, wherein the seed-specific promoter is *PvPhas* and comprises the sequence of SEQ ID NO: 18, or a sequence at least 90% identical thereto.
- [990] 55. The vector of embodiment 53, wherein the seed-specific promoter is *GmSeed2* and comprises the sequence of SEQ ID NO: 19, or a sequence at least 90% identical thereto.
- [991] 56. The vector of any one of embodiments 51-55, wherein the 5' untranslated region is selected from the group consisting of *Arc5'UTR* and *glnBIUTR*.
- [992] 57. The vector of embodiment 56, wherein the 5' untranslated region is *Arc5'UTR* and comprises the sequence of SEQ ID NO: 20, or a sequence at least 90% identical thereto.
- [993] 58. The vector of any one of embodiments 51-57, wherein the expression cassette comprises a 3' untranslated region.
- [994] 59. The vector of embodiment 58, wherein the 3' untranslated region is *Arc5-1* and comprises SEQ ID NO: 21, or a sequence at least 90% identical thereto.
- [995] 60. The vector of any one of embodiments 51-59, wherein the terminator sequence is a terminator isolated or derived from a gene encoding Nopaline synthase, *Arc5-1*, an Extensin, *Rb7*

matrix attachment region, a Heat shock protein, Ubiquitin 10, Ubiquitin 3, and M6 matrix attachment region.

[996] 61. The vector of embodiment 60, wherein the terminator sequence is isolated or derived from a *Nopaline synthase* gene and comprises the sequence of SEQ ID NO: 22, or a sequence at least 90% identical thereto.

[997] 62. A plant comprising the recombinant fusion protein of any one of embodiments 26-44 or the nucleic acid of any one of embodiments 45-47.

[998] 63. A method for stably expressing a recombinant fusion protein in a plant, the method comprising: a) transforming a plant with a plant transformation vector comprising an expression cassette comprising: a sequence encoding a fusion protein, wherein the fusion protein comprises an unstructured milk protein, and a structured animal protein; and b) growing the transformed plant under conditions wherein the recombinant fusion protein is expressed in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant.

[999] 64. The method of embodiment 63, wherein the unstructured milk protein is κ -casein.

[1000] 65. The method of embodiment 63 or 64, wherein the structured animal protein is β -lactoglobulin.

[1001] 66. A food composition comprising the recombinant fusion protein of any one of embodiments 26-44.

[1002] 67. A method for making a food composition, the method comprising: expressing the recombinant fusion protein of any one of embodiments 26-44 in a plant; extracting the recombinant fusion protein from the plant; optionally, separating the milk protein from the structured animal protein or the structured plant protein; and creating a food composition using the milk protein or the fusion protein.

[1003] 68. The method of embodiment 67, wherein the plant stably expresses the recombinant fusion protein.

[1004] 69. The method of embodiment 68, wherein the plant expresses the recombinant fusion protein in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant.

[1005] 70. The method of any one of embodiments 67-69, wherein the plant is soybean.

[1006] 71. The method of any one of embodiments 67-70, wherein the food composition comprises the structured animal or plant protein.

[1007] 72. The method of any one of embodiments 67-71, wherein the milk protein and the structured animal or plant protein are separated from one another in the plant cell, prior to extraction.

[1008] 73. The method of any one of embodiments 67-71, wherein the milk protein is separated from the structured animal or plant protein after extraction, by contacting the fusion protein with an enzyme that cleaves the fusion protein.

[1009] 74. A food composition produced using the method of any one of embodiments 67-73.

Embodiment Set Number 16: Modulation of Post-Translational Modifications by Modifying the Amino Acid Sequence of a Milk Protein

[1010] 1. A recombinant milk protein, wherein the amino acid sequence of the milk protein is modified to promote addition of one or more post-translational modifications in a plant cell.

[1011] 2. The recombinant milk protein of embodiment 1, wherein the milk protein is expressed in a plant, and wherein the milk protein comprises one or more post-translational modifications that are not present in a non-modified milk protein expressed in the same type of plant.

[1012] 3. The recombinant milk protein of embodiment 1, wherein the milk protein is expressed in a plant in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant.

[1013] 4. The recombinant milk protein of any one of embodiments 1-3, wherein the milk protein is a casein protein selected from α -S1 casein, α -S2 casein, β -casein, κ -casein, and para- κ -casein.

[1014] 5. The recombinant milk protein of any one of embodiments 1-4, wherein the milk protein is κ -casein or para- κ -casein.

[1015] 6. The recombinant milk protein of any one of embodiments 1-4, wherein the milk protein is β -casein.

[1016] 7. The recombinant milk protein of any one of embodiments 1-4, wherein the milk protein is β -lactoglobulin.

[1017] 8. The recombinant milk protein of any one of embodiments 1-7, wherein the one or more post-translational modifications are selected from glycosylation, phosphorylation, lipidation, ubiquitylation, nitrosylation, methylation, acetylation, amidation, prenylation, alkylation, gamma-carboxylation, biotinylation, oxidation, and sulfation.

[1018] 9. A nucleic acid encoding the recombinant milk protein of any one of embodiments 1-8.

Embodiment Set Number 17: Modulation of Post-Translational Modifications (PTMs) by Expressing One or More Enzymes which Add/Remove PTMs

[1019] 1. A method for stably expressing a milk protein in a plant, the method comprising: transforming the plant with a sequence encoding the milk protein and a sequence encoding a kinase.

[1020] 2. The method of embodiment 1, wherein the milk protein is a casein protein selected from the group consisting of: α -S1 casein, α -S2 casein, β -casein, κ -casein, para- κ -casein.

[1021] 3. The method of embodiment 1 or 2, wherein the milk protein is fused to a second protein.

[1022] 4. The method of any one of embodiments 1-3, wherein the kinase is a kinase in the 20C family.

[1023] 5. The method of any one of embodiments 1-3, wherein the kinase that phosphorylates Ser-X-Glu/pSer motifs.

[1024] 6. The method of any one of embodiments 1-3, wherein the kinase is a Fam20C kinase, or a fragment or variant thereof.

[1025] 7. The method of any one of embodiments 1-3, wherein the kinase comprises SEQ ID NO: 821, or a sequence at least 90% or 95% identical thereto.

[1026] 8. The method of any one of embodiments 1-3, wherein the kinase comprises amino acids 94-586 of SEQ ID NO: 821, or a sequence at least 90% or 95% identical thereto.

[1027] 9. The method of any one of embodiments 1-8, wherein the sequence encoding the milk protein and the sequence encoding the kinase are in the same vector.

[1028] 10. The method of embodiment 9, wherein the vector is a binary vector.

Embodiment Set Number 18: Fusion of a Milk Protein to a Glycoprotein Tag

[1029] 1. A method for stably expressing a milk protein in a plant, the method comprising: transforming the plant with a sequence encoding a milk protein fused to a glycoprotein tag.

