
Declarations under Rule 4.17:  
— as to applicant’s entitlement to apply for and be granted a patent (Rule 4.17(b))  
— as to the applicant’s entitlement to claim the priority of the earlier application (Rule 4.17(in))

Published:  
— without international search report and to be republished upon receipt of that report (Rule 48.2(g))  
— with sequence listing part of description (Rule 5.1(a))

Title: TOLERANCE THERAPEUTIC FOR TREATING POLYPEPTIDE INDUCED ALLERGY

FIG. 1

Antigen shaft head

His-6

Abstract: The present disclosure is directed to compositions comprising one or more components of a polypeptide allergen combined with a reovirus-derived targeting protein, and related methods for the generation of tolerance against the polypeptide allergens.
TOLERANCE THERAPEUTIC FOR TREATING POLYPEPTIDE INDUCED ALLERGY

CROSS-REFERENCE TO RELATED APPLICATION

This application claims the benefit of U.S. Provisional Application No. 62/161,6318, filed February 13, 2015, the disclosure of which is incorporated herein by reference in its entirety.

STATEMENT REGARDING SEQUENCE LISTING

The sequence listing associated with this application is provided in text format in lieu of a paper copy and is hereby incorporated by reference into the specification. The name of the text file containing the sequence listing is VRTC155292_ST25.txt. The text file is 67KB; was created on February 9, 2016; and is being submitted via EFS-Web with the filing of the specification.

SUMMARY

This summary is provided to introduce a selection of concepts in a simplified form that are further described below in the Detailed Description. This summary is not intended to identify key features of the claimed subject matter, nor is it intended to be used as an aid in determining the scope of the claimed subject matter.

In one aspect, the disclosure provides an isolated fusion protein comprising a reovirus-derived targeting polypeptide and at least one allergen polypeptide. In one embodiment, the reovirus-derived targeting polypeptide comprises the protein sigma polypeptide (ρσΙ), or a functional portion or a derivative thereof. In one embodiment, the functional portions of the ρσΙ include the head domain, trimerization domain, sialic acid binding domain, and/or the shaft domain of the ρσΙ protein, or any derivative thereof.

In one embodiment, the at least one allergen polypeptide is a food allergen, an environmental allergen, an autoantigen, and/or biological therapeutic, or is derived therefrom.

In one embodiment, the food allergen is from a ground nut, tree nut, milk, gluten, egg, fish, shellfish, and the like. In one embodiment, the food allergen is from a peanut (Arachis hypogaea) and the allergen polypeptide is Arah2, Arah6, Arah1, Arah3, Arah4, Arah5, and the like.

-1-
Arah5, Arah7, Arah8, Arah9, Arah10, Arah11, Arah12, or is derived therefrom. In one embodiment, the food allergen is from gluten and the allergen polypeptide is a prolamin (such as a α-gliadin, β-gliadin, γ-gliadin, co-gliadin, hordein, secalin, zein, kafirin, avenin), glutenin, or is derived therefrom. In one embodiment, the food allergen is from milk and the allergen polypeptide is alpha S1-casein, alpha S2-casein, β-lactoglobulin, β-casein, κ-casein, or is derived therefrom. In one embodiment, the food allergen is from egg and the allergen polypeptide is ovomucoid, ovotransferrin, lysozyme, livetin, apovitellin, phosvitin, or is derived therefrom. In one embodiment, the food allergen is from fish and the allergen polypeptide is Che ag, Lop pi, Gelatin/Ore a, Parvalbumin/Sebm, Ore a1 (Oreochromis aurea; blue tulapia), Sebml, Sarsa 1.0101, Albumin/Oncm 1 (rainbow trout/Onorhynchus mykiss), glyceraldehyde-3-phosphate dehydrogenase, or is derived therefrom.

In one embodiment, the environmental allergen is from an animal or insect, such as dust mite, bee, wasp, cat, dog, and the like, or plant, such as ragweed, grass, tree, and the like. In one embodiment, the environmental allergen is from a dust mite and the allergen polypeptide is Derpl through Derp23, Derfl through Derf33, Eurml, 2, 3, 4, or 14, Derml, or is derived therefrom. In one embodiment, the environmental allergen is from cat and the allergen polypeptide is a secretoglobin such as Feldl, a lipocalin such as Feld4, an albumin such as Feld2, a cystatin such as Feld3, IgA such as Feld5w, or is derived therefrom. In one embodiment, the environmental allergen is from ragweed and the allergen polypeptide is Ambal through Ambal 1, Ambp5, Ambt5, or is derived therefrom. In one embodiment, the environmental allergen is from tree, such as birch, alder, and ash, and the allergen polypeptide is Betv1, Betv2, Betv3, Betv4, Betv6, Betv7, Alngl, Alng4, Frael, or is derived therefrom.

In one embodiment, the autoantigen is transglutaminase, myelin-associated glycoprotein (MAG), CNS-specific myelin oligodendrocyte glycoprotein (MOG), myelin basic protein (MBP), proteolipid protein (PLP), zinc transporter-8 (ZnT8), glutamic decarboxylase 65 (GAD65), glutamic decarboxylase 67 (GAD67), preproinsulin, proinsulin, insulin, tyrosine phosphatase like autoantigen, insulinoma antigen-2 (IA-2; ICA512, PTPRN), IA-2b (Phogrin, PTPRN2), islet cell antigen-69 (ICA69), chromogranin A, islet amyloid polypeptide (ppIAPP), heat shock protein 60 (hsp60), or is derived therefrom.
In one embodiment, the allergen polypeptide is derived from a protein therapeutic, such as an antibody CDR or, for example, erythropoietin. In one embodiment, the at least one allergen polypeptide comprises an MHC Class I epitope and/or an MHC Class II epitope.

In one embodiment, the targeting polypeptide is separated from the at least one allergen polypeptide by a linker. In one embodiment, the fusion protein comprises at least two allergen polypeptides. In one embodiment, the at least two allergen polypeptides are separated by a linker.

In another aspect, the disclosure provides a pharmaceutical composition comprising the isolated fusion protein described herein and a pharmaceutically acceptable carrier. In one embodiment, the composition is formulated for oral or intranasal administration.

In another aspect, the disclosure provides a nucleic acid, or a vector comprising the nucleic acid, wherein the nucleic acid comprises a sequence encoding the isolated fusion protein described herein. In a further aspect, the disclosure provides a cultured cell transfected or comprising the vector described herein.

In another aspect, the disclosure provides a method for inducing tolerance to a polypeptide allergen, comprising administering to a subject in need thereof a pharmaceutically effective amount of the isolated fusion protein described herein, wherein the isolated fusion protein comprises polypeptide derived from the polypeptide allergen. In one embodiment, the method consists of administering a single dose of the effective amount of the isolated fusion polypeptide. In another embodiment, the method comprises administering two or more doses of the effective amount of the isolated fusion polypeptide. In one embodiment, the effective amount of the isolated fusion polypeptide comprises less than 100mg, 75mg, 50mg, 25mg, 20mg, 15mg, 10mg, 9mg, 8mg, 7mg, 6mg, 5mg, 4mg, 3mg, 2mg, 1.5mg, or 1mg of the isolated fusion polypeptide.

BACKGROUND

Allergies against foods and environmental factors are a major health concern worldwide. An allergy is a hypersensitivity of the immune system to particular antigens (also referred to as "allergens"), which can result in uncomfortable and potentially dangerous immune reactions that can cause severe swelling, rhinitis, bronchoconstriction, edema, hypotension, digestive distress, hives, and itchy sensations. The range of severity
can vary greatly from mere discomfort, to inducement of vomiting, asphyxiation, coma and even death.

Potential allergens can be derived from a variety of sources, such as food, plants, chemicals and environmental antigens. Strategies to address allergies include avoidance of the allergen, induction of tolerance (i.e., preventing the hypersensitive reaction when exposed), and ameliorating the response once it occurs.

As one example of the breadth and severity of allergies in a population, it is estimated that more than 1% of the US population (~ 3 million people) suffer from peanut or tree nut allergies. Approximately half of the 30,000 food allergy-related emergency room visits each year, including 100-150 deaths, are due to peanut allergies. Unlike many food allergies, reaction to peanuts persists throughout adulthood in approximately 80% of individuals. Taken together these numbers indicate that peanut allergy represents the most prevalent and severe form of food allergy. Although there has been progress in developing oral desensitization procedures for peanut allergies, the regimens require a gradual increase in exposure over approximately 12 months or more and are not applicable to severely allergic individuals because of potential anaphylactic responses. In addition, the responses observed with these existing oral desensitization regimens are not long-lasting, and the patients' allergic response to peanut allergy returns shortly after stopping the oral administration of allergen. As a result, the vast majority of sufferers rely on strict avoidance and epinephrine administration if exposed. However, because peanuts are such a common food source, the risk of exposure is always a concern, particularly in children.

Accordingly, a need remains for a simple and effective approach to address allergic responses to various allergens, such as polypeptide allergens. The present disclosure addresses this and related needs by providing a strategy to induce tolerance to polypeptide allergens by incorporating the allergen, or one or more components of the allergen in a fusion protein, with a reovirus-derived targeting protein.

DESCRIPTION OF THE DRAWINGS

The foregoing aspects and many of the attendant advantages of this invention will become more readily appreciated as the same become better understood by reference to the following detailed description, when taken in conjunction with the accompanying drawings, wherein:
FIGURE 1 is a schematic representation of an exemplary Arah2-psl protein that includes a tolerogen/antigen, Arah2 and psl (shaft and head), and a 6 histidine-tag for purification.

FIGURE 2 is an image of an immunoblot of the purified Arah2-psl protein. The immunoblot was stained with anti-psl rabbit serum followed by an anti-rabbit HRP secondary. Lane 1, purified Arah2-psl; lane 2, crude yeast lysate; lane 3, recombinant psl protein; lane 4, MWM standard.

FIGURES 3A and 3B illustrate Arah2-psl protein activity. (A) Arah2-psl binding to L cells. Cells were stained with Arah2-psl followed by either normal rabbit serum (Dark = MFI = 32) or anti-psl rabbit serum (Gray = MFI = 1877) and FITC-labeled anti-rabbit. (B) Arah2-psl binding to HeLa cells. Cells were stained with Arah2-psl followed by either normal rabbit serum (Dark = MFI = 26) or anti-psl rabbit serum (Gray MFI = 1420) and FITC-labeled anti-rabbit.

FIGURE 4 is a schematic illustration of an exemplary protocol for establishing a peanut allergy model in mice.

DETAILED DESCRIPTION

The present disclosure is generally directed to tolerance therapeutics and related methods that can induce tolerance to polypeptide allergens.

The gut and the nasopharynx constitute major regions of the body that first contact many antigens and allergens from the environment, such as food-borne or ambient, air-borne allergens. The epithelial layer that covers the Gut Associated Lymphoid Tissue (GALT) and Nasopharyngeal Associated Lymphoid Tissue (NALT) regions contains a subpopulation of microfold cells (M cells) specialized to sample environmental antigens and present them to the adjacent immune cells. A number of studies now indicate that the M cells in the GALT and NALT regions play an important role in the generation of either an immune response or a tolerance response to a given antigen.

Reoviruses are segmented, double-stranded RNA viruses that infect humans via mucosal surfaces and can cause both enteric and respiratory infections. To initiate infection, it has been demonstrated that reoviruses first bind to the surface of M cells. Specifically, a reovirus cell adhesin protein, protein sigma ("σσι"), has been shown to interact with at least two host receptors via separate binding domains. The head domain
binds with a component of tight junctions, whereas sequences contained within the fibrous tail domain bind terminal $\sigma$-linked sialic acid residues on host cells.

