IL-6 BINDING PROTEINS

CROSS-REFERENCE TO RELATED PATENT APPLICATIONS

[01] The present patent application claims benefit of priority to U.S. Provisonal Patent Application No. 60/699,137, filed July 13, 2005, which is incorporated by reference for all purposes.

BACKGROUND OF THE INVENTION

[02] Interleukin-6 (IL-6) has a central role in regulating the immune response and inflammation. IL-6 has a range of biological activities, acting in an endocrine, autocrine, and paracrine fashion on a diverse number of target cells. The major sources of IL-6 are monocytes. Fibroblasts and endothelial cells along with T cells and B cells, keratinocytes, mesangial cells, and several tumor cells produce IL-6. Its synthesis is induced by IL-1, IL-2, TNF-β and interferons. IL-6 is inhibited by IL-4, IL-10, and IL-13. IL-6 thus plays an important role in immune regulation and inflammation as well as autoimmune diseases.

[03] IL-6 has the ability to stimulate the final stages of B lymphocyte maturation, whereby B lymphocytes differentiate into mature plasma cells and secrete immunoglobulins (Ig). In addition, IL-6 induces T cell growth and cytotoxic T cell differentiation through augmentation of IL-2 receptor expression and IL-2 production. IL-6 synergizes with other cytokines to support bone marrow stem cell maturation, is a neutrophil activator and stimulates the production of platelets from megakaryocytes. It is also a potent inducer of terminal macrophage and osteoclast differentiation. IL-6 also shares several activities with IL-1 and TNF, including the induction of pyrexia and the production of acute phase proteins such as serum amyloid A, CRP, alpha 1 antitrypsin, fibrinogen, and haptoglobulin.

[04] IL-6 mediates several unique anti-inflammatory effects. Where both IL-1 and TNF-α induce synthesis of each other, as well as IL-6, IL-6 terminates this upregulatory inflammatory cascade and inhibits IL-1 and TNF-α synthesis. IL-6 is also important as a cell growth factor that can induce proliferation of epidermal keratinocytes and mesangial cells and has an active role in mesangial proliferative glomerulonephritis.

[05] The receptor for IL-6 (IL-6R) belongs to the type I cytokine receptor superfamily and consists of two chains, IL-6R (80 kDa glycoprotein) and gp130. IL-6R is the ligand specific binding component while gp130 is a shared receptor component responsible for transmitting intracellular signals of IL-6 related cytokines such as leukemia inhibitory factor (LIF), ciliary neurotropic factor, oncostatin M, and IL-11.

[06] IL-6 binding to the receptor leads to dimerization of gp130, resulting in the activation of gp130 associated kinase JAK1 and subsequently tyrosine phosphorylation of gp130. Whereas gp130 is ubiquitously expressed, IL-6R is more restricted. Both IL-6R and gp130 have soluble forms. The complex of IL-6 and soluble IL-6R can act on cells expressing only gp130. This is potentially an important proinflammatory mechanism for soluble receptors.

BRIEF SUMMARY OF THE INVENTION

[07] The present invention provides a polypeptide comprising a monomer domain that binds to (e.g., selectively binds) IL-6. In some embodiments, the polypeptide comprises a monomer domain that binds to IL-6, wherein the monomer domain is an LDL receptor class A monomer domain comprising either of the following sequences:

CXXXQFQCXSG[W,Y]C[I,L,F]XXXW[V,L]CDGXNDCEDXSDEXXC (SEQ ID NO:2)

or

CX[A,P]X[E,Q]FXCXXGQC[I,L]PXXXXCDGX[D,N]DCXDXSDEXXC (SEQ ID NO:3),

wherein X is any amino acid and amino acids in brackets indicate alternative amino acids at a specified position.

[08] In some embodiments, the polypeptide comprises CXXXQFQCXSG[W,Y]C[I,L,F]XXXW[V,L]CDGXNDCEDXSDEXXC (SEQ ID NO:2).

[09] In some embodiments, the polypeptide comprises CX[A,P]X[E,Q]FXCXXGQC[I,L]PXXXXCDGX[D,N]DCXDXSDEXXC (SEQ ID NO:3).

[10] In some embodiments, the polypeptide comprises at least one and no more than six monomer domains that bind IL-6.

[11] In some embodiments, the polypeptide comprises at least two monomer domains that bind IL-6. In some embodiments, the polypeptide comprises a first monomer domain with binding specificity for IL-6 and a second monomer domain with binding specificity for IL-6; wherein:

the first monomer domain comprises

CXXXQFQCXSG[W,Y]C[I,L,F]XXXW[V,L]CDGXNDCEDXSDEXXC (SEQ ID NO:2); and

the second monomer domain comprises

CX[A,P]X[E,Q]FXCXXGQC[I,L]PXXXXCDGX[D,N]DCXDXSDEXXC (SEQ ID NO:3).

[12] In some embodiments, the polypeptide comprises:

CX[A,P]X[E,Q]FXCXXGQC[I,L]PXXXXCDGX[D,N]DCXDXSDEXXCX$_n$CXXXQFQCX SG[W,Y]C[I,L,F]XXXW[V,L]CDGXNDCEDXSDEXXC (SEQ ID NO:4), wherein n is between 0 and 100. In some embodiments, n is 3-15, 1-20, 0-30, or 0-50.

[13] In some embodiments, the polypeptide comprises any of the following sequences (SEQ ID NOS:5-13):

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCRADEFQCG
SGRCVPADWLCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTQHT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCQANEFTCS
SGNCLSPALLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQRWVCDGEND
CEDSSDEANCAGSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDEASCQPPT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEADCAAPT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCEDDSDEESCPPHTCLPDQFQCGSGYCLSQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCESDQFQCD
SGHCVPLNWVCDGVNDCVDGSDEAGCPPRTCLPGQFQCGSGYCISQRWVCDGEND
CEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT;
or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCGASQFQCG
SGYCISQRWVCDGENDCEDSSDEANCAGSVPTCLPSEFQCRDSKICLSADWLCDGVN
DCADNSDEEGCPPPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETNCPART.

[14] In some embodiments, the polypeptide comprises any of the following sequences (SEQ ID NOS:14-22):

CLPDQFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRW
VCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEE
DCTEHT;

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHT;

CRADEFQCGSGRCVPADWLCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHT;

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTQHT;

CQANEFTCSSGNCLSPALLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQR
WVCDGENDCEDSSDEANCAGSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDE
ASCQPPT;

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EADCAAPT;

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPHTCLPDQFQCGSGYCLSQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHT;

CESDQFQCDSGHCVPLNWVCDGVNDCVDGSDEAGCPPRTCLPGQFQCGSGYCISQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHT; or

CGASQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCLPSEFQCRDSKICLS
ADWLCDGVNDCADNSDEEGCPPPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSD
ETNCPART.

[15] In some embodiments, the polypeptide comprises one of the IL-6-binding monomer, dimer or trimer sequences displayed herein.

[16] In some embodiments, the polypeptide comprises three monomer domains that bind IL-6.

[17] In some embodiments, the polypeptide further comprises a monomer domain that has a binding specificity for a blood factor, thereby increasing the serum half-life of the polypeptide when the polypeptide is injected into an animal compared to the serum half-life of a polypeptide lacking the blood factor-binding monomer domain. In some embodiments, the blood factor is serum albumin, an immunoglobulin or an erythrocyte.