[1030] 2. The method of embodiment 1, wherein the milk protein is a casein protein selected from the group consisting of: α -S1 casein, α -S2 casein, β -casein, κ -casein, para- κ -casein.

[1031] 3. The method of embodiment 1 or 2, wherein the milk protein is fused to a second protein.

[1032] 4. The method of any one of embodiments 1-3, wherein the glycoprotein tag is isolated or derived from a hydroxyproline (Hyp)-rich glycoprotein (GRGP).

[1033] 5. The method of any one of embodiments 1-3, wherein the glycoprotein tag comprises the M domain of CD45.

[1034] 6. The method of any one of embodiments 1-3, wherein the glycoprotein tag is an (SP)11 tag.

[1035] 7. The method of any one of embodiments 1-3, wherein the glycoprotein tag comprises SEQ ID NO: 825, or a sequence at least 90% or 95% identical thereto.

[1036] 8. The method of any one of embodiments 1-3, wherein the glycoprotein tag comprises SEQ ID NO: 827, or a sequence at least 90% or 95% identical thereto.

[1037] 9. The method of any one of embodiments 1-8, wherein the sequence encoding the milk protein and the sequence encoding the kinase are in the same vector.

[1038] 10. The method of embodiment 9, wherein the vector is a binary vector.

Embodiment Set Number 19: Reducing the Expression of One or More Proteases in a Plant Cell

[1039] 1. A plant cell for expressing recombinant milk proteins, wherein expression of one or more proteases is knocked down or knocked out in the cell.

[1040] 2. The plant cell of embodiment 1, wherein expression of the one or more proteases is knocked down or knocked out using a gene editing technology or base editing technology.

[1041] 3. The plant cell of embodiment 1, wherein expression of the one or more proteases is knocked down or knocked out using RNA interference.

[1042] 4. The plant cell of embodiment 1, wherein the one or more proteases is a cysteine protease, a serine protease, or an aspartyl protease.

[1043] 5. A transgenic plant comprising the plant cell of any one of embodiments 1-4.

[1044] 6. A method for stably expressing a recombinant milk protein in a plant, the method comprising: (i) reducing expression of one or more proteases in the plant, (ii) transforming the plant with a plant transformation vector comprising an expression cassette encoding a recombinant milk protein or the fusion protein comprising the recombinant milk protein, (iii) growing the transformed plant under conditions wherein the recombinant milk protein is expressed in an amount of 1% or higher per total weight of soluble protein extractable from the plant.

Embodiment Set Number 20: Food Composition Comprising a Milk Protein Derived from a Fusion Protein

[1045] 1. A food composition comprising the recombinant milk protein derived from a fusion protein of any one of the embodiment sets above.

[1046] 2. A method for making a food composition, the method comprising: expressing the recombinant fusion protein of any one of the embodiment sets above; extracting the recombinant

fusion protein from the plant; optionally, separating the first protein from the second protein; and creating a food composition using the milk protein or the fusion protein.

[1047] 3. The method of embodiment 2, wherein the plant stably expresses the recombinant fusion protein.

[1048] 4. The method of embodiment 2 or 3, wherein the plant expresses the recombinant fusion protein in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant.

[1049] 5. The method of any one of embodiments 2-4, wherein the plant is soybean.

[1050] 6. The method of any one of embodiments 2-5, wherein the food composition comprises the first protein and the second protein.

[1051] 7. The method of embodiment 6, wherein the first protein and the second protein are separated from one another in the plant cell, prior to extraction.

[1052] 8. The method of embodiment 6, wherein the first protein and the second protein are separated after extraction, by contacting the fusion protein with an enzyme that cleaves the fusion protein.

[1053] 9. A food composition produced using the method of any one of embodiments 2-8.

[1054] 10. A food composition comprising a first or second protein, wherein the first or second protein is derived from the fusion protein of any one of the embodiment sets above.

Embodiment Set Number 21: A Solid-Phase Protein Stabilized-Emulsion Comprising a Recombinant Casein Protein

[1055] 1. A solid phase, protein-stabilized emulsion comprising at least one recombinant casein protein selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein and alpha-S2-casein; wherein the emulsion has at least one of the following characteristics: a firmness of at least 150 grams, as determined by compressing a cylindrical-shaped sample of the emulsion having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C; a melting point of about 35°C to about 100°C; or ability to stretch to at least 3 cm in length without breaking, as determined by

heating a 100 gram mass of the emulsion to a temperature of about 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[1056] 2. The solid phase, protein-stabilized emulsion of embodiment 1, wherein the recombinant casein protein is plant-expressed.

[1057] 3. The solid phase, protein stabilized emulsion of embodiment 1, wherein the recombinant casein protein is yeast-expressed or bacterial-expressed.

[1058] 4. The solid phase, protein stabilized emulsion of any one of embodiments 1-3, wherein the recombinant casein protein is derived from a fusion protein.

[1059] 5. The solid phase, protein stabilized emulsion of embodiment 4, wherein the fusion protein comprises a first and a second protein.

[1060] 6. The solid phase, protein stabilized emulsion of embodiment 5, wherein the first protein comprises β -Casein and the second protein comprises a milk protein.

[1061] 7. The solid phase, protein stabilized emulsion of embodiment 5, wherein the first protein comprises β -Casein and the second protein comprises a non-milk protein.

[1062] 8. The solid phase, protein stabilized emulsion of embodiment 6, wherein the milk protein is selected from the group consisting of β -lactoglobulin, casein, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, and immunoglobulin.

[1063] 9. The solid phase, protein stabilized emulsion of embodiment 8, wherein the milk protein is β -lactoglobulin.

[1064] 10. The solid phase, protein stabilized emulsion of embodiment 8, wherein the milk protein is casein, and wherein the casein is selected from the group consisting of: α -S1 Casein, α -S2 Casein, β -Casein, κ -Casein, and para- κ -Casein.

[1065] 11. The solid phase, protein stabilized emulsion of embodiment 10, wherein the milk protein is β -Casein.

[1066] 12. The solid phase, protein-stabilized emulsion of any one of embodiments 1-11, wherein the emulsion comprises at least one lipid and at least one salt.

[1067] 13. The solid phase, protein-stabilized emulsion of any one of embodiments 1-5, wherein the emulsion comprises at least two plant-expressed casein proteins each selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein and alpha-S2-casein.

[1068] 14. The solid phase, protein-stabilized emulsion of any one of embodiments 1-5, wherein the emulsion comprises at least three plant-expressed casein proteins each selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein and alpha-S2-casein.

[1069] 15. The solid phase, protein-stabilized emulsion of any one of embodiments 1-14, wherein the emulsion comprises at least one additional mammalian or plant protein that is not a casein protein.