In view of the above, preliminary studies have been performed to assess the ability of reovirus attachment proteins, such as $\rho\sigma\text{I}$, to serve as targeting proteins to assist the delivery of antigenic payloads to M cells. For example, it has been demonstrated that administration of a recombinant fusion protein combining the reovirus $\rho\sigma\text{I}$ protein with the full ovalbumin (OVA) protein (OVA-$\rho\sigma\text{I}$) reduced OVA induction of serum Ig, IFN-gamma, IL-2 and IL-17 levels, while increasing IL-10 and IL-4 in an IL-10 dependent fashion. Imaging studies demonstrated that the OVA-$\rho\sigma\text{I}$ specifically binds to the mucosa surface. Immune cells isolated from the mice were characterized, revealing an induction of anti-inflammatory cytokines and an increase of suppressive regulatory T-cells (Tregs) even with a single dose of OVA-$\rho\sigma\text{I}$ fusion protein. See Rynda, A., et al., "Low-dose tolerance is mediated by the microfold cell ligand, reovirus protein signal," J. Immunol. 750:5187-5200 (2008); and Suzuki, H., et al., "Ovalbumin-protein sigma 1 M-cell targeting facilitates oral tolerance with reduction of antigen-specific CD4+ T cells," Gastroenterology 735:917-925 (2008), each incorporated herein by reference in its entirety. In additional studies, the OVA-$\rho\sigma\text{I}$ was further modified to include antigens, i.e., proteolipid protein (PLP) and myelin oligodendrocyte glycoprotein (MOG), that normally induce an autoimmune reaction, EAE, in a murine model. The incorporation of these self-antigens into an OVA-$\rho\sigma\text{I}$ construct, either by addition to replacement of the OVA component, results in diminished EAE. These results indicate the potential use of the $\rho\sigma\text{I}$ to treat autoimmune diseases. See Rynda, A., et al., "IL-28 supplants requirement for T(reg) cells in protein signal-mediated protection against murine experimental autoimmune encephalomyelitis (EAE)," PLoS One 5:e8720 (2010); and Rynda-Apple, A., et al., "Active immunization using a single dose immunotherapeutic abates established EAE via IL-10 and regulatory T cells," Eur. J. Immunol. 7:313-323 (2011), each incorporated herein by reference in its entirety.

However, these preliminary studies are limited to using whole OVA antigen and/or MOG antigen fused to $\rho\sigma\text{I}$ to induce tolerance to these specific antigens in specially designed murine models. These studies do not address whether the reovirus $\rho\sigma\text{I}$ can be fused generally to any allergenic polypeptide (including other intact whole polypeptide allergens or protein fragments and derivatives thereof) to effectively induce tolerance to the source of that allergenic polypeptide. Therefore, these studies do not
inform as to whether the reovirus \( \rho \sigma \) can be used generally to target any food-borne or air-borne protein allergens to the M cells and functionally induce tolerance in such a way as to ameliorate a subject's allergic reactions to normal exposure of the protein source. The prior studies also do not instruct as to what structural characteristics of the intended allergenic protein are required to actually obtain some level of tolerance. For example, is the full-length antigen/allergen required, or can the \( \rho \sigma \)-based fusion protein incorporate only a fragment of the full-length antigen/allergen. If so, what fragment(s) is/are preferred for optimized tolerance induction? Can multiple fragments be incorporated in the fusion for enhanced effect? What fragment(s) is/are preferred for optimized protein expression from a cell expression system? What fragment(s) is/are preferred for optimized protein solubility for therapeutic administration? Can the performance of the fusion protein be modulated by inserting and/or manipulating a polypeptide linker? Can the performance of the fusion protein be improved by selectively designing fusion proteins that incorporate allergen polypeptide (or polypeptide fragments) to provide a multivalent fusion protein against distinct allergens? If so, what design format is preferable? Is glycosylation of the allergen crucial to the induction of tolerance? Also, which specific combinations of full length or polypeptide fragments are required to effectively treat an individual? Such questions require additional characterization of the reovirus fusion proteins to establish their utility as tolerance-inducing platform.

To address the extensive morbidity associated with hypersensitivity to allergens and autoantigens, such as, e.g., peanut or gluten allergens, the present disclosure addresses studies that provide new insight into reagents and therapeutic approaches that efficiently induce tolerance to polypeptide allergens.

In accordance with the foregoing, the present disclosure provides an isolated fusion protein comprising a reovirus-derived targeting polypeptide and at least one allergen polypeptide.

As a preliminary matter, as used herein the terms "protein" and "polypeptide" generally refer to a macromolecule of multiple amino acids linked by peptide (amide) bonds. As used herein, an "amino acid" refers to any of the naturally occurring amino acids found in proteins, D-stereoisomers of the naturally occurring amino acids (e.g., D-threonine), unnatural amino acids, and chemically modified amino acids. Each of these categories of amino acids is not mutually exclusive. \( \alpha \)-Amino acids comprise a carbon atom to which is bonded an amino group, a carboxyl group, a hydrogen atom, and
a distinctive group referred to as a "side chain." The side chains of naturally occurring amino acids are well-known in the art and include, for example, hydrogen (e.g., as in glycine), alkyl (e.g., as in alanine, valine, leucine, isoleucine, proline), substituted alkyl (e.g., as in threonine, serine, methionine, cysteine, aspartic acid, asparagine, glutamic acid, glutamine, arginine, and lysine), arylalkyl (e.g., as in phenylalanine and tryptophan), substituted arylalkyl (e.g., as in tyrosine), and heteroarylmethyl (e.g., as in histidine).

The following abbreviations are typically used for the 20 canonical, naturally occurring canonical amino acids: alanine (Ala; A), asparagine (Asn; N), aspartic acid (Asp; D), arginine (Arg; R), cysteine (Cys; C), glutamic acid (Glu; E), glutamine (Gln; Q), glycine (Gly; G), histidine (His; H), isoleucine (Ile; I), leucine (Leu; L), lysine (Lys; K), methionine (Met; M), phenylalanine (Phe; F), proline (Pro; P), serine (Ser; S), threonine (Thr; T), tryptophan (Trp; W), tyrosine (Tyr; Y), and valine (Val; V).

Noncanonical amino acids (that is, those that are not naturally found in proteins) are also known in the art, as set forth in, for example, Mol. Cell. Biol., 9:2574 (1989); J. Amer. Chem. Soc, 772:401-4030 (1990); J. Amer. Chem. Soc, 56:1280-1283 (1991); J. Amer. Chem. Soc, 773:9276-9286 (1991), each reference incorporated herein in its entirety. β- and γ-amino acids are known in the art and are also contemplated herein as noncanonical amino acids. Several methods are known in the art for incorporating noncanonical (or non-naturally-occurring) amino acid residues into proteins. For example, an in vitro system can be employed wherein nonsense mutations are suppressed using chemically aminoacylated suppressor tRNAs. Methods for synthesizing amino acids and aminoacylating tRNA are known in the art.

The polypeptide can also have chemically modified amino acids, which refers to an amino acid whose side chain has been chemically modified. For example, a side chain may be modified to comprise a signaling moiety, such as a fluorophore or a radiolabel. A side chain may be modified to comprise a new functional group, such as a thiol, carboxylic acid, or amino group. Post-translationally modified amino acids are also included in the definition of chemically modified amino acids.

Finally, persons of ordinary skill in the art will readily appreciate that the polypeptide can encompass altered polymer structures, such as a type of peptidomimetic where a canonical chemical aspect of the polypeptide is modified. As used herein, the term "peptidomimetic" refers to compounds whose essential elements (pharmacophore) mimic a natural peptide or polypeptide in 3D space, and which retain the ability to
interact with the biological target (e.g., a receptor) and produce the same biological effect as an unmodified, canonical polypeptide structure. However, peptidomimetics are designed to circumvent some of the problems associated with a natural peptide: e.g., stability against proteolysis (duration of activity) and poor bioavailability. Certain other properties, such as receptor selectivity or potency, often can be substantially improved. The structural modifications that result in peptidomimetics are well-known and have been described elsewhere. See, e.g., Vagner, J., et al., "Peptidomimetics, a synthetic tool of drug discovery," *Curr. Opin. Chem. Biol.* 72(3):292-296 (2008), incorporated herein by reference in its entirety.

As used herein, the term "isolated" in the context of an isolated fusion protein, indicates that the fusion protein has been produced through human intervention and has been substantially separated from the materials co-existing in the protein production environment, such as the intra-cellular organelles and proteins in a cell culture system. In contrast, a naturally expressed protein in cell is not "isolated." Furthermore, the term "fusion" in the context of a fusion protein indicates that the overall protein or polypeptide contains a nonnaturally occurring polypeptide sequence. Typically, a fusion protein combines to two or more existing polypeptides or polypeptide fragments, from the same or different source proteins, in a chimeric polymer where the polypeptides (or fragments) do not naturally occur together in that manner. Methods of producing fusion proteins are well known. For example, nucleic acids encoding the different polypeptide components of the fusion protein can be generated and amplified using PCR and assembled into an expression vector in the same reading frame to produce a fusion gene. The expression vector can be transformed into any appropriate expression system, such as prokaryotic or eukaryotic cells, which can then express the protein. See, e.g., such standard references as Coligan, Dunn, Ploegh, Speicher and Wingfield "Current Protocols in Protein Science" (1999), Volume I and II (John Wiley & Sons Inc.); Sambrook et al., "Molecular Cloning: A Laboratory Manual" (1989), 2nd Edition (Cold Spring Harbor Laboratory Press); and Prescott, Harley and Klein "Microbiology" (1999), 3rd Edition (WBC McGraw-Hill), each incorporated herein by reference. One exemplary approach for creating fusion proteins is described in more detail in the below examples. In another embodiment, the fusion protein can be created by linking two or more existing polypeptide fragments. For example, the reovirus-derived targeting polypeptide component (e.g., sigma polypeptide (ρσΐ), homologs thereof, or functional portions thereof as described below) can be
produced separately from the allergen polypeptide. Each of these separate components can be generated or obtained independently from one another by any known and conventional technique. The components can subsequently be fused or linked to one another by chemical means. For example, each component can have complementary linker components such that they will form strong mutual bonds, thereby linking their respective components to produce the fusion protein. The linker moieties can be homobifunctional or heterobifunctional. An illustrative, nonlimiting example of such chemical linker constructs include having one component (e.g., targeting polypeptide component) include biotin and the other component (e.g., allergen polypeptide) include strep-avidin, or vice versa. The biotin and strep-avidin moieties will form high-affinity bonds, thereby linking, or "fusing", the components to result in the fusion protein. Other common linking chemistries can also be used, such as, for example, gluteraldehyde, and the like.

The reovirus-derived targeting polypeptide component of the fusion protein can comprise the reovirus protein sigma polypeptide (ρστ), homologs thereof, or functional portions or derivatives thereof. As used in this context, the term "functional" refers to the ability for the one or more combined portions of the ρστ polypeptide to induce some degree of tolerance to an allergen polypeptide fused thereto. Without being bound to any particular theory, this functionality likely requires the ability of the one or more combined portions of the ρστ polypeptide to bind to the target M cells in the mucosa sufficiently to transfer the allergen polypeptide thereto. The structure and sequence of the reovirus has been previously described. See, e.g., Turner, D.L., et al., "Site-directed mutagenesis of the C-terminal portion of reovirus protein sigma 1: evidence for a conformation-dependent receptor binding domain," Virology 75(5):219-227 (1992); Nibert, M.L., et al., "Infectious subvirion particles of reovirus type 3 Dealing exhibit a loss in infectivity and contain a cleaved sigma 1 protein," J. Virol. 59:5057-5067 (1995); and Lee, P.W. and Leone, G., "Reovirus protein sigma 1: from cell attachment to protein oligomerization and folding mechanisms," Bioessays 7(5):199-206 (1994); Barton, E.S., et al., "Utilization of Sialic Acid as a Coreceptor Enhances Reovirus Attachment by Multistep Adhesion Strengthening," J. Biol. Chem. 275(5):2200-2211 (2000); Fraser, R. D. B., et al., "Molecular Structure of the Cell-Attachment Protein of Reovirus: Correlation of Computer-Processed Electron Micrographs with Sequence-Based Predictions," J. Virol. 64:2990-3000 (1990); Nibert, M. L., et al. "Structure of the Reovirus Cell-Attachment Protein: A
Model for the Domain Organization of pSl, "J. Virol. 64:2916-2989 (1990), each of which is incorporated herein by reference in its entirety.