[18] In some embodiments, the monomer domain that has a binding specificity for a blood factor binds to immunoglobulin (IgG) and the monomer domain that has a binding specificity for a blood factor is an LDL receptor class A monomer domain comprising a sequence selected from the following:
CXSSGRCIPXXWVCDGXXDCRDXSDE (SEQ ID NO:23), and
CXSSGRCIPXXWLCDGXXDCRDXSDE (SEQ ID NO:24),
wherein X is any amino acid.

[19] In some embodiments, the monomer domain that has a binding specificity for a blood factor binds to immunoglobulin (IgG) and the monomer domain that has a binding specificity for a blood factor is an LDL receptor class A monomer domain comprising the following sequence:
[EQ]FXCRX[ST]XRC[IV]XXXW[ILV]CDGXXDCXD[DN]SDE (SEQ ID NO:25),
wherein X is any amino acid and amino acids in brackets are alternative amino acids at a single position.

[20] In some embodiments, the monomer domain that has a binding specificity for a blood factor comprises
CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENC (SEQ ID NO:26).

[21] In some embodiments, the monomer domain is between 35 to 45 amino acids. In some embodiments, the monomer comprises an amino acid sequence in which: at least 10% of the amino acids in the sequence are cysteine; and/or at least 25% of the amino acids are non-naturally-occurring amino acids.

[22] The present invention also provides methods for identifying a polypeptide that binds to (e.g., selectively binds) IL-6. In some embodiments, the method comprises,

screening a library of polypeptides for affinity to IL-6; and

selecting a polypeptide comprising at least one monomer domain that binds to IL-6,

wherein the monomer domain:

is a non-naturally-occurring monomer domain;

comprises at least one disulfide bond; and

binds to an ion.

[23] In some embodiments, the selected polypeptide comprises an LDL receptor A domain monomer domain comprising at least one of:

Family 1 motif (SEQ ID NO:27):
CX[AS][DN]EFXCXXGXC[ILV]XXAXXCDGV[DN]DCXDNSDEXXC[AST]XX[AST]

Family 2 motif (SEQ ID NO:28):
CXXXQFXC[DN]NGXCVPXXWXCDG[DE][DN]DCXD[DN]SDEXXCXXP[AST]

Family 3 motif (SEQ ID NO:29):
CXXXQFQCXSG[W,Y]C[I,L,F]XXXW[V,L]CDGXNDCEDXSDEXXC

Family 4 motif (SEQ ID NO:30):
CXXX[EQ]FXCXXGXC[ILV]XXXXXCDGX[DN]DCXDXSDEXX(XX)C

Family 5 motif (SEQ ID NO:31):
CX[A,P]X[E,Q]FXCXXGQC[I,L]PXXXXCDGX[D,N]DCXDXSDEXXC

Family 6 motif (SEQ ID NO:32):
CXXX[EQ]FXCX(X)SXXC[ILV]XX[DN]W[ILV]CDGVNDCXDXSDEXGCXXXT

Family 7 motif (SEQ ID NO:33):
CL[AS]X[EQ]FXCXNGXCIP[KQR]XWVCDGDNDC[ILV]DXSDE(P)XXCX[EKQ]XT.

[24] In some embodiments, the monomer domain comprises an amino acid sequence in which at least 10% of the amino acids in the sequence are cysteine; and/or at least 25% of the amino acids are non-naturally-occurring amino acids.

[25] In some embodiments, the method further comprises

linking the monomer domain in the selected polypeptide to a second monomer domain to form a library of multimers, each multimer comprising at least two monomer domains;

screening the library of multimers for the ability to bind to IL-6; and

selecting a multimer that binds IL-6.

[26] In some embodiments, the method further comprises

linking the monomer domain in the selected polypeptide to a second monomer domain to form a library of multimers, each multimer comprising at least two monomer domains;

screening the library of multimers for the ability to bind to a target molecule other than IL-6; and

selecting a multimer that binds to the target molecule.

[27] In some embodiments, the method further comprises a step of mutating at least one monomer domain, thereby providing a library comprising mutated monomer domains.

[28] In some embodiments, the library of monomer domains is expressed as a phage display, ribosome display or cell surface display.

[29] In some embodiments, the polypeptide comprises at least two monomer domains and the monomer domains are linked by a linker. In some embodiments,

the linker is a peptide linker. In some embodiments, the linker is between 4 to 12 amino acids long.

[30] In some embodiments, the monomer domain comprises an amino acid sequence in which at least 10% of the amino acids in the sequence are cysteine; and/or at least 25% of the amino acids are non-naturally-occurring amino acids.

[31] The present invention also provides polynucleotides encoding a polypeptide comprising a monomer domain that binds to (e.g., selectively binds) IL-6, wherein the monomer domain:

is a non-naturally-occurring monomer domain consisting of 30 to 50 amino acids;

comprises at least one disulfide bond.

[32] In some embodiments, the polynucleotides encode a polypeptide comprising one of the following sequences:

Family 1 motif (SEQ ID NO:27):
CX[AS][DN]EFXCXXGXC[ILV]XXAXXCDGV[DN]DCXDNSDEXXC[AST]XX[AST]

Family 2 motif (SEQ ID NO:28):
CXXXQFXC[DN]NGXCVPXXWXCDG[DE][DN]DCXD[DN]SDEXXCXXP[AST]

Family 3 motif (SEQ ID NO:29):
CXXXQFQCXSG[W,Y]C[I,L,F]XXXW[V,L]CDGXNDCEDXSDEXXC

Family 4 motif (SEQ ID NO:30):
CXXX[EQ]FXCXXGXC[ILV]XXXXXCDGX[DN]DCXDXSDEXX(XX)C

Family 5 motif (SEQ ID NO:31):
CX[A,P]X[E,Q]FXCXXGQC[I,L]PXXXXCDGX[D,N]DCXDXSDEXXC

Family 6 motif (SEQ ID NO:32):
CXXX[EQ]FXCX(X)SXXC[ILV]XX[DN]W[ILV]CDGVNDCXDXSDEXGCXXXT

Family 7 motif (SEQ ID NO:33):
CL[AS]X[EQ]FXCXNGXCIP[KQR]XWVCDGDNDC[ILV]DXSDE(P)XXCX[EKQ]XT.

[33] In some embodiments, the polynucleotides encode a polypeptide comprising one of the following sequences (SEQ ID NOS:5-13)

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCRADEFQCG
SGRCVPADWLCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTQHT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCQANEFTCS
SGNCLSPALLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQRWVCDGEND
CEDSSDEANCAGSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDEASCQPPT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEADCAAPT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCEDDSDEESCPPHTCLPDQFQCGSGYCLSQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCESDQFQCD
SGHCVPLNWVCDGVNDCVDGSDEAGCPPRTCLPGQFQCGSGYCISQRWVCDGEND
CEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT;
or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCGASQFQCG
SGYCISQRWVCDGENDCEDSSDEANCAGSVPTCLPSEFQCRDSKICLSADWLCDGVN
DCADNSDEEGCPPPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETNCPART.

