IL-7 Variants with Reduced Immunogenicity

Cross Reference to Related Applications

[0001] This application claims priority to and the benefit of U.S. Provisional Patent Application No. 60/634,470, filed on December 9, 2004, the entire disclosure of which is incorporated by reference herein.

Field of the Invention

[0002] The invention relates generally to IL-7 molecules modified to reduce their immunogenicity. These molecules include also fusion proteins comprising said modified IL-7 molecules and immunoglobulin molecules or portions thereof, especially corresponding Fc fusion proteins.

Background

[0003] Cytokines are stimulators of the immune system and are thus useful as drugs. For example, interferon-alpha (IFN-α), interferon-beta (IFN-β), interleukin-2 (IL-2), and granulocyte/macrophage-colony stimulating factor (GM-CSF) are all approved drugs used to treat viral infections, cancer, immune system misregulation such as autoimmune disease, and to promote recovery of the immune system after cancer chemotherapy. Unfortunately, these proteins can stimulate an immune response against themselves, causing patients to develop antibodies against the therapeutic protein. These antibodies can also inhibit function of the same protein endogenously produced within the patient, resulting in potential long-term consequences for patient health.

[0004] Interleukin-7 is a cytokine that promotes survival and proliferation of T-cells, B-cells, and other immune cells. It is also potentially a therapeutic protein to treat patients whose immune systems have been damaged by cancer chemotherapy, HIV infection, or other diseases, disorders, or chemical exposures. However, based on its immunostimulatory properties, therapeutically administered IL-7 is expected to induce an antibody response against itself. Therefore, there is a need in the art for improved versions of IL-7 that are less immunogenic, but that retain the property of stimulating the immune system.

CONFIRMATION COPY

Summary of the Invention

[0005] The present invention is directed to interleukin-7 (IL-7) which has been modified to reduce its immunogenicity in comparison to wild-type IL-7. More specifically, the IL-7 proteins of the invention are modified to remove potential T-cell epitopes. As a result, IL-7 proteins, including immunoglobulin –IL7 fusion proteins, preferably Fc-IL7 fusion proteins, of the invention have improved biological properties compared to wild-type IL-7.

[0006] Accordingly, in one aspect, the invention features a polypeptide at least 80% identical to a human IL-7 moijety or an active portion thereof, comprising an amino acid substitution at one or more residues corresponding to Gln22, Leu24, Ile30, Phe39, Met54, Phe57, Arg58, Ala60, Leu63, Lys68, Met69, Leu77, Ile88, Val96, Leu104, Leu128, Met147, Thr149, or Lys150. These amino acid modifications can be used singly or in combination to reduce an anti-IL-7 T-cell response. Thus, the invention encompasses IL-7 moieties with for example, one, at least two, at least four, or at least eight amino acid modifications at positions selected from Gln22, Leu24, Ile30, Phe39, Met54, Phe57, Arg58, Ala60, Leu63, Lys68, Met 69, Leu77, Ile88, Val96, Leu104, Leu128, Met147, Thr149, and Lys150. In one embodiment, the IL-7 moiety incorporates one, two, three, four, five or more of the following substitutions: Gln22Asp, Leu24Asp, Ile30Thr, Phe39Pro, Met54Ala, Phe57Lys, Phe57Asn, Arg58Asp, Ala60Ser, Arg61Glu, Leu77Asp, Leu104Ser, Leu104Val, Leu128Ala, Leu128Val, Leu128Pro, Leu128Ser, Met147Lys, Thr149Ser, or Lys150Stop.

[0007] In one embodiment, the polypeptide contains a substitution or substitutions at one or more at Phe39, Phe57, Leu77, and Leu128. In a further embodiment, the polypeptide has one or more of substitutions Phe39Pro, Phe57Asn, Leu77Asp, and Leu128Ser. In another embodiment, the polypeptide includes the substitutions Phe39Pro, Phe57Asn, Leu77Asp, and Leu128Ser, while in a further embodiment, the polypeptide includes the substitutions Phe39Pro, Phe57Asn, and Leu128Ser.

Preferred substitutions according to this invention are at postions:

Phe39, or Phe57, or Leu77, or Leu128;

Phe39 and Phe57, or Phe39 and Leu77, or Phe57 and Leu77,or Phe39 and Leu128, or Phe57 and Leu128, or Leu77 and Leu128;

Phe39 and Phe 57 and Leu77, or Phe39 and Phe57 and Leu128, or

Phe57 and Leu77 and Leu128, or Phe39 and Leu77 and Leu128;
Phe39 and Phe 57 and Leu77 and Leu128.
Preferred specific substitutions according to this invention are:
Phe39Pro, or Phe57Lys, or Leu77Asp, or Leu128Ser;
Phe39Pro and Phe57Lys, or Phe39Pro and Leu77Asp, or Phe57Lys and Leu77Asp, or Phe39Pro and Leu128Ser, or Phe57Lys and Leu128Ser, or Leu77Asp and Leu128Ser;
Phe39Pr and Phe 57Lys and Leu77Asp, or Phe39Pro and Phe57Lys and Leu128Ser, or
Phe57Lys and Leu77Asp and Leu128Ser, or Phe39Pro and Leu77Asp and Leu128Ser;
Phe39Pro and Phe 57Lys and Leu77Asp and Leu128Ser.

[0008] In certain embodiments of the invention, the polypeptide with at least 80% identity with a human IL-7 moiety further comprises an immunoglobulin (Ig) moiety, such as a human Ig moiety. In one embodiment, the Ig moiety is IgG2. In some embodiments, the Ig moiety is an Fc portion. The invention also relates to a cell comprising a nucleic acid sequence encoding a polypeptide modified according to the invention. In one embodiment, the cell is a prokaryotic cell.

[0009] In a further embodiment, the polypeptide has at least 90% identity to a human IL-7 moiety or an active portion thereof, while in another embodiment, the polypeptide has at least 95% identity to a human IL-7 moiety or an active portion thereof.

[0010] The invention also features a method of treating a patient comprising administering a therapeutically effective amount of a polypeptide of the invention to, for example, a patient diagnosed with cancer or HIV. In one embodiment, the invention provides for administration of between about 0.01 and about 10 mg/kg/day or between 0.01 and 10.00 mg/kg/day of a polypeptide of the invention.

Brief Description of the Drawings

[0011] Figure 1 depicts the amino acid sequence for human IL-7. The signal sequence is shown in bold. Also depicted in bold and italics is a stretch of eighteen amino acids which can be deleted from the IL-7 sequence.

[0012] Figure 2 depicts the amino acid sequence for cow IL-7. The signal sequence is shown in bold.

[0013] Figure 3 depicts the amino acid sequence for sheep IL-7. The signal sequence is shown in bold.

[0014] Figure 4 depicts the amino acid sequence of an exemplary deimmunized human IL-7 wherein T cell epitope sequences have been modified.

[0015] Figure 5 depicts the amino acid sequence of a bacterially produced deimmunized human IL-7.

[0016] Figure 6 depicts the nucleic acid sequence encoding mature IL-7 incorporating codons for the mutations F39P, F57N, and L128S.

[0017] Figure 7 depicts the amino acid sequence of mature IL-7 with the mutations F39P, F57N and L128S.

[0018] Figure 8 depicts the nucleic acid sequence encoding mature IL-7 incorporating codons for the mutations F39P, F57N, L77D, and L128S.

[0019] Figure 9 depicts the amino acid sequence of mature IL-7 with the mutations F39P, F57N, L77D and L128S.

[0020] Figure 10 depicts the nucleic acid sequence encoding bacterially produced deimmunized IL-7 (bDel-IL-7), codon-optimized for *E. coli* with codons for the amino acid mutations K68D, M69D, I88T, V96G.

[0021] Figure 11 depicts the nucleic acid sequence encoding a mature IL-7 variant, deimmunized IL-7 (Del-IL-7), with codons for the amino acid mutations K68D, M69D, I88T, V96G.

[0022] Figure 12 depicts the amino acid sequence of Fcγ1-IL-7, where the Fc portion consists of a γ1 hinge, γ1 CH2, and γ1 CH3 region.

[0023] Figure 13 depicts the amino acid sequence of human Fcγ2(h)(FN>AQ)-IL-7, which is an Fc portion with a γ1 hinge, a γ2 CH2 domain, and a γ2 CH3 domain. The Fc portion incorporates the amino acid mutations F296A and N297Q.

[0024] Figure 14 depicts the amino acid sequence of human Fcγ1-(linker1)-IL-7, which is an Fc portion consisting of a γ1 hinge, CH1, and CH2 region connected to an IL-7 moiety via a polypeptide linker of amino acid sequence GGGGSGGGGSGGGGS.

[0025] Figure 15 depicts the amino acid sequence of human Fcγ1(YN>AQ)-(linker 2)-IL-7, which is a γ1 Fc portion with a γ1 hinge, CH1, and CH2 region,

incorporating the mutations Y296A and N297Q, connected to an IL-7 moiety via a polypeptide linker of amino acid sequence GGGGSGGGG.

[0026] Figure 16 depicts the amino acid sequence of human Fcγ1(YN>AQ,d)-(linker 2)-IL-7, which is a γ1 Fc portion with a γ1 hinge, CH1, and CH2 domains, incorporating the mutations Y296A and N297Q, as well deletion of the C-terminal lysine and preceding glycine of the Fc moiety. The Fc portion is connected to an IL-7 moiety via a polypeptide linker of amino acid sequence GGGGSGGGG.

[0027] Figure 17 depicts the nucleic acid sequence of Fcγ1, an Fc portion with a hinge, CH1 domain and CH2 domain, all of IgG1.

[0028] Figure 18 depicts the nucleic acid sequence of Fcγ1(YN>AQ), which is an Fc portion with a hinge, CH1 domain and CH2 domain, all of IgG1. The Fc portion incorporates the mutations Tyr296Ala and Asn297Gln.

[0029] Figure 19 depicts the nucleic acid sequence of Fcγ2(h), which is an Fc portion with an IgG1 hinge and IgG2 CH2 and CH3 domains.

[0030] Figure 20 depicts the nucleic acid sequence of Fcγ2(h)(FN>AQ), which is an Fc portion with an IgG1 hinge and IgG2 CH2 and CH3 domains. The Fc portion incorporates the mutations F296A and N297Q.

[0031] Figure 21 depicts the amino acid sequence of mature human deimmunized IL-7.1, wherein the IL-7 incorporates the substitutions L24D, M54A, F57K, A60S, R61E, M147K, T149S, and deletes residues K150, E151, and H152.

[0032] Figure 22 depicts the nucleic acid sequence encoding the amino acid sequence of Figure 21.

[0033] Figure 23 depicts the amino acid sequence of mature human deimmunized IL-7.2 which incorporates the substitutions D76N, L77D, T87Q, I88T, V96G, L119S, L128V, M147K, T149S, and deletes K150, E151, and H152.

[0034] Figure 24 depicts the nucleic acid sequence encoding the amino acid sequence of Figure 23.

[0035] Figure 25 depicts the amino acid sequence of mature human deimmunized IL-7.3 which incorporates the substitutions L24D, I30T, F39P, M54A, F57K, A60S, R61E, M68D, N69D, L77D, T87Q, I88T, V96G, L119S, L128A, M147K, T149S, and deletes K150, E151, and H152.

[0036] Figure 26 depicts the nucleic acid sequence encoding the amino acid sequence of Figure 25.

[0037] Figure 27 depicts the nucleic acid sequence encoding the linker sequence GGGGSGGGG followed by mature human IL-7 containing the amino acid substitution F39P, F57N, and L128S (PNS), and which contains flanking restriction sites Xma I and Xho I at the 5' and 3' ends respectively..

[0038] Figure 28 depicts the nucleic acid sequence of mature huFcγ2(h)(FN>AQ)(linker2)-IL-7(F39P, F57N, L77D, L128S) which is a human Fc portion with an IgG1 hinge, and IgG2 CH2 and CH3 domains incorporating the mutations F296A and N297Q connected to the N-terminus of a human IL-7 moiety incorporating the mutations F39P, F57N, L77D, and L128S. The Fc portion and the IL-7 moiety are connected by a linker sequence GGGGSGGGG.

[0039] Figure 29 depicts the nucleic acid sequence of mature huFcγ2(h)-(linker2)-IL-7(F39P, F57N, L77D, L128S) which is a human Fc portion with an IgG1 hinge, and IgG2 CH2 and CH3 domains connected to the N-terminus of a human IL-7 moiety incorporating the mutations F39P, F57N, L77D and L128S. The Fc portion and the IL-7 moiety are connected by a linker sequence GGGGSGGGG.

[0040] Figure 30 depicts the nucleic acid sequence of mature huFcγ2(h)-(linker2)-IL-7(F39P, F57N, L128S) which is a human Fc portion with an IgG1 hinge, and IgG2 CH2 and CH3 domains connected to the N-terminus of a human IL-7 moiety incorporating the mutations F39P, F57N, and L128S. The Fc portion and the IL-7 moiety are connected by a linker sequence GGGGSGGGG.

[0041] Figure 31 depicts the amino acid sequence of mature huFcγ2(h)(FN>AQ)-(linker2)-IL-7(F39P, F57N, L128S), which is a human Fc portion with a hinge of IgG1 and CH2 and CH3 domains of IgG2. The Fc portion contains the mutations F296A and N297Q. The Fc portion is linked to the IL-7 moiety via a linker of sequence GGGGSGGGG. The IL-7 moiety contains the mutations F39P, F57N, and L128S.

[0042] Figure 32 depicts the amino acid sequence of mature huFcγ2(h)(FN>AQ)-(linker2)-IL-7(F39P, F57N, L77D, L128S), which is a human Fc portion with a hinge of IgG1 and CH2 and CH3 domains of IgG2. The Fc portion contains the mutations F296A and N297Q. The Fc portion is linked to the IL-7 moiety via a linker of sequence GGGGSGGGG. The IL-7 moiety contains the mutations F39P, F57N, L77D and L128S.

[0043] Figure 33 depicts the amino acid sequence of mature huFcγ2(h)-(linker2)-IL-7(F39P, F57N, L77D, L128S), which is a human Fc portion with a hinge of IgG1 and CH2 and CH3 domains of IgG2. The Fc portion is linked to the IL-7 moiety via a linker of sequence GGGGSGGGG. The IL-7 moiety contains the mutations F39P, F57N, L77D and L128S.

[0044] Figure 34 depicts the amino acid sequence of mature huFcγ2(h) (linker2)-IL-7(F39P, F57N, L128S), which is a human Fc portion with a hinge of IgG1 and CH2 and CH3 domains of IgG2. The Fc portion is linked to the IL-7 moiety via a linker of sequence GGGGSGGGG. The IL-7 moiety contains the mutations F39P, F57N, L77D and L128S.