[1070] 16. The solid phase, protein-stabilized emulsion of embodiment 2, wherein the plant-expressed casein protein is expressed in a soybean plant.

[1071] 17. The solid phase, protein-stabilized emulsion of any one of embodiments 1-16, wherein the emulsion has a pH of about 5.2 to about 5.9.

[1072] 18. The solid phase, protein-stabilized emulsion of any one of embodiments 1-17, wherein the emulsion does not contain an organoleptically functional amount of beta-lactoglobulin.

[1073] 19. A solid phase, protein-stabilized emulsion comprising one plant-expressed casein protein selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein; wherein the emulsion does not contain any additional casein proteins; wherein the emulsion has at least one of the following characteristics: a firmness of at least 150 grams, as determined by compressing a cylindrical-shaped sample of the emulsion having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C; a melting point of about 35°C to about 100°C; or ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the emulsion to a temperature of about 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[1074] 20. The solid phase, protein-stabilized emulsion of embodiment 19, wherein the emulsion further comprises at least one lipid and at least one salt.

[1075] 21. The solid phase, protein-stabilized emulsion of embodiment 19 or 20, wherein the plant-expressed casein protein is expressed in soybean plant.

[1076] 22. The solid phase, protein-stabilized emulsion of any one of embodiments 19-21, wherein the plant-expressed casein protein is derived from a fusion protein.

[1077] 23. The solid phase, protein stabilized-emulsion of any one of embodiments 19-22, wherein the emulsion has a pH of about 5.2 to about 5.9.

[1078] 24. The solid phase, protein-stabilized emulsion of any one of embodiments 19-23, wherein the emulsion does not contain an organoleptically functional amount of beta-lactoglobulin.

[1079] 25. A solid phase, protein-stabilized emulsion comprising: a plant-expressed casein protein selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein; and plant-expressed beta-lactoglobulin; wherein the ratio of the casein protein to the beta-lactoglobulin is about 8:1 to about 1:2.

[1080] 26. The solid phase, protein-stabilized emulsion of embodiment 25, wherein the emulsion has at least one of the following characteristics: a firmness of at least 150 grams, as determined by compressing a cylindrical-shaped sample of the emulsion having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C; a melting point of about 35°C to about 100°C; or ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the emulsion to a temperature of about 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[1081] 27. The solid phase, protein-stabilized emulsion of embodiment 25 or 26, wherein the emulsion comprises at least at least one additional mammalian or plant protein that is not a casein protein.

[1082] 28. The solid phase, protein-stabilized emulsion of any one of embodiments 25-27, wherein the ratio of the casein protein to the beta-lactoglobulin is about 2:1.

[1083] 29. The solid phase, protein-stabilized emulsion of any one of embodiments 25-28, wherein the emulsion has a pH of about 5.2 to about 5.9.

[1084] 30. The solid phase, protein-stabilized emulsion of any one of embodiments 25-29, wherein the plant-expressed casein protein is derived from a fusion protein.

[1085] 31. A solid-phase protein-stabilized emulsion comprising about 8% (w/v) to about 25% (w/v) total protein, one or more lipids, and one or more salts; wherein at least 4% of the total protein comprises casein proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein; wherein at least 20% to 100% of the casein protein is kappa casein; wherein the emulsion has at least one of the following characteristics: a firmness of at least 150 grams, as determined by compressing a cylindrical-shaped sample of the emulsion having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C; a melting point of about 35°C to about 100°C; or ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the emulsion to a temperature of about 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[1086] 32. The solid-phase protein-stabilized emulsion of embodiment 31, wherein the kappa casein is expressed in a plant.

[1087] 33. The solid phase, protein-stabilized emulsion of any one of embodiments 31-32, wherein the kappa casein is derived from a fusion protein.

[1088] 34. The solid phase, protein-stabilized emulsion of any one of embodiments 31-33, wherein the emulsion has a pH of about 5.2 to about 5.9.

[1089] 35. The solid phase, protein-stabilized emulsion of any one of embodiments 31-34, wherein the composition comprises only one, only two, only three, or only four casein proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein.

[1090] 36. The solid phase, protein-stabilized emulsion of any one of embodiments 31-35, wherein the emulsion does not contain an organoleptically functional amount of beta-lactoglobulin.

[1091] 37. A solid-phase protein-stabilized emulsion comprising about 8% to about 25% total protein, one or more lipids, and one or more salts; wherein at least 4% of the total protein comprises casein proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein; wherein at least 20% to 100% of the casein protein is para-kappa casein; wherein the emulsion has at least one of the following characteristics: a firmness of at least 150 grams, as determined by compressing a cylindrical-shaped sample of the emulsion having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C; a melting point of about 35°C to about 100°C;

or ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the emulsion to a temperature of about 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[1092] 38. The solid-phase protein-stabilized emulsion of embodiment 37, wherein the para-kappa casein is expressed in a plant.

[1093] 39. The solid phase, protein-stabilized emulsion of embodiment 37 or 38, wherein the para-kappa casein is derived from a fusion protein.

[1094] 40. The solid-phase protein-stabilized emulsion of any one of embodiments 37-39 wherein the para-kappa casein is produced without the use of any enzyme that cleaves kappa-casein to para-kappa casein.

[1095] 41. The solid phase, protein-stabilized emulsion of any one of embodiments 37-40, wherein the emulsion has a pH of about 5.2 to about 5.9.

[1096] 42. The solid phase, protein-stabilized emulsion of any one of embodiments 37-41, wherein the composition comprises only one, only two, only three, or only four casein proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein.

[1097] 43. The solid phase, protein-stabilized emulsion of any one of embodiments 37-42, wherein the emulsion does not contain an organoleptically functional amount of beta-lactoglobulin.

[1098] 44. A solid-phase protein-stabilized emulsion comprising about 8% to about 25% total protein, one or more lipids, and one or more salts; wherein at least 4% of the total protein comprises casein proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein; wherein at least 50% to 100% of the casein protein is beta-casein; wherein the emulsion has at least one of the following characteristics: a firmness of at least 150 grams, as determined by compressing a cylindrical-shaped sample of the emulsion having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C; a melting point of about 35°C to about 100°C; or ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the emulsion to a temperature of about 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[1099] 45. The solid-phase protein-stabilized emulsion of embodiment 44, wherein the beta-casein is expressed in a plant.

[1100] 46. The solid phase, protein-stabilized emulsion of any one of embodiments 44-45, wherein the plant-expressed casein protein is derived from a fusion protein.

[1101] 47. The solid phase, protein-stabilized emulsion of any one of embodiments 44-46, wherein the emulsion has a pH of about 5.2 to about 5.9.

[1102] 48. The solid phase, protein-stabilized emulsion of any one of embodiments 44-47, wherein the composition comprises only one, only two, only three, or only four casein proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein.