[34] In some embodiments, the polynucleotides encode a polypeptide comprising one of the following sequences (SEQ ID NOS:14-22):

CLPDQFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRW
VCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEE
DCTEHT;

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHT;

CRADEFQCGSGRCVPADWLCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHT;

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTQHT;

CQANEFTCSSGNCLSPALLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQR
WVCDGENDCEDSSDEANCAGSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDE
ASCQPPT;

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EADCAAPT;

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPHTCLPDQFQCGSGYCLSQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHT;

CESDQFQCDSGHCVPLNWVCDGVNDCVDGSDEAGCPPRTCLPGQFQCGSGYCISQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHT; or

CGASQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCLPSEFQCRDSKICLS
ADWLCDGVNDCADNSDEEGCPPPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSD
ETNCPART.

[35] In some embodiments, the polynucleotide encoding a monomer domain with binding specificity for IL-6 is optimized for codon usage of a cell (e.g., a bacterial, plant or mammalian cell). For example,

ATGAAACACTCTGGTGGTTCGTGTCACCCTACCGGCCAATTTAGATGCCGCAGTA
GCGGCAGATGCGTTTCACCGACCTGGGTTTGCGACGGCGATAATGATTGTGGGG
ACAACTCGGATGAAGAAAACTGTAGCGCGCCAGCATCTGAACCGCCGGGTTCCC
TGTGTCTGCCGGACCAGTTCCGGTGCGGCAACGGCCAATGCATTCCACTGGACTG
GGTTTGCGACGGCGTTAATGATTGTCCGGACGACTCGGATGAAGAAGGCTGTCC
ACCGCGTACGTGTGCGCCGAGCCAGTTCCAGTGCGGCAGCGGCTACTGCATTTCA
CAGCGCTGGGTTTGCGACGGCGAAAATGATTGTGAGGACGGCTCGGATGAAGCA
AACTGTGCAGGCTCCGTACCTACGTGTCCGTCGGACGAGTTCCGGTGCAGAAACG
GCCGCTGCATTCCACGGGCCTGGCGTTGCGACGGCGTAAATGATTGTGCGGACA
ACTCGGATGAAGAAGACTGTACAGAGCATACGTCATTGCAGTAG (SEQ ID NO:34)

encodes the IL-6 binding trimer

MKHSGGSCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLC
LPDQFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRW
VCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEE
DCTEHTSLQ* (SEQ ID NO:35). However,

ATGAAACACTCTGGTGGTAGCTGTCACCCTACCGGCCAATTTCGTTGCCGCAGTA
GCGGCCGTTGCGTTTCACCGACCTGGGTTTGCGACGGCGATAATGATTGTGGCGA
CAACAGCGATGAAGAAAACTGTAGCGCGCCAGCATCTGAACCGCCGGGTTCCCT
GTGTCTGCCGGACCAGTTCCGTTGCGGCAACGGCCAATGCATTCCACTGGACTGG
GTTTGCGACGGCGTTAATGATTGTCCGGACGACTCGGATGAAGAAGGCTGTCCAC
CGCGTACGTGTGCGCCGAGCCAGTTCCAGTGCGGCAGCGGCTACTGCATTAGCCA
GCGCTGGGTTTGCGACGGCGAAAATGATTGTGAGGACGGCTCGGATGAAGCAAA
CTGTGCAGGCTCCGTACCTACGTGTCCGTCGGACGAGTTCCGCTGCCGTAACGGC
CGCTGCATTCCACGTGCCTGGCGCTGCGACGGCGTAAATGATTGTGCGGACAACA

GCGATGAAGAAGACTGTACCGAGCATACGTCATTGCAGTAA (SEQ ID NO:36) is a polynucleotide that also encodes MKHSGGSCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLC LPDQFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRW VCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEE DCTEHTSLQ* (SEQ ID NO:35), but has been codon optimized for bacterial expression.

DEFINITIONS

[36] Unless otherwise indicated, the following definitions supplant those in the art.

[37] As used herein, use of parentheses in a sequence motif indicates that that the positions within the parentheses may be present or absent (e.g., "([ekq])" indicates that the position is absent or either E, K, or Q may be at that position). When more than one "x" is used in parentheses (e.g., "(xx)"), each x represents a possible position. Thus "(xx)" indicates that zero, one or two amino acids may be at that position(s), where each amino acid is independently selected from any amino acid.

[38] The terms "monomer domain" or "monomer" are used interchangeably and herein refer to a discrete region found in a protein or polypeptide. A monomer domain forms a native three-dimensional structure in solution in the absence of flanking native amino acid sequences. Monomer domains of the invention will often bind to a target molecule. For example, a polypeptide that forms a three-dimensional structure that binds to a target molecule is a monomer domain. As used herein, the term "monomer domain" does not encompass the complementarity determining region (CDR) of an antibody.

[39] The term "loop" refers to that portion of a monomer domain that is typically exposed to the environment by the assembly of the scaffold structure of the monomer domain protein, and which is involved in target binding. The present invention provides three types of loops that are identified by specific features, such as, potential for disulfide bonding, bridging between secondary protein structures, and molecular dynamics (i.e., flexibility). The three types of loop sequences are a cysteine-defined loop sequence, a structure-defined loop sequence, and a B-factor-defined loop sequence.

[40] As used herein, the term "cysteine-defined loop sequence" refers to a subsequence of a naturally occurring monomer domain-encoding sequence that is bound at each end by a cysteine residue that is conserved with respect to at least one other naturally occurring monomer domain of the same family. Cysteine-defined loop sequences are

identified by multiple sequence alignment of the naturally occurring monomer domains, followed by sequence analysis to identify conserved cysteine residues. The sequence between each consecutive pair of conserved cysteine residues is a cysteine-defined loop sequence. The cysteine-defined loop sequence does not include the cysteine residues adjacent to each terminus. Monomer domains having cysteine-defined loop sequences include the LDL receptor A-domains, EGF-like domains, sushi domains, Fibronectin type 1 domains, and the like. Thus, for example, in the case of LDL receptor A-domains represented by the consensus sequence, $CX_6CX_4CX_6CX_5CX_8C$ (SEQ ID NO:37), wherein X_6, X_4, X_5, and X_8 each represent a cysteine-defined loop sequence comprising the designated number of amino acids.

[41] The term "B-factor-defined loop sequence" refers to a subsequence of at least three amino acid residues of a monomer-domain encoding sequence in which the B-factors for the alpha carbons in the B-factor-defined loop are among the 25% highest alpha carbon B factors in the entire monomer domain. Typically the average alpha-carbon B-factor for the subsequence is at least about 65. As used herein, the term "B-factor" (or "temperature factor" or "Debye-Waller factor") is derived from X-ray scattering data. The B-factor is a factor that can be applied to the X-ray scattering term for each atom, or for groups of atoms, that describes the degree to which electron density is spread out B-factors employed in the practice of the present invention may be either isotropic or anisotropic. The term "average alpha-carbon B-factor" refers to:

$$(\sum_{i=1}^{n} \text{B-factor}_{C\alpha i}) / n$$

where n corresponds to the number of residues in the loop, and is at least 3, and $\text{B-factor}_{C\alpha i}$ is the B-factor for the alpha carbon of amino acid residue i of the loop.