[0045] Figure 35 is an amino acid sequence alignment of IL-7 proteins from human, chimpanzee, baboon, macaque, bovine, pig, sheep, rat, and murine sources.

[0046] Figure 36 depicts Fc-IL-7 plasma concentrations in μg/ml for both test mice and control mice administered Fc-IL-7 subcutaneously.

[0047] Figure 37 shows values for the average fold change in plasma Fc-IL-7 concentrations between day 0 and day 2, and between day 2 and 4 in test mice administered Fc-IL-7 subcutaneously (SC).

[0048] Figure 38 depicts the average organ weights of organs taken from test mice sacrificed on day 7 compared to the average organ weights of mice in the control group.

[0049] Figure 39 depicts a comparison of the frequency of granulocyte Gr-1+ cells in cells/μL in the peripheral blood of mice from the control group, 0.5 mg/kg dosage group, 5.0 mg/kg dosage group, and 25 mg/kg dosage group on day 7.

[0050] Figure 40 depicts a comparison of the frequency of CD19+ cells in cells/μL in the peripheral blood of mice from the control group, 0.5 mg/kg dosage group, 5.0 mg/kg dosage group, and 25 mg/kg dosage group on day 7.

[0051] Figure 41 depicts a comparison of the frequency of CD4+ cells in cells/μL in the peripheral blood of test from the control group, 0.5 mg/kg dosage group, 5.0 mg/kg dosage group, and 25 mg/kg dosage group on day 7.

[0052] Figure 42 depicts a comparison of the frequency of CD8+ cells in cells/μL in the peripheral blood of test from the control group, 0.5 mg/kg dosage group, 5.0 mg/kg dosage group, and 25 mg/kg dosage group on day 7.

[0053] Figure 43 depicts the activity of Fc-IL-7 as compared to wild type IL-7 based on incorporation of tritiated thymidine in counts per minute versus IL-7/Fc-IL-7 concentration in a standard cell proliferation assay.

Detailed Description of the Invention

[0054] The invention is directed to IL-7 proteins that have reduced immunogenicity as compared to wild-type IL-7, as well as methods for making and using such proteins. More specifically, the invention provides mutations within IL-7 moieties that have the effect of reducing the immunogenicity of IL-7 itself, primarily by removing T-cell epitopes within IL-7 that may stimulate to an immune response. The invention also encompasses fusion proteins incorporating IL-7 moieties modified according to the teachings of the invention.

[0055] T-cell epitopes can be identified by a variety of computer and non-computer methods, including predictions based on structure-based computer modeling or by synthesis of peptides and testing for binding to specific MHC Class II molecules or in an immunogenicity assay. According to the invention, a potential T-cell epitope is a sequence that, when considered as an isolated peptide, is predicted to bind to an MHC Class II molecule or an equivalent in a non-human species. A potential T-cell epitope is defined without consideration of other aspects of antigen processing, such as the efficiency of protein uptake into antigen-presenting cells, the efficiency of cleavage at sites in an intact protein to yield a peptide that can bind to MHC Class II, and so on. Thus, the set of T-cell epitopes that are actually presented on MHC Class II after administration of a protein to an animal is a subset of the potential T-cell epitopes. According to the invention, a T-cell epitope is an epitope on a protein that interacts with an MHC class II molecule. Without wishing to be bound by theory, it is understood that a T-cell epitope is an amino acid sequence in a protein that failed to undergo the negative T-cell selection process during T-cell development and therefore will be expected to be presented by an MHC Class II molecule and recognized by a T-cell receptor.

[0056] B-cell epitopes are also identified by a variety of computer and non-computer methods, including predictions based on structure-based computer modeling or by synthesis of peptides and testing for binding to specific B-cell antigen receptor molecules or in an immunogenicity assay. According to the

invention, a potential B-cell epitope is a sequence that, when considered as an isolated peptide, is predicted to bind to a B-cell antigen receptor or an equivalent in a non-human species. A B-cell epitope is an epitope that does bind or is recognized by a B-cell antigen receptor and is a subset of potential B-cell epitopes.

[0057] The invention provides methods related to reducing the immunogenicity of IL-7. According to one embodiment of the invention, potential non-self T-cell epitopes are identified in sequences of IL-7. For example, potential non-self T-cell epitopes are identified by computational methods based on modeling peptide binding to MHC Class II molecules. Substitutions are then made such that the ability of peptides containing potential T-cell epitopes to bind to MHC Class II is reduced or eliminated. This process of identifying and modifying peptides which bind to MHC Class II is termed "de-immunization" and the resultant modified protein molecules are termed "de-immunized."

[0058] According to the invention, MHC Class II binding can be removed in situations where a protein is to be produced in bacteria or in an organism that does not generate a mammalian glycosylation pattern, such as yeast or insect cells.

[0059] The invention provides non-computer methods for reducing or eliminating the number of T-cell epitopes in IL-7 without requiring elaborate computer simulations or protein three-dimensional structures. In one embodiment, a method of the invention takes advantage of the fact that a core segment of nine amino acids interacts with both the MHC class II molecule as well as the T-cell receptor during antigen presentation. The most N-terminal amino acid, the "anchor" position, binds to a deep pocket within the MHC class II molecule. One of the following amino acids is typically present at the anchor position, which is important for binding to an MHC class II molecule: leucine, valine, isoleucine, methionine, phenylalanine, tyrosine and tryptophan. According to the invention, an additional 2 to 3 amino acids adjacent to the core 9 amino acids also affect the interaction with MHC molecules.

[0060] A general method of the invention includes mutating any leucines, valines, isoleucines, methionines, phenylalanines, tyrosines or tryptophans that occur in IL-7. In one embodiment, one or more of these amino acids in a candidate T-cell epitope is mutated to a threonine, an alanine or a proline,

thereby retaining some of the hydrophobic nature of the amino acid that is replaced. In further embodiments of the invention, one or more of the above-mentioned amino acids is deleted from a candidate T-cell epitope or potential T-cell epitope, or replaced with an appropriate amino acid analog. According to the invention, if an amino acid is deleted to destroy a potential T-cell epitope, care should be taken not to generate a new T-cell epitope that includes amino acids near the deletion.

[0061] Thus, the invention provides nucleic acid sequences and proteins that are useful in construction of less immunogenic IL-7 proteins. Specifically, the invention provides proteins with mutations of leucines, valines, isoleucines, methionines, phenylalanines, tyrosines, or tryptophans. Any aliphatic or aromatic residue (leucine, valine, isoleucine, methionine, phenylalanine, tryptophan or tyrosine) presents a high risk of creating an MHC binding peptide with the amino acid in the first position (anchor position) that binds the pocket of the MHC molecule. Therefore, substitution of any of the above-mentioned amino acids, with an amino acid that is not one of the above-mentioned amino acids, or with alanine, proline, or threonine, will remove a candidate T-cell epitope.

[0062] The proteins can be human proteins with sequences that generally correspond to sequences found in the human body. The invention also provides nucleic acid sequences encoding such proteins. The nucleic acid sequences for this aspect of the invention may exist as plasmids, PCR-generated fragments, or nucleic acids produced by chemical synthesis.

[0063] As used herein, the term "interleukin-7" or "IL-7" means IL-7 polypeptides and derivatives and analogs thereof having substantial amino acid sequence identity to wild-type mature mammalian IL-7. For example, IL-7 refers to an amino acid sequence of a recombinant or non-recombinant polypeptide having an amino acid sequence of: i) a native or naturally-occurring allelic variant of an IL-7 polypeptide, ii) a biologically active fragment of an IL-7 polypeptide, iii) a biologically active polypeptide analog of an IL-7 polypeptide, or iv) a biologically active variant of an IL-7 polypeptide.

[0064] IL-7 polypeptides modified according to the invention can be derived from any species, *e.g.*, human, cow or sheep. IL-7 nucleic acid and amino acid sequences are well known in the art. For example, the human IL-7 amino acid sequence has a Genbank accession number of NM 000880 (SEQ ID NO:1) and

is shown in Figure 1; the mouse IL-7 amino acid sequence has a Genbank accession number of NM 008371; the rat IL-7 amino acid sequence has a Genbank accession number of AF 367210; the cow IL-7 amino acid sequence has a Genbank accession number of NM 173924 (SEQ ID NO:2) and is shown in Figure 2; and the sheep IL-7 amino acid sequence has a Genbank accession number of U10089 (SEQ ID NO:3) and is shown in Figure 3. The signal sequence for each of the polypeptide species is shown in bold in each of the figures and is typically not included where the IL-7 portion is fused C-terminal to the carrier protein.

[0065] In addition, in Figure 35, an alignment of various mammalian IL-7 sequences is shown. IL-7 from non-human primates is generally more than 90% identical to human IL-7. Although the murine IL-7 sequence is the most divergent from the human IL-7 sequence, with less than 70% identity, it is nevertheless capable or activating the human IL-7 receptor. Therefore, IL-7 moieties from a range of species are particularly useful in accordance with the teachings of the invention.

[0066] A "variant" of an IL-7 protein is defined as an IL-7 amino acid sequence that is altered by one or more amino acids as compared to wild-type IL-7. The variant can have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties, *e.g.*, replacement of leucine with isoleucine. More rarely, a variant can have "nonconservative" changes, *e.g.*, replacement of a glycine with a tryptophan. Similar minor variations can also include amino acid deletions or insertions, or both.

[0067] Variant IL-7 proteins also include polypeptides that have at least about 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity with wild-type IL-7. To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino acid or nucleic acid sequence). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (*i.e..*, % homology = (# of identical positions/total # of positions)times 100). The determination of percent homology between two sequences can be accomplished using a mathematical

algorithm. A non-limiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul, (1990) Proc. Natl. Acad. Sci. USA, 87:2264-68, modified as in Karlin and Altschul, (1993) Proc. Natl. Acad. Sci. USA, 90:5873-77. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul *et al.*, (1990) J. Mol. Biol., 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score=100, wordlength=12. BLAST protein searches can be performed with the XBLAST program, score=50, wordlength=3. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.*, (1997) Nucleic Acids Research, 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used.

[0068] Furthermore, the invention also includes IL-7 fusion proteins wherein the IL-7 moiety contains a deletion and which retain comparable activity compared to the corresponding unmodified IL-7 fusion proteins. For example, the invention provides a form of Ig-IL-7 or IL-7 in which the IL-7 moiety contains an eighteen amino acid internal deletion corresponding to the sequence VKGRKPAALGEAQPTKSL (SEQ ID NO:25), of wild-type human IL-7. (See SEQ ID NO:1). In addition, the invention provides an active form of IL-7 wherein Lys150 is deleted. Glu151and His152 may also be deleted in conjunction with Lys150 while still leaving an active form of IL-7.

[0069] Throughout this application, the positions of amino acid residues in the IL-7 sequence are given with reference to the mature human IL-7 protein. For example, the cysteine in the N-terminal sequence MDCDIEGK...(SEQ ID NO:22) of bacterially produced human IL-7 protein, which includes a start methionine, is still referred to as Cys2.

Modifying IL-7 Proteins

[0070] One aspect of the invention derives from the insight that IL-7 produced by bacterial expression will not contain post-translational modifications that are characteristic of eukaryotes, such as mammals. For example, IL-7 contains three predicted N-linked glycosylation sites at positions 70, 91, and 116. In an Fc-IL-7 fusion protein expressed in mammalian cells, the asparagines at positions 70 and 91 are glycosylated, while the asparagine at position 116 is not. It is likely that IL-7 endogenously produced in the human body is also N-glycosylated, at least at

positions 70 and 91, and possibly at position 116. These N-linked glycosylations are not present in bacterially produced IL-7, and represent sequences that might be recognized by the human immune system as "non-self," *i.e.*, not normally present in the human body. As such, the invention encompasses deimmunizing these potential epitope regions on IL-7 to reduce the immunogenicity of IL-7 and related proteins.

[0071] According to the invention, T-cell epitopes are present in IL-7 that include positions 70 and 91, as described in Table 1. The epitopes shown in Table 1 are defined in terms of a minimal 9-mer peptide, with the strong MHC Class II anchor residue in the first position.

Table 1

T cell epitopes including position 70	T cell epitopes including position 91
LRQFLKMNS (SEQ ID NO:4)	ILLNCTGQV (SEQ ID NO:7)
FLKMNSTGD (SEQ ID NO:5)	LLNCTGQVK (SEQ ID NO:8)
LKMNSTGDF (SEQ ID NO:6)	
MNSTGDFDL (SEQ ID NO:9)	

[0072] According to the invention, one method for reducing the immunogenicity of bacterially produced IL-7 is to introduce one or more of the following mutations: Leu63Ala, Leu63Val, Leu63Pro, Leu63Thr, Lys68Asp, Met69Asp, Lys68Glu, Met69Glu, Ile88Thr, Ile88Ala, Ile88Val, and Val96Gly. Other mutations may be introduced at positions 63, 68, 69, 88 and/or 94. Some mutations are particularly useful in combination, such as the pairs Lys68Asp coupled with Met69Asp and/or Ile88Thr coupled with Val96Gly.

[0073] When these mutations are introduced into IL-7 or a fusion protein comprising IL-7, the resulting mutant protein generally has enough IL-7 biological activity to be useful as a therapeutic protein. In fact, the biological activity of the IL-7 moiety is at least 10%, 20%, 50%, 70%, 80%, 90%, 95%, 99% or 100% in comparison to the biological activity of wild type IL-7. Activity of the IL-7 of the invention can be tested in an *in vitro* or *in vivo* assay. Example 9 shows an assay for testing biological activity of the IL-7 variants of the invention.

[0074] In addition, the mutations generally allow proper folding of the IL-7 moiety so that a pure protein, largely free of high-molecular weight aggregates

and incorrectly disulfide-bonded forms, may be isolated. However, the folding and biological activity that results from any particular combination should be tested, for example as illustrated in the Examples, to verify that the desired activity is obtained.

[0075] According to the invention, an alternative strategy for reducing the immunogenicity of bacterially produced IL-7 is to alter Asn70 and Asn91 to aspartic acid. Without wishing to be bound by theory, the mutation of Asn70 and Asn91 to aspartic acid may be useful for the following reasons.

[0076] The immunogenicity of an exogenously administered therapeutic protein is mediated, in part, through the presentation of T-cell epitopes derived from the therapeutic protein. Such presentation is thought to occur through the following mechanism. A therapeutic protein is taken up by an antigen-presenting cell (APC), such as a dendritic cell, macrophage, or B-cell by endocytosis. The protein is transported into a series of vesicles termed endosomes, including the early, middle and late endosomes. In these vesicles, the environment becomes progressively more harsh and less favorable for extracellular, disulfide-bonded proteins that may be stably folded at neutral pH. Proteases termed cathepsins degrade internalized proteins into small peptides. A proportion of these protein fragments then become bound by MHC Class II proteins which transport the fragments to the cell surface as MHC Class II/peptide complexes. Such complexes are recognized by T-cell receptors on CD4+ T-cells.