[1103] 49. The solid phase, protein-stabilized emulsion of any one of embodiments 44-48, wherein the emulsion does not contain an organoleptically functional amount of beta-lactoglobulin.

[1104] 50. A solid-phase protein-stabilized emulsion comprising about 8% to about 25% total protein, one or more lipids, and one or more salts; wherein at least 4% of the total protein comprises casein proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein; wherein at least 50% to 100% of the casein protein is alpha-S1-casein; wherein the emulsion has at least one of the following characteristics: a firmness of at least 150 grams, as determined by compressing a cylindrical-shaped sample of the emulsion having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C; a melting point of about 35°C to about 100°C; or ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the emulsion to a temperature of about 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[1105] 51. The solid-phase protein-stabilized emulsion of embodiment 50, wherein the alpha-S1-casein is expressed in a plant.

[1106] 52. The solid phase, protein-stabilized emulsion of any one of embodiments 50-51, wherein the alpha-S1-casein is derived from a fusion protein.

[1107] 53. The solid phase, protein-stabilized emulsion of any one of embodiments 50-52, wherein the emulsion has a pH of about 5.2 to about 5.9.

[1108] 54. The solid phase, protein-stabilized emulsion of any one of embodiments 50-53, wherein the composition comprises only one, only two, only three, or only four casein proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein.

[1109] 55. The solid phase, protein-stabilized emulsion of any one of embodiments 50-54, wherein the emulsion does not contain an organoleptically functional amount of beta-lactoglobulin.

[1110] 56. A solid-phase protein-stabilized emulsion comprising about 8% to about 25% total protein, one or more lipids, and one or more salts; wherein at least 4% of the total protein comprises casein proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein; wherein at least 20% to 100% of the casein protein is alpha-S2-casein; wherein the emulsion has at least one of the following characteristics: a firmness of at least 150 grams, as determined by compressing a cylindrical-shaped sample of the emulsion having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C; a melting point of about 35°C to about 100°C; or ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the emulsion to a temperature of about 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[1111] 57. The solid-phase protein-stabilized emulsion of embodiment 56, wherein the alpha-S2-casein is expressed in a plant.

[1112] 58. The solid phase, protein-stabilized emulsion of any one of embodiments 56-57, wherein the plant-expressed casein protein is derived from a fusion protein.

[1113] 59. The solid phase, protein-stabilized emulsion of any one of embodiments 56-58, wherein the emulsion has a pH of about 5.2 to about 5.9.

[1114] 60. The solid phase, protein-stabilized emulsion of any one of embodiments 56-59, wherein the composition comprises only one, only two, only three, or only four casein proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein.

[1115] 61. The solid phase, protein-stabilized emulsion of any one of embodiments 56-60, wherein the emulsion does not contain an organoleptically functional amount of beta-lactoglobulin.

Embodiment Set Number 22: Alternative Dairy Compositions Comprising One or more Recombinant Casein Proteins

[1116] 1. An alternative dairy composition comprising one or more recombinant casein proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein; wherein the alternative dairy composition has at least one of the following characteristics: a firmness of at least 150 grams, as determined by compressing a cylindrical-shaped sample of the alternative dairy composition having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C; a melting point of about 35°C to about 100°C; or ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the alternative dairy composition to a temperature of about 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[1117] 2. The alternative dairy composition of embodiment 1, wherein the composition further comprises at least one lipid and at least one salt.

[1118] 3. The alternative dairy composition of embodiment any one of embodiments 1-2, wherein the composition further comprises at least one additional mammalian or plant protein that is not a casein protein.

[1119] 4. The alternative dairy composition of any one of embodiments 1-3, wherein the one or more recombinant casein proteins are expressed in a plant.

[1120] 5. The alternative dairy composition of embodiment 4, wherein the one or more recombinant casein proteins are expressed in a soybean plant.

[1121] 6. The alternative dairy composition of any one of embodiments any one of embodiments 1-5, wherein the one or more recombinant casein proteins are derived from one or more fusion proteins.

[1122] 7. The alternative dairy composition of embodiment 6, wherein one of the one or more fusion proteins comprises a first and a second protein.

[1123] 8. The alternative dairy composition of embodiment 7, wherein the first protein comprises β -Casein and the second protein comprises a milk protein.

[1124] 9. The alternative dairy composition of embodiment 7, wherein the first protein comprises β -Casein and the second protein comprises a non-milk protein.

[1125] 10. The alternative dairy composition of embodiment 8, wherein the milk protein is selected from the group consisting of β -lactoglobulin, casein, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, and immunoglobulin.

[1126] 11. The alternative dairy composition of embodiment 8, wherein the milk protein is β -lactoglobulin.

[1127] 12. The alternative dairy composition of embodiment 8, wherein the milk protein is casein, and wherein the casein is selected from the group consisting of: α -S1 Casein, α -S2 Casein, β -Casein, κ -Casein, and para- κ -Casein.

[1128] 13. The alternative dairy composition of embodiment 8, wherein the milk protein is β -Casein.

[1129] 14. The alternative dairy composition of any one of embodiments 1-13, wherein the composition has a pH of about 5.2 to about 5.9.

[1130] 15. The alternative dairy composition of any one of embodiments 1-9, wherein the composition does not contain an organoleptically functional amount of beta-lactoglobulin.

[1131] 16. An alternative dairy composition comprising one or more recombinant casein proteins, one or more lipids; and one or more salts; wherein the alternative dairy composition does not contain an organoleptically functional amount of beta-lactoglobulin; wherein the alternative dairy composition has at least one of the following characteristics: a firmness of at least 150 grams, as determined by compressing a cylindrical-shaped sample of the alternative dairy composition having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C; a melting point of about 35°C to about 100°C; or ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the alternative dairy composition to a temperature of about 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[1132] 17. The alternative dairy composition of embodiment 16, wherein the composition comprises at least one additional mammalian or plant protein that is not a casein protein.

[1133] 18. The alternative dairy composition of embodiment 16 or 17, wherein the one or more recombinant casein proteins are expressed in a plant.

[1134] 19. The alternative dairy composition of clam 18, wherein the one or more recombinant casein proteins are expressed in a soybean plant.

[1135] 20. The alternative dairy composition of any one of embodiments 16-19, wherein the one or more recombinant casein proteins are derived from one or more fusion proteins.

[1136] 21. The alternative dairy composition of any one of embodiments 16-20, wherein the composition has a pH of about 5.2 to about 5.9.