[42] The term "multimer" is used herein to indicate a polypeptide comprising at least two monomer domains. The separate monomer domains in a multimer can be joined together by a linker. A multimer is also known as a combinatorial mosaic protein or a recombinant mosaic protein.

[43] The term "family" and "family class" are used interchangeably to indicate proteins that are grouped together based on similarities in their amino acid sequences. These similar sequences are generally conserved because they are important for the function of the protein and/or the maintenance of the three dimensional structure of the protein. Examples of such families include the LDL Receptor A-domain family, the EGF-

like family, and the like. Additionally, related sequences that bind to the same target molecule can be divided into families based on common sequence motifs.

[44] The term "ligand," also referred to herein as a "target molecule," encompasses a wide variety of substances and molecules, which range from simple molecules to complex targets. Target molecules can be proteins, nucleic acids, lipids, carbohydrates or any other molecule capable of recognition by a polypeptide domain. For example, a target molecule can include a chemical compound (i.e., non-biological compound such as, e.g., an organic molecule, an inorganic molecule, or a molecule having both organic and inorganic atoms, but excluding polynucleotides and proteins), a mixture of chemical compounds, an array of spatially localized compounds, a biological macromolecule, a bacteriophage peptide display library, a polysome peptide display library, an extract made from a biological materials such as bacteria, plants, fungi, or animal (e.g., mammalian) cells or tissue, a protein, a toxin, a peptide hormone, a cell, a virus, or the like. Other target molecules include, e.g., a whole cell, a whole tissue, a mixture of related or unrelated proteins, a mixture of viruses or bacterial strains or the like. Target molecules can also be defined by inclusion in screening assays described herein or by enhancing or inhibiting a specific protein interaction (i.e., an agent that selectively inhibits a binding interaction between two predetermined polypeptides).

[45] The term "linker" is used herein to indicate a moiety or group of moieties that joins or connects two or more discrete separate monomer domains. The linker allows the discrete separate monomer domains to remain separate when joined together in a multimer. The linker moiety is typically a substantially linear moiety. Suitable linkers include polypeptides, polynucleic acids, peptide nucleic acids and the like. Suitable linkers also include optionally substituted alkylene moieties that have one or more oxygen atoms incorporated in the carbon backbone. Typically, the molecular weight of the linker is less than about 2000 daltons. More typically, the molecular weight of the linker is less than about 1500 daltons and usually is less than about 1000 daltons. The linker can be small enough to allow the discrete separate monomer domains to cooperate, e.g., where each of the discrete separate monomer domains in a multimer binds to the same target molecule via separate binding sites. Exemplary linkers include a polynucleotide encoding a polypeptide, or a polypeptide of amino acids or other non-naturally occurring moieties. The linker can be a portion of a native sequence, a variant thereof, or a synthetic sequence. Linkers can comprise, e.g., naturally occurring, non-naturally occurring amino acids, or a combination of both.

[46] The term "separate" is used herein to indicate a property of a moiety that is independent and remains independent even when complexed with other moieties, including for example, other monomer domains. A monomer domain is a separate domain in a protein because it has an independent property that can be recognized and separated from the protein. For instance, the ligand binding ability of the A-domain in the LDLR is an independent property. Other examples of separate include the separate monomer domains in a multimer that remain separate independent domains even when complexed or joined together in the multimer by a linker. Another example of a separate property is the separate binding sites in a multimer for a ligand.

[47] As used herein, "directed evolution" refers to a process by which polynucleotide variants are generated, expressed, and screened for an activity (e.g., a polypeptide with binding activity) in a recursive process. One or more candidates in the screen are selected and the process is then repeated using polynucleotides that encode the selected candidates to generate new variants. Directed evolution involves at least two rounds of variation generation and can include 3, 4, 5, 10, 20 or more rounds of variation generation and selection. Variation can be generated by any method known to those of skill in the art, including, e.g., by error-prone PCR, gene recombination, chemical mutagenesis and the like.

[48] The term "shuffling" is used herein to indicate recombination between non-identical sequences. In some embodiments, shuffling can include crossover via homologous recombination or via non-homologous recombination, such as via cre/lox and/or flp/frt systems. Shuffling can be carried out by employing a variety of different formats, including for example, *in vitro* and *in vivo* shuffling formats, in silico shuffling formats, shuffling formats that utilize either double-stranded or single-stranded templates, primer based shuffling formats, nucleic acid fragmentation-based shuffling formats, and oligonucleotide-mediated shuffling formats, all of which are based on recombination events between non-identical sequences and are described in more detail or referenced herein below, as well as other similar recombination-based formats. The term "random" as used herein refers to a polynucleotide sequence or an amino acid sequence composed of two or more amino acids and constructed by a stochastic or random process. The random polynucleotide sequence or amino acid sequence can include framework or scaffolding motifs, which can comprise invariant sequences.

[49] The term "pseudorandom" as used herein refers to a set of sequences, polynucleotide or polypeptide, that have limited variability, so that the degree of residue

variability at some positions is limited, but any pseudorandom position is allowed at least some degree of residue variation.

[50] The terms "polypeptide," "peptide," and "protein" are used herein interchangeably to refer to an amino acid sequence of two or more amino acids.

[51] The term "amino acid" refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, *e.g.*, hydroxyproline, γ-carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, *i.e.*, an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, *e.g.*, homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs have modified R groups (*e.g.*, norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. "Amino acid mimetics" refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that functions in a manner similar to a naturally occurring amino acid.

[52] "Conservative amino acid substitution" refers to the interchangeability of residues having similar side chains. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulfur-containing side chains is cysteine and methionine. Conservative amino acids substitution groups include: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine.

[53] The phrase "nucleic acid sequence" refers to a single or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases read from the 5' to the 3' end or an analog thereof.

[54] The term "encoding" refers to a polynucleotide sequence encoding one or more amino acids. The term does not require a start or stop codon. An amino acid sequence can be encoded in any one of six different reading frames provided by a polynucleotide sequence.

[55] The term "promoter" refers to regions or sequence located upstream and/or downstream from the start of transcription that are involved in recognition and binding of RNA polymerase and other proteins to initiate transcription.

[56] A "vector" refers to a polynucleotide, which when independent of the host chromosome, is capable of replication in a host organism. Examples of vectors include plasmids. Vectors typically have an origin of replication. Vectors can comprise, e.g., transcription and translation terminators, transcription and translation initiation sequences, and promoters useful for regulation of the expression of the particular nucleic acid.

[57] The term "recombinant" when used with reference, *e.g.*, to a cell, or nucleic acid, protein, or vector, indicates that the cell, nucleic acid, protein or vector, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration of a native nucleic acid or protein, or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found within the native (nonrecombinant) form of the cell or express native genes that are otherwise abnormally expressed, under-expressed or not expressed at all.

[58] The phrase "specifically (or selectively) binds" to a polypeptide, when referring to a monomer or multimer, refers to a binding reaction that can be determinative of the presence of the polypeptide in a heterogeneous population of proteins (e.g., a cell or tissue lysate) and other biologics. Thus, under standard conditions or assays used in antibody binding assays, the specified monomer or multimer binds to a particular target molecule above background (e.g., 2X, 5X, 10X or more above background) and does not bind in a significant amount to other molecules present in the sample.