[0077] In the case of peptides deriving from foreign proteins, presentation of an MHC Class II/peptide complex may stimulate an immune response. However, in the case of peptides deriving from self proteins, there are multiple mechanisms by which T-cells recognizing MHC Class II/peptide complexes are deleted or prevented from activating an immune response.

[0078] With the preceding two paragraphs taken as background, it is important to consider how an N-glycosylated protein would be processed in the endosome. Such a protein could be degraded into N-linked oligosaccharide-containing peptides that could bind to MHC Class II molecules. According to an insight of the invention, the endosome also contains an endoglycosidase that sometimes removes the oligosaccharide from asparagine, and in doing so, converts the asparagine into aspartic acid. Thus, self protein sequences that contain asparagine-linked oligosaccharides may be presented by MHC Class II

as peptides containing the asparagine linked to an oligosaccharide, or as corresponding peptides containing aspartic acid instead of asparagine.

[0079] As part of the invention, it is also recognized that this strategy for reducing the immunogenicity of mammalian proteins that are expressed in bacteria may be applied in a general manner. Specifically, the substitution of aspartic acid for asparagine at a site of N-linked glycosylation generally has the effect of reducing the immunogenicity of a mammalian protein that is expressed in a prokaryote.

[0080] The invention contains additional mutations that reduce the immunogenicity of IL-7 and IL-7-containing fusion proteins when expressed in either bacterial or mammalian cells. These mutations include those listed in Table 2 below. An IL-7 or IL-7 containing fusion protein may comprise one or more of these mutations. For example, in one embodiment, IL-7 is modified to incorporate one or more of L24D, M54A, F57K, A60S, R61E, M147K, and T149S, with K150, E151 and H152 being deleted. In another embodiment, IL-7 is modified to incorporate one or more of D76N, L77D, T87Q, I88T, V96G, L119S, M147K, and T149S, with K150, E151 and H152 being deleted. In a further embodiment, IL-7 can be modified to incorporate one or more of L24D, I30T, F39P, M54A, F57K, A60S, R61E, M68D, N69D, L77D, T87Q, I88T, V96G, L119S, L128A, M147K, and T149S, with K150, E151 and H152 being deleted.

[0081] In another embodiment, an IL-7 molecule or an IL-7 containing fusion protein may include mutations to one or more of residues 39, 57, 77 and/or 128 of IL-7. For example, IL-7 in one embodiment, includes a mutation at residue 39. In another embodiment, IL-7 includes a mutation at residue 57. In a further embodiment, IL-7 includes mutations at both residues 39 and 57. In yet another embodiment, IL-7 includes mutations at residues 39, 57 and 128, while in another embodiment, IL-7 includes mutations at residues 39, 57 and 77. In yet another embodiment, IL-7 includes mutations at residues 39, 57, 77 and 128. In a further embodiment, the phenylalanine residue at position 39 is replaced by a proline residue (F39P). In another embodiment, the phenylalanine residue at 57 is replaced by an asparagines residue (F57N). In another embodiment, the leucine residue at position 77 is replaced by aspartic acid (L77D). In yet another embodiment, the leucine residue at position 128 is replaced by serine (L128S).

Table 2

Initial position in mature human IL-7	Substitution
Gln22	Asp
Leu24	Asp
Ile30	Thr
Phe39	Pro
Met54	Ala
Phe57	Lys, Asn
Arg58	Asp
Ala60	Ser
Arg61	Glu
Leu63	Ala, Val, Pro
Lys68	Asp
Met69	Asp
Leu77	Asp
Ile88	Thr
Val96	Gly
Leu104	Ser, Val
Leu128	Ala, Val, Pro, Ser
Met147	Lys
Thr149	Ser
Lys150	Stop

Verification of the Reduced Immunogenicity of the Proteins of the Invention

[0082] To check that a mutation of the invention has indeed resulted in reduced immunogenicity, standard experimental tests, which are well known in the art, may be employed. For example, a T-cell stimulation assay may be used (*e.g.* Jones *et al.*, (2004), J. Interferon Cytokine Res., 24:560). In such an assay, human peripheral blood mononuclear cells (PBMCs) are obtained and cultured according to standard conditions. After an optional pre-stimulation, a peptide corresponding to a potential MHC Class II epitope is added to the culture of PBMCs; the PBMCs are further incubated, and at a later time tritiated thymidine is added. The peptide may be a minimal 9-mer, or may have about 10 to 15 or more amino acids. After further incubation of the cells, incorporation of tritiated thymidine into DNA is then measured by standard techniques.

[0083] The T-cell stimulation assay is thought to work by the following mechanisms. First, if a peptide is used as a stimulator, the peptide must first bind to an MHC Class II molecule present on a cell among the PBMCs. Second, the

MHC Class II/peptide complex must interact productively with a T-cell receptor on a CD4+ T-cell. If the test peptide is unable to bind sufficiently tightly to an MHC Class II molecule, no signal will result. If the peptide is able to bind an MHC Class II molecule and there are T-cells expressing an appropriately rearranged T-cell receptor capable of recognizing a particular MHC Class II/peptide complex, a signal should result. However, if such T-cells have been deleted as a result of a negative selection process, no signal will result. These mechanisms are considered relevant to the immunogenicity of a protein sequence, as inferred from the stimulation or lack of stimulation by a given peptide.

[0084] If recognizing T-cells are present in very low numbers in the PBMC population for stochastic reasons relating to failure of an appropriate T-cell receptor to take place or proliferation of other, unrelated T-cells followed by homeostasis of the T-cell population, there may also be no signal even though a signal is expected. Thus, false negative results may occur. Based on these considerations, it is important to use a large number of different sources of PBMCs and to test these samples independently. It is also generally useful to test PBMCs from an ethnically diverse set of humans, and to determine the MHC Class II alleles present in each PBMC population.

[0085] The standard T-cell assay has the disadvantage that the tritium incorporation signal is often only two-fold greater than the background incorporation. The proteins and peptides of the invention may also be tested in a modified T-cell assay in which, for example, purified CD4+ T-cells and purified dendritic cells are co-cultured in the presence of the test peptide, followed by exposure to tritiated thymidine and then assayed for tritiated thymidine incorporation. This second assay has the advantage that tritiated thymidine incorporation into irrelevant cells, such as CD8+ T-cells, is essentially eliminated and background is thus reduced.

[0086] A third assay involves the testing of a candidate protein with reduced immunogenicity in an animal such as a primate. Such an assay would generally involve the testing of an entire IL-7 protein or IL-7-containing fusion protein in which the IL-7 moiety had been designed by testing individual component peptides for potential immunogenicity in a cell-based assay such as one described above. Once such a candidate IL-7-containing protein is designed and expressed, the protein is tested for immunogenicity by injection into an animal.

[0087] Injection of the modified IL-7-containing protein is generally performed in the same manner as the anticipated route of delivery during therapeutic use in humans. For example, intradermal, subcutaneous, intramuscular, intraperitoneal injection or intravenous infusion may be used. If more than one administration is used, the administrations may be by different routes.

[0088] For immunogenicity testing purposes, it may be useful to coadminister an adjuvant to increase the signal and minimize the number of animals that need to be used. If an adjuvant is used, it is possible to use an adjuvant lacking a protein component, such as non-coding DNA with unmethylated CpG dinucleotides, bacterial lipid A, N-formyl methionine, or other bacterial non-protein components. Without wishing to be bound by theory, the rationale for avoiding protein-containing adjuvants is that other proteins may provide T-cell epitopes that will ultimately contribute to an antibody response against the candidate protein.

[0089] After one or more administrations of the candidate IL-7-containing protein, the presence of anti-IL-7 antibodies is tested according to standard techniques, such as the ELISA method. It is found that the altered IL-7-containing molecules of the invention induce antibody formation less frequently, and to a lesser extent, than corresponding molecules containing normal human IL-7.

[0090] Many of the proteins of the invention alter surface residues of IL-7. It is contemplated that the proteins of the invention, while being less immunogenic than corresponding proteins containing human IL-7, may still occasionally induce formation of antibodies. Because the B-cell epitopes of the proteins of the invention are generally different from those of unmodified IL-7, antibodies to the proteins of the invention will generally not cross-react with endogenous IL-7, and formation of antibodies to the proteins of the invention will have no long-term consequences for the health of the patient.

Fc-IL-7 Fusion Proteins

[0091] A key aspect of the invention is that IL-7 modified according to the invention may be fused to a carrier protein to create a fusion protein. In one embodiment, the carrier protein is disposed towards the N-terminus of the fusion protein and the IL-7 is disposed towards the C-terminus. In another embodiment,

the IL-7 is disposed towards the N-terminus of the fusion protein and the carrier protein is disposed towards the C-terminus.

[0092] The carrier protein can be any polypeptide covalently fused to the IL-7 protein. In one embodiment, the carrier protein is albumin, for example, human serum albumin. The albumin moiety may be fused to the C-terminal or N-terminal end of the IL-7 moiety. In another embodiment, the carrier protein is an immunoglobulin (Ig) moiety, such as an Ig heavy chain. The Ig chain may be derived from IgA, IgD, IgE, IgG, or IgM. According to the invention, the Ig moiety may be an intact antibody and may direct the IL-7 fusion protein to specific target sites in the body. Fusion proteins making use of antibody targeting are known to those in the art.

[0093] In one embodiment, the Ig moiety comprises an Fc region. As used herein, "Fc portion" encompasses domains derived from the constant region of an immunoglobulin, such as a human immunoglobulin, including a fragment, analog, variant, mutant or derivative of the constant region. Suitable immunoglobulins include IgG1, IgG2, IgG3, IgG4, and other classes. The constant region of an immunoglobulin is defined as a naturally-occurring or synthetically-produced polypeptide homologous to the immunoglobulin C-terminal region, and can include a hinge, a CH2 domain, a CH3 domain, or a CH4 domain, separately or in any combination. In the present invention, the Fc portion typically includes at least a CH2 domain. For example, the Fc portion can include hinge-CH2-CH3. Alternatively, the Fc portion can include all or a portion of the hinge region, the CH2 domain and/or the CH3 domain. Methods for making Fc-IL-7 fusion proteins are disclosed in U.S. Provisional Patent Application No. 60/533,406.

[0094] The constant region of an immunoglobulin is responsible for many important antibody functions including Fc receptor (FcR) binding and complement fixation. There are five major classes of heavy chain constant region, classified as IgA, IgG, IgD, IgE, and IgM. For example, IgG is separated into four γ subclasses: γ 1, γ 2, γ 3, and γ 4, also known as IgG1, IgG2, IgG3, and IgG4, respectively.

[0095] IgG molecules interact with multiple classes of cellular receptors, including three classes of Fcγ receptors (Fc γ R) specific for the IgG class of antibody, namely FcγRI, FcγRII, and FcγRIII. The important sequences for the binding of IgG to the FcγR receptors have been reported to be located in the CH2

and CH3 domains. The serum half-life of an antibody is influenced by the ability of that antibody to bind to an Fc receptor (FcR). Similarly, the serum half-life of immunoglobulin fusion proteins is also influenced by the ability to bind to such receptors (Gillies *et al.*, (1999) Cancer Res. 59:2159-66). Compared to those of IgG1, CH2 and CH3 domains of IgG2 and IgG4 have biochemically undetectable or reduced binding affinity to Fc receptors. It has been reported that immunoglobulin fusion proteins containing CH2 and CH3 domains of IgG2 or IgG4 had longer serum half-lives compared to the corresponding fusion proteins containing CH2 and CH3 domains of IgG1 (U.S. Patent No. 5,541,087; Lo *et al.*, (1998) Protein Engineering, 11:495-500). Accordingly, in certain embodiments of the invention, CH2 and CH3 domains are derived from an antibody isotype with reduced receptor binding affinity and effector functions, such as, for example, IgG2 or IgG4.

[0096] The hinge region is normally located C-terminal to the CH1 domain of the heavy chain constant region. In the IgG isotypes, disulfide bonds typically occur within this hinge region, permitting the final tetrameric molecule to form. This region is dominated by prolines, serines and threonines. When included in the present invention, the hinge region is typically at least homologous to the naturally-occurring immunoglobulin region that includes the cysteine residues to form disulfide bonds linking the two Fc moieties. Representative sequences of hinge regions for human and mouse immunoglobulins are known in the art and can be found in Borrebaeck, C. A. K., ed., (1992) Antibody Engineering, A Practical Guide, W. H. Freeman and Co. Suitable hinge regions for the present invention can be derived from IgG1, IgG2, IgG3, IgG4, and other immunoglobulin classes.

[0097] The IgG1 hinge region has three cysteines, two of which are involved in disulfide bonds between the two heavy chains of the immunoglobulin. These same cysteines permit efficient and consistent disulfide bonding formation of an Fc portion. Therefore, a hinge region of the present invention in one embodiment is derived from IgG1, such as human IgG1. When the IgG1 hinge is used, the first cysteine can be mutated to another amino acid, such as serine.

[0098] The IgG2 isotype hinge region has four disulfide bonds that tend to promote oligomerization and possibly incorrect disulfide bonding during secretion in recombinant systems. A suitable hinge region can be derived from an IgG2

hinge. In one embodiment, the first two cysteines of the IgG2 hinge are mutated to another amino acid.

[0099] The hinge region of IgG4 is known to form interchain disulfide bonds inefficiently. However, a suitable hinge region for the present invention can be derived from the IgG4 hinge region, and can contain a mutation that enhances correct formation of disulfide bonds between heavy chain-derived moieties (Angal *et al.*, (1993) Mol. Immunol., 30:105-8).

[00100] In accordance with the present invention, the Fc portion can contain CH2 and/or CH3 and/or CH4 domains and a hinge region that are derived from different antibody isotypes, *i.e.*, a hybrid Fc portion. For example, in one embodiment, the Fc portion contains CH2 and/or CH3 domains derived from IgG2 or IgG4 and a mutant hinge region derived from IgG1. Alternatively, a mutant hinge region from another IgG subclass is used in a hybrid Fc portion. For example, a mutant form of the IgG4 hinge that allows efficient disulfide bonding between the two heavy chains can be used. A mutant hinge can also be derived from an IgG2 hinge in which the first two cysteines are each mutated to another amino acid. Such hybrid Fc portions facilitate high-level expression and improve the correct assembly of the Fc-IL-7 fusion proteins. Assembly of such hybrid Fc portions is known in the art and has been described in U.S. Published Patent Application No. 2003-0044423.