[1137] 22. An alternative dairy composition comprising: a recombinant casein protein selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein; and a recombinant beta-lactoglobulin; wherein the ratio of the casein protein to the beta-lactoglobulin is about 8:1 to about 1:2; wherein the alternative dairy composition has at least one of the following characteristics: a firmness of at least 150 grams, as determined by compressing a cylindrical-shaped sample of the alternative dairy composition having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C; a melting point of about 35°C to about 100°C; or ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the alternative dairy composition to a temperature of about 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[1138] 23. The alternative dairy composition of embodiment 22, wherein the composition comprises at least one additional mammalian or plant protein that is not a casein protein.

[1139] 24. The alternative dairy composition of embodiment 22 or 23, wherein recombinant casein protein is expressed in a plant.

[1140] 25. The alternative dairy composition of clam 24, wherein recombinant casein protein is expressed in a soybean plant.

[1141] 26. The alternative dairy composition of any one of embodiments 22-25, wherein the recombinant casein protein is derived from a fusion protein.

[1142] 27. The alternative dairy composition of any one of embodiments 22-26, wherein the composition has a pH of about 5.2 to about 5.9.

[1143] 28. An alternative dairy composition comprising kappa-casein and essentially no para-kappa casein, wherein the alternative dairy composition has at least one of the following characteristics: a firmness of at least 150 grams, as determined by compressing a cylindrical-shaped sample of the alternative dairy composition having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C; a melting point of about 35°C to about 100°C; or ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the alternative dairy composition to a temperature of about 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[1144] 29. The alternative dairy composition of embodiment 28, wherein the composition comprises at least one additional mammalian or plant protein that is not a casein protein.

[1145] 30. The alternative dairy composition of embodiment 28 or 29, wherein the kappa casein is recombinant.

[1146] 31. The alternative dairy composition of any one of embodiments 28-30, wherein the kappa casein is expressed in a plant.

[1147] 32. The alternative dairy composition of embodiment 31, wherein the kappa casein is expressed in a soybean plant.

[1148] 33. The alternative dairy composition of any one of embodiments 28-32, wherein the kappa casein is derived from a fusion protein.

[1149] 34. The alternative dairy composition of any one of embodiments 28-33, wherein the composition has a pH of about 5.2 to about 5.9.

[1150] 35. The alternative dairy composition of any one of embodiments 28-33, wherein the composition does not contain an organoleptically functional amount of beta-lactoglobulin.

[1151] 36. An alternative dairy composition comprising one to four of the milk proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein; wherein the alternative dairy composition has at least one of the following characteristics: a firmness of at

least 150 grams, as determined by compressing a cylindrical-shaped sample of the alternative dairy composition having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C; a melting point of about 35°C to about 100°C; or ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the alternative dairy composition to a temperature of about 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[1152] 37. The alternative dairy composition of embodiment 36, wherein at least one milk protein is recombinant.

[1153] 38. The alternative dairy composition of embodiment 36, wherein the at least one milk protein is plant-expressed.

[1154] 39. The alternative dairy composition of embodiment 38, wherein the at least one milk protein is expressed in a soybean plant.

[1155] 40. The alternative dairy composition of embodiment 37, wherein the at least one milk protein is yeast- or bacterial-expressed.

[1156] 41. The alternative dairy composition of any one of embodiments 36-40, wherein at least one milk protein is derived from a fusion protein.

[1157] 42. The alternative dairy composition of any one of embodiments 36-41, wherein the alternative dairy composition comprises one of the milk proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein.

[1158] 43. The alternative dairy composition of any one of embodiments 36-41, wherein the alternative dairy composition comprises two of the milk proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein.

[1159] 44. The alternative dairy composition of any one of embodiments 36-41, wherein the alternative dairy composition comprises three of the milk proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein.

[1160] 45. The alternative dairy composition of any one of embodiments 36-41, wherein the alternative dairy composition comprises four of the milk proteins selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein.

[1161] 46. The alternative dairy composition of any one of embodiments 36-45, wherein the composition has a pH of about 5.2 to about 5.9.

[1162] 47. The alternative dairy composition of any one of embodiments 36-46, wherein the composition does not contain an organoleptically functional amount of beta-lactoglobulin.

[1163] 48. An alternative dairy composition comprising 2 to 4 casein proteins; wherein the alternative dairy composition has at least one of the following characteristics: a firmness of at least 150 grams, as determined by compressing a cylindrical-shaped sample of the alternative dairy composition having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C; a melting point of about 35°C to about 100°C; or ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the alternative dairy composition to a temperature of about 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[1164] 49. The alternative dairy composition of embodiment 48, wherein the alternative dairy composition does not contain an organoleptically functional amount of beta-lactoglobulin.

[1165] 50. The alternative dairy composition of embodiment 48 or 49, wherein the casein proteins are selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein.

[1166] 51. The alternative dairy composition of any one of embodiments 48-50, wherein the composition comprises at least one lipid and at least one salt.

[1167] 52. The alternative dairy composition of any one of embodiments 48-51, wherein the composition has a pH of about 5.2 to about 5.9.

[1168] 53. The alternative dairy composition of any one of embodiments 48-52, wherein at least one of the casein proteins is derived from a fusion protein.

[1169] 54. An alternative dairy composition comprising one to four plant-expressed recombinant milk proteins, wherein the alternative dairy composition comprises three or more organoleptic properties similar to a dairy composition selected from the group consisting of taste, appearance, mouthfeel, structure, texture, density, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess, and emulsification.

[1170] 55. The alternative dairy composition of embodiment 54, wherein the plant-expressed milk proteins are selected from beta lactoglobulin, kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein.

[1171] 56. The alternative dairy composition of embodiment 54 or 55, wherein the composition is a milk composition.

[1172] 57. The alternative dairy composition of embodiment 54 or 55, wherein the composition is a cream composition.

[1173] 58. The alternative dairy composition of embodiment 54 or 55, wherein the composition is a yogurt composition.

[1174] 59. The alternative dairy composition of embodiment 54 or 55, wherein the composition is an ice cream composition.

[1175] 60. The alternative dairy composition of embodiment 54 or 55, wherein the composition is a frozen custard composition.

[1176] 61. The alternative dairy composition of embodiment 54 or 55, wherein the composition is a frozen desert composition.

[1177] 62. The alternative dairy composition of embodiment 54 or 55, wherein the composition is a crème fraîche composition.

[1178] 63. The alternative dairy composition of embodiment 54 or 55, wherein the composition is a curd composition.

[1179] 64. The alternative dairy composition of embodiment 54 or 55, wherein the composition is a cottage cheese composition.

[1180] 65. The alternative dairy composition of embodiment 54 or 55, wherein the composition is a cream cheese composition.

[1181] 66. The alternative dairy composition of any one of embodiments 54-65, wherein at least one of the plant-expressed recombinant milk proteins is derived from a fusion protein.

[1182] 67. An alternative dairy food composition comprising: a recombinant beta-casein protein, and least one lipid, wherein the alternative dairy food composition does not comprise an organoleptically functional amount of beta-lactoglobulin.