The reovirus-derived targeting polypeptide component can include less than the full length of ρσι̇ polypeptide, but can contain functional fragments or derivatives of fragments, or fusions of non-contiguous fragments thereof, so long as the protein retains the ability to target the overall fusion protein to M-cells and induce some degree of tolerance to the allergen polypeptide fused thereto. Domains of the ρσι̇ that contribute the targeting functionality of the fusion protein include (from C-terminus to N-terminus) the head domain, the trimerization domain, the sialic acid binding domain, and the shaft domain. Although a truncated ρσι̇ could be constructed, the truncated ρσι̇ would preferentially still comprise the head domain, which binds with a component of tight junctions on cells, as well as the sequences contained in the tail domain, which bind terminal a-linked sialic acid residues on host cells. These components are typically required for the induction of tolerance. (Zlotkowska, D., et al., "Loss of Sialic Acid Binding Domain Redirects Protein σι̇ to Enhance M Cell-Directed Vaccination," PLoS One 7:e36182 (2012)). Typically, fusions will incorporate the chosen allergen polypeptide(s) at the C-terminal end of the ρσι̇ polypeptide (or fragment thereof) so as to avoid interfering with the ability of the head domain to bind to the mucosal cell receptors.

As used herein, the term "derivative thereof" refers to any ρσι̇ protein or functional portion thereof that has one or more amino acid additions, substitutions, or deletions, with respect to a reference ρσι̇ protein or functional portion thereof that has substantially equivalent or enhanced functionality. For instance, the ρσι̇ could incorporate various mutations from a reference ρσι̇ sequence, such as in the head domain, that increases the binding avidity of the ρσι̇ or functional portion thereof to the M cell.

In one embodiment, the fusion protein comprises one allergen polypeptide. In this context, an allergen polypeptide is any stretch of contiguous amino acids in a polypeptide molecule that stimulates an immune response in a vertebrate, where the immune response has a negative impact on the health, comfort, and well-being of the vertebrate subject. The polypeptide can be the full length protein of a known allergen or antigen. Alternatively, the polypeptide can be "derived from" a source allergen or antigen. In this regard, the term "derived from" indicates that the allergen polypeptide component of the fusion protein can be the result of some procession of the source allergen protein. For example, the allergen polypeptide can be a fragment of the source allergen protein where
one or more end portions of the full-length source proteins have been removed. In another embodiment, the allergen polypeptide can itself be a fusion of non-contiguous sections of the source allergen protein, where an internal portion(s) have been removed. It will be appreciated that the remaining portions of the source allergen protein can be oriented in the allergen polypeptide in a contiguous orientation, or, alternatively, can be separated by a linker moiety.

In another embodiment, the fusion protein comprises a plurality (i.e., more than one) allergen polypeptides. In this context, reference to multiple fusion polypeptides as distinct components implies that the polypeptides are, or are derived from, distinct source allergen proteins. The source allergen proteins themselves can be from the same overall source (e.g., two distinct source proteins from peanut), or from different sources (e.g., a source protein from peanut and a source protein from walnut, fish, gluten, dust mites, and the like). The plurality of allergen polypeptides can be in any relative orientation, including being N-terminal or C-terminal to the ρσ$i$ component of the fusion protein, or chemically linked through an amino acid side chain, as described above.

The multiple components of the fusion protein (e.g., the targeting polypeptide, or subcomponents thereof, e.g., domains of ρσ$i$, and the one or more allergen polypeptides, and potential subcomponents thereof) can be disposed in adjoining, contiguous sequence. Alternatively, one or more of the proximate components can be joined by a linker moiety, which would be disposed between the components and covalently attached to each. The linker moiety can be a synthetic polypeptide sequence, which is typically between about four and about 40 amino acids in length. The linker preferably provides an attachment between the otherwise proximate components in the fusion providing sufficient space and flexibility such that each component can freely assume its natural three-dimensional configuration without requiring significant adjustment for the configuration assumed by the proximate component. Accordingly, such linkers are typically designed to avoid significant formation of rigid secondary structures that could reduce the flexibility or distance provided between the proximate components. Thus, the linker is designed to provide a linear or alpha-helical structure. Such linkers are commonly used and are well-understood in the art. As an illustrative, non-limiting example, the linker can comprise the amino acid sequence GlyArgProGly (SEQ ID NO:1). In other embodiments, the linker is a non-polypeptide chemical linker, as known in the art. For example, as described above, the linker moiety can be homobifunctional or heterobifunctional.
Examples include strep-avidin/biotin and a crosslinker, such as a thiol or an amide-linker system, as used in antibody technologies.

Exemplary allergens and allergen sources that are useful for the allergen polypeptide are now described. A large number of defined allergens are known to the artisan. Online data bases which provide the approved nomenclature for many known allergens and provide links to known nucleic acid and amino acid sequences are available, including for example, the allergenonline data base provided by the University of Nebraska-Lincoln and the official allergen nomenclature website approved by the World Health Organization and the International Union of Immunological Societies Allergen Nomenclature Subcommittee.

The allergen polypeptide of the present disclosure can be generally a food allergen, an environmental allergen, an autoantigen, and/or a biological therapeutic. Moreover, the allergen polypeptide can be derived from any of the sources in the above categories. In this context, the allergen polypeptide integrated into the fusion protein can be a full-length allergen protein found in the allergen source, or can be a subcomponent, or a fusion of multiple subcomponents, of the full-length protein.

Food allergens (and their general food sources) are well-known and many protein components of each allergen have been identified and characterized. For example, illustrative and non-limiting sources of food allergens include various fruits (such as mango and strawberries), garlic, fish, shellfish, meats, milk, peanuts and other legumes or ground nuts, tree nuts (such as almonds, Brazil nuts, cashews, chestnuts, filberts/hazelnuts, macadamia nuts, pecans, pistachios, pine nuts, and walnuts), soy, oats, gluten, and egg.

To further illustrate, in peanut (Arachis hypogaeae) allergy there are approximately twelve proteins identified by reactive serum IgE that are, thus, identified as being allergenic. These are referred to by the following abbreviations and in parenthesis a subtype designation and/or a general database identifier (GI ), which database identifiers are incorporated herein by reference: Arah2 (.0201 GI|26245447; .0101 GI|3 1322014), Arah6 (GI|5923742; GI|17225991, Arah1 (GI|1168390; GI|1168391), Arah3 (.0101 GI|3703107; .0201 GI|5712199), Arah4 (renamed Arah3.0201), Arah5 (GI|5902098; GI|43182555; GI|284810529), Arah7 (GI|5931948; GI|158121995), Arah8 (GI|37499626; GI|145904610; GI|169786740; or GI|110676574), Arah9 (.0101 GI|161087230; .0201 GI|161610580), Arah10 (.0101 GI|1 3200509; .0102 GI|52001239), Arah 1 (.0101
GI|71040655), and Arahl2 (.0101 GI|160623326). Of these, antibodies to Arahl, 2, and 6 are typically detected in more than 80 to 90% of allergic individuals. Using cell based degranulation assays, it has been reported that removal of Arah2 and Arah6 from whole peanut extract (WPE) reduces the antigenicity by 90%. Because these two proteins are closely related (approximately 60% homology), it has been proposed that a therapy generating a robust tolerance response to either protein would be expected to significantly improve the lives of the majority of peanut allergy sufferers. Thus, it is hypothesized herein that Arah2 represents the best single antigen for developing a ρσγ targeted tolerance therapeutic to treat individuals with peanut allergy. However, as described above, the efficacy of any fusion protein can potentially be improved to treat unresponsive patients by adding another one or two other allergen polypeptides with an Arah2 fusion protein, or by developing additional fusion proteins that contain other major peanut allergens, such as for example, Arahl and Arah6, and using a combination therapy.

In some embodiments, the food allergen is from gluten. Several protein allergens from gluten are known and have been characterized and are encompassed by the present disclosure. For example, the allergen polypeptide can be a prolamin from wheat (*Triticum aestivum*), barley (*Hordeum vulgare*), oats (*Avena sativa*), rye (*Secale cereal*), corn (*Zea mays*) or sorghum (*Sorghum bicolor*) and can include, for example α-gliadin, β-gliadin, γ-gliadin, ω-gliadin, hordein, secalin, zein, kafirin, avenin; a glutenin, or can be derived therefrom. The prolamin can include any one of the proteins, protein isoforms, or fragments thereof. These are referred to by the following abbreviations and in parenthesis a subtype designation and/or a general database identifier (GIj), which database identifiers are incorporated herein by reference: *Triticum aestivum* omega 5 gliadin (trial 9) GI|73912496, GI|208605344, GI|208605348, GI|208605346, GI|508732623; *Triticum aestivum* γ gliadin (trial20) GI|508732621, GI|170702, GI|170708.., GI|170736, Gij|170738, GI|1063270, GI|62484809, GI|508732621; *Triticum aestivum* α/β gliadin, for example tria21 and tria25 (GIJ|283476402, GI|21755, GI|21757, GI|21761, GI|170710, GI|170712, GI|170718, GI|170726, GI|170728 and the like); *Hordeum vulgare* γ hordein 3, for example, horv20 (GI|288709); *Secale cereal* γ seculin, for example, secc20 (GIj75 198759, GI|75 198753); *Avena sativa* avenin, for example, GI|166555, Gij|166553, GI|166557, Gij|166551, GI|398616299; *Sorghum bicolor* kafirin, for example, Gij|21 174, Gij|125|167. *Zea mays* zein, for example, Gij|168701, Gij|168699,
GIj468517, GIj468515, and the like. All data base identifiers are each incorporated herein by reference.

In some embodiments, the food allergen is from milk. Several allergens from milk are known and have been characterized and are encompassed by the present disclosure. For example, the allergen polypeptide can be alpha S1-casein, for example from Bos taurus, GI|162794, and GI|30794348, alpha S2-casein from Bos taurus, for example, bsdIO GI|27806963, β-lactoglobulin from Bos domesticus, for example, bsd5 GI|520; β-casein from Bos taurus, for example, bsd1 GI|9420764504, GI|162797, GI|162805, GI|459292; κ-casein from Bos taurus, for example, bsd12 GI|162811. GI|27881412, or can be derived therefrom.

In some embodiments, the food allergen is from egg. Several allergens from egg are known and have been characterized and are encompassed by the present disclosure. For example, the allergen polypeptide can be ovomucoid from Gallus gallus for example galdl GI|124757, GI|209979542 or gald2 GI|63052, GI|129293, ovotransferrin from Gallus gallus, for example, gald3 GI|757851, GI|1351295, lysozyme from Gallus gallus, for example gald4 GI|126608, GI|212279, GI|63581, livetin (chicken serum albumin) from Gallus gallus, for example gald5 GI|63748, apovitellin, phosvitin, or can be derived therefrom. Fragments of ovalbumin comprising tolerogenic epitopes are also considered a composition of the present disclosure.