[00101] In some embodiments, the Fc portion contains amino acid modifications that generally extend the serum half-life of an Fc fusion protein. Such amino acid modifications include mutations substantially decreasing or eliminating Fc receptor binding or complement fixing activity. For example, the glycosylation site within the Fc portion of an immunoglobulin heavy chain can be removed. In IgG1, the glycosylation site is Asn297 within the amino acid sequence Gln-Tyr-Asn-Ser (SEQ ID NO:26). In other immunoglobulin isotypes, the glycosylation site corresponds to Asn297 of IgG1. For example, in IgG2 and IgG4, the glycosylation site is the asparagine within the amino acid sequence Gln-Phe-Asn-Ser (SEQ ID NO:28). Accordingly, a mutation of Asn297 of IgG1 removes the glycosylation site in an Fc portion derived from IgG1. In one embodiment, Asn297 is replaced with Gln. In other embodiments, the tyrosine within the amino acid sequence Gln-Tyr-Asn-Ser (SEQ ID NO:26) is further mutated to eliminate a potential non-self T-cell epitope resulting from asparagine

mutation. For example, the amino acid sequence Gln-Tyr-Asn-Ser (SEQ ID NO:26) within an IgG1 heavy chain can be replaced with a Gln-Ala-Gln-Ser (SEQ ID NO:27) amino acid sequence.

[00102] Similarly, in IgG2 or IgG4, a mutation of asparagine within the amino acid sequence Gln-Phe-Asn-Ser (SEQ ID NO:28) removes the glycosylation site in an Fc portion derived from IgG2 or IgG4 heavy chain. In one embodiment, the asparagine is replaced with a glutamine. In other embodiments, the phenylalanine within the amino acid sequence Gln-Phe-Asn-Ser (SEQ ID NO:28) is further mutated to eliminate a potential non-self T-cell epitope resulting from asparagine mutation. For example, the amino acid sequence Gln-Phe-Asn-Ser (SEQ ID NO:28) within an IgG2 or IgG4 heavy chain can be replaced with a Gln-Ala-Gln-Ser (SEQ ID NO:27) amino acid sequence. Other mutations that are useful in reducing Fc receptor binding are disclosed in U.S. Patent Application No. 09/256,156.

[00103] It has also been observed that alteration of amino acids near the junction of the Fc portion and the non-Fc portion can dramatically increase the serum half-life of the Fc fusion protein. (U.S. Published Patent Application No. 2002-0147311). Accordingly, the junction region of an Fc-IL-7 or IL-7-Fc fusion protein of the present invention can contain alterations that, relative to the naturally-occurring sequences of an immunoglobulin heavy chain and IL-7, lie within about 10 amino acids of the junction point. These amino acid changes can cause an increase in hydrophobicity by, for example, changing the C-terminal lysine of the Fc portion to a hydrophobic amino acid such as alanine or leucine. (See *e.g.* SEQ ID NO:34). In yet another embodiment of the invention, the C-terminal lysine and preceding glycine of the Fc portion is deleted. (See *e.g.* SEQ ID NO:35).

[00104] In other embodiments, the Fc portion contains amino acid alterations of the Leu-Ser-Leu-Ser segment near the C-terminus of the Fc portion of an immunoglobulin heavy chain. The amino acid substitutions of the Leu-Ser-Leu-Ser (SEQ ID NO:29) segment eliminate potential junctional T-cell epitopes. In one embodiment, the Leu-Ser-Leu-Ser (SEQ ID NO:29) amino acid sequence near the C-terminus of the Fc portion is replaced with an Ala-Thr-Ala-Thr (SEQ ID NO:30) amino acid sequence. In other embodiments, the amino acids within the Leu-Ser-Leu-Ser (SEQ ID NO:29) segment are replaced with other amino acids

such as glycine or proline. Detailed methods of generating amino acid substitutions of the Leu-Ser-Leu-Ser (SEQ ID NO:29) segment near the C-terminus of an IgG1, IgG2, IgG3, IgG4, or other immunoglobulin class molecules, as well as other exemplary modifications for altering junctional T-cell epitopes, have been described in U.S. Published Patent Application No. 2003-0166877.

[00105] In one embodiment, a spacer or linker peptide is inserted between the carrier protein and the IL-7 protein. The spacer or linker peptide can be non-charged or non-polar or hydrophobic. The length of a spacer or linker peptide is between 1 and about 100 amino acids, or between 1 and about 50 amino acids, or between 1 and about 25 amino acids, or between 1 and about 15 amino acids. In one embodiment, the spacer contains a sequence $(G_4S)_n$, where n is less than 10. In another embodiment, the linker sequence is GGGGSGGGG (SEQ ID NO:67). In yet another embodiment, the spacer contains a motif that is recognized as an N-linked glycosylation site. In another embodiment of the invention, the carrier protein and the IL-7 fusion protein are joined via a spacer or linker peptide. In an alternative embodiment of the invention, the carrier protein and IL-7 fusion protein are separated by a synthetic spacer, for example a PNA spacer. The spacer can be non-charged, or non-polar or hydrophobic.

Production of IL-7 fusion proteins

[00106] Fusion proteins containing IL-7 modified according to the teachings of the invention can be synthesized by the non-limiting methods described herein. Assays useful for testing pharmacokinetic activities of fusion proteins containing IL-7 modified according to the invention in *in vivo* animal models are also described herein.

[00107] The IL-7 fusion proteins of the invention can be produced using recombinant expression vectors known in the art. The term "expression vector" refers to a replicable DNA construct used to express DNA which encodes the desired IL-7 fusion protein and which includes a transcriptional unit comprising an assembly of (1) genetic element(s) having a regulatory role in gene expression, for example, promoters, operators, or enhancers, operatively linked to (2) a DNA sequence encoding the desired IL-7 fusion protein which is transcribed into mRNA and translated into protein, and (3) appropriate transcription and translation initiation and termination sequences. The choice of promoter and other regulatory elements generally varies according to the intended host cell.

[00108] The nucleic acid encoding the IL-7 fusion protein is transfected into a host cell using recombinant DNA techniques. In the context of the present invention, the foreign DNA includes a sequence encoding the inventive proteins. Suitable host cells include prokaryotic, yeast or higher eukaryotic cells. In one embodiment, the host is a prokaryotic organism.

[00109] The recombinant IL-7 fusion proteins can be expressed in yeast hosts, such as from *Saccharomyces* species, such as *S. cerevisiae*. Yeast of other genera such as *Pichia* or *Kluyveromyces* may also be employed. Yeast vectors will generally contain an origin of replication from a yeast plasmid or an autonomously replicating sequence (ARS), a promoter, DNA encoding the IL-7 fusion protein, sequences for polyadenylation and transcription termination and a selection gene. Suitable promoter sequences in yeast vectors include the promoters for metallothionein, 3-phosphoglycerate kinase or other glycolytic enzymes, such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-4-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase and glucokinase.

[00110] Various mammalian or insect cell culture systems can be employed to express recombinant protein. Baculovirus systems for production of proteins in insect cells are well known in the art. Examples of suitable mammalian host cell lines include NS/0 cells, L cells, C127, 3T3, Chinese hamster ovary (CHO), HeLa, and BHK cell lines. Additional suitable mammalian host cells include CV-1 cells (ATCC CCL70) and COS-7 cells both derived from monkey kidney. Another suitable monkey kidney cell line, CV-1/EBNA, was derived by transfection of the CV-1 cell line with a gene encoding Epstein-Barr virus nuclear antigen-1 (EBNA-1) and with a vector containing CMV regulatory sequences (McMahan *et al.*, (1991), EMBO J., 10:2821). The EBNA-1 gene allows for episomal replication of expression vectors, such as HAV-EO or pDC406, that contain the EBV origin of replication.

[00111] Mammalian expression vectors may comprise non-transcribed elements such as an origin of replication, a suitable promoter and enhancer linked to the gene to be expressed, and other 5' or 3' flanking nontranscribed sequences, and 5' or 3' nontranslated sequences, such as necessary ribosome binding sites, a poly-adenylation site, splice donor and acceptor sites, and

transcriptional termination sequences. Commonly used promoters and enhancers are derived from Polyoma, Adenovirus 2, Simian Virus 40 (SV40), and human cytomegalovirus. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early and late promoter, enhancer, splice, and polyadenylation sites may be used to provide the other genetic elements required for expression of a heterologous DNA sequence.

[00112] When secretion of the IL-7 fusion protein from the host cell is desired, the expression vector may comprise DNA encoding a signal or leader peptide. In the present invention the native signal sequence of IL-7 can be used, or alternatively, a heterologous signal sequence may be added, such as the signal sequence from interleukin-4.

[00113] The present invention also provides a process for preparing the recombinant proteins of the present invention including culturing a host cell transformed with an expression vector comprising a DNA sequence that encodes the IL-7 fusion protein under conditions that promote expression. The desired protein is then purified from culture media or cell extracts. For example, supernatants from expression systems that secrete recombinant protein into the culture medium can be first concentrated using a commercially available protein concentration filter, for example, an Amicon or Millipore Pellicon ultrafiltration unit. Following the concentration step, the concentrate can be applied to a suitable purification matrix, as known in the art.

[00114] An "isolated" or "purified" IL-7 fusion protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the IL-7 fusion protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of IL-7 fusion protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of IL-7 fusion protein having less than about 30% (by dry weight) of non-IL-7 fusion protein (also referred to herein as a "contaminating protein"), less than about 20% of non-IL-7 fusion protein, less than about 10% of non-IL-7 fusion protein, or less than about 5% non-IL-7 fusion protein. When the IL-7 fusion protein or biologically active portion thereof is

purified from a recombinant source, it is, in one embodiment, substantially free of culture medium, *i.e..*, culture medium represents less than about 20%, less than about 10%, or less than about 5% of the volume of the protein preparation.

[00115] The term "substantially pure Ig-IL-7 fusion protein" or "substantially pure IL-7 fusion protein" refers to a preparation in which the IL-7 comprising fusion protein constitutes at least 60%, 70%, 80%, 90%, 95% or 99% of the proteins in the preparation.

Methods of Treatment Using IL-7 Proteins

[00116] The IL-7 proteins, including fusion proteins, of the invention are useful in treating immune deficiencies and in accelerating the natural reconstitution of the immune system that occurs, for example, after diseases or treatments that are immunosuppressive in nature. For example, IL-7 proteins can be used to treat viral infections, immune disorders, and to enhance the growth (including proliferation) of specific cell types. Moreover, the IL-7 proteins can be in the treatment of cancers such as bladder cancer, lung cancer, brain cancer, breast cancer, skin cancer, and prostate cancer. In one example, it is useful to treat patients who have undergone one or more cycles of chemotherapy with IL-7 proteins as described above to help their immune cells replenish. Alternatively, it is also useful to administer the IL-7 proteins described above to patients with HIV, the elderly, patients receiving a transplant or other patients with suppressed immune system function.

Administration

[00117] Both the IL-7 and IL-7 fusion proteins of the invention can be incorporated into a pharmaceutical composition suitable for administration. Such compositions typically comprise IL-7 or an IL-7 fusion protein and a pharmaceutically-acceptable carrier. As used herein the language "pharmaceutically-acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art.

[0118] A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation),

transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

[0119] Medicaments that contain the IL-7 proteins of the invention can have a concentration of 0.01 to 100% (w/w), though the amount varies according to the dosage form of the medicaments.

[0120] Administration dose depends on the body weight of the patients, the seriousness of the disease, and the doctor's opinion. However, it is generally advisable to administer between about 0.01 to about 10 mg/kg body weight a day, about 0.02 to about 2 mg/kg/day in case of injection, or about 0.5 mg/kg/day. The dose can be administered once or several times daily according to the seriousness of the disease and the doctor's opinion.

[0121] Compositions of the invention are useful when co-administered with one or more other therapeutic agents, for example, a molecule also known to be useful to replenish blood cells. For example, the molecule may be erythropoietin which is known to be used to replenish red blood cells, G-CSF which is used to replenish neutrophils or GM-CSF which is used to replenish granulocytes and macrophages.

[0122] Aspects of invention are further illustrated by the following examples.

Examples

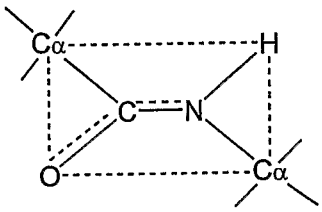
Example 1: Identification of T-cell epitopes by computational methods

[0123] According to the invention, epitopes of IL-7 can be modified using methods for introducing mutations into proteins to modulate their interaction with the immune system. These methods are similar to those disclosed in U.S. Published Patent Application No. 2003-0166877. According to the invention, known methods in the art that can be adapted according to the invention include

those described in the prior art (WO 92/10755 and WO 96/40792 (Novo Nordisk), EP 0519 596 (Merck & Co.), EP 0699 755(Centro de Immunologia Moelcular), WO 98/52976 and WO 98/59244 (Biovation Ltd.) or related methods.

[0124] Advantageous mutant proteins, however, can be obtained if the identification of said epitopes is realized by the following method which is described herewith in detail and applied to IL-7. There are a number of factors that play important roles in determining the total structure of a protein, polypeptide or immunoglobulin. First, the peptide bond, *i.e..*, that bond which joins the amino acids in the chain together, is a covalent bond. This bond is planar in structure, essentially a substituted amide. An "amide" is any of a group of organic compounds containing the grouping -CONH-.

[0125] The planar peptide bond linking Cα of adjacent amino acids may be



represented as depicted below:

[0126] Because the O=C and the C-N atoms lie in a relatively rigid plane, free rotation does not occur about these axes. Hence, a plane schematically depicted by the interrupted line is sometimes referred to as an "amide" or "peptide plane" plane wherein lie the oxygen (O), carbon (C), nitrogen (N), and hydrogen (H) atoms of the peptide backbone. At opposite corners of this amide plane are located the Cα atoms. Since there is substantially no rotation about the O=C and C-N atoms in the peptide or amide plane, a polypeptide chain thus comprises a series of planar peptide linkages joining the Cα atoms.

[0127] A second factor that plays an important role in defining the total structure or conformation of a polypeptide or protein is the angle of rotation of each amide plane about the common Cα linkage. The terms "angle of rotation" and "torsion angle" are hereinafter regarded as equivalent terms. Assuming that the O, C, N, and H atoms remain in the amide plane (which is usually a valid assumption, although there may be some slight deviations from planarity of these atoms for some conformations), these angles of rotation define the N and R

polypeptide's backbone conformation, *i.e..*, the structure as it exists between adjacent residues. These two angles are known as ϕ and ψ. A set of the angles ϕ_i, ψ_i, where the subscript i represents a particular residue of a polypeptide chain, thus effectively defines the polypeptide secondary structure. The conventions used in defining the ϕ, ψ angles, *i.e..*, the reference points at which the amide planes form a zero degree angle, and the definition of which angle is ϕ, and which angle is ψ, for a given polypeptide, are defined in the literature. (See, e.g,, Ramachandran *et al.*, (1968), Adv. Prot. Chem. 23:283-437, at pages 285-94).