[1183] 68. The alternative dairy food composition of embodiment 67, wherein the recombinant beta-casein protein confers on the alternative dairy food composition one or more characteristics of a dairy food product selected from the group consisting of: taste, aroma, appearance, handling, mouthfeel, density, structure, texture, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess and emulsification.

[1184] 69. The alternative dairy food composition of embodiment 67 or 68, wherein the composition does not comprise any additional casein proteins.

[1185] 70. The alternative dairy food composition of embodiment 67 or 68, wherein the composition comprises at least one additional casein protein.

[1186] 71. The alternative dairy food composition of embodiment 70, wherein at least 50% by weight of the total casein protein in the composition is beta-casein.

[1187] 72. The alternative dairy food composition of embodiment 70, wherein at least 75% by weight of the total casein protein in the composition is beta-casein.

[1188] 73. The alternative dairy food composition of embodiment 70, wherein at least 90% by weight of the total casein protein in the composition is beta-casein.

[1189] 74. The alternative dairy food composition of any one of embodiments 70-73, wherein the at least one additional casein protein is selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein and alpha-S2-casein.

[1190] 75. The alternative dairy food composition of any one embodiment 70-73, wherein the at least one additional casein protein is kappa-casein or para-kappa casein.

[1191] 76. The alternative dairy food composition of any one of embodiments 67-75, wherein the recombinant beta-casein is plant-expressed.

[1192] 77. The alternative dairy food composition of embodiment 76, wherein the recombinant beta-casein is expressed in a soybean.

[1193] 78. The alternative dairy food composition of any one of embodiments 70-77, wherein all caseins in the composition are plant-expressed.

[1194] 79. The alternative dairy food composition of any one of embodiments 67-78, wherein the composition comprises a fusion protein comprising the recombinant beta-casein.

[1195] 80. The alternative dairy food composition any one of embodiments 67-79, wherein the recombinant beta-casein protein confers on the alternative dairy food composition two or more characteristics of a dairy food product selected from the group consisting of: taste, aroma, appearance, handling, mouthfeel, density, structure, texture, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess and emulsification.

[1196] 81. The alternative dairy food composition of any one of embodiments 67-80, wherein the composition is a milk composition, a cream composition, a yogurt composition, an ice cream composition, a frozen custard composition, a frozen dessert composition, a crème fraiche composition, a curd composition, a cottage cheese composition, or a cream cheese composition.

[1197] 82. The alternative dairy food composition of any one of embodiments 67-81, wherein the composition comprises at least one lipid and at least one salt.

[1198] 83. The alternative dairy food composition any one of embodiments 67-82, wherein the composition comprises calcium.

[1199] 84. The alternative dairy food composition of embodiment 83, wherein the composition comprises calcium at a concentration of about 0.1% to about 2% by weight.

[1200] 85. The alternative dairy food composition any one of embodiments 67-84, wherein the composition has a pH of about 4 to about 8.

Embodiment Set Number 23: Colloidal Suspensions Comprising One or More Recombinant Casein Proteins

[1201] 1. A colloidal suspension comprising: one to four plant-expressed recombinant milk proteins, wherein the recombinant milk proteins comprise between 0.5% (w/v) to 15% (w/v) of the composition; and ash; wherein the colloidal suspension has at least one, at least two, or at least three characteristics that are substantially similar to bovine milk selected from taste, appearance, mouthfeel, structure, texture, density, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess, and emulsification.

[1202] 2. The colloidal suspension of embodiment 1, wherein the plant-expressed milk proteins are selected from beta-lactoglobulin, kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein.

[1203] 3. The colloidal suspension of any one of embodiments 1-2, wherein at least one of the plant-expressed recombinant milk proteins is derived from a fusion protein.

[1204] 4. A colloidal suspension comprising: one casein protein, wherein the casein protein comprises between 0.5% (w/v) to 15% (w/v); and ash; wherein the colloidal suspension has at least one, at least two, or at least three characteristics that are substantially similar to bovine milk selected from taste, appearance, mouthfeel, structure, texture, density, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess, and emulsification.

[1205] 5. The colloidal suspension of embodiment 4, wherein the casein protein is selected from beta lactoglobulin, kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein, and alpha-S2-casein.

[1206] 6. The colloidal suspension of embodiment 4 or 5, wherein the casein protein is beta-casein.

[1207] 7. The colloidal suspension of any one of embodiments 4-6, wherein the casein protein is plant-expressed.

[1208] 8. The colloidal suspension of any one of embodiments 4-7, wherein the casein protein is derived from a fusion protein.

[1209] 9. A method of making an alternative dairy composition comprising processing the colloidal suspension of any one of embodiments 1-8.

[1210] 10. An alternative dairy composition produced from the method of embodiment 9.

[1211] 11. The alternative dairy composition of embodiment 10, wherein the alternative dairy composition is a cream composition, a yogurt composition, a cheese composition, an ice cream composition, a frozen custard composition, a frozen desert composition, a crème fraîche composition, a curd composition, a cottage cheese composition, or a cream cheese composition.

[1212] 12. A colloidal suspension comprising: recombinant beta-casein protein, and at least one lipid; wherein the suspension does not contain an organoleptically functional amount of beta-lactoglobulin.

[1213] 13. The colloidal suspension of embodiment 12, wherein the suspension is a non-Newtonian fluid.

[1214] 14. The colloidal suspension of embodiment 12 or 13, which is characterized as a shear thinning fluid with an apparent viscosity greater than 10 centipoise, at a shear rate of 1 sec^{-1} .

[1215] 15. The colloidal suspension of any one of embodiments 12-14, wherein the suspension is an aqueous suspension.

[1216] 16. The colloidal suspension of any one of embodiments 12-15, wherein the suspension does not comprise any additional casein proteins.

[1217] 17. The colloidal suspension of any one of embodiments 12-15, wherein the composition comprises at least one additional casein protein.

[1218] 18. The colloidal suspension of embodiment 17, wherein at least 80% by weight of the total casein protein in the composition is beta-casein.

[1219] 19. The colloidal suspension of embodiment 17 or 18, wherein the at least one additional casein protein is selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein and alpha-S2-casein.

[1220] 20. The colloidal suspension of embodiment 17 or 18, wherein the at least one additional casein protein is kappa-casein or para-kappa casein.

[1221] 21. The colloidal suspension of any one of embodiments 12-20, wherein the recombinant beta-casein is plant-expressed.

[1222] 22. The colloidal suspension of any one of embodiments 12-21, wherein the composition comprises a fusion protein comprising the recombinant beta-casein.

Embodiment Set Number 24: Cheese Compositions

[1223] 1. A cheese composition comprising para-kappa-casein produced without the use of any enzyme that cleaves kappa-casein to para-kappa casein.