In some embodiments, the food allergen is from fish. Several allergens from fish are known and have been characterized and are encompassed by the present disclosure. For example, the allergen polypeptide can be Che ag, Lop pi, Gelatin/Ore a, parvalbumin from ocean perch Sebastes marinus, for example Sebml.0101 GI|242253959 or Sebml.0201 GI|242253961; parvalbumin from talapia Oreal, parvalbumin from Pacific pilchard Sardinops sagax Sarsal.0101 GI|193247972, parvalbumin from rainbow trout Oncorhynchus mykiss oncml GI|288559140, glyceraldehyde-3-phosphate dehydrogenase from a number of fish species, or can be derived therefrom.

Environmental allergens (and their general sources) are well-known and many protein components of many environmental allergen sources have been identified and characterized. For example, illustrative and non-limiting sources of environmental allergens include mold, fungus; pollen from trees, grasses, and ragweed; dust mites; glycoproteins in animal dander (e.g., from cat and dog); in insect stings (e.g., bee and wasp); other animal (e.g., reptile) venoms; and other animal allergens known in the art.
In some embodiments, the environmental allergen is from a house dust mite. Several allergens from dust mites are known and have been characterized and are encompassed by the present disclosure. For example, the allergen polypeptide can be from *Dermatophagoides pteronyssinus*, including for example, Derp1 through Derp23, from *Dermatophagoides farinae*, including for example, Derfl through Derf33; from *Euroglyphus maynei*, including for example, (Eurml (GI|3941388, incorporated by reference herein), Eurm2 (GI|3941386, incorporated by reference herein), Eurm3 (GI|42004421, incorporated by reference herein), Eurm4 (GI|5059164, incorporated by reference herein), Eurml4 (GI|6492307, incorporated by reference herein); from *Dermatophagoides microceras*, including for example, Derml (GI|127205, incorporated by reference herein), or can be derived therefrom.

In some embodiments, the environmental allergen is from a cat (*Felis domesticas*). Several allergens from cats are known and have been characterized and are encompassed by the present disclosure. For example, the allergen polypeptide can be a secretoglobin such as Feldl (chain 1 GI|1364212, GI|1364213, GI|163825, GI|169655, GI|14326420; chain 2 GI|395407, GI|163823, each incorporated by reference herein), a lipocalin such as Feld4 (GI|45775300, incorporated by reference herein), an albumin such as Feld2 (GI|886485, incorporated by reference herein), a cystatin such as Feld3 (GI|17939981, incorporated by reference herein), IgA such as Feld5w, or can be derived therefrom.

Plants that produce allergy inducing pollen are typically anemophilous (i.e., have their pollen dispersed by wind) and include ragweed, oak, birch, hickory, alder, ash, and pecan trees, and summer grasses. In some embodiments, the environmental allergen is from a tree. Several allergens from trees are known and have been characterized and are encompassed by the present disclosure. For example, the allergen polypeptide can be Betvl (for example, GI|320545, GI|534900, GI|1321716, GI|1321722, each incorporated by reference herein), Betv2 (for example, GI|157830684, GI|166953, each incorporated by reference herein), Betv3 (GI|488605, incorporated by reference herein), Betv4 (GI|809536, incorporated by reference herein), Betv6 (GI|10764491, incorporated by reference herein), or Betv7 (GI|21886603, incorporated by reference herein) from the European White Birch *Betula pendula*; Alngl (GI|261407, incorporated by reference herein), or Alng4 (GI|3319651, incorporated by reference herein) from the alder *Alnus glutinosa*; Frael (GI|33327133, GI|56122438, GI|34978692, each incorporated by reference herein) from the European ash *Fraxinus excelsior*, or can be derived therefrom.
In other embodiments, the environmental allergen is from ragweed \textit{(Ambrosia artemisiifolia, Ambrosia psilostachya or Ambrosia trifida)}. Several allergens from ragweed are known and have been characterized and are encompassed by the present disclosure. For example, the allergen polypeptide can be Ambal through Ambal 1 (GI|166435, GI|166437, GI|302127812, GI|166411, GI|166443, GI|302127814, GI|302127816, GI|166445, GI|302127824, GI|166447, GI|302127828, GI|416636, GI|291197394, GI|1916292, GI|62249502, GI|62249512, GI|62249470, GI|62249481, GI|62249491, GI|558482540, each incorporated by reference herein) from \textit{Ambrosia artemisiifolia}; Ambp5 (GI|5 15953, GI|5 15955, each incorporated by reference herein) from \textit{Ambrosia psilostachya}; Ambt5 (GI|17680, incorporated by reference herein) from \textit{Ambrosia trifida}, or can be derived therefrom.

Many autoantigens that can cause autoimmune diseases have been identified and characterized and are encompassed by the present disclosure as allergen polypeptides. For example, the autoantigen can be selected from the non-limiting list of a transglutaminase, myelin-associated glycoprotein (MAG; GI|62205282, incorporated by reference herein), CNS-specific myelin oligodendrocyte glycoprotein (MOG; GI|984147, GI|793839, each incorporated by reference herein), myelin basic protein (MBP, GI|1184244, GI|307161, GI|307162, GI|307160, each incorporated by reference herein), proteolipid protein (PLP, GI|41393531, incorporated by reference herein), Zinc transporter-8 (ZnT8, a chain GI|64762489, b chain GI|289803013, GI|289803009, GI|289803007, GI|289803003, each incorporated by reference herein), glutamic decarboxylase 65 (GAD65, GI|352216, incorporated by reference herein), glutamic decarboxylase 67 (GAD67, GI|1352213, GI|385451, each incorporated by reference herein), preproinsulin (GI|758088, GI|389620191, GI|631226408), proinsulin, insulin, tyrosine phosphatase-like autoantigen, insulinoma antigen-2 (IA-2; ICA512, PTPRN; GI|2499754, incorporated by reference herein), IA-2b (Phogrin, PTPRN2; GI|47939489, incorporated by reference herein), islet cell antigen-69 (ICA69; GI|20141584, incorporated by reference herein), chromogranin A (GI|180527, incorporated by reference herein), islet amyloid polypeptide (ppIAPP; GI|4557655, incorporated by reference herein), and heat shock protein 60 (hsp60; GI|77702086, incorporated by reference herein), or can be derived therefrom.

Additionally, allergen polypeptides can be from biological (i.e., protein-based) therapeutic compositions. For example, portions of humanized antibodies such as the
CDRs have been shown to elicit immune responses and, thus, the induction of tolerance to such a therapeutic is desired to maintain the utility of such compositions. Another example is recombinant erythropoietin and other cytokines and therapeutic hormones can elicit immune responses. In addition, other therapeutic proteins can elicit immune responses including for example, growth hormone, interferons, monoclonal antibody therapeutic products, for example Remicade®, Humira®, Simboni®, and the like. Accordingly, the allergen polypeptide can be any of such biological (i.e., proteinaceous) composition, or can be derived therefrom.

Amino acid sequences of illustrative, non-limiting fusion protein constructs that incorporate different exemplary fusion proteins are provided in Table 1.

Table 1: Illustrative fusion protein constructs

<table>
<thead>
<tr>
<th>Fusion name</th>
<th>Allergy/disorder</th>
<th>Amino acid sequence</th>
<th>Comment</th>
<th>SEQ ID NO:</th>
</tr>
</thead>
<tbody>
<tr>
<td>064</td>
<td>Peanut</td>
<td>MGRQQWELOGDRCQSOLERANLRPCEOHLMQ KJQRDESYGRDPYSPQDSPQDPDPRDPYSP SPYDRRGAGSQHQCERCCNELNEFNNQRCMCE ALQQIMENQSDRQLQROEQQOFKRELNLNPQ CGLRAPOCRNELVESGGRDRY GRPGMDPRLREEV VRLIIALTSNGASLSKIGEIGSRLSEALEKTSIQHTILRITQGLDDANKRIIALEQSRSDDLVA VivitJCLASADGSGGAPQYMSKNNIVIEEQQVDGLRL RVEGGGSITHSNKWPMAITVSYPRTSFT</td>
<td>An Arah2-ps1 fusion protein; the N-terminal Arah2 domain is tethered to the C-terminal ps1 domain by the GRPG linker (underlined)</td>
<td>7</td>
</tr>
<tr>
<td>X64</td>
<td>Peanut</td>
<td>MGKSPYRTENPCACORCLSCSQEPDDLQKAC ESRTCKLEYDPDVYDTGATNQHRPPGERTGRQ PGYDDDRQPRREEGGRWGPAEPEREEREDW RQPREDWRPSPHQPPKRKIREGREGEEQTWGTPGS EYREETSNNPPFYPSSRFSTRYGNQNRIVLQR FDORSQFQNLONHRIVQIEARPNTLVLPKHADA DINLVIQQOQATVTAVGNNRKSFLNDEGHALLRI PSGFISYILNRHDNOQLRVAKISMPVNTPGQEDF FPASSRDQQSYLYQFSSRTLEAAFNEFIEIRVL LEENAGGEQEEREQRRRSTRSDDNEGIVKVSKSE HVQELTKHAKSVSKGSSEETYNTPINLDRGPEPDLSNFFGROLFEVKPKKKNPQLDDMMLTCEIKEG</td>
<td>An Arah1-Arah2-Arah3-Arah9-ps1 fusion protein; the N-terminal, concatenated Arah1-Arah2-Arah3-Arah9 domain is tethered to the C-terminal ps1 domain by the GRPG linker (underlined)</td>
<td>8</td>
</tr>
<tr>
<td>Fusion name</td>
<td>Allergy/disorder</td>
<td>Amino acid sequence</td>
<td>Comment</td>
<td>SEO ID NO:</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---------------------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>ALMLPHFSKAMVIVWNKGTGNLLEVAVRKEQQQQQREREQWEEEDDEEEEEEGSNGRVRVRRTARLKEGDVFIMPAAHVPNMARKSHELHLLGFGINAENNHRIPLAGDKDNVIDQIEKQAKDLAPFSGSEQVKEKLIKNQRESHFVSAARPQSQPSSPEKEQEEENQOGKGLPLSILKAFNRRQWELQGRRRCQSLERANLRPCQQLMQKQRDSDSYGRDPSQPSQDPSPQDPRRDPYSPSPYDRRGAGSSQHRERCCNELNEFENQRCMEALQQPMENQSDRLQRORQQQEQFQKRELRLNPQQCGLRAPOCREDLAVSGGDRDYIRSFQQPGEACQFOQRNLNAPRDPDNRESEGYYETWPNNQEFECAOVALSRVLRNRRLNPFPYSNAPQEIIFIQQQGRGYFLIFQPCPRHYEPHTCQRGRSRQQPRPLRLQGEDQSQQORDSHQKVHRDFEGDLIAVPTGAVFALYNHHDRTDWAASLTDTNNNDNQLDQFPRRFNLAGNTEQFLRYYQOSRQSRRSLPYSPSYSPSQPRQTEEREFSRPGQHRSRRERAQEEEMEGNNIFSGFTPEFLEQAFQVDDRQIVQLRGETEESEEGIAVTGVRGLRILSPDRKRRADEEEDYYDEEYYEDDEDRRRRGRSRGRRNGIETECTASAKKNGNRSPDNPYAGSSKLANTANDLNLRLIRLWPSAEYGNLRNALFVAYHNTNAHIYYLRGRHAVAQWSDNSNGNRYVEDEELQEGHVLWQPQNFAVAGKQSENFEYVFAKTDSPSIAVLNAGESVINDLEEWANSGLRQEOARQLKNNPPK£FVPSQSPRAVAISCGQVSNALAPCIFPLTKGAPPAPACCSSVRGLLLGALRTRADQAAACNCLAAAGSRLGLNQGNAALPRGCVSVIPYKISTSTNCATIKFPRGPMEDXRREERWLHIIALTSDNGAISKGLSERVSALERTSQHIHSDTILRITQGLDDANKRIALEQSRDLVAVSVDASQALIRLESSIGALTWNGLDSSVTQLGARVGQLETGLAEQIVDHDNIVARVDTAEARINEGSTLELSTLTLRTVSIQAFDESIRSTLERTAVSTAGASPLSRNRTMMLGNDTLGLSIINLAIRLPQNTGLN IQGGGLQFRFNTDQQFIVNNDLTLKPGTVDNSRIGATEQSYVASAVTPLRNSTKVLDMFIDTSELEINSSGQLVRSTSPNLRYPAPIVASG GIGMSPNYRFQSWMIGIVSYSGSGLNWRVQVNSDIFIVDDYIYICLPAFDGFSIADGGDSLNLNFVTLGGLPPLLTTGDEPAFHNDDWTYGAQTVAILSGGAQPGYMSKLNWVEQWDQVLRLREVEGGISHTSNKWPAMTVSYSRSFT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**MS3** Multiple Sclerosis