[0128] The method can be applied to any protein, and is based in part upon the discovery that in humans the primary Pocket 1 anchor position of MHC Class II molecule binding grooves has a well designed specificity for particular amino acid side chains. The specificity of this pocket is determined by the identity of the amino acid at position 86 of the beta chain of the MHC Class II molecule. This site is located at the bottom of Pocket 1 and determines the size of the side chain that can be accommodated by this pocket. Marshall, J. Immunol., (1994), 152:4946-4956. If this residue is a glycine, then all hydrophobic aliphatic and aromatic amino acids (hydrophobic aliphatics being: valine, leucine, isoleucine, methionine and aromatics being: phenylalanine, tyrosine and tryptophan) can be accommodated in the pocket, with a preference being for the aromatic side chains. If this pocket residue is a valine, then the side chain of this amino acid protrudes into the pocket and restricts the size of peptide side chains that can be accommodated such that only hydrophobic aliphatic side chains can be accommodated. Therefore, in an amino acid residue sequence, wherever an amino acid with a hydrophobic aliphatic or aromatic side chain is found, there is the potential for a MHC Class II restricted T-cell epitope. If the side-chain is hydrophobic aliphatic, however, it is approximately twice as likely to be associated with a T-cell epitope than an aromatic side chain (assuming an approximately even distribution of Pocket 1 types throughout the global population).

[0129] An exemplary computational method profiles the likelihood of peptide regions of IL-7 to contain T-cell epitopes as follows: (1) The primary sequence of a peptide segment of predetermined length is scanned, and all hydrophobic aliphatic and aromatic side chains present are identified. (2) The hydrophobic

aliphatic side chains are assigned a value greater than that for the aromatic side chains; preferably about twice the value assigned to the aromatic side chains, *e.g.*, a value of 2 for a hydrophobic aliphatic side chain and a value of 1 for an aromatic side chain. (3) The values determined to be present are summed for each overlapping amino acid residue segment (window) of predetermined uniform length within the peptide, and the total value for a particular segment (window) is assigned to a single amino acid residue at an intermediate position of the segment (window), preferably to a residue at about the midpoint of the sampled segment (window). This procedure is repeated for each sampled overlapping amino acid residue segment (window). Thus, each amino acid residue of the peptide is assigned a value that relates to the likelihood of a T-cell epitope being present in that particular segment (window). (4) The values calculated and assigned as described in Step 3, above, can be plotted against the amino acid coordinates of the entire amino acid residue sequence being assessed. (5) All portions of the sequence which have a score of a predetermined value, *e.g.*, a value of 1, are deemed likely to contain a T-cell epitope and can be modified, if desired.

[0130] This particular aspect of the present invention provides a general method by which T-cell epitopes of IL-7 can be described. Modifications to the peptide in these regions have the potential to modify the MHC Class II binding characteristics.

[0131] According to another aspect of the present invention, T-cell epitopes can be predicted with greater accuracy by the use of a more sophisticated computational method which takes into account the interactions of peptides with models of MHC Class II alleles.

[0132] The computational prediction of T-cell epitopes present within a peptide according to this particular aspect contemplates the construction of models of at least 42 MHC Class II alleles based upon the structures of all known MHC Class II molecules and a method for the use of these models in the computational identification of T-cell epitopes, the construction of libraries of peptide backbones for each model in order to allow for the known variability in relative peptide backbone alpha carbon (Cα) positions, the construction of libraries of amino-acid side chain conformations for each backbone dock with each model for each of the 20 amino-acid alternatives at positions critical for the

interaction between peptide and MHC Class II molecule, and the use of these libraries of backbones and side-chain conformations in conjunction with a scoring function to select the optimum backbone and side-chain conformation for a particular peptide docked with a particular MHC Class II molecule and the derivation of a binding score from this interaction.

[0133] Models of MHC Class II molecules can be derived via homology modeling from a number of similar structures found in the Brookhaven Protein Data Bank ("PDB"). These may be made by the use of semi-automatic homology modeling software (Modeller *et al.*, (1993), J. Mol. Biol., 234:779-815) which incorporates a simulated annealing function, in conjunction with the CHARMm force-field for energy minimization (available from Molecular Simulations Inc., San Diego, Ca.). Alternative modeling methods can be utilized as well.

[0134] Other computational methods which use libraries of experimentally derived binding data of each amino-acid alternative at each position in the binding groove for a small set of MHC Class II molecules (Marshall *et al.*, (1995), Biomed. Pept. Proteins Nucleic Acids, 1(3):157-162) are known, as are yet other computational methods which use similar experimental binding data in order to define the binding characteristics of particular types of binding pockets within the groove, again using a relatively small subset of MHC Class II molecules, and then 'mixing and matching' pocket types from this pocket library to artificially create further 'virtual' MHC Class II molecules (Sturniolo *et al.*, (1999), Nat. Biotech, 17(6): 555-561. Both methods suffer the major disadvantage that, due to the complexity of the assays and the need to synthesize large numbers of peptide variants, only a small number of MHC Class II molecules can be experimentally scanned. Therefore the first method can only make predictions for a small number of MHC Class II molecules. The second method also makes the assumption that a pocket lined with similar amino-acids in one molecule will have the same binding characteristics when in the context of a different Class II allele and suffers further disadvantages in that only those MHC Class II molecules can be 'virtually' created which contain pockets contained within the pocket library. Using the modeling approach described herein, the structure of any number and type of MHC Class II molecules can be deduced, therefore alleles can be specifically selected to be representative of the global population. In addition, the number of MHC Class II molecules scanned can be increased by making further

models further than having to generate additional data via complex experimentation.

[0135] The use of a backbone library allows for variation in the positions of the Cα atoms of the various peptides being scanned when docked with particular MHC Class II molecules. This is again in contrast to the alternative computational methods described above which rely on the use of simplified peptide backbones for scanning amino-acid binding in particular pockets. These simplified backbones are not likely to be representative of backbone conformations found in 'real' peptides leading to inaccuracies in prediction of peptide binding. The present backbone library is created by superposing the backbones of all peptides bound to MHC Class II molecules found within the Protein Data Bank and noting the root mean square (RMS) deviation between the Cα atoms of each of the eleven amino-acids located within the binding groove. While this library can be derived from a small number of suitable available mouse and human structures (currently 13), in order to allow for the possibility of even greater variability, the RMS figure for each C"- α position is increased by 50%. The average Cα position of each amino-acid is then determined and a sphere drawn around this point whose radius equals the RMS deviation at that position plus 50%. This sphere represents all allowed Cα positions.

[0136] Working from the Cα with the least RMS deviation (that of the amino-acid in Pocket 1 as mentioned above, equivalent to Position 2 of the 11 residues in the binding groove), the sphere is three-dimensionally gridded, and each vertex within the grid is then used as a possible location for a Cα of that amino-acid. The subsequent amide plane, corresponding to the peptide bond to the subsequent amino-acid is grafted onto each of these Cαs and the ϕ and ψ angles are rotated step-wise at set intervals in order to position the subsequent Cα. If the subsequent Cα falls within the 'sphere of allowed positions' for this Cα than the orientation of the dipeptide is accepted, whereas if it falls outside the sphere then the dipeptide is rejected. This process is then repeated for each of the subsequent Cα positions, such that the peptide grows from the Pocket 1 Cα 'seed', until all nine subsequent Cαs have been positioned from all possible permutations of the preceding Cαs. The process is then repeated once more for

the single Cα preceding pocket 1 to create a library of backbone Cα positions located within the binding groove.

[0137] The number of backbones generated is dependent upon several factors: The size of the 'spheres of allowed positions'; the fineness of the gridding of the 'primary sphere' at the Pocket 1 position; the fineness of the step-wise rotation of the ϕ and ψ angles used to position subsequent Cαs. Using this process, a large library of backbones can be created. The larger the backbone library, the more likely it will be that the optimum fit will be found for a particular peptide within the binding groove of an MHC Class II molecule. Inasmuch as all backbones will not be suitable for docking with all the models of MHC Class II molecules due to clashes with amino-acids of the binding domains, for each allele a subset of the library is created comprising backbones which can be accommodated by that allele. The use of the backbone library, in conjunction with the models of MHC Class II molecules creates an exhaustive database consisting of allowed side chain conformations for each amino-acid in each position of the binding groove for each MHC Class II molecule docked with each allowed backbone. This data set is generated using a simple steric overlap function where a MHC Class II molecule is docked with a backbone and an amino-acid side chain is grafted onto the backbone at the desired position. Each of the rotatable bonds of the side chain is rotated step-wise at set intervals and the resultant positions of the atoms dependent upon that bond noted. The interaction of the atom with atoms of side-chains of the binding groove is noted and positions are either accepted or rejected according to the following criteria: the sum total of the overlap of all atoms so far positioned must not exceed a pre-determined value. Thus the stringency of the conformational search is a function of the interval used in the step-wise rotation of the bond and the pre-determined limit for the total overlap. This latter value can be small if it is known that a particular pocket is rigid; however, the stringency can be relaxed if the positions of pocket side-chains are known to be relatively flexible. Thus allowances can be made to imitate variations in flexibility within pockets of the binding groove. This conformational search is then repeated for every amino-acid at every position of each backbone when docked with each of the MHC Class II molecules to create the exhaustive database of side-chain conformations.

[0138] A suitable mathematical expression is used to estimate the energy of binding between models of MHC Class II molecules in conjunction with peptide ligand conformations which have to be empirically derived by scanning the large database of backbone/side-chain conformations described above. Thus a protein is scanned for potential T-cell epitopes by subjecting each possible peptide of length varying between 9 and 20 amino-acids (although the length is kept constant for each scan) to the following computations: An MHC Class II molecule is selected together with a peptide backbone allowed for that molecule and the side-chains corresponding to the desired peptide sequence are grafted on. Atom identity and interatomic distance data relating to a particular side-chain at a particular position on the backbone are collected for each allowed conformation of that amino-acid (obtained from the database described above). This is repeated for each side-chain along the backbone and peptide scores derived using a scoring function. The best score for that backbone is retained and the process repeated for each allowed backbone for the selected model. The scores from all allowed backbones are compared and the highest score is deemed to be the peptide score for the desired peptide in that MHC Class II model. This process is then repeated for each model with every possible peptide derived from the protein being scanned, and the scores for peptides versus models are displayed.

[0139] In the context of the present invention, each ligand presented for the binding affinity calculation is an amino-acid segment selected from a peptide or protein as discussed above. Thus, the ligand is a selected stretch of amino acids about 9 to 20 amino acids in length derived from a peptide, polypeptide or protein of known sequence. The terms "amino acids" and "residues" are hereinafter regarded as equivalent terms. The ligand, in the form of the consecutive amino acids of the peptide to be examined grafted onto a backbone from the backbone library, is positioned in the binding cleft of an MHC Class II molecule from the MHC Class II molecule model library via the coordinates of the C"-α atoms of the peptide backbone and an allowed conformation for each side-chain is selected from the database of allowed conformations. The relevant atom identities and interatomic distances are also retrieved from this database and used to calculate the peptide binding score. Ligands with a high binding affinity for the MHC Class II binding pocket are flagged as candidates for site-directed mutagenesis. Amino-

acid substitutions are made in the flagged ligand (and hence in the protein of interest) which is then retested using the scoring function in order to determine changes which reduce the binding affinity below a predetermined threshold value. These changes can then be incorporated into the protein of interest to remove T-cell epitopes.

[0140] Binding between the peptide ligand and the binding groove of MHC Class II molecules involves non-covalent interactions including, but not limited to: hydrogen bonds, electrostatic interactions, hydrophobic (lipophilic) interactions and van der Waal's interactions. These are included in the peptide scoring function as described in detail below. It should be understood that a hydrogen bond is a non-covalent bond which can be formed between polar or charged groups and consists of a hydrogen atom shared by two other atoms. The hydrogen of the hydrogen donor has a positive charge where the hydrogen acceptor has a partial negative charge. For the purposes of peptide/protein interactions, hydrogen bond donors may be either nitrogens with hydrogen attached or hydrogens attached to oxygen or nitrogen. Hydrogen bond acceptor atoms may be oxygens not attached to hydrogen, nitrogens with no hydrogens attached and one or two connections, or sulphurs with only one connection. Certain atoms, such as oxygens attached to hydrogens or imine nitrogens (*e.g.* C=NH) may be both hydrogen acceptors or donors. Hydrogen bond energies range from 3 to 7 Kcal/mol and are much stronger than van der Waal's bonds, but weaker than covalent bonds. Hydrogen bonds are also highly directional and are at their strongest when the donor atom, hydrogen atom and acceptor atom are co-linear. Electrostatic bonds are formed between oppositely charged ion pairs and the strength of the interaction is inversely proportional to the square of the distance between the atoms according to Coulomb's law. The optimal distance between ion pairs is about 2.8Å. In protein/peptide interactions, electrostatic bonds may be formed between arginine, histidine or lysine and aspartate or glutamate. The strength of the bond will depend upon the pKa of the ionizing group and the dielectric constant of the medium although they are approximately similar in strength to hydrogen bonds.

[0141] Lipophilic interactions are favorable hydrophobic-hydrophobic contacts that occur between the protein and the peptide ligand. Usually, these will occur between hydrophobic amino acid side chains of the peptide buried within the

pockets of the binding groove such that they are not exposed to solvent. Exposure of the hydrophobic residues to solvent is highly unfavorable since the surrounding solvent molecules are forced to hydrogen bond with each other forming cage-like clathrate structures. The resultant decrease in entropy is highly unfavorable. Lipophilic atoms may be sulphurs which are neither polar nor hydrogen acceptors and carbon atoms which are not polar.

[0142] van der Waal's bonds are non-specific forces found between atoms which are 3- 4Å apart. They are weaker and less specific than hydrogen and electrostatic bonds. The distribution of electronic charge around an atom changes with time and, at any instant, the charge distribution is not symmetric. This transient asymmetry in electronic charge induces a similar asymmetry in neighboring atoms. The resultant attractive forces between atoms reaches a maximum at the van der Waal's contact distance but diminishes very rapidly at about 1Å to about 2Å. Conversely, as atoms become separated by less than the contact distance, increasingly strong repulsive forces become dominant as the outer electron clouds of the atoms overlap. Although the attractive forces are relatively weak compared to electrostatic and hydrogen bonds (about 0.6 Kcal/mol), the repulsive forces in particular may be very important in determining whether a peptide ligand may bind successfully to a protein.