[59] The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same. "Substantially identical" refers to two or more nucleic acids or polypeptide sequences having a specified percentage of amino acid residues or nucleotides that are the same (*i.e.*, 60% identity, optionally 65%, 70%, 75%, 80%, 85%, 90%, or 95% identity over a specified region, or, when not specified, over the entire sequence), when compared and aligned for maximum correspondence over a comparison window, or designated region as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. Optionally, the identity or substantial identity exists over a region that is at least about 50 nucleotides in length, or over a region that is 100 to 500 or 1000 or more nucleotides or amino acids in length, or if not specified, over the length of a sequence. The present invention provides polypeptides comprising IL-6 binding sequences,

wherein the IL-6 binding sequences are substantially identical (e.g., at least 65%, 70%, 75%, 80%, 85%, 90%, or 95% identical) to at least one IL-6 binding sequences as provided herein, including but not limited to, any of SEQ ID NOS:27-33.

[60] A polynucleotide or amino acid sequence is "heterologous to" a second sequence if the two sequences are not linked in the same manner as found in naturally-occurring sequences. For example, a promoter operably linked to a heterologous coding sequence refers to a coding sequence which is different from any naturally-occurring allelic variants. The term "heterologous linker," when used in reference to a multimer, indicates that the multimer comprises a linker and a monomer that are not found in the same relationship to each other in nature (*e.g.*, they form a non-naturally occurring fusion protein).

[61] A "non-naturally-occurring amino acid" in a protein sequence refers to any amino acid other than the amino acid that occurs in the corresponding position in an alignment with a naturally-occurring polypeptide with the lowest smallest sum probability where the comparison window is the length of the monomer domain queried and when compared to a naturally-occurring sequence in the non-redundant ("nr") database of Genbank using BLAST 2.0 as described herein.

[62] "Percentage of sequence identity" is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

[63] The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence over a comparison window, or designated region as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. Such sequences are then said to be "substantially identical." The present invention provides polypeptides that bind to IL-6 and that comprise a sequence substantially identical to any of the full-length sequences described herein. This definition also refers to the complement of a test sequence. Optionally, the

identity exists over a region that is at least about 50 amino acids or nucleotides in length, or more preferably over a region that is 75-100 amino acids or nucleotides in length.

[64] For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

[65] A "comparison window", as used herein, includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, *e.g.*, by the local homology algorithm of Smith and Waterman (1970) *Adv. Appl. Math.* 2:482c, by the homology alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity method of Pearson and Lipman (1988) *Proc. Nat'l. Acad. Sci. USA* 85:2444, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (*see, e.g.,* Ausubel *et al.*, *Current Protocols in Molecular Biology* (1995 supplement)).

[66] One example of a useful algorithm is the BLAST 2.0 algorithm, which is described in Altschul *et al.* (1990) *J. Mol. Biol.* 215:403-410, respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the

parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) or 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (*see* Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

[67] The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (*see, e.g.*, Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5787). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001.

BRIEF DESCRIPTION OF THE DRAWINGS

[68] Figure 1 schematically illustrates the alignment of partial amino acid sequence from a variety of the LDL-receptor class A-domains (SEQ ID NOS:38-50) to demonstrate the conserved cysteines. The connectivity of cysteines in the three disulfide bonds of the folded domain is illustrated schematically on the consensus sequence (SEQ ID NO:51). Residues whose side-chains contribute to calcium binding are designated with an asterisk in the consensus sequence.

[69] Figure 2, panel A schematically illustrates an example of an A-domain. Panel A schematically illustrates conserved amino acids in an A-domain (SEQ ID NO:52) of about 40 amino acids long. The conserved cysteine residues are indicated by C, and the conserved negatively charged amino acids are indicated by a circle with a minus ("-") sign. Circles with an "H" indicate conserved hydrophobic residues. Panel B schematically

illustrates two folded A-domains connected via a linker. Panel B also indicates two calcium binding sites, dark circles with Ca^{+2}, and three disulfide bonds within each folded A-domain for a total of 6 disulfide bonds.

[70] Figure 3 indicates some of the ligands recognized by naturally-occurring members of the LDL-receptor family, which include inhibitors, proteases, protease complexes, vitamin-carrier complexes, proteins involved in lipoprotein metabolism, non-human ligands, antibiotics, viruses, and others.

[71] Figure 4 schematically illustrates a general scheme for identifying monomer domains that bind to a ligand, isolating the selected monomer domains, creating multimers of the selected monomer domains by joining the selected monomer domains in various combinations and screening the multimers to identify multimers comprising more than one monomer that binds to a ligand.

[72] Figure 5 is a schematic representation of another selection strategy (guided selection). A monomer domain with appropriate binding properties is identified from a library of monomer domains. The identified monomer domain is then linked to monomer domains from another library of monomer domains to form a library of multimers. The multimer library is screened to identify a pair of monomer domains that bind simultaneously to the target. This process can then be repeated until the optimal binding properties are obtained in the multimer.

[73] Figure 6 depicts an alignment of A domains (SEQ ID NOS:53-269). At the top of the figure, small letters (a-q) indicate conserved residues.

[74] Figure 7 illustrates various possible antibody-monomer or multimer conformations. In some embodiments, the monomer or multimer replaces the Fab fragment of the antibody.

[75] Figure 8 depicts a possible conformation of a multimer of the invention comprising at least one monomer domain that binds to a half-life extending molecule and other monomer domains binding to one or optionally two or more target molecules. In the Figure, two monomer domains bind to two first target molecules. Optionally, the two monomer domains can bind to different sites on one first target molecule (not depicted).

[76] Figure 9 shows possible designs of polypeptides of the invention that prevent IL-6/IL-6R interaction.

[77] Figure 10 illustrates the serum half-life in monkeys of monomers that bind to IgG.

DETAILED DESCRIPTION OF THE INVENTION

I. INTRODUCTION

[78] The present invention provides for non-naturally-occurring proteins that bind to IL-6. Generally, the proteins of the present invention comprise a domain that binds to IL-6. These domains may be readily identified using a variety of polypeptide scaffolds to generate a plurality of polypeptide variants and then selecting a variant that binds to IL-6. The present invention therefore also provides for selecting a protein that binds to IL-6. Proteins that bind IL-6 are useful, e.g., for treating, e.g., autoimmune and inflammatory diseases, cancer, metabolic/endocrine diseases, cardiovascular disease, and neurological diseases. The polypeptides of the invention are also useful to detect and quantify IL-6. The amino acid sequence of IL-6 is known and is displayed in SEQ ID NO:1.

[79] IL-6 functions by binding the IL-6 receptor (IL-6R). Accordingly, in some embodiments, the polypeptides of the invention comprise at least one monomer domain that inhibits binding of IL-6 to its receptor, i.e., that acts as an antagonist to prevent IL-6 from binding and/or activating IL-6R.

[80] While the present invention provides for polypeptides comprising single domains, multimers of the domains may also be synthesized and used. In some embodiments, all of the domains of the multimer bind IL-6. In some of these embodiments, each of the domains are identical and bind to the same portion (i.e., "epitope") of IL-6. In other embodiments, at least some of the domains in the multimer bind to different portions of IL-6. In yet other embodiments, at least some of the domains of the polypeptide bind to a molecule or molecules other than IL-6 (*e.g.*, a blood factor such as serum albumin, immunoglobulin, or erythrocytes).