<table>
<thead>
<tr>
<th>Amino acid sequence</th>
<th>Comment</th>
<th>SEO ID NO:</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGQFVP1GPRHPIRALVGDEVELPCRISPGKNATGMEVGWYRPFSRWLYRNQKQDDQGQAEPYRGBTLLKDAIEGKVTLRIRNRVRFSDEGGGFTCFFRDHYQEEAAMEKLVEDPFLYVPSGVLVLLAVLPULLQITTILQFLCQLYRLRGKRAEJENLHRTFDPHFLRVPWCKITLFLVPIVPLGPLVALICYNWLLHRLAGQFEELNRNASOKRPSRHSKLYLATSTMDHARHGFPRHRDTGILDSIGRFGGDGRAPKRGSGKDSHIPARTAHYSPLPSKHGRTQDENPVVHFKNINVTTRTPPSOGKRGLSLSRFSGWAGEQPQPGFYGYGRASDYSKASHGFKGVDGAQHTLSKIFKLGGRDSSRSGMPALFCCGHEALTGETKLEYTFSSNYQDYEYLFYTTGAVRQOFJDYKTICGLKSATVTGGQHCLGKWLGHDPKFVGITNTWTCQSIAFPSK</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

An MOG-MBP-PLP (extracellular regions)-pol protein; the N-terminal, concatenated MOG-MBP-PLP (extracellular domain) is tethered to the C-terminal ρη domain by the GRPG linker (underlined).
<table>
<thead>
<tr>
<th>Fusion name</th>
<th>Allersv/disorder</th>
<th>Amino acid sequence</th>
<th>Comment</th>
<th>SEO ID NO:</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSASIGSLCDARVYGLPWNAAFPGKVCSSNLSSICGRPGMDPLREFWRLIALTSNDAGSLPKGLESRRVSALETSQIHDSTLIRIQGGLDANKRIAIEQSRDDLVASVSDAQLAISRELESSIGLQTVVNGLDSVTQLGARVQLETGLAEVRHDMLVARVIDAERNIGSLTELSTLRLRVTASIAAGPLS RKNNRTMLGLNDGLTLSGNLARLPNTGGLNINQNGGQFRFRONTIQVVINNLKTLKTSFDISNRGATEQSYVSAVSTPRLRNSSTKVLDMIDSSSTLEINSQGGTVRTSTPNLRYPADVGGIGMSPNYNRFQSMWIGVYSYGGLNWVQVNSDFIVDDYIIHICPLAFDGFSIA DGGDLSNFVTGLPLPLTGTDEPAFHHDWTVTQAIGLSGGAPQYMSKNLWDDQGVLRLRVEGGGITHSNKWPAMTVSYPRTF</td>
<td>A six T-cell epitope (MOG, MBP, and PLP)-ρζ fusion protein; the N-terminal, domain of 6 concatenated epitopes (MOG, MBP, and PLP) is tethered to the C-terminal ρζ domain by the GRP linkage (underlined)</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MGQFRVYGRPHIRALVGDEVLEMGEVGYWYPFPRSVRVIHYRNXGDKENPWHFKNIVTPRTGPDVQAQGTLSKIKFLGRDSRSGBPMTGKEKLTFTSFKNODYEHCMLKWLHGNDFKVGITGRPGMDPLREEWRLIIALTSDAGLSKGLERSVSALSEKTSIQHSDTLRITIQGLDANKRIAEQSRDDLVASVSAQLAISRELESSIGLQTWNGLDSSVTQLGARVQLETGLAEVRHDMLVARVIDAERNIGSLTELSTLRLRVTASIAAGPLSRKNNRTMLGLNDGLTLSGNLARLPNTGGLNINQNGGQFRFRONTIQVVINNLKTLKTSFDISNRGATEQSYVSAVSTPRLRNSSTKVLDMIDSSSTLEINSQGGTVRTSTPNLRYPADVGGIGMSPNYNRFQSMWIGVYSYGGLNWVQVNSDFIVDDYIIHICPLAFDGFSIA DGGDLSNFVTGLPLPLTGTDEPAFHHDWTVTQAIGLSGGAPQYMSKNLWDDQGVLRLRVEGGGITHSNKWPAMTVSYPRTF</td>
<td>A human proinsulin-ρζ fusion protein; the N-terminal, human proinsulin domain is tethered to the C-terminal ρζ domain by the GRP linkage (underlined)</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MFVNHQICGLSHLVEALYLVREGGRFFYTPKRTREAEDLQVQVGLGGPGAGSLQPLALEGSLQKRGVEOCCTSICSLYOLENCGNRPMDPREREWLIALTSNDAGLSKGLERSVSALEKTSIQHSDTSLRITQGLDANKRIAEQSRDDLVASVSAQLAISRELESSIGLQTWNGLDSSVTQLGARVQLETGLAEVRHDMLVARVIDAERNIGSLTELSTLRLRVTASIAAGPLSRKNNRTMLGLNDGLTLSGNLARLPNTGGLNINQNGGQFRFRONTIQVVINNLKTLKTSFDISNRGATEQSYVSAVSTPRLRNSSTKVLDMIDSSSTLEINSQGGTVRTSTPNLRYPADVGGIGMSPNYNRFQSMWIGVYSYGGLNWVQVNSDFIVDDYIIHICPLAFDGFSIA DGGDLSNFVTGLPLPLTGTDEPAFHHDWTVTQAIGLSGGAPQYMSKNLWDDQGVLRLRVEGGGITHSNKWPAMTVSYPRTF</td>
<td>A human proinsulin-ρζ fusion protein; the N-terminal, human proinsulin domain is tethered to the C-terminal ρζ domain by the GRP linkage (underlined)</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEFLELTYLVNDAKAMKYAFTLESVELQQKPVNDQCPRRPREELESGLGMYCHSGSKPTEGANEYAYAKWKLCSASAICIFMIAEWGGHIAGSLAWTDAAHLLIDTSLFSLFLWSKPPSKRLTGFWHRAELGALLSILCIWWTGVLVYLAERLLYPDYQ</td>
<td>A ZnT8-GAD65-human proinsulin-ρζ fusion protein; the N-terminal,</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fusion name</td>
<td>Aller/s disorder</td>
<td>Amino acid sequence</td>
<td>Comment</td>
<td>SEO ID NO:</td>
</tr>
<tr>
<td>-------------</td>
<td>------------------</td>
<td>---------------------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>IQATVMIHVSSCAAVANIVLTLWHTQRCLGHNHKEVQAANASVRAFVHALGDLQSFVSLISALITYFKPEYKIAIDPCTIFISILVLASTITLKDPSILMEGVPSKLNYSVGVKELIADVGLVSVHLSIHWSTLMQVILS</td>
<td>ZincT8-GAD65- proinsulin domain is tethered to the C-terminal ρσ domain by the GRPG linker (underlined)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLFCEDPCDASCPSGFWSFSGEDSGDSENPGTARAWRCVQAVQFTGGGIGNKLCALLYGDAEKAPEASGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SQQPPRAAARKAAACACDQKPCSCSKVDVNYAFLHTARDLPCADGERPRTLAFLODVIMILLQYWKSFDRSTKVIDFHYPNELLQEQYNYWELADQPNLEEILMHCQTTLKYAIKTHGPFRYFNQLSTGLDMVGLAADDVLTSTANTMFTYEIAPFIVLLEYVTLLKMKREIGWPNGSSGDPGSPGGAISNYMYAMMARFKMPEVK</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VEGKAALPLRLIAFTSEHSHFLKKGAAALGIHTDSVILICDERRMGPSDLERRLAEAKQKGFPFLVSATGGTTYVYAFDPPLLAVADICKKYKIMHDVDAAWGGGLLSMRKHKWKLSGVERANSVTWNPMKHGPSLPQCSALLVREEGLMQCNQMHASYLFFQDKH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YIDLSYDTGKDALKCGRHVDVFKLWLMRAGKTTFGEAHDVKCLELAEYLNYIKNREGYEMVFDFGKPQHTNVCFWYIPPSLRTLEDNEERMSRLSKVAPVISARKMEYGTMTMVSYQPLDGKVFNNFRMVISNPAATHQDIDFLLEEEIRLGDQLFNVQHLCSGSHVEALYLVCGERGFFYPPTKRREAELQVQGVELPGGPA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSLQPQALEGLQSKRGIVEQCTSCISLYQULENYCNGRPQMDPRLREERWRLIALTSDNGASLKGLEVSVALEKTSQHSTDLRITQGGLDAKNIKIALEQS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDDVLASVSDAQALAIRSLESSIGALTQWNLDSLSTVQLARVGQLQFTGLARVHDHLVARVTDATERNILTTLSLTTLRTLQFOSTDTERIATV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAGAPLIRKNNRTMGMCLUDLTSGNMLAIRLGHNTLNOQGNGQQRFFNQDFQVQNNLTLKTTTV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSINSPIGATEQSYVASAVTPRLNSSTKVLDMILSDSTIELSSQETVRSTSPNLYRIADVSGGIGMSPYNFRFQSMWIGIVSYSGLSNWRVQVNSDFIVDPHYHIICLPADGFAPAAGDGLSLNFLVTGGLPLLWTG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM2 Dust Mite</td>
<td>MRPSIKPFEYWKKAFNKSYATFEDDEAARKNFLESVYQNSNGAINHLDLSLDEFFKRNFLMSAEAFHLEKTQFDMALNANTCNSINGNAPEALDRQHSMRTVTPIRMQGGCGSCWAFSGGAATESAYLARYNQSL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DLAEQELVDACSQHCHGDITPRGIEYIQHNGWQESYYRYVAREQSCRRPANQRFGINSNCYQIPPNKIREAPQTHSAIIVIGKDLDAFRHYDGRTIQIRDNGYPNYHAVNVGYSNAGQVDYWIVRNSWDTNWGDNQGYYFAANIDLMIMIEPPWILKFLALASSLVSTVYARPASIKTFEEKFKAFKNYATVEEVEARKNFLESLYVEANKGAITHLDSLDEEFKRNLYLMSAEAEELQKTQFLNAETSACRINSNVNPSELDDLRLTRTVTPIRMQGCGSCWAFSGVATESTAYLAVNSLDSELEQVLDCAQHQCHGDITPRGIEYIQQNGWERSYPYVAREQCRPRPSQH</td>
<td>ADerpl-Defl fusion protein; the N-terminal, Derpl-Defl domain is tethered to the C-terminal ρσ domain by the GRPG linker (underlined)</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fusion name</td>
<td>Allergy/disorder</td>
<td>Amino acid sequence</td>
<td>Comment</td>
<td>SEO ID NO:</td>
</tr>
<tr>
<td>------------</td>
<td>----------------</td>
<td>---------------------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>GT7</td>
<td>Celiac Disease</td>
<td>YGISNYCQYPPDVQJREALQTHAIAVIGKDLRAFQHYDGRITIQHNDGYQPQNYHA/NIVGVGYSTQGDDYWIVRNSWTDGSGYGFQAGNMLMMIEOYPYWIMGRPGMDPRLREEWRLIAATLDNGASLSKGLERSVSALEKTQIHSITLIRITQGDLDANKRIIALEQSRDDLAVSVDQQLAISRELESSIGALQTWTNGLDSSVTQLGARVQGQLETGLAEVRDHNLRVARVTDARNISSLTTESLTLTRVSTIQADFESRILTERTAVTSAGAPSIRNRMRTMGLDGSLGTTSRPRNLRYPIADVSGGIGMPNYRFQSMWIGIVSYSGSGLNRVQVSNSDIFIVDDYIHICLPADFGSIAADGDSLNFVTVGLLPPLTGTDEPAFHNVDVTVYGAQTVAIAGLSGGAPQYMSKLNWVEQWQDGVLRLRVEGGSITHSNKSWPAMTVSYPRTFT</td>
<td>A seven T cell gluten epitope-pref fusion protein; the N-terminal, seven T cell gluten epitope (concatenated epitopes from α-gliadin, ω5-gliadin, Hordein, and Secalin) domain is tethered to the C-terminal pref domain by the GRPG linker (underlined)</td>
<td>14</td>
</tr>
</tbody>
</table>