[0143] In one embodiment, the Böhm scoring function (SCORE1 approach) is used to estimate the binding constant. (Böhm, H.J., (1994), J. Comput. Aided Mol. Des., 8(3):243-256) which is hereby incorporated in its entirety). In another embodiment, the scoring function (SCORE2 approach) is used to estimate the binding affinities as an indicator of a ligand containing a T-cell epitope (Böhm, H.J., (1998), J. Comput. Aided Mol. Des., 12(4):309-323) which is hereby incorporated in its entirety). However, the Böhm scoring functions as described in the above references are used to estimate the binding affinity of a ligand to a protein where it is already known that the ligand successfully binds to the protein and the protein/ligand complex has had its structure solved, the solved structure being present in the Protein Data Bank ("PDB"). Therefore, the scoring function has been developed with the benefit of known positive binding data. In order to allow for discrimination between positive and negative binders, a repulsion term must be added to the equation. In addition, a more satisfactory estimate of binding energy is achieved by computing the lipophilic interactions in a pairwise

manner rather than using the area based energy term of the above Böhm functions. Therefore, in one embodiment, the binding energy is estimated using a modified Böhm scoring function. In the modified Böhm scoring function, the binding energy between protein and ligand (ΔG_{bind}) is estimated considering the following parameters: The reduction of binding energy due to the overall loss of translational and rotational entropy of the ligand (ΔG_0); contributions from ideal hydrogen bonds (ΔG_{hb}) where at least one partner is neutral; contributions from unperturbed ionic interactions (ΔG_{ionic}); lipophilic interactions between lipophilic ligand atoms and lipophilic acceptor atoms (ΔG_{lipo}); the loss of binding energy due to the freezing of internal degrees of freedom in the ligand, *i.e..*, the freedom of rotation about each C-C bond is reduced (ΔG_{rot}); the energy of the interaction between the protein and ligand (E_{VdW}). Consideration of these terms gives equation 1:

$$(\Delta G_{bind})=(\Delta G_0)+(\Delta G_{hb} x N_{hb})+(\Delta G_{ionic} x N_{ionic})+(\Delta G_{lipo} x N_{lipo})+(\Delta G_{rot}+N_{rot})+ (E_{VdW}).$$

where N is the number of qualifying interactions for a specific term and, in one embodiment, ΔG_0, ΔG_{hb}, ΔG_{ionic}, ΔG_{lipo} and ΔG_{rot} are constants which are given the values: 5.4, -4.7, -4.7, -0.17, and 1.4, respectively.

[0144] The term N_{hb} is calculated according to equation 2:

$$N_{hb} = \sum_{h\text{-}bonds} f(\Delta R, \Delta\alpha) \times f(N_{neighb}) \times f_{pcs}$$

[0145] $f(\Delta R, \Delta\alpha)$ is a penalty function which accounts for large deviations of hydrogen bonds from ideality and is calculated according to equation 3:

$$f(\Delta R, \Delta\text{-}\alpha) = f1(\Delta R) \times f2(\Delta\alpha)$$

where: $f1(\Delta R) = 1$ if $\Delta R <= TOL$

or $= 1 - (\Delta R - TOL)/0.4$ if $\Delta R <= 0.4 + TOL$

or $= 0$ if $\Delta R > 0.4 + TOL$

and: $f2(\Delta\alpha) = 1$ if $\Delta\alpha < 30°$

or $= 1-(\Delta\alpha - 30)/50$ if $\Delta\alpha <= 80°$

or $= 0$ if $\Delta\alpha > 80°$

[0146] TOL is the tolerated deviation in hydrogen bond length = 0.25Å;

ΔR is the deviation of the H-O/N hydrogen bond length from the ideal value = 1.9Å;

$\Delta\alpha$ is the deviation of the hydrogen bond angle $\angle_{N/O\text{-}H..O/N}$ from its idealized value of 180°.

[0147] $f(N_{neighb})$ distinguishes between concave and convex parts of a protein surface and therefore assigns greater weight to polar interactions found in

pockets rather than those found at the protein surface. This function is calculated according to equation 4 below:

$f(N_{neighb}) = (N_{neighb}/N_{neighb,0})^{\alpha}$ where $\alpha = 0.5$.

[0148] N_{neighb} is the number of non-hydrogen protein atoms that are closer than 5Å to any given protein atom.

[0149] $N_{neighb,0}$ is a constant = 25

[0150] f_{pcs} is a function which allows for the polar contact surface area per hydrogen bond and therefore distinguishes between strong and weak hydrogen bonds and its value is determined according to the following criteria:

f_{pcs}= ß when $A_{polar}/N_{HB} < 10$ Å^2

or f_{pcs}= I when $A_{polar}/N_{HB} > 10$ Å^2

A_{polar} is the size of the polar protein-ligand contact surface

N_{HB} is the number of hydrogen bonds

ß is a constant whose value = 1.2

[0151] For the implementation of the modified Böhm scoring function, the contributions from ionic interactions, ΔG_{ionic}, are computed in a similar fashion to those from hydrogen bonds described above since the same geometry dependency is assumed.

[0152] The term N_{lipo} is calculated according to equation 5 below:

$N_{lipo} = \sum_{lL} f(r_{lL})$

[0153] $f(r_{lL})$ is calculated for all lipophilic ligand atoms, l, and all lipophilic protein atoms, L, according to the following criteria:

$f(r_{lL}) = 1$ when $r_{lL} <= R1$ $f(r_{lL}) = (r_{lL} - R1)/(R2-R1)$ when $R2 < r_{lL} > R1$

$f(r_{lL}) = 0$ when $r_{lL} >= R2$

where: $R1 = r_l^{vdw} + r_L^{vdw} + 0.5$

and $R2 = R1 + 3.0$

and r_l^{vdw} is the van der Waal's radius of atom l

and r_L^{vdw} is the van der Waal's radius of atom L

[0154] The term N_{rot} is the number of rotable bonds of the amino acid side chain and is taken to be the number of acyclic sp^3 - sp^3 and sp^3 - sp^2 bonds. Rotations of terminal $-CH_3$ or $-NH_3$ are not taken into account.

[0155] The final term, E_{VdW}, is calculated according to equation 6 below:

$E_{VdW} = \varepsilon_1\varepsilon_2((r_1^{vdw} + r_2^{vdw})^{12}/r^{12} - (r_1^{vdw} + r_2^{vdw})^6/r^6)$, where:

ε_1 and ε_2 are constants dependent upon atom identity;

$r_1^{vdw} + r_2^{vdw}$ are the van der Waal's atomic radii; and

r is the distance between a pair of atoms.

[0156] With regard to equation 6, in one embodiment, the constants ε_1 and ε_2 are given the atom values: C: 0.245, N: 0.283, O: 0.316, S: 0.316, respectively (*i.e.*. for atoms of Carbon, Nitrogen, Oxygen and Sulfur, respectively). With regards to equations 5 and 6, the van der Waal's radii are given the atom values C: 1.85, N: 1.75, O: 1.60, S: 2.00Å.

[0157] It should be understood that all predetermined values and constants given in the equations above are determined within the constraints of current understandings of protein ligand interactions with particular regard to the type of computation being undertaken herein.

[0158] As described above, the scoring function is applied to data extracted from the database of side-chain conformations, atom identities, and interatomic distances. For the purposes of the present description, the number of MHC Class II molecules included in this database is 42 models plus four solved structures. It should be apparent from the above descriptions that the modular nature of the construction of the computational method of the present invention means that new models can simply be added and scanned with the peptide backbone library and side-chain conformational search function to create additional data sets which can be processed by the peptide scoring function as described above. This allows for the repertoire of scanned MHC Class II molecules to easily be increased, or structures and associated data to be replaced if data are available to create more accurate models of the existing alleles.

[0159] The present prediction method can be calibrated against a data set comprising a large number of peptides whose affinity for various MHC Class II molecules has previously been experimentally determined. By comparison of calculated versus experimental data, a cut of value can be determined above which it is known that all experimentally determined T-cell epitopes are correctly predicted.

[0160] It should be understood that, although the above scoring function is relatively simple compared to some sophisticated methodologies that are available, the calculations are performed extremely rapidly. It should also be understood that the objective is not to calculate the true binding energy *per se* for each peptide docked in the binding groove of a selected MHC Class II protein. The underlying objective is to obtain comparative binding energy data as an aid to predicting the location of T-cell epitopes based on the primary structure (*i.e.*. amino acid sequence) of a selected protein. A relatively high binding energy or a binding energy above a selected threshold value would suggest the presence of a T-cell epitope in the ligand. The ligand may then be subjected to at least one round of amino-acid substitution and the binding energy recalculated. Due to the rapid nature of the calculations, these manipulations of the peptide sequence can be performed interactively within the program's user interface on cost-effectively available computer hardware. Major investment in computer hardware is thus not required.

[0161] It would be apparent to one skilled in the art that other available software could be used for the same purposes. In particular, more sophisticated software which is capable of docking ligands into protein binding-sites may be used in conjunction with energy minimization. Examples of docking software are: DOCK (Kuntz *et al.*, (1982), J. Mol. Biol., 161:269-288), LUDI (Böhm, H.J., (1994), J. Comput Aided Mol. Des., 8:623-632) and FLEXX (Rarey *et al.*, (1995), ISMB, 3:300-308). Examples of molecular modeling and manipulation software include: AMBER (Tripos) and CHARMm (Molecular Simulations Inc.). The use of these computational methods would severely limit the throughput of the method of this invention due to the lengths of processing time required to make the necessary calculations. However, it is feasible that such methods could be used as a 'secondary screen' to obtain more accurate calculations of binding energy for peptides which are found to be 'positive binders' via the method of the present invention. The limitation of processing time for sophisticated molecular mechanic or molecular dynamic calculations is one which is defined both by the design of the software which makes these calculations and the current technology limitations of computer hardware. It may be anticipated that, in the future, with the writing of more efficient code and the continuing increases in speed of computer processors, it may become feasible to make such calculations within a

more manageable time-frame. Further information on energy functions applied to macromolecules and consideration of the various interactions that take place within a folded protein structure can be found in: Brooks *et al.*, (1983), J. Comput. Chem., 4:187-217 and further information concerning general protein-ligand interactions can be found in: Dauber-Osguthorpe *et al.*, (1988), Proteins, 4(1):31-47. Useful background information can also be found, for example, in Fasman, G.D., ed., Prediction of Protein Structure and the Principles of Protein Conformation, Plenum Press, New York, ISBN: 0-306 4313-9.

Example 2: *In vitro* analysis of IL-7 derived peptides as potential CD4+ T helper cell epitopes by unfractionated PBMC cultures

[0162] Based on *in silico* predictions that sequences surrounding N-linked glycosylation sites of the IL-7 protein are immunogenic, peptides encompassing these regions, spanning for example Leu63 to Ser71 (LRQFLKMNS SEQ ID NO:4) or Ile88 to Val96 (ILLNCTGQV SEQ ID NO:7) in mature human IL-7 protein, are analyzed for their immunogenicity, which is measured by their ability to induce T-cell proliferation *in vitro*. In essence, PBMCs isolated from human blood are incubated with individual overlapping 15-mer peptides, and proliferative responses are measured by ^{3}H-thymidine incorporation. In principle, T-cells within the mixture of PBMCs will only proliferate if they recognize individual peptide-MHC complexes on autologous APCs (antigen presenting cells), and thus proliferation is an indication of peptide immunogenicity.

[0163] For example, 15-mer peptides that are staggered by three amino acids and span the region from, for example, Met54 to Leu80 in human IL-7 are synthesized (Pepscan Systems, Netherlands), resuspended in DMSO (Sigma Chemical, St. Louis, MO, U.S.A.), and used at a final concentration of 5 μM in 0.5% DMSO in culture media.

[0164] PBMCs are isolated from peripheral blood from healthy donors by Ficoll-Hypaque gradient centrifugation, and are stored frozen in liquid nitrogen. In addition, each PBMC sample is HLA typed, using a SSP PCR typing kit (Bio-Synthesis, Lewisville, TX) on DNA isolated with a QiaAmp Tissue Kit (Qiagen, Valencia, CA).

[0165] In a typical proliferation assay, each of the overlapping 15-mer peptides is assayed in sextuplicate PBMC cultures derived from 40 naïve donors.

Briefly, 2 x10^5 PBMCs, thawed rapidly before use, are mixed with 5 μM of each peptide and incubated at 37°C in 5% CO_2 for 7 days. As a positive control, samples are incubated with the tetanus toxin derived peptide MQYIKANSKFIGI (SEQ ID NO:15), whereas negative control samples are incubated with 0.5% DMSO. During the last 12 hours of incubation the cultures are pulsed with of [methyl-^{3}H]thymidine (0.5 μCi/well) (NEN Life Science Products, Boston, MA), the cultures are harvested onto filter mats and thymidine incorporation is measured as counts per minute (CPM) using a Wallac microplate beta top plate scintillation counter (Perkin Elmer, Boston, MA). The stimulatory index for each peptide is calculated by dividing the CPM value of a given peptide divided by the CPM value obtained from negative controls.

[0166] It is found that the stimulatory index of the positive control peptide is significantly greater than 1 (the stimulatory index of the negative control), and peptides containing the core sequence LRQFLKMNS (SEQ ID NO:4), FLKMNSTGD (SEQ ID NO:5) or LKMNSTGDF (SEQ ID NO:6), such as, for example peptides ARKLRQFLKMNSTGD (SEQ ID NO:10), LRQFLKMNSTGDFDL (SEQ ID NO:11), or FLKMNSTGDFDLHLL (SEQ ID NO:12), have an increase in stimulatory index. Therefore peptide sequences such as LRQFLKMNS (SEQ ID NO:4) or LKMNSTGDF (SEQ ID NO:6) indeed represent potential T-cell epitopes.

[0167] A similar analysis is performed with a series of 15-mer peptides encompassing a region defined by the core peptides ILLNCTGQV (SEQ ID NO:7) and LLNCTGQVK (SEQ ID NO:8), and it is found that these peptide sequences as well represent potential T-cell epitopes.