II. MONOMERS

[81] Monomer domains can be polypeptide chains of any size. In some embodiments, monomer domains have about 25 to about 500, about 30 to about 200, about 30 to about 100, about 35 to about 50, about 35 to about 100, about 90 to about 200, about 30 to about 250, about 30 to about 60, about 9 to about 150, about 100 to about 150, about 25 to about 50, or about 30 to about 150 amino acids. Similarly, a monomer domain of the present invention can comprise, e.g., from about 30 to about 200 amino acids; from about 25 to about 180 amino acids; from about 40 to about 150 amino acids; from about 50 to about 130 amino acids; or from about 75 to about 125 amino acids. Monomer domains can typically maintain a stable conformation in solution, and are often heat stable, e.g., stable at 95° C for at least 10

minutes without losing binding affinity. Sometimes, monomer domains can fold independently into a stable conformation. In one embodiment, the stable conformation is stabilized by ions (e.g., such as metal or calcium ions). The stable conformation can optionally contain disulfide bonds (e.g., at least one, two, or three or more disulfide bonds). The disulfide bonds can optionally be formed between two cysteine residues. In some embodiments, monomer domains, or monomer domain variants, are substantially identical to the sequences exemplified.

A. IL-6 Binders

[82] In some aspects, the invention provides monomer domains that bind to (e.g., selectively binds) a IL-6 polypeptide or a portion thereof. A portion of a polypeptide can be, e.g., at least 5, 10, 15, 20, 30, 50, 100, or more contiguous amino acids of the polypeptide. In some embodiments, the monomers of the invention will bind to IL-6 with a affinity stronger than 10^{-3} M (e.g, 10^{-4} M, etc.). In some embodiments, the affinity is stronger than 10^{-4} M, 10^{-5} M, 10^{-6} M, 10^{-7} M, or 10^{-8} M.

[83] A large number of IL-6 binding sequences having an A domain scaffold were generated. As described in detail in the examples, seven families (*i.e.*, Families 1-7, or "Fam 1-7") of monomer domains that bind to IL-6 have been identified. The consensus motifs generated based on these families indicate common amino acid residues between IL-6 binders. Those of skill in the art will appreciate that positions where there is no consensus (marked with an "X") can be any amino acid. In some embodiments, the amino acid at "X" positions will be selected from amino acids in the analogous position of one of the exemplified IL-6 binders either from the same family or a different family.

[84] Family 1 has the following consensus motif (SEQ ID NO:27):

CX[AS] [DN] EFXCXXGXC [ILV] XXAXXCDGV [DN] DCXDNSDEXXC [AST] XX [AST]

[85] Exemplary sequences comprising the IL-6 Family 1 motif are displayed in the examples. References to IL-6 binding monomers or multimers encompass each Family 1 sequence exemplified in the examples.

[86] Family 2 has the following consensus motif (SEQ ID NO:28):

CXXXQFXC [DN] NGXCVPXXWXCDG [DE] [DN] DCXD [DN] SDEXXCXXP [AST]

[87] Exemplary sequences comprising the IL-6 Family 2 motif are displayed in the examples. References to IL-6 binding monomers or multimers encompass each Family 2 sequence exemplified in the examples.

[88] Family 3 has the following consensus motif (SEQ ID NO:29):

CXXXQFQCXSG [W, Y] C [I, L, F] XXXW [V, L] CDGXNDCEDXSDEXXC

[89] Exemplary sequences comprising the IL-6 Family 3 motif are displayed in the examples. References to IL-6 binding monomers or multimers encompass each Family 3 sequence exemplified in the examples.

[90] Family 4 has the following consensus motif (SEQ ID NO:30):

CXXX[EQ]FXCXXGXC[ILV]XXXXXCDGX[DN]DCXDXSDEXX(XX)C

[91] Exemplary sequences comprising the IL-6 Family 4 motif are displayed in the examples. References to IL-6 binding monomers or multimers encompass each Family 4 sequence exemplified in the examples.

[92] Family 5 has the following consensus motif (SEQ ID NO:31):

CX[A,P]X[E,Q]FXCXXGQC[I,L]PXXXXCDGX[D,N]DCXDXSDEXXC

[93] Exemplary sequences comprising the IL-6 Family 5 motif are displayed in the examples. References to IL-6 binding monomers or multimers encompass each Family 5 sequence exemplified in the examples.

[94] Family 6 has the following consensus motif (SEQ ID NO:32):

CXXX[EQ]FXCX(X)SXXC[ILV]XX[DN]W[ILV]CDGVNDCXDXSDEXGCXXXT

[95] Exemplary sequences comprising the IL-6 Family 6 motif are displayed in the examples. References to IL-6 binding monomers or multimers encompass each Family 6 sequence exemplified in the examples.

[96] Family 7 has the following consensus motif (SEQ ID NO:33):

CL[AS]X[EQ]FXCXNGXCIP[KQR]XWVCDGDNDC[ILV]DXSDE(P)XXCX[EKQ]XT

[97] Exemplary sequences comprising the IL-6 Family 7 motif are displayed in the examples. References to IL-6 binding monomers or multimers encompass each Family 7 sequence exemplified in the examples.

[98] As described herein, monomer domains can be linked to form multimers that bind the same target. The present invention provides polypeptides comprising both homo- (i.e., at least two identical monomers) and hetero- (i.e., different monomers that bind the same target) multimers. Exemplary dimers and trimers are displayed in the examples.

[99] Several IL-6 trimer family motifs were identified based on similarity between sequences. *See* Example 7. The consensus motifs generated based on these families indicate common amino acid residues between IL-6 trimers. Those of skill in the art will appreciate that positions where there is no consensus (marked with an "X") can be any amino acid. In some embodiments, the amino acid at "X" positions will be selected from amino acids in the analogous position of one of the exemplified IL-6 binders either from the same family or a different family.

[100] Exemplary IL-6-binding trimers include the following (SEQ ID NOS:270-278):

CLPDQFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQC GSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGV NDCADNSDEEDCTEHT;

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQC GSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGV NDCADNSDEEDCTEHT;

CRADEFQCGSGRCVPADWLCDGVNDCADNSDETGCPTRTCLPSQFQC GSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGV NDCADNSDEEDCTEHT;

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQC GSGYCLPQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGV NDCADNSDEEDCTQHT;

CQANEFTCSSGNCLSPALLCDGVDDCGDNSDETYCASPAPTCLPSQFQ CGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCQPGEFTCESGRCLPLRWLCDG EDDCVDSSDEASCQPPT;

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQC GSGYCLPQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGV NDCADNSDEADCAAPT;

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPHTCLPDQFQC GSGYCLSQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGV NDCADNSDEEDCTEHT;

CESDQFQCDSGHCVPLNWVCDGVNDCVDGSDEAGCPPRTCLPGQFQC GSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGV NDCADNSDEEDCTEHT; and

CGASQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCLPSEFQ CRDSKICLSADWLCDGVNDCADNSDEEGCPPPTCLPDQFQCGSGYCLPQAWVCDGE NDCEDGSDETNCPART.