[1224] 2. A substantially transparent plant-based cheese composition.

[1225] 3. A cheese composition comprising a recombinant beta-casein protein; wherein the cheese composition has at least one of the following characteristics: a firmness of at least 150 grams, as determined by compressing a cylindrical-shaped sample of the cheese composition having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C; a melting point of about 35°C to about 100°C; or ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the cheese composition to a temperature of about 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[1226] 4. The cheese composition of embodiment 141, wherein the composition does not comprise any additional casein proteins.

[1227] 5. The cheese composition of embodiment 141, wherein the composition comprises at least one additional casein protein.

[1228] 6. The cheese composition of embodiment 143, wherein at least 80% by weight of the total casein protein in the composition is beta-casein.

[1229] 7. The cheese composition of embodiment 143, wherein at least 90% by weight of the total casein protein in the composition is beta-casein.

[1230] 8. The cheese composition of embodiment 143, wherein at least 95% by weight of the total casein protein in the composition is beta-casein.

[1231] 9. The cheese composition of embodiment 143, wherein the at least one additional casein protein is selected from kappa-casein, para-kappa-casein, beta-casein, alpha-S1-casein and alpha-S2-casein.

[1232] 10. The cheese composition of embodiment 143, wherein the at least one additional casein protein is kappa-casein.

[1233] 11. The cheese composition of embodiment 143, wherein the at least one additional casein protein is para-kappa casein.

[1234] 12. The cheese composition of any one of embodiments 141-149, wherein the recombinant beta-casein is plant-expressed.

[1235] 13. The cheese composition of embodiment 150, wherein the recombinant beta-casein is expressed in a soybean.

[1236] 14. The cheese composition of any one of embodiments 143-149, wherein all caseins in the composition are plant-expressed.

[1237] 15. The cheese composition of any one of embodiments 141-152, wherein the recombinant casein protein is derived from a fusion protein.

[1238] 16. The cheese composition of any one of embodiments 141-153, wherein the composition does not contain an organoleptically functional amount of beta-lactoglobulin.

[1239] 17. The cheese composition of any one of embodiments 141-154, wherein the composition has the ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the composition at a temperature of 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[1240] 18. The cheese composition of any one of embodiments 141-155, wherein the composition has the ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the composition at a temperature of 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass; and a firmness of at least 150 grams, as determined by compressing a cylindrical-shaped sample of the cheese composition having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C.

[1241] 19. The cheese composition of any one of embodiments 141-156, wherein the composition comprises at least one lipid and at least one salt.

[1242] 20. The cheese composition of any one of embodiments 141-157, wherein the composition comprises calcium.

[1243] 21. The cheese composition of embodiment 158, wherein the composition comprises calcium at a concentration of about 0.01% to about 2% by weight.

[1244] 22. The cheese composition of any one of embodiments 141-159, wherein the composition has a pH of about 5.2 to about 5.9.

[1245] 23. The cheese composition of any one of embodiments 141-160, wherein the composition comprises at least one organoleptic properties similar to cheese selected from the group consisting of taste, appearance, mouthfeel, structure, texture, density, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess, and emulsification.

[1246] 24. A method of making the cheese composition of embodiments 141-161, the method comprising expressing the recombinant beta-casein protein in a plant, extracting the beta-casein from the plant, and combining the beta-casein with at least one lipid and/or salt.

[1247] 25. A cheese composition comprising a recombinant beta-casein protein; wherein the cheese composition has ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the composition at a temperature of 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

[1248] 26. The cheese composition of embodiment 163, wherein the composition does not comprise any additional casein proteins.

[1249] 27. The cheese composition of embodiment 163, wherein the composition comprises at least one additional casein protein, and wherein at least 80% by weight of the total casein protein in the composition is beta-casein.

[1250] 28. The cheese composition of embodiment 165, wherein the at least one additional casein protein is kappa-casein or para-kappa casein.

[1251] 29. The cheese composition of any one of embodiments 163-166, wherein the recombinant beta-casein is plant-expressed.

[1252] 30. The cheese composition of any one of embodiments 165-167, wherein the recombinant casein protein is derived from a fusion protein.

[1253] 31. The cheese composition of any one of embodiments 163-168, wherein the composition has at least one of the following characteristics: a firmness of at least 150 grams, as determined by compressing a cylindrical-shaped sample of the cheese composition having a height of 3 cm and a diameter of 3 cm to a height of 1.5 cm at 5°C; or a melting point of about 35°C to about 100°C.

[1254] 32. A method of making the cheese composition of any one of embodiments 163-169, the method comprising expressing the recombinant beta-casein protein in a plant, extracting the beta-casein from the plant, and combining the beta-casein with at least one lipid and/or salt.

CLAIMS

What is claimed is:

1. A recombinant fusion protein comprising:
 - (i) a first milk protein; and
 - (ii) a second milk protein.
2. The recombinant fusion protein of claim 1, wherein at least one of the first milk protein and the second milk protein is α -S1 casein, α -S2 casein, β -casein, κ -casein, para- κ -casein, β -lactoglobulin, α -lactalbumin, lysozyme, lactoferrin, lactoperoxidase, serum albumin, or an immunoglobulin.
3. The recombinant fusion protein of claim 1, wherein at least one of the first milk protein and the second milk protein is β -lactoglobulin.
4. The recombinant fusion protein of claim 1, wherein at least one of the first milk protein and the second milk protein is α -S1 casein, α -S2 casein, β -casein, κ -casein, or para- κ -casein.
5. The recombinant fusion protein of claim 1, wherein:
 - i) the first milk protein is α -S1 casein, α -S2 casein, β -casein, κ -casein, or para- κ -casein; and
 - ii) the second milk protein is α -S1 casein, α -S2 casein, β -casein, κ -casein, or para- κ -casein.
6. The recombinant fusion protein of claim 1, wherein at least one of the first milk protein and the second milk protein is κ -casein and comprises the sequence of SEQ ID NO: 4, or a sequence at least 90% identical thereto.
7. The recombinant fusion protein of claim 1, wherein at least one of the first milk protein and the second milk protein is para- κ -casein and comprises the sequence of SEQ ID NO: 2, or a sequence at least 90% identical thereto.

8. The recombinant fusion protein of claim 1, wherein at least one of the first milk protein and the second milk protein is β -casein and comprises the sequence of SEQ ID NO: 6, or a sequence at least 90% identical thereto.
9. The recombinant fusion protein of claim 1, wherein at least one of the first milk protein and the second milk protein is α -S1 casein and comprises the sequence SEQ ID NO: 8, or a sequence at least 90% identical thereto.
10. The recombinant fusion protein of claim 1, wherein at least one of the first milk protein and the second milk protein is α -S2 casein and comprises the sequence SEQ ID NO: 84, or a sequence at least 90% identical thereto.
11. The recombinant fusion protein of claim 1, wherein the first milk protein and the second milk protein are different proteins.
12. The recombinant fusion protein of claim 1, wherein the first milk protein and the second milk protein are the same proteins.
13. The recombinant fusion protein of any one of claims 1-12, which is plant-expressed.
14. The recombinant fusion protein of claim 13, which is expressed in a soybean plant.
15. The recombinant fusion protein of any one of claims 1-14, which comprises a protease cleavage site.
16. The recombinant fusion protein of claim 15, wherein the protease cleavage site is a chymosin cleavage site.
17. A nucleic acid encoding the recombinant fusion protein of any one of claims 1-16.