Example 3: Mapping of CD4$^+$ T helper cell epitopes using differentiated human dendritic cells (DCs) *in vitro*

[0168] Mature dendritic cells (DCs) are potent antigen-presenting cells (APCs) that present antigenic peptides or whole proteins to T-cells efficiently. Isolated DCs, pulsed with antigenic peptides *in vitro*, are used to induce primary T-cell responses that can be measured in *in vitro* proliferation assays. Differentiated DCs are generated, for example by the following procedure: first, human monocytes are generated by allowing PBMCs to adhere to plastic tissue culture flasks or by purifying CD14$^+$ PBMCs with magnetically labeled antibodies

(Miltenyi Biotec, Auburn, CA). The purified monocytes (0.5 to 1.5 x 10^6 cells/ml) are then cultured in AIM V media (GIBCO BRL, Grand Island, NY, U.S.A.) containing 1000 U/ml GM-CSF (Endogen; Woburn, MA) and 500 U/ml IL-4 (Endogen; Woburn, MA) for 3 days. Subsequently, these immature DCs are pulsed with 5 μg/ml of experimental or control peptides and further incubated with a combination of 1000 U/ml TNF-α (Endogen; Woburn, MA), 1000 U/ml GM-CSF and 500 U/ml IL-4 for another 48 hours. Mature DCs are monitored by high surface expression levels of $CD80^+$, $CD86^+$ and HLA-DR.

[0169] These mature antigen-pulsed DCs are irradiated with 4200 Rads and are used in a proliferation assay with purified autologous $CD4^+$ T-cells. ($CD4^+$ T-cells are purified with magnetically labeled antibodies (Miltenyi Biotec, Auburn, CA), using frozen PBMC aliquots from the same donor that provides monocytes for the *in vitro* DC differentiation.) In a typical assay, antigen-pulsed DCs (2 × 10^5/mL) are incubated together with autologous $CD4^+$ T-cells (2 × 10^6 cells/mL) in round-bottomed 96-well plates at 37°C in 5% CO_2 for 7 days. [methyl-3H]thymidine (NEN Life Science Products, Boston, MA) is added during the last 12 hours of incubation at 0.5 μCi/well, the samples are harvested, lysed onto glass filters, and ^{3}H-thymidine incorporation is measured in a scintillation counter.

[0170] 15-mer peptides, as described for Example 1, are tested in this assay and compared to reference peptides and other controls. It is found that this assay is more sensitive than the assay described in Example 1, allowing better differentiation between the ability of individual peptides to induce T-cell proliferation. It is found, that IL-7 peptides containing the core sequences LRQFLKMNS (SEQ ID NO:4), FLKMNSTGD (SEQ ID NO:5), LKMNSTGDF (SEQ ID NO:6), ILLNCTGQV(SEQ ID NO:7) or LLNCTGQVK (SEQ ID NO:8) do indeed induce significant T-cell proliferation and therefore these sequences do represent potential T-cell epitopes.

Example 4: *In vitro* analysis of de-immunizing amino acid substitutions in IL-7

[0171] Amino acid substitutions in the peptide regions described above which are considered to render the IL-7 protein less immunogenic are tested in *in vitro* assays, as described in Example 1 and Example 2. For example, the variant IL-7 peptide encompassing the sequence LRQFLDDNS (SEQ ID NO:13) is expected to generate a significantly decreased T-cell proliferative response compared to

the wild-type parental peptide encompassing LRQFLKMNS (SEQ ID NO:4). Similarly, the variant IL-7 peptide encompassing the sequence TLLNCTGQG (SEQ ID NO:14) is expected to generate a significantly decreased T-cell proliferative response compared to the wild-type parental peptide encompassing ILLNCTGQV (SEQ ID NO:7)

[0172] A series of IL-7 derived 15-mer peptides is synthesized that encompass the variant IL-7 sequences LRQFLDDNS (SEQ ID NO:13) or TLLNCTGQG (SEQ ID NO:14), as described in Example 1. In addition, the variant IL-7 proteins, or fusion proteins containing variant IL-7, which include substitutions of the invention are produced either in a prokaryotic or eukaryotic expression system (Del-IL-7's). For example, a variant IL-7 is produced that includes the amino acid substitutions K68D, M69D, I88T and V94G. (In addition, the prokaryotically produced IL-7 proteins include a start methionine.) These peptides and purified proteins, and their parental counterparts, are tested in their ability to induce T-cell proliferation, in assays with whole PBMC cultures as described in Example 1 or by pulsing human DCs as described in Example 2. 25 μg/ml of protein is used to stimulate PBMCs or to pulse DCs.

[0173] It is found that in general, peptides derived from the variant IL-7 sequences have a significantly reduced ability to induce T-cell proliferation compared to the corresponding peptides derived from the wild-type human IL-7 protein. Therefore, these variant peptide sequences are much poorer potential T-cell epitopes. Likewise, the bacterially produced variant IL-7 protein also has a reduced ability to induce T-cell proliferation than wild-type IL-7, indicating that these mutated regions may be significant contributors to the immunogenicity of prokaryotically produced IL-7.

EXAMPLE 5: Analysis of IL-7 derived peptides as potential B-cell epitopes

[0174] For bacterially-produced, unglycosylated human IL-7 protein, sequences surrounding N-linked glycosylation sites of IL-7 may be recognized as "non-self" by the human immune system, and elicit an antibody response. Essentially, to assess if these sequences represent linear B-cell epitopes, peptides spanning these sequences are used to immunize rabbits, and the reactivity of resulting antibodies toward bacterially-produced native human IL-7

and denatured human IL-7 is tested. As a further control, a native eukaryotically-produced glycosylated huFc-IL-7 fusion protein is used.

[0175] Methods and materials to raise polyclonal antibodies against a specific peptide antigen in, for example, rabbits, and their subsequent purification are generally known to those skilled in the art, and references thereto may be found, for example, in: Antibodies: A Laboratory Manual, E. Harlow and D. Lane, Cold Spring Harbor Press.

[0176] Briefly, in one example, a peptide containing the core sequence FLKMNSTGD (SEQ ID NO:5), such as LRQFLKMNSTGDFDL[C] (SEQ ID NO: 18) or [C]LRQFLKMNSTGDFDL (SEQ ID NO:19) is coupled via an added terminal cysteine to three different carrier proteins, keyhole limpet hemocyanin (KLH, EMD Biosciences, San Diego, CA), BSA (EMD Biosciences, San Diego, CA), and ovalbumin (Pierce, Rockford, IL) using a coupling agent such as SMCC (Pierce, Rockford, IL), and multiple rabbits are immunized by successive injections with each of the peptide conjugate in the presence of adjuvant. The immune response is boosted with further injections of one of the peptide-carrier conjugates at monthly intervals, and the resulting antiserum from each rabbit is affinity-purified over a Sulfo-Link column (Pierce, Rockford, IL) to which the peptide is coupled, and the antibody is further concentrated over a hydroxyapatite column (Bio-Rad Laboratories, Hercules, CA).

[0177] The purified antibodies are tested against bacterially-produced native human IL-7, denatured human IL-7 and a eukaryotically-produced glycosylated huFc-IL-7 fusion protein in an ELISA assay, following standard procedures. Briefly, ELISA plates, coated with the purified protein preparations, are incubated with the test antibody samples, the plates are washed, incubated with a secondary antibody such as horseradish peroxidase-coupled anti-rabbit IgG, washed again and incubated with a chromogenic substrate solution to indicate the concentration of bound antibody.

[0178] Similarly, antibodies are raised to other peptides encompassing the ...MNSTG...(SEQ ID NO:20) glycosylation site (at Asn70) in human IL-7, or to peptides encompassing the ...LNCTG...(SEQ ID NO:21) glycosylation site (at Asn91). Using this approach, it is found that generally the denatured bacterially-produced human IL-7 is well recognized by the antibodies and that the glycosylated huFc-IL-7 fusion protein is not. Peptides that give rise to antibodies

reacting with the bacterially-produced native human IL-7 protein indicate that a linear B-cell epitope at the glycosylation site is recognized. It is further found that, in a cell-based proliferation assay as described in Example 9, the antibodies raised against peptides of this Example have the effect of inhibiting IL-7-stimulated cell proliferation. This result indicates that these antibodies have neutralizing activity.

EXAMPLE 6: Construction of human IL-7 variants that lack potential T-cell epitopes

[0179] Nucleic acids are constructed that encode versions of human IL-7 variants either suitable for bacterial expression or suitable for eukaryotic expression, for instance as a fusion protein. For example, nucleic acids encoding a mature human IL-7 variant containing the substitutions K68D, M69D, I88T, and V96G (DeI-IL-7 SEQ ID NO:16) are constructed, using standard methods familiar to those skilled in the art. Figure 11 shows an example of such a DNA sequence encoding a mature IL-7 variant, DeI-IL-7, with codon substitutions of amino acid residues K68D, M69D, I88T, and V96G (SEQ ID NO:24).

[0180] For bacterial expression, the protein sequence of DeI-IL-7 including a start methionine (bDeI-IL-7 SEQ ID NO:17) is reverse-translated using a codon bias appropriate for optimal *E. coli* expression. The resulting nucleic acid sequence is further adapted to include desired (or to exclude undesired) features such as a stop codon or restriction sites, and sequences are added that facilitate cloning into a bacterial expression vector, for example an appropriate vector from the pET series (EMD Biosciences, San Diego, CA). The nucleic acid sequence is synthesized by total gene synthesis (Blue Heron Biotechnology, Bothell, WA) and inserted into the expression vector. An example of a DNA sequence encoding bDeI-IL-7, codon-optimized for *E. coli* with codon substitutions of amino acid residues K68D, M69D, I88T, V96G, is shown in Figure 10 (SEQ ID NO:23).

[0181] For eukaryotic expression as a huFc-DeI-IL-7 fusion protein, the nucleic acid sequence of the mature human IL-7 is modified to incorporate codons for desired amino acid mutations of the invention as described above. (See *e.g.* SEQ ID NO:24). The sequence is further adapted to incorporate flanking sequences with unique restriction sites for insertion as a Xma I / Xho I fragment in-frame into a pdCs-huFc expression vector encoding the hinge, CH2

and CH3 region of IgG1 (see Lo *et al.*, (1998), Protein Engineering 11:495), and is synthesized by total gene synthesis (Blue Heron Biotechnology, Bothell, WA). The synthetic Xma I / Xho I DeI-IL-7 fragment is then cloned into the pdCs-huFc vector, yielding an expression plasmid encoding huFc-DeI-IL-7. Other IL-7 and Fc-IL-7 variants of the invention can be produced by similar methods.

[0182] Specifically, nucleic acids encoding the human deimmunized Fc-IL-7 fusion proteins huFcγ2(h)(FN>AQ)-(linker2)-IL-7(PNS) and huFcγ2(h)(FN>AQ)-(linker2)-IL-7(PNDS) were generated as follows. huFcγ2(h)(FN>AQ)-(linker2)-IL-7(PNS) is a human Fc-IL-7 fusion protein comprising the N-terminus of human IL-7 genetically fused to the C-terminus of a human IgG2 Fc domain with an IgG1 hinge via a linker sequence GGGGSGGGG. The Fc portion contains the mutations Phe296Ala and Asn297Gln. The IL-7 portion contains the mutations F39P, F57N and L128S. huFcγ2(h)(FN>AQ)-(linker2)-IL-7(PNDS) is the same as huFcγ2(h)(FN>AQ)-(linker2)-IL-7(PNS), but for containing an additional mutation in the IL-7 moiety, L77D. The sequence also contains codons for the mutation of the LSLS sequence near the C-terminus of the Fc portion to be replaced by ATAT. In addition, the nucleic acid sequence includes a codon to replace the C-terminal lysine of the Fc portion with an alanine residue.

[0183] A nucleic acid of the sequence presented in Figure 27 was synthesized de novo (Blue Heron Biotechnology, Bothell, WA), which encodes the linker sequence GGGGSGGGG followed by mature human IL-7 containing the amino acid substitutions F39P, F57N, and L128S (IL-7(PNS)) and which contains flanking restriction sites Xma I and Xho I at the 5'- and 3' ends, respectively. This purified Xma I / Xho I fragment was ligated to a likewise digested and purified vector fragment of the pdCs-huFc series, pdC10-huFcγ2(h)(FN>AQ), generating a plasmid encoding huFcγ2(h)(FN>AQ)-(linker2)-IL-7(PNS). Lo *et al.*, (1998), Protein Engineering 11:495. The coding sequence was ascertained by sequencing.

[0184] The further introduction of the substitution L77D into IL-7(PNS) was performed by standard PCR mutagenesis methods, using mutagenic primers M(s) (5'-TGACTTTGATGACCACCTGTTAAAAGTTTC-3' (SEQ ID NO: 50); mutated codon underlined) and M(a) (5'- AACAGGTGGTCATCAAAGTCACCAGTGC-3') (SEQ ID NO:51). Briefly,

separate PCR reactions were performed on a plasmid template containing (linker2)-IL-7(PNS), one with M(s) and the downstream primer 5' – CTCGAGTCAGTGTTCTTTAGTGCCCATC – 3' (SEQ ID NO:52), the other with M(a) and the upstream primer 5' – CCCGGGTGCTGGAGGTGGAGGATCAGGTG- 3' (SEQ ID NO:53), the PCR fragments were purified and combined as the template for a secound round of PCR using again the upstream primer 5' – CCCGGGTGCTGGAGGTGGAGGATCAGGTG- 3'(SEQ ID NO:54) and the downstream primer 5' – CTCGAGTCAGTGTTCTTTAGTGCCCATC – 3'(SEQ ID NO:55). The resultant purified fragment was inserted into a TA cloning vector pCR2.1 ((Invitrogen, Carlsbad, CA), and its sequence was confirmed. An Xma I / Xho I fragment encoding (linker2)-IL-7(PNDS) was excised, and ligated to a likewise digested and purified vector fragment of the pdCs-huFc series, pdC10-huFcγ2(h)(FN>AQ), generating a plasmid encoding huFcγ2(h)(FN>AQ)-(linker2)-IL-7(PNDS).

[0185] Similarly, plasmids encoding variants of these fusion proteins differing in the Fc moiety are obtained; for example a plasmid encoding huFcγ2(h)(linker2)-IL-7(PNDS) is obtained by ligating an Xma I / Xho I fragment encoding (linker2)-IL-7(PNDS) to a likewise digested and purified vector fragment of the pdCs-huFc series, pdC10-huFcγ2(h).

EXAMPLE 7: Expression and purification of IL-7 variants

[0186] For eukaryotic expression of the huFc-Del-IL-7 fusion protein, electroporation is used to introduce the DNA encoding the fusion protein into a mouse myeloma NS/0 cell line. To perform electroporation NS/0 cells are grown in Dulbecco's modified Eagle's medium supplemented with 10% heat-inactivated fetal bovine serum, 2 mM glutamine and penicillin/streptomycin. About $5x10^6$ cells are washed once with PBS and resuspended in 0.5 ml PBS. 10 µg of linearized plasmid DNA for huFc-Del-IL-7 is then incubated with the cells in a Gene Pulser Cuvette (0.4 cm electrode gap, BioRad) on ice for 10 min. Electroporation is performed using a Gene Pulser (BioRad, Hercules, CA) with settings at 0.25 V and 500 µF. Cells are allowed to recover for 10 min on ice, after which they are resuspended in growth medium and plated onto two 96 well plates.