It will be appreciated that the representative allergen polypeptides and their sources are non-limiting examples and that any known allergen polypeptide is encompassed by the present disclosure. Further, as described above, the particular allergen polypeptide or polypeptides incorporated into the disclosed fusion protein need not be the full length polypeptide from the allergen source, but instead may be "derived therefrom". In some embodiments, the polypeptide is a subcomponent, such as a fragment or fusion of multiple fragments, of the full-length source protein. The incorporation of such derivatives can be advantageous for purposes of production of the fusion protein. In this regard, recombinant expression of the fusion protein can be more efficient for smaller overall proteins, or can be enhanced with the exclusion of...
particularly problematic domains of the source protein. Furthermore, in some instances the resulting fusion protein will be more effective at inducing tolerance because the fusion protein contains the one or more critical antigens/epitopes while excluding other domains that may diminish the tolerization effect.

In instances where the intended source allergen is not incorporated in its entirety, but rather fragments thereof are used as the allergen polypeptide in the fusion protein, the selection of fragments as the allergen polypeptide can be made based on various parameters. For instance, the allergen polypeptide preferably comprises an MHC Class I and/or MHC Class II epitope (also referred to as a T cell epitope). Such epitopes are short, linear lengths of polypeptides that MHC molecules can process and present to T cells. Cells in the mucosa, such as in the GALT and the NALT regions, express both MHC Class I and II, and can play a role in tolerization to antigens. Epitopes presented by MHC class I molecules are typically peptides between 8 and 11 amino acids in length, whereas MHC class II molecules present longer peptides, 13-17 amino acids in length.

Accordingly, the allergen polypeptide will typically comprise at least 8 amino acids. However, it will be appreciated that the polypeptide can be much larger, limited only by the ability of the expression or synthesis system to produce the final fusion protein. Specific MHC epitopes can be readily predicted from the selected source allergen protein sequence. As indicated, the lengths of the typical MHC epitopes are known.

Furthermore, MHC Class I and MHC Class II epitopes have characteristic anchor points that rely on generalized sequence patterns. Thus, algorithms exist to predict the MHC epitopes from a source sequence. See, e.g., Koren, E., et al., "Clinical validation of the "in silico" prediction of immunogenicity of a human recombinant therapeutic protein," *Clinical Immunol.* 124:26-32 (2007), incorporated herein by reference. Many useful applications are available on the world wide web to apply various prediction algorithms to provided source sequences. For example, the Immune Epitope Database (IEDB) and Analysis Resource provides a website at the address iedb.org, which funded by a contract from the National Institute of Allergy and Infectious Diseases. This resource offers easy searching of experimental database with data characterizing known T cell epitopes (presented via MHC) as studied in humans, non-human primates, and other animal species. Epitopes involved in allergy, autoimmunity, and transplant are included. This resource also hosts tools to assist in the prediction and analysis of B cell and T cell epitopes. With the application of such an algorithm to any of source allergen protein
sequence, such as the illustrative source allergen proteins described above, a person of ordinary skill in the art can readily select the best epitope(s) to include in the one or more allergen polypeptide(s) that is ultimately incorporated into the fusion protein.

As an example, Arah2 peptides containing dominant CD4+ T cell epitopes are known in the art. See for example, Prickett S. R., et al., "Arah2 Peptides Containing Dominant CD4+ T-cell Epitopes: Candidates for a Peanut Allergy Therapeutic," J. Allergy Clin. Immunol. 727:608-615 (2011) and Glasllo I. N., et al., "Characterization of the T-cell Epitopes of a Major Peanut Allergen, Ara h 2," Allergy (50):35-40 (2005), incorporated herein in their entirety. Prickett et al. disclose five dominant CD4+ T-cell epitopes including aa32-44 (SQRERANLRPCEQ; SEQ ID NO:2), aa37-47 (ANLRPCEQHLM; SEQ ID NO:3), aa91-102 (ELNEFENNQRCM; SEQ ID NO:4), aa95-107 (FENNQRCM; SEQ ID NO:5), and aa28-141 (RELRLNPQQCGLRA, SEQ ID NO:6). In combination these epitopes were presented by HLA-DR, HLA-DP and HLA-DQ molecules and recognized by T cells from all of the subjects tested. Any fusion polypeptide of the present disclosure would include at least one and likely more than one T cell epitope.

T cell peptide epitopes are also known for a-gliadin and include, for example, and not by limitation, a 33 amino acid sequence comprising aa56-88 to contain six partly overlapping copies of three DQ2-restricted T cell epitopes. See, for example, Shan, L., et al., "Structural Basis for Gluten Intolerance in Celiac Sprue," Science 297:2275-2279 (2002) and Qiao, S.W., et al., "Antigen Presentation to Celiac Lesion-Derived T Cells of a 33-mer Gliadin Peptide Naturally Formed by Gastrointestinal Digestion," J. Immunol. 773:1757-1759 (2004).

The fusion protein can also include various tags that can assist the expression, production, or later analysis (e.g., visualization) thereof. Such tags are well-known and are commonly used in the art during the production of recombinant fusion proteins. Tags can be attached at the N- or C-terminus of the antigen construct but are usually placed at the N-terminal end. Examples of tags are: NusA, thioredoxin, maltose binding protein, small ubiquitin-like molecules (Sumo-tag), and His-repeats. If desired, to facilitate removal of the tag during purification, a unique protease site can be inserted between the tag and the fusion protein per se. Such protease sites may include those for thrombin, factor Xa, enterokinase, PreScission™, Sumo™. Alternatively, removal of the tag may be achieved via inclusion of an intein sequence between the tag and the fusion protein per
Inteins are self-cleaving proteins and in response to a stimulus (e.g., lowered pH) are capable of self-splicing at the junction between the intein and the antigen construct, thus eliminating the need for the addition of specific proteases. Examples of inteins include domains derived from Mycobacterium tuberculosis (RecA), and Pyrococcus horikoshii (RadA) (Fong, et al., Trends Biotechnol. 25:272-279 (2010)).

To facilitate purification, the fusion protein can include one or more purification tags to enable specific chromatography steps (e.g., metal ion chelating, affinity chromatography) to be included in the purification processes. Such purification tags can, for example, include: repeat histidine residues (e.g., 6-10 histidine residues), maltose binding protein, glutathione S-transferase; and streptavidin. These tags can be attached at the N- and/ or C-terminus of the polypeptide antigens of the invention. To facilitate removal of such tags during purification, protease sites and/ or inteins (examples above) can be inserted between the polypeptide and the purification tag(s).

The fusion protein can also include a visualization tag. For example, this tag can include portions of proteins that are known to provide a detectable signal, such as fluorescence. Alternatively, any tag herein can provide an epitope for specific recognition and binding by a detectably labeled antibody or antibody fragment, or any other molecule capable of emitting detectable light or energy. Exemplary tags that can provide a detectable signal include GFP, any of the numerous related GFP variants known in the art to similarly fluoresce upon stimulation, such as blue fluorescent protein, cyan fluorescent protein, and yellow fluorescent protein, mCherry, and the like. The visualization tag can also serve as an epitope for binding and isolation of the fusion protein.

In another aspect, the present disclosure provides a pharmaceutical composition comprising the isolated fusion protein described herein. The pharmaceutical composition can also comprise pharmaceutically acceptable carriers, stabilizers, excipients, and other additives to provide an appropriate formulation for the preferred route of administration, as is familiar in the art. Generally, oral and intranasal routes of administration are addressed herein, but other known routes of administration are contemplated as well. An exemplary formulation for intranasal administration can include components to facilitate inhalation and delivery to the mucosal surface. For example, such formulations can include aerosols, particulates, and the like. In general, the goal for particle size for inhalation is about 1 µm or less. Such formulation can be delivered by in the form of an
aerosol spray. Oral formulations may be liquid (for example, syrups, solutions, or suspensions), or solid (for example, powders, pills, tablets, or capsules). For solid compositions, conventional non-toxic solid carriers can include pharmaceutical grades of mannitol, lactose, starch, or magnesium stearate. Actual methods of preparing such dosage forms are known, or will be apparent, to those of ordinary skill in the art. Solid formulations for oral administration can also comprise known binding agents, fillers, lubricants, disintegrants, or wetting agents. The dose form can also be coated. Liquids for oral administration can contain additional additives such as suspending agents, emulsifiers, non-aqueous vehicles, and preservatives. In another aspect, the disclosure provides a nucleic acid encoding the isolated fusion protein described herein.

As used herein, the term "nucleic acid" refers to any polymer molecule that comprises multiple nucleotide subunits (i.e., a polynucleotide). Nucleic acids encompassed by the present disclosure can include deoxyribonucleotide polymer (DNA), ribonucleotide polymer (RNA), cDNA or a synthetic nucleic acid known in the art.

Nucleotide subunits of the nucleic acid polymers can be naturally occurring or artificial or modified. A nucleotide typically contains a nucleobase, a sugar, and at least one phosphate group. The nucleobase is typically heterocyclic. Canonical nucleobases include purines and pyrimidines and more specifically adenine (A), guanine (G), thymine (T) (or typically in RNA, uracil (U) instead of thymine (T)), and cytosine (C)). The sugar is typically a pentose sugar. Suitable sugars include, but are not limited to, ribose and deoxyribose. The nucleotide is typically a ribonucleotide or deoxyribonucleotide. The nucleotide typically contains a monophosphate, diphosphate, or triphosphate. These are generally referred to herein as nucleotides or nucleotide residues to indicate the subunit. Without specific identification, the general terms nucleotides, nucleotide residues, and the like, are not intended to imply any specific structure or identity. The nucleotides can also be synthetic or modified.

In another aspect, the disclosure provides vectors comprising the nucleic acid sequences described herein, such as a vector comprising a nucleic acid sequence encoding the polypeptide described above. Such vectors are useful for the recombinant expression of the fusion protein in a cell-based expression system. Such expression systems are well-known in the art, and include cell strains optimized for recombinant expression of genes associated with specific vectors parameters. For example, any vector
described herein can further comprise a promoter sequence to facilitate expression of the nucleic acid encoding the fusion protein in the intended cellular expression system. Any appropriate promoter can be used, such as a constitutive promoter or inducible promoter, appropriate for the expression system to be used, as known in the art. For example, an inducible promoter can comprise an acetamide-inducible promoter. Additionally, the vector can also include selectable markers, such as antibiotic or toxin resistance genes, that will confer protection against such applied agents. In this manner, cells that are successfully transformed with the operational vector can be retained in culture and the non-transformed cells in the system can be removed.

Also provided are cultured cells transfected with any vector described herein, or progeny thereof, wherein the cell is capable of expressing a fusion protein, as described above. The cell can be prokaryotic, such as \textit{E. coli}, or eukaryotic, such as insect or mammalian.