18. The nucleic acid of claim 17, which is codon-optimized for expression in a plant.
19. The nucleic acid of claim 18, which is codon-optimized for expression in a soybean.
20. An expression vector comprising the nucleic acid molecule of any one of claims 17-19.
21. A host cell comprising the nucleic acid of any one of claims 17-19 or the expression vector of claim 20.
22. The host cell of claim 21, wherein the cell is a plant cell, bacterial cell, fungal cell, or mammalian cell.
23. The host cell of claim 22, wherein the plant cell is a soybean cell.
24. A plant stably transformed with the nucleic acid of any one of claims 17-19 or the expression vector of claim 20.
25. The plant of claim 24, wherein the fusion protein is expressed in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant.
26. A method for making a fusion protein, the method comprising:
 - (a) transforming a host cell with the nucleic acid of any one of claims 17-19 or the expression vector of claim 20; and
 - (b) growing the transformed host cell under conditions wherein the fusion protein is expressed.
27. The method of claim 26, which comprises co-expressing in the host cell a protein capable of forming a protein body.

28. The method of claim 27, wherein the protein capable of forming a protein body is a prolamin selected from a gliadin, a hordein, a secalin, a zein, a kafirin, or an avenin.
29. The method of claim 26, which comprises expressing a kinase in the host cell.
30. The method of claim 26, wherein expression of one or more proteases is knocked down or knocked out in the cell.
31. The method of any one of claims 26-30, wherein the fusion protein is expressed in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant.
32. A transgenic plant comprising the recombinant fusion protein of any one of claims 1-16, the nucleic acid of any one of claims 17-20, or the expression vector of claim 20.
33. The transgenic plant of claim 32, which is a soybean plant.
34. A method for stably expressing the recombinant fusion protein of any one of claims 1-16 in a plant, the method comprising:
- (i) transforming a plant with a plant transformation vector comprising an expression cassette comprising a nucleic acid molecule encoding the fusion protein; and
 - (ii) growing the transformed plant under conditions wherein the recombinant fusion protein is expressed.
35. The method of claim 34, wherein the fusion protein is expressed in an amount of 1% or higher per total protein weight of soluble protein extractable from the plant.
36. A seed processing composition, comprising the fusion protein of any one of claims 1-16.
37. A food composition comprising the fusion protein of any one of claims 1-16.

38. The food composition of claim 37, which is selected from the group consisting of cheese and processed cheese products, yogurt and fermented dairy products, directly acidified counterparts of fermented dairy products, cottage cheese dressing, frozen dairy products, frozen desserts, desserts, baked goods, toppings, icings, fillings, low-fat spreads, dairy-based dry mixes, soups, sauces, salad dressing, geriatric nutrition, creams and creamers, analog dairy products, follow-up formula, baby formula, infant formula, milk, dairy beverages, acid dairy drinks, smoothies, milk tea, butter, margarine, butter alternatives, growing up milks, low-lactose products and beverages, medical and clinical nutrition products, protein/nutrition bar applications, sports beverages, confections, meat products, analog meat products, meal replacement beverages, weight management food and beverages, cultured buttermilk, sour cream, yogurt, skyr, leben, lassi, kefir, powder containing a milk protein, and low-lactose products.

39. The food composition of claim 37, which is a cheese composition.

40. The food composition of claim 39, wherein the cheese composition has the ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the composition at a temperature of 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

41. The food composition of claim 39 or 40, comprising a total amount of casein protein; wherein about 32% to 100% by weight of the total amount of casein protein in the food composition is beta-casein.

42. A method of making a food composition, comprising combining the fusion protein of any one of claims 1-16 into a food composition.

43. An alternative dairy food composition comprising:
i) the recombinant fusion protein of any one of claims 1-16; and
ii) at least one lipid,

wherein the recombinant fusion protein confers on the alternative dairy food composition one or more characteristics of a dairy food product selected from the group consisting of: taste, aroma, appearance, handling, mouthfeel, density, structure, texture, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess and emulsification.

44. The alternative dairy food composition of claim 43, wherein the alternative dairy food composition does not comprise any other milk proteins.

45. The alternative dairy food composition of claim 43 or 44, which comprises calcium at a concentration of about 0.01 to about 2% by weight.

46. The alternative dairy food composition of any one of claims 43-45, comprising a total amount of casein protein;

wherein about 32% to 100% by weight of the total amount of casein protein in the food composition is beta-casein.

47. The alternative dairy food composition of any one of claims 43-46, wherein the composition has a pH of about 5.2 to about 5.9.

48. The alternative dairy food composition of any one of claims 43-47, which is selected from the group consisting of cheese and processed cheese products, yogurt and fermented dairy products, directly acidified counterparts of fermented dairy products, cottage cheese dressing, frozen dairy products, frozen desserts, desserts, baked goods, toppings, icings, fillings, low-fat spreads, dairy-based dry mixes, soups, sauces, salad dressing, geriatric nutrition, creams and creamers, analog dairy products, follow-up formula, baby formula, infant formula, milk, dairy beverages, acid dairy drinks, smoothies, milk tea, butter, margarine, butter alternatives, growing up milks, low-lactose products and beverages, medical and clinical nutrition products, protein/nutrition bar applications, sports beverages, confections, meat products, analog meat products, meal replacement beverages, weight management food and beverages, cultured

buttermilk, sour cream, yogurt, skyr, leben, lassi, kefir, powder containing a milk protein, and low-lactose products.

49. The alternative dairy food composition of any one of claims 43-47, which is a cheese composition.

50. A solid phase, protein-stabilized emulsion comprising the fusion protein of any one of claims 1-16, wherein the emulsion has the ability to stretch to at least 3 cm in length without breaking, as determined by heating a 100 gram mass of the emulsion to a temperature of about 225°C for 4 minutes and cooling to about 90°C and pulling with a fork placed beneath the mass.

51. A colloidal suspension comprising the fusion protein of any one of claims 1-16; wherein the colloidal suspension has at least one, at least two, or at least three characteristics that are substantially similar to bovine milk selected from taste, appearance, mouthfeel, structure, texture, density, elasticity, springiness, coagulation, binding, leavening, aeration, foaming, creaminess, and emulsification.