[0187] Stably transfected clones are selected by their growth in the presence of 100 nM methotrexate (MTX), which is added to the growth medium two days post-transfection. The cells are fed every 3 days two to three more times, and MTX-resistant clones appear in 2 to 3 weeks. Supernatants from clones are assayed by anti-Fc ELISA to identify clones that produced high amounts of the IL-7 fusion protein. High producing clones are isolated and propagated in growth medium containing 100 nM MTX. Typically, a serum-free growth medium, such as H-SFM or CD medium (Life Technologies), is used.

[0188] A standard purification of Fc-containing fusion proteins is performed based on the affinity of the Fc protein moiety for Protein A. Briefly, NS/0 cells expressing the fusion protein, such as huFc-Del-IL-7, are grown in tissue culture medium and the supernatant containing the expressed protein is collected and loaded onto a pre-equilibrated Fast Flow Protein A Sepharose column. The column is then washed extensively with buffer (such as 100 mM sodium phosphate, 150 mM NaCl at neutral pH). Bound protein is eluted at a low pH (pH 2.5 - 3) in same buffer as above and fractions are immediately neutralized.

[0189] Bacterial expression and purification of bDel-IL-7 is performed essentially as described by Cosenza *et al.* for bacterially-produced IL-7 (see Cosenza *et al.*, (1997) JBC, 272:32995). In essence, bDel-IL-7 is isolated from inclusion bodies, denatured and refolded. Briefly, bacterial expression cultures transformed with the expression vector encoding, for example, bDel-IL-7 are grown to mid-log phase and recombinant protein expression is induced. Following induction, the bacteria are harvested and lysed by sonication, and inclusion bodies are isolated in buffer A (50 mM Tris HCl (7.5), 5 mM EDTA, 20% sucrose). After extensive washes, the inclusion bodies are resuspended in a guanidine denaturation buffer (50 mM Tris-HCl (pH8.0), 5 M guanidine HCl, 5 mM EDTA), briefly sonicated and reduced in 6 mM DTT. Denatured bDel-IL-7 protein is then further purified by denaturing size exclusion HPLC. The protein is then refolded in refolding buffer (50 mM glycine, 30 mM NaOH, 0.4 M L-arginine, 1 mM DTT, pH 10), dialyzed into a phosphate buffer, and further purified by size exclusion HPLC.

EXAMPLE 8: Biochemical analysis of IL-7 variants

[0190] The effect of the introduced mutations on the integrity of IL-7 proteins is assessed by routine reducing and non-reducing SDS-PAGE analysis and size exclusion chromatography.

[0191] For example, the fusion protein huFc-Del-IL-7, expressed from NS/0 cells, is captured on Protein A Sepharose beads (Repligen, Needham, MA) from the tissue culture medium into which it is secreted, and eluted by boiling in protein sample buffer, with or without a reducing agent such as β-mercaptoethanol. The sample is fractionated by SDS-PAGE and the protein bands are visualized by Coomassie staining. It is expected that a fusion protein containing IL-7 mutations that sufficiently interfere with proper folding is more likely to show degradation products by SDS-PAGE.

[0192] Purified huFc-Del-IL-7 is also analyzed by size exclusion chromatography (SEC) to assess the extent to which the fusion protein is aggregated. Briefly, the cell culture supernatant is loaded onto a pre-equilibrated Fast-Flow Protein A Sepharose column, the column is washed extensively in a physiological buffer (such as 100 mM Sodium Phosphate, 150 mM NaCl at neutral pH), and the bound protein is eluted at about pH 2.5 to 3 in same salt buffer as above. Fractions are immediately neutralized, peak fractions are pooled, and an aliquot is fractionated over an analytical SEC column.

EXAMPLE 9: *In vitro* activity of IL-7 variants

[0193] To determine whether the IL-7 variants containing mutations of the invention retain their cytokine activity *in vitro* cellular proliferation bioassays are performed. Human PBMC (Peripheral Blood Mononuclear Cells) are activated by PHA-P to produce cells which are responsive to IL-7. Proliferation is measured in a standard thymidine incorporation assay.

[0194] For example, the cytokine activity of huFc-Del-IL-7 and bDel-IL-7 is determined. Briefly, PBMC's are first incubated for five days with 10 microgram/ml PHA-P, cells are washed and then incubated in medium with huFc-Del-IL-7 or bDel-IL-7, prepared as a dilution series, for a total of 48 hours. During the final 12 hours, the samples are pulsed with 0.3 μCi of [methyl-3H]thymidine (Dupont-NEN-027). Cells are then washed extensively, harvested and lysed onto glass filters. ^{3}H-thymidine incorporated into DNA is measured in a scintillation

counter. As a standard, wild type hulL-7 protein, obtained from R&D Systems (Minneapolis, MN), or obtained from the National Institute for Biological Standards and Control (NIBSC), is assayed.

[0195] An ED50 value of cell proliferation for huFc-Del-IL-7 or bDel-IL-7 is obtained from plotting a dose response curve according to standard techniques, and determining the protein concentration that results in half-maximal response.

Example 10: Induction of anti-human IL-7 antibodies in monkeys by wild-type IL-7 and IL-7 variants

[0196] It is known that bacterial-derived wild-type human IL-7 administered to monkeys often results in neutralizing anti-human IL-7 antibody titers (Storek *et al.*, (2003), Blood, 101:4209; Fry *et al.*, (2003), Blood, 101:2294). Thus, the propensity of prokaryotically produced variant IL-7 and wild-type IL-7 proteins, as well as eukaryotically produced fusion proteins containing wild-type or variant IL-7 polypeptides, to induce neutralizing antibodies in nonhuman primates is assessed. In a typical experiment, rhesus macaques are injected with 40 μg/kg of the protein samples subcutaneously once a day for four weeks. For example, the protein samples are commercially available prokaryotically-produced IL-7 (PeproTech, Rocky Hill, NJ), the prokaryotically produced variant IL-7(K68D, M69D, I88T, V94G), and the equivalent Fc-IL-7 fusion proteins produced in a mammalian expression system. At regular intervals, serum is obtained from the animals, and serum concentrations of antibodies against human IL-7 are measured by ELISA using human IL-7 coated 96 well plates (Nunc, Naperville, IL). Typically, serial dilutions of each serum sample are added to each well in triplicate for two hours, washed with 0.05% Tween (Tween 20) in PBS and blocked with 1% BSA/1% goat serum in PBS. To each sample a horseradish peroxidase-conjugated anti-macaque IgG is added (1:60,000 in sample buffer), incubated at 37°C for 2 hr, and the plate is washed 8 times with 0.05% Tween in PBS. Samples are then assayed using the colorimetric substrate solution OPD (o-phenylenediamine dihydrochloride) by measuring the OD at 490 nm, subtracting the background OD reading at 650 nm.

[0197] It is found that prokaryotically produced wild-type IL-7 protein indeed gives rise to high anti-IL-7 antibody titers. In contrast, the antibody titers of the prokaryotically produced variant IL-7 gives rise to significantly lower titers of anti

IL-7 antibodies. It is also found that the differences in the levels of anti IL-7 antibody titers produced by animals administered mammalian produced wild-type and variant Fc-IL-7 fusion proteins, (with mutations around the N-linked glycosylation sites) are not as pronounced. This result may indicate that the lack of glycosylation at these sites in the prokaryotically produced proteins contributes to the immunogenicity of these proteins.

Example 11: Acute tolerability of Fc-IL-7 in immunocompetent mice

[0198] The Fc-IL-7 fusion protein huFcγ2(h)(N>Q)(linker2)-huIL-7 was prepared according to the method described in Example No. 6, purified, then formulated in 50mM phosphate, 150mM sodium chloride pH 7.00, 0.05% (v/v) Tween 80. The protein concentrations of diluted solutions were determined using the absorbance at 280nm and the theoretical extinction coefficient of 0.98 mg/OD_{280}, based on the known protein sequence. For dosing mice, an aliquot of each sample was removed from stock vials and diluted with 0.9% saline within one hour of dosing.

[0199] C57B1/6 mice (Charles River Laboratories, Wilmington, MA), 17 weeks of aged were divided into groups of 2 mice each and were administered the Fc-IL-7 fusion proteins subcutaneously for 5 consecutive days. Groups received dosages of either 0.5, 5.0, 25 mg/kg or the vehicle control each day. All mice survived the treatment through day 7, at which point the mice were sacrificed.

[0200] Fc-IL-7 plasma levels were determined by obtaining blood samples from the retro-orbital sinus 6 hours after dosing on days 0, 2, 4, and 7. Blood samples were collected in tubes containing heparin to prevent clotting. Cells were removed by centrifugation and the concentration of intact Fc-IL-7 fusion protein in the plasma was measured using standard ELISA procedures. Plasma levels of Fc-IL-7 in $\mu g/\mu l$ for test mice are shown in Figure 36. The plasma of mice showed a dose dependent increase in Fc-IL-7 concentration at all time points tested and further increased following each dose. However, as shown in Figure 37, the magnitude of the increase lessened after each dosing.

[0201] The functional activity of Fc-IL-7 was confirmed by measuring increases in B cells and T cells on day 7 following the initiation of dosing. Since IL-7 boosts the production of immune effector cells such as B cells and T cells,

the cellularity and weight of the spleen is expected to increase. Mice were sacrificed on day 7 and organs were removed and weighed. Figure 38 shows the average organ weights on day 7. As expected, spleen weight increased 3 to 5 fold 1 week after the initial dose. Lung weights increased 2 fold following the 2 higher doses of Fc-IL-7 due to lymphotcytic infiltration. No weight changes were observed in the kidney or liver.

[0202] The response of B cells (CD19+, CD4+, CD8+ and granulocytes (Gr-1+)) in all groups was observed on day 7. Fig. 39 shows the frequency of Gr-1+ cells in the peripheral blood of two mice in each group. As the data shows, granulocytes were generally unresponsive to Fc-IL-7. Fig. 40 shows the frequency of CD19+ cells in the peripheral blood of two mice in each group. Fig. 41 shows the frequency of CD4+ cells in the peripheral blood of two mice in each group, while Fig. 42 shows the frequency of CD8+ cells in the peripheral blood of two mice in each group. The increases in B cell (Fig. 40 and T cell (Figs. 41 and 42 numbers were maximal for the 5 mg/kg dosage group with each mouse tested showing significantly increased T cell and B cell numbers over the control group and the 0.5 mg/kg dosage group. However, T cell numbers either declined or increased for mice in the 25 mg/kg dosage group. All measurements of cells are shown in cells per μL of blood.

Example 12. Assessment of Human Fc-IL-7 Activity

[0203] The biological activity of the Fc-IL-7 fusion protein tested in Example 11 was measured by tritiated thymidine uptake in a standard cell proliferation assay using peripheral blood mononuclear cell (PBMC) PHA blasts according to the method described in Yokota *et al.*, (1986), Proc. Natl. Acad. Sci. USA, 83:5894; and Stern *et al.*, (1990), Proc. Natl. Acad. Sci. USA, 87:6808-6812, with human IL-7 used as a standard. As shown in Figure 43, cellular proliferation as measured by the uptake of tritiated thymidine for the Fc-IL-7 molecule is similar to that of the standard NIBSC human IL-7 (World Health Organization), indicating that the activity of the Fc-IL-7 molecule is similar to wild-type human IL-7.

Equivalents

[0204] The invention may be embodied in other specific forms without departing from the spirit or essential characteristics thereof. The foregoing

embodiments are therefore to be considered in all respects illustrative rather than limiting on the invention described herein. Scope of the invention is thus indicated by the appended claims rather than by the foregoing description, and all changes which come within the meaning and range of equivalency of the claims are intended to be embraced therein.

Patent Claims

1. A polypeptide comprising a modified human IL-7 molecule or an active portion thereof having a T-cell epitope modified to reduce an anti-IL-7 T-cell response, wherein said modified IL-7 comprises an amino acid substitution at one or more residues of the wild-type IL-7 corresponding to Gln22, Leu24, Ile30, Phe39, Met54, Phe57, Arg58, Ala60, Arg61, Leu63, Lys68, Met69, Leu77, Ile88, Val96, Leu104, Leu128, Met147, Thr149, or Lys150.

2. A polypeptide of claim 1 comprising at least two of said amino acid substitutions.

3. A polypeptide of claim 1 comprising at least four of said amino acid substitutions.

4. A polypeptide of any of the claims 1 - 3, wherein the substitution is a substitution at one or more of residues Phe39, Phe57, Leu77, and Leu128.

5. A polypeptide of any of the claims 1 - 4, comprising one or more of the following substitutions: Gln22Asp, Leu24Asp, Ile30Thr, Phe39Pro, Met54Ala, Phe57Lys, Phe57Asn, Arg58Asp, Ala60Ser, Arg61Glu, Leu77Asp, Leu104Ser, Leu104Val, Leu128Ala, Leu128Val, Leu128Pro, Leu128Ser Met147Lys, Thr149Ser, or Lys150Stop.

6. A polypeptide of claim 5, wherein the polypeptide comprises one or more of substitutions Phe39Pro, Phe57Asn, Leu77Asp, and Leu128Ser.

7. A polypeptide of claim 6, wherein the polypeptide comprises the substitutions Phe39Pro, Phe57Asn, Leu77Asp, and Leu128Ser.

8. A polypeptide according any of the claims 1 – 7, wherein said polypeptide is at least 80% identical to human wild-type Il-7.

9. A polypeptide according to any of the claims 1 – 8, wherein said polypeptide further comprises an immunoglobulin molecule or a portion thereof fused to said modified IL-7 molecule or active fragment thereof.

10. A polypeptide of claim 9, wherein said immunoglobulin molecule or portion thereof is fused via its C-terminal to the N-terminal of said modified IL-7 molecule or active fragment thereof.

11. A polypeptide of claim 9 or 10, wherein said immunoglobulin molecule is a human immunoglobulin.

12. A polypeptide of any of the claims 9 – 11, wherein said immunoglobulin molecule is IgG2.

13. A polypeptide of any of the claims 9 – 12, wherein said immunoglobulin is an Fc portion.

14. A nucleic acid molecule encoding a polypeptide according to any one of claims 1-13.

15. An expression vector comprising a nucleic acid molecule of claim 14.

16. A prokaryotic cell comprising an expression vector of claim 15.

17. Use of a polypeptide according to any of the claims 1 – 13 for the manufacture of a medicament for the treatment of cancer diseases or HIV.

18. Use of claim 17, wherein the effective amount of the polypeptide administered to the patient is between 0.01 and 10 mg/kg/day.

19. A method of treating a patient suffering from cancer diseases or HIV by administering to said patient a therapeutically effective amount of a polypeptide according to any one of claims 1-13.