In another aspect, the present disclosure provides a method for inducing tolerance to a protein allergen. The method comprises administering a pharmaceutically effective amount of the isolated fusion protein or the pharmaceutical composition, as described herein, to a subject in need thereof. The fusion protein comprises a polypeptide derived from the protein allergen to which tolerance is desired. Therefore, the fusion protein need not necessarily comprise the entire protein allergen. It is preferable, however, that the fusion protein, and specifically the allergen polypeptide, comprises the most reactive epitopes of the protein allergen to induce a more comprehensive tolerance to the allergen.

In some embodiments, the method consists of administering a single dose of the effective amount of the isolated fusion polypeptide. In other embodiments, the method can further comprise a second, third, fourth, or more additional administrations. In embodiments with multiple administrations, each administration need not contain the same dose. Furthermore, in some embodiments, each administration need not contain the same fusion protein, but can contain additional or different allergen polypeptide(s).

Illustrative, non-limiting effective doses of isolated fusion polypeptide include less than about 100mg, 75mg, 50mg, 25mg, 20mg, 15mg, 10mg, 9mg, 8mg, 7mg, 6mg, 5mg, 4mg, 3mg, 2mg, 1.5mg, 1mg, 750 \mu\text{g}, 500 \mu\text{g}, 250 \mu\text{g}, 100 \mu\text{g}, 75 \mu\text{g}, 50 \mu\text{g}, or 25 \mu\text{g}, or any number or range therein.

In another aspect, the disclosure provides a method for screening a subject to provide a personalized fusion protein to maximize the tolerization to an allergen or
autoantigen by the individual. The method includes obtaining peripheral blood mononuclear cells (PBMCs) from the subject. This can involve affirmatively obtaining a blood sample and isolating the PBMCs. The isolated PBMCs are contacted with an isolated candidate antigen, either whole or a substantial fragment (portion) thereof. The PBMCs are monitored for T cell proliferation. In some embodiments, PBMC fractions can be exposed separately to a panel of candidate allergens/antigens, or a panel of different fragments of one or more candidate allergens/antigens. The antigen/allergen, or fragment thereof, that elicits a strong proliferation of T cells in the proliferation assay is chosen for inclusion in the fusion protein to be administered to the subject from whom the PBMCs were obtained. As an example, a patient with multiple sclerosis (MS) can be tested for an appropriate therapeutic fusion protein. PBMCs can be exposed to myelin basic protein and myelin oligodendrocyte glycolprotein (MOG), fragments thereof, various fusions of fragments thereof, or any other known antigen that is suspected to contribute to MS. The antigens that elicit the greatest T cell proliferation can be incorporated into a therapeutic fusion protein, as described herein, for an enhanced treatment personalized to the unique characteristics of the patient's own PBMC population. As another example, PBMCs from a patient suffering from a peanut allergy can be exposed to various known proteins from peanut, fragments thereof, or fusions of various fragments thereof. The reactivity of the PBMCs against the panel peanut allergens can be monitored in a T cell proliferation assay, and only the antigen polypeptide(s) eliciting a high reactivity with the PBMCs can be incorporated into one or more fusion protein constructs, as described herein. Accordingly, the patient will only receive one or more fusion protein constructs incorporating the most highly reactive allergen polypeptides for that subject.

It is noted that, as used herein, the use of the term "or" in the claims means "and/or" unless explicitly indicated to refer to alternatives only or the alternatives are mutually exclusive, although the disclosure supports a definition that refers to only alternatives and "and/or."

Following long-standing patent law, the words "a" and "an," when used in conjunction with the word "comprising" in the claims or specification, denotes one or more, unless specifically noted.

The practice of the present disclosure employs, unless otherwise indicated, conventional immunological and molecular biological techniques and pharmacology
within the skill of the art. Such techniques are well-known to the skilled worker, and are explained fully in the literature. See, e.g., Coligan, Dunn, Ploegh, Speicher and Wingfield "Current Protocols in Protein Science" (1999), Volume I and II (John Wiley & Sons Inc.); Sambrook et al, "Molecular Cloning: A Laboratory Manual" (1989), 2nd Edition (Cold Spring Harbor Laboratory Press); and Prescott, Harley and Klein "Microbiology" (1999), 4th Edition (WBC McGraw-Hill). Additionally, such considerations as routes of administration, antigen dose, number, frequency of administrations, and appropriate formulations are all matters of optimization within the scope of the ordinary skill in the art.

All publications, patents, and patent applications cited herein, whether supra or infra, are hereby incorporated by reference in their entirety. However, publications mentioned herein are cited for the purpose of describing and disclosing the protocols, reagents, and the like, which are reported in the publications and which might be used in connection with the invention.

Unless the context clearly requires otherwise, throughout the description and the claims, the words "comprise," "comprising," and the like are to be construed in an inclusive sense as opposed to an exclusive or exhaustive sense; that is to say, in the sense of "including, but not limited to." Words using the singular or plural number also include the plural and singular number, respectively. Additionally, the words "herein," "above," and "below," and words of similar import, when used in this application, shall refer to this application as a whole and not to any particular portions of the application.

Disclosed are materials, compositions, and components that can be used for, in conjunction with, in preparation for, or are products of the disclosed methods and compositions. It is understood that, when combinations, subsets, interactions, groups, etc., of these materials are disclosed, each of various individual and collective combinations is specifically contemplated, even though specific reference to each and every single combination and permutation of these compounds may not be explicitly disclosed. This concept applies to all aspects of this disclosure including, but not limited to, steps in the described methods. Thus, specific elements of any foregoing embodiments can be combined or substituted for elements in other embodiments. Additionally, it is understood that the embodiments described herein can be implemented using any suitable material such as those described elsewhere herein or as known in the art.
The following examples provide illustrative, non-limiting descriptions of experimental approaches for creating, testing, and using the disclosed tolerance therapeutic.

EXAMPLE 1

Introduction: This example describes an exemplary approach for producing a fusion protein that can induce tolerance to polypeptide antigen, such as a peanut-derived polypeptide allergen.

Experimental Design: A cDNA encoding peanut allergen, Arah2, was synthesized with appropriate restriction sites and cloned into a yeast expression vector generating the Arah2-pol fusion protein (also referred to herein as fusion 064; see Table 1). The construct includes a poly-histidine tag for affinity purification (see, e.g., FIGURE 1). After initial characterization, expression was scaled up and material was purified and characterized by SDS-PAGE and Western blot using antibodies to ρσι, c-myc, and Arah2. Functional activity of arah2-pol was demonstrated in vitro using both L-cell and HeLa cells binding assays.

Vector Construction: The complete sequence encoding Arah2 fused to ρσι and optimized for yeast codon optimization was synthesized by Life Sciences (Gene Art) and cloned into the P. pastoris expression vector pICZ. The sub-cloning places the expression of the fusion protein under control of the alcohol oxidase 1 promoter allowing induction by methanol. In addition, the vector incorporates carboxy terminal poly-histidine and myc epitope tags. (See, e.g., Stanley, J.S., et al. "Identification and mutational analysis of the immunodominant IgE binding epitopes of the major peanut allergen Arah2," Arch. Biochem. Biophys. 342:244-253 (1997), incorporated herein by reference in its entirety).

Protein Expression and Purification: A yeast clone expressing the Arah2-pol fusion protein was expanded in minimal medium (yeast nitrogen base + amino acids) containing glycerol as the sole carbon source. Large-scale cultures were inoculated in minimal medium with 0.5% methanol as the carbon source and incubated for 72 hours at 20°C, with additional methanol added at 24 and 48 hours. Cells were collected by centrifugation, washed once with PBS and stored at -80°C. Cell pellets were thawed and resuspended in one-tenth the original culture volume of cold lysis buffer (8M urea, 300mM NaCl, 10mM Imidazole, 6mM 2-mercaptoethanol, 1% Triton-x 100). The cell
suspension was mixed with and equal volume of glass beads (0.5 micron diameter) and processed in a BioSpec Bead Beater using an ice jacket. Cells were disrupted with eight to ten one minute bursts with one minute cooling intervals. The cell lysate was collected from the glass beads and centrifuged at high speed. The cleared supernatant was applied to an immobilized metal affinity resin (HisPur immobilized cobalt resin, Thermo Fisher) to purify the fusion protein. Cobalt resin was equilibrated in the lysis buffer by washing with three bed volumes. Crude lysate was applied to the resin incubated on ice for 30 minutes with frequent mixing. The resin was washed with three washes of two times the bed volume of lysis buffer, followed by seven washes with lysis buffer containing 0.01% Triton-X 100. The bound protein was eluted with three washes of 1 bed volume of elution buffer (8M urea, 500mM NaCl, 10mM Imidazole, 6mM 2-mercaptoethanol, 0.01% Triton-x 100). Purified protein was dialyzed against at least two 1000-fold volumes of refolding buffer (100mM Arginine, 10% glycerol, 5mM reduced glutathione, 0.5M NaCl, in phosphate buffered saline). At various stages, the fusion protein was analyzed by western blot (anti-ψσ, anti-myc, and anti-Arah2) and SDS PAGE followed by staining with coomassie blue. Arah2-pol protein purification was >4 mg/L under these conditions. See FIGURE 2.

(3x104 cells) were incubated with or without 20μg Arah2-pol for 30 minutes on ice, and following wash, rabbit polyclonal anti-ρσ1 or commercially available normal rabbit serum was incubated for 30 minutes on ice. Following wash, FITC-labeled goat-anti-rabbit IgG (Jackson ImmunoResearch Laboratories) was incubated for 30 minutes on ice. Following wash, cells were analyzed using flow cytometry to confirm the functional activity of Arah2-pol via head region (binding to L-cells) and tail sialic acid binding region (HeLa cells) binding. See FIGURES 3A and 3B.

EXAMPLE 2

Introduction: This example describes an exemplary approach for determining the optimal oral dose of a fusion protein produced as described in Example 1.

Experimental Design: Determination of the optimal oral dose of a fusion protein comprising the peanut allergen polypeptide Arah2 is described. The study can also include control animals that are dosed with the individual fusion components, Arah2 and ρσ1, to demonstrate that such proteins do not generate efficacy at the highest dose of Arah2-pol fusion protein used.

1) Establishment of the mouse peanut model: The whole peanut extract (WPE) is prepared from steam blanched raw peanuts as previously described. See e.g., Kroghsbo, S., et al., "Assessment of the Sensitizing Potential of Processed Peanut Proteins in Brown Norway Rats: Roasting Does Not Enhance Allergenicity," PLoS One 9:e96475 (2014), incorporated herein by reference in its entirety. To sensitized (S) the animals, mice (5/group) are dosed by intragastric (IG) gavage with 6mg WPE and 15μg cholera toxin (CT) per mouse on three consecutive days (0, 1, 2) followed by weekly doses on days 7, 14, and 28 (see FIGURE 4). Subsequently, mice are treated (T) with either a control reagent or Arah2-pol fusion protein. On day 35, mice can receive the WPE challenge (C) in one of three routes of administration, oral, peripheral and systemic, to measure the induction of tolerance against Arah2. For oral challenge, mice can receive 15mg of WPE IG, and serum is collected 24 hours later for analysis. For peripheral challenge, mice are injected with 10μg of WPE in the left ear and PBS in the right ear, and ear swelling is measured. For systemic challenge, 1 mg of WPE can be administered IP, and mice are evaluated for anaphylaxis.

2) Determination of the optimal oral dose of the Arah2-pol fusion protein: Groups of five mice can be treated orally with either PBS or increasing doses of the
Arah2-pol fusion protein at 10, 50, 100, and 500μg per mouse and characteristics can be observed, such as levels of anti-PNA IgE, levels of anti-PNA IgG, levels of Histamine, degree of ear swelling, Anaphylactic Score, and change in body temperature, as described in more detail below.

Serum IgE/IgG/histamine: Serum peanut-specific IgE can be measured by sandwich ELISA. For example, 96-well plate Maxi-Sorp plates is coated with 2μg/ml purified rat anti-mouse IgE Ab (BD Pharmingen) in PBS overnight at 4°C. Coated plates are then washed and blocked with 10% normal serum 1% BSA/PBS/0.05% Tween 20 for 1 hour at 37°C. After washing, serial diluted serum samples are added and incubated for 2 hours at room temperature. Subsequently, biotin conjugated CPE is added and incubated 1 hour at 37°C. After washing, HRP-streptavidin is added and incubated 30 minutes at room temperature, followed by a tetramethylbenzidine substrate. The reaction can be stopped with 2M H₂SO₄ and absorbance is measured at 450nm. Results can be expressed relative to placebo-treated controls.

Serum peanut-specific IgG can be measured by a modified direct ELISA. For example, 96 well plates are coated with 20μg/ml WPE over night at 4°C, washed, and blocked with 1% BSA/PBS/0.05% Tween 20 for 1 hour at 37°C. After washing, serial diluted serum samples is added and incubated for 2 hour at room temperature and washed. Biotinylated anti-mouse IgG is added, incubated for 1 hour at 37°C and washed. The remaining assay can be run as for IgE detection. Results can be expressed as Log2 of end point titer.

The levels of serum histamine can be measured using a commercially available ELISA kit following the manufacturer protocols.

Ear swelling: Ear thickness can be measured 3 hours after challenge using a digital micrometer. Swelling can be calculated by subtracting the thickness of the PBS treated ear from the ear injected with WPE.

Anaphylactic clinical score: The anaphylactic score can be determined using the 0-5 criteria score as outlined in Table 2. Body temperature can be measured with a rectal thermometer 40 minutes after challenge.
Table 2: Anaphylactic clinical scoring

<table>
<thead>
<tr>
<th>Score</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No Clinical Symptoms</td>
</tr>
<tr>
<td>1</td>
<td>Repetitive mouth/ear scratching and ear canal digging with hind legs</td>
</tr>
<tr>
<td>2</td>
<td>Decreased activity; self-isolation, puffiness around eyes and/or mouth</td>
</tr>
<tr>
<td>3</td>
<td>Periods of motionless for more than 1 min; lying prone on stomach</td>
</tr>
<tr>
<td>4</td>
<td>No response to whisker stimuli; reduced or no response to prodding</td>
</tr>
<tr>
<td>5</td>
<td>Endpoint: tremor; convulsion; death</td>
</tr>
</tbody>
</table>

**Expected Outcome:** It is expected that the optimal effective dose of Arah2-pol fusion protein will be 5C^g. It is also anticipated that significant induction of tolerance to WPE will be observed across all parameters evaluated: serum IgE, histamine, DTH and systemic anaphylaxis.

**Alternative Strategies:** A peanut allergy model in mice is well-established. If needed, the sensitization protocol described above can be adjusted by varying the number and size of the WPE dose, as well as the amount of CT. In addition, it is possible that one administration of Arah2-pol fusion protein is not sufficient to induce statistically significant tolerance at any doses tested. If this is the case, the number of Arah2-pol fusion protein treatments at the planned doses will be extended to determine the optimal dose.

**EXAMPLE 3**

**Introduction:** This example describes an exemplary approach for validating that oral administration of a fusion protein, such as Arah2-pol fusion protein, to induce tolerance to a polypeptide allergen, such as a peanut allergen, provides optimal efficacy over nasal administration.

**Experimental Design:** While oral administration of Arah2-pol fusion protein is a primary approach, it is possible that intranasal administration will be more effective at inducing tolerance. Thus, the efficacy of Arah2-pol fusion protein following oral and intranasal administration can be directly compared. The same protocols and assays as described above in Example 2. After sensitization, mice (5 per group) can be administered PBS and Arah2-pol fusion protein either orally or intranasally at the
optimal dose determined in Example 2. Oral, peripheral, and systemic challenge can be carried out, and IgE, histamine, DTH, and anaphylaxis (change in body temperature) responses will be measured as described above for each treatment. This will generate an efficacy data set that will allow the direct comparison of the two routes.

Expected outcome: It is expected that oral administration will be as effective as intranasal administration, supporting the preferred route of administration. However, if intranasal treatment is significantly more effective at inducing tolerance to WPE than oral administration, the h 2-ρσ1 fusion protein can be further developed specifically for intranasal administration according to known techniques.

Statistical Analysis: Data can be expressed as mean ± SEM. Significant differences between 2 groups can be determined using an unpaired 2-tailed t test, and differences among multiple data sets can be determined by ANOVA with standard post-hoc testing. Analysis can be carried out using the appropriate software, such as PRIZM (Graphpad).

Summary: Expected completion of these Examples of this application will establish a novel, pre-clinical-stage, therapeutic product for treatment of peanut allergy. Furthermore, the success of this work will establish the applicability of this treatment approach for other allergies, as well as strategies to induce tolerance to self-antigens in patients suffering from debilitating autoimmune diseases.

While the preferred embodiments of the compositions and methods for tolerization have been illustrated and described, it will be appreciated that various changes can be made therein without departing from the spirit and scope of the invention.
CLAIMS

The embodiments of the invention in which an exclusive property or privilege is claimed are defined as follows:

1. An isolated fusion protein comprising a reovirus-derived targeting polypeptide and at least one allergen polypeptide.

2. The isolated fusion protein of Claim 1, wherein the reovirus-derived targeting polypeptide comprises the protein sigma polypeptide (ρστ), or functional portions or derivatives thereof.

3. The isolated fusion protein of Claim 1, wherein the functional portions of the ρστ include the head domain, trimerization domain, sialic acid binding domain, and/or the shaft domain of the ρστ protein, or any derivative thereof.

4. The isolated fusion protein of Claim 1, wherein the at least one allergen polypeptide is a food allergen, an environmental allergen, an autoantigen, and/or biological therapeutic, or is derived therefrom.

5. The isolated fusion protein of Claim 4, wherein the food allergen is from a ground nut, tree nut, milk, gluten, egg, fish, shellfish, and the like.

6. The isolated fusion protein of Claim 5, wherein the food allergen is from a peanut and the allergen polypeptide is Arah2, Arah6, Arah1, Arah3, Arah4, Arah5, Arah7, Arah8, Arah9, Arah10, Arah1, Arah2, or is derived therefrom.

7. The isolated fusion protein of Claim 5, wherein the food allergen is from gluten and the allergen polypeptide is a prolamin (such as a a-gliadin, β-gliadin, γ-gliadin, ω-gliadin, hordein, secalin, zein, kafirin, avenin), glutenin, or is derived therefrom.

8. The isolated fusion protein of Claim 5, wherein the food allergen is from milk and the allergen polypeptide is alpha S1-casein, alpha S2-casein, b-lactoglobulin, b-casein, k-casein, or is derived therefrom.
9. The isolated fusion protein of Claim 5, wherein the food allergen is from egg and the allergen polypeptide is ovomucoid, ovotransferrin, lysozyme, livetin, apovitillin, phosvitin, or is derived therefrom.

10. The isolated fusion protein of Claim 5, wherein the food allergen is from fish and the allergen polypeptide is Che ag, Lop pi, Gelatin/Ore a, Parvalbumin/Seb m, Ore al, Seb ml, Sar sal. 0101, Albumin/One ma, glyceraldehyde-3-phosphate dehydrogenase, or is derived therefrom.

11. The isolated fusion protein of Claim 4, wherein the environmental allergen is from an animal or insect, such as dust mite, bee, wasp, cat, dog, and the like, or plant, such as ragweed, grass, tree, and the like.

12. The isolated fusion protein of Claim 11, wherein the environmental allergen is from dust mite and the allergen polypeptide is Derpl through Derp23, Derfl through Derf33, Eurml, 2, 3, 4, or 14, Derml, or is derived therefrom.

13. The isolated fusion protein of Claim 11, wherein the environmental allergen is from cat and the allergen polypeptide is a secretoglobin such as Feldl, a lipocalin such as Feld4, an albumin such as Feld2, a cystatin such as Feld3, IgA such as Feld5w, or is derived therefrom.

14. The isolated fusion protein of Claim 11, wherein the environmental allergen is from ragweed and the allergen polypeptide is Ambal through Ambal 1, Ambp5, Ambt5, or is derived therefrom.

15. The isolated fusion protein of Claim 11, wherein the environmental allergen is from tree, such as birch, alder, and ash, and the allergen polypeptide is Betvl, Betv2, Betv3, Betv4, Betv6, Betv7, Alngl, Alng4, Frael, or is derived therefrom.

16. The isolated fusion protein of Claim 4, wherein the autoantigen is transglutaminase, myelin-associated glycoprotein (MAG), CNS-specific myelin oligodendrocyte glycoprotein (MOG), myelin basic protein (MBP), proteolipid protein (PLP), Zinc transporter-8 (ZnT8), Glutamic decarboxylase 65 (GAD65), Glutamic decarboxylase 67 (GAD67), Preproinsulin, proinsulin, insulin, Tyrosine phosphatase like autoantigen, insulinoma antigen-2 (IA-2; ICA512, PTPRN), IA-2b (Phogrin, PTPRN2),
Islet cell antigen-69 (ICA69), Chromogranin A, Islet amyloid polypeptide (ppIAPP), Heat shock protein 60 (hsp60), or is derived therefrom.

17. The isolated fusion protein of Claim 4, wherein the allergen polypeptide is derived from a protein therapeutic, such as an antibody CDR or erythropoietin.

18. The isolated fusion protein of Claim 1, wherein the at least one allergen polypeptide comprises an MHC Class I epitope and/or an MHC Class II epitope.

19. The isolated fusion protein of Claim 1, wherein the targeting polypeptide is separated from the at least one allergen polypeptide by a linker.

20. The isolated fusion protein of Claim 1, wherein the fusion protein comprises at least two allergen polypeptides.

21. The isolated fusion protein of Claim 20, wherein the at least two allergen polypeptides are separated by a linker.

22. A pharmaceutical composition comprising the isolated fusion protein of Claim 1 and a pharmaceutically acceptable carrier.

23. The pharmaceutical composition of Claim 22, wherein the composition is formulated for oral or intranasal administration.

24. A nucleic acid comprising a sequence encoding the isolated fusion protein of Claim 1.

25. A vector comprising the nucleic acid of Claim 24.

26. A cultured cell transfected with the vector of Claim 25.

27. A method for inducing tolerance to a polypeptide allergen, comprising administering to a subject in need thereof a pharmaceutically effective amount of the isolated fusion protein of Claim 1, wherein the isolated fusion protein comprises a polypeptide derived from the polypeptide allergen.

28. The method of Claim 27, wherein the method consists of administering a single dose of the effective amount of the isolated fusion polypeptide.
29. The method of Claim 27, wherein the method comprises of administering two or more doses of the effective amount of the isolated fusion polypeptide.

30. The method of Claim 27, wherein the effective amount of the isolated fusion polypeptide of Claim 1, comprises less than about 100mg, 75mg, 50mg, 25mg, 20mg, 15mg, 10mg, 9mg, 8mg, 7mg, 6mg, 5mg, 4mg, 3mg, 2mg, 1.5mg, or 1mg, of the isolated fusion polypeptide.
FIG. 2
**FIG. 3A**

- Normal Rabbit Serum
- anti-pol Rabbit Serum

**FIG. 3B**

- Normal Rabbit Serum
- anti-pol Rabbit Serum
S : Sensitization - 6mg Oral WPE
T : Therapy - PBS, ps1, ara h 2, VTC-064
C : Challenge - Oral, Peripheral, or Systemic

**FIG. 4**