B. IgG binders and serum half-life extension

[101] The invention further provide monomer domains that bind to (e.g., selectively binds) a blood factor (*e.g.*, serum albumin, immunoglobulin, or erythrocytes). In some embodiments, the monomers of the invention will bind to a blood factor with an affinity

of less than 10^{-3} M (i.s., binds stronger than an affinity of 10^{-3} M). In some embodiments, the affinity is less than 10^{-4} M, 10^{-5} M, 10^{-6} M, 10^{-7} M, or 10^{-8} M.

[102] In some embodiments, the the monomer domains bind to an immunoglobulin polypeptide or a portion thereof.

[103] Two families (*i.e.*, A domain IgG Families 2 and 3) of monomer domains that bind to immunoglobulin have been identified.

[104] IgG Family 2 has the following consensus motif (SEQ ID NO:279):

[EQ]FXCRX[ST]XRC[IV]XXXW[ILV]CDGXXDCXD[DN]SDE

[105] Exemplary sequences comprising the IgG Family 2 motif are displayed in the examples. References to IgG binding monomers or multimers encompass each Family 2 sequence exemplified in the examples.

[106] IgG Family 3 has either of the two following consensus motifs (SEQ ID NOS:280 and 281):

CXSSGRCIPXXWVCDGXXDCRDXSDE; or

CXSSGRCIPXXWLCDGXXDCRDXSDE

[107] Exemplary sequences comprising the IgG Family 3 motif are displayed in the examples. References to IgG binding monomers or multimers encompass each Family 3 sequence exemplified in the examples.

[108] Monomer domains that bind to red blood cells (RBC) or serum albumin (CSA) are described in U.S. Patent Publication No. 2005/0048512, and include, e.g. SEQ ID NOS:282-285:

RBCA **C**RSSQFQ**C**NDSRI**C**IPGRWR**C**DGDND**C**QDGSDETG**C**GDSHILPFSTPGPST
RBCB **C**PAGEFP**C**KNGQ**C**LPVTWL**C**DGVND**C**LDGSDEKG**C**GRPGPGATSAPAA
RBC11 **C**PPDEFP**C**KNGQ**C**IPQDWL**C**DGVND**C**LDGSDEKD**C**GRPGPGATSAPAA
CSA-A8 **C**GAGQFP**C**KNGH**C**LPLNLL**C**DGVND**C**EDNSDEPSEL**C**KALT

[109] The present invention provides a method for extending the serum half-life of a protein, including, e.g., a multimer of the invention or a protein of interest in an animal. The protein of interest can be any protein with therapeutic, prophylactic, or otherwise desirable functionality. This method comprises first providing a monomer domain that has been identified as a binding protein that specifically binds to a half-life extender such as a blood-carried molecule or cell, such as serum albumin (e.g., human serum albumin), IgG, red blood cells, etc. The half-life extender-binding monomer is then covalently linked to another monomer domain that has a binding affinity for the protein of interest (e.g., IL-6 or a different target). This complex formation results in the half-life extension protecting the

multimer and/or bound protein(s) from proteolytic degradation and/or other removal of the multimer and/or protein(s) and thereby extending the half-life of the protein and/or multimer. One variation of this use of the invention includes the half-life extender-binding monomer covalently linked to the protein of interest. The protein of interest may include a monomer domain, a multimer of monomer domains, or a synthetic drug. Alternatively, monomers that bind to either immunoglobulins or erythrocytes could be generated using the above method and could be used for half-life extension.

[110] The half-life extender-binding multimers are typically multimers of at least two domains, chimeric domains, or mutagenized domains (i.e., one that binds to IL-6 and one that binds to the blood-carried molecule or cell). Suitable domains include all of those described herein, that are further screened and selected for binding to a half-life extender. The half-life extender-binding multimers are generated in accordance with the methods for making multimers described herein, using, for example, monomer domains pre-screened for half-life extender -binding activity. The serum half-life of a molecule can be extended to be, e.g., at least 1, 2, 3, 4, 5, 10, 20, 30, 40, 50, 60, 70 80, 90, 100, 150, 200, 250, 400, 500 or more hours.

[111] The half-life extending monomer domain can be linked to the N-, C-terminus or both of a target molecule (e.g., IL-6-binding)-binding monomer domain. Exemplary sequences of polypeptides that bind to IL-6 and IgG include the following (SEQ ID NOS:286-339):

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRTCLPDQFRCGNGQCIPLD
WVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRWVCDGENDCEDGSDEAN
CAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRTCLPDQFRCGNGQCIPLD
WVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEA
NCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRTCRADEFQCGSGRCVPAD
WLCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEAN
CAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRTCLPDQFRCGNGQCIPLD
WVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDCEDGSDEA
NCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTQHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRTCQANEFTCSSGNCLSPA
LLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDE
ANCAGSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDEASCQPPT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRTCLPDQFRCGNGQCIPLD
WVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDCEDGSDEA
NCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEADCAAPT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRTCLPDQFRCGNGQCIPLD
WVCDGVNDCEDDSDEESCPPHTCLPDQFQCGSGYCLSQRWVCDGENDCEDGSDEA
NCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRTCESDQFQCDSGHCVPLN
WVCDGVNDCVDGSDEAGCPPRTCLPGQFQCGSGYCISQRWVCDGENDCEDGSDEA
NCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRTCGASQFQCGSGYCISQR
WVCDGENDCEDSSDEANCAGSVPTCLPSEFQCRDSKICLSADWLCDGVNDCADNSD
EEGCPPPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETNCPART

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCKPRTCLPDQFRCGNGQCIPL
DWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRWVCDGENDCEDGSDEA
NCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCKPRTCLPDQFRCGNGQCIPL
DWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDE
ANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCKPRTCRADEFQCGSGRCVP
ADWLCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDE
ANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCKPRTCLPDQFRCGNGQCIPL
DWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDCEDGSDE
ANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTQHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCKPRTCQANEFTCSSGNCLSP
ALLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSD
EANCAGSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDEASCQPPT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCKPRTCLPDQFRCGNGQCIPL
DWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDCEDGSDE
ANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEADCAAPT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCKPRTCLPDQFRCGNGQCIPL
DWVCDGVNDCEDDSDEESCPPHTCLPDQFQCGSGYCLSQRWVCDGENDCEDGSDE
ANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCKPRTCESDQFQCDSGHCVP
LNWVCDGVNDCVDGSDEAGCPPRTCLPGQFQCGSGYCISQRWVCDGENDCEDGSD
EANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCKPRTCGASQFQCGSGYCIS
QRWVCDGENDCEDSSDEANCAGSVPTCLPSEFQCRDSKICLSADWLCDGVNDCADN
SDEEGCPPPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETNCPART

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCSAPASEPPGSLCLPDQFRCGN
GQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRWVCDGENDCE
DGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCSAPASEPPGSLCLPDQFRCGN
GQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCE
DGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCSAPASEPPGSLCRADEFQCGS
GRCVPADWLCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCE
DGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCSAPASEPPGSLCLPDQFRCGN
GQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDCE
DGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTQHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCSAPASEPPGSLCQANEFTCSSG
NCLSPALLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQRWVCDGENDCE
DSSDEANCAGSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDEASCQPPT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCSAPASEPPGSLCLPDQFRCGN
GQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDCE
DGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEADCAAPT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCSAPASEPPGSLCLPDQFRCGN
GQCIPLDWVCDGVNDCEDDSDEESCPPHTCLPDQFQCGSGYCLSQRWVCDGENDCE
DGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCSAPASEPPGSLCESDQFQCDSG
HCVPLNWVCDGVNDCVDGSDEAGCPPRTCLPGQFQCGSGYCISQRWVCDGENDCE
DGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCSAPASEPPGSLCGASQFQCGS
GYCISQRWVCDGENDCEDSSDEANCAGSVPTCLPSEFQCRDSKICLSADWLCDGVN
DCADNSDEEGCPPPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETNCPART

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCRADEFQCG
SGRCVPADWLCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTQHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCQANEFTCS
SGNCLSPALLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQRWVCDGEND
CEDSSDEANCAGSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDEASCQPPT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEADCAAPT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCEDDSDEESCPPHTCLPDQFQCGSGYCLSQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCESDQFQCD
SGHCVPLNWVCDGVNDCVDGSDEAGCPPRTCLPGQFQCGSGYCISQRWVCDGEND
CEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCGASQFQCG
SGYCISQRWVCDGENDCEDSSDEANCAGSVPTCLPSEFQCRDSKICLSADWLCDGVN
DCADNSDEEGCPPPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETNCPART

CLPDQFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRW
VCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEE
DCTEHTCLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRT

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHTCLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRT

CRADEFQCGSGRCVPADWLCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHTCLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRT

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTQHTCLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRT

CQANEFTCSSGNCLSPALLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQR
WVCDGENDCEDSSDEANCAGSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDE
ASCQPPTCLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRT

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EADCAAPTCLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRT

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPHTCLPDQFQCGSGYCLSQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHTCLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRT

CESDQFQCDSGHCVPLNWVCDGVNDCVDGSDEAGCPPRTCLPGQFQCGSGYCISQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHTCLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRT

CGASQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCLPSEFQCRDSKICLS
ADWLCDGVNDCADNSDEEGCPPPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSD
ETNCPARTCLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRT

CLPDQFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRW
VCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEE
DCTEHTCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSL

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHTCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSL

CRADEFQCGSGRCVPADWLCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHTCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSL

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTQHTCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSL

CQANEFTCSSGNCLSPALLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQR
WVCDGENDCEDSSDEANCAGSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDE
ASCQPPTCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSL

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EADCAAPTCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSL

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPHTCLPDQFQCGSGYCLSQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHTCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSL

CESDQFQCDSGHCVPLNWVCDGVNDCVDGSDEAGCPPRTCLPGQFQCGSGYCISQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHTCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSL

CGASQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCLPSEFQCRDSKICLS
ADWLCDGVNDCADNSDEEGCPPPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSD
ETNCPARTCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSL

C. Discussion of Monomer Domains

[112] Monomer domains that are particularly suitable for use in the practice of the present invention are cysteine-rich domains comprising disulfide bonds. Cysteine-rich domains employed in the practice of the present invention typically do not form an α helix, a β sheet, or a β-barrel structure. Typically, the disulfide bonds promote folding of the domain into a three-dimensional structure. Usually, cysteine-rich domains have at least two disulfide bonds, more typically at least three disulfide bonds. In some embodiments, at least 5, 10, 15 or 20% of the amino acids in a monomer domain are cysteines.

[113] Domains can have any number of characteristics. For example, in some embodiments, the domains have low or no immunogenicity in an animal (e.g., a human). Domains can have a small size. In some embodiments, the domains are small enough to penetrate skin or other tissues. Domains can have a range of *in vivo* half-lives or stabilities.

[114] Illustrative monomer domains suitable for use in the practice of the present invention include, e.g., an EGF-like domain, a Kringle-domain, a fibronectin type I domain, a fibronectin type II domain, a fibronectin type III domain, a PAN domain, a Gla domain, a SRCR domain, a Kunitz/Bovine pancreatic trypsin Inhibitor domain, a Kazal-type serine protease inhibitor domain, a Trefoil (P-type) domain, a von Willebrand factor type C domain, an Anaphylatoxin-like domain, a CUB domain, a thyroglobulin type I repeat, LDL-receptor class A domain, a Sushi domain, a Link domain, a Thrombospondin type I domain, an Immunoglobulin-like domain, a C-type lectin domain, a MAM domain, a von Willebrand factor type A domain, a Somatomedin B domain, a WAP-type four disulfide core domain, a F5/8 type C domain, a Hemopexin domain, an SH2 domain, an SH3 domain, a Laminin-type EGF domain, a C2 domain, and other such domains known to those of ordinary skill in the art, as well as derivatives and/or variants thereof.

[115] In some embodiments, suitable monomer domains (e.g. domains with the ability to fold independently or with some limited assistance) can be selected from the families of protein domains that contain β-sandwich or β-barrel three dimensional structures as defined by such computational sequence analysis tools as Simple Modular Architecture Research Tool (SMART), *see* Shultz *et al.*, *SMART: a web-based tool for the study of genetically mobile domains*, (2000) Nucleic Acids Research 28(1):231-234) or CATH (*see*

Pearl *et.al.*, *Assigning genomic sequences to CATH*, (2000) Nucleic Acids Research 28(1):277-282).

[116] In another embodiment, monomer domains of the present invention include domains other than a fibronectin type III domain, an anticalin domain and a Ig-like domain from CTLA-4. Some aspects of these domains are described in WO01/64942 entitled "Protein scaffolds for antibody mimics and other binding proteins" by Lipovsek et al., published on September 7, 2001, WO99/16873 entitled "Anticalins" by Beste et al., published April 8, 1999 and WO 00/60070 entitled "A polypeptide structure for use as a scaffold" by Desmet, et al., published on October 12, 2000.

[117] As described *supra*, monomer domains are optionally cysteine rich. Suitable cysteine rich monomer domains include, e.g., the LDL receptor class A domain ("A-domain") or the EGF domain. The monomer domains can also have a cluster of negatively charged residues.

[118] Other features of monomer domains can include the ability to bind ligands or the ability to bind an ion (e.g., Ca^{2+} binding by the LDL receptor A-domain). Monomer domains that bind ions to maintain their secondary structure include, e.g., A domain, EGF domain, EF Hand (e.g., such as those found in present in calmodulin and troponin C), Cadherin domain, C-type lectin, C2 domain, Annexin, Gla-domain, Trombospondin type 3 domain, all of which bind calcium, and zinc fingers (e.g., C2H2 type C3HC4 type (RING finger), Integrase Zinc binding domain, PHD finger, GATA zinc finger, FYVE zinc finger, B-box zinc finger), which bind zinc. Without intending to limit the invention, it is believed that ion-binding provides stability of secondary structure while providing sufficient flexibility to allow for numerous binding conformations depending on primary sequence.

[119] As described herein, monomer domains may be selected for the ability to bind to targets other than the target that a homologous naturally occurring domain may bind. Thus, in some embodiments, the invention provides monomer domains (and multimers comprising such monomers) that do not bind to the target or the class or family of target proteins that a substantially identical naturally occurring domain may bind.

[120] Characteristics of a monomer domain can include the ability to fold independently and the ability to form a stable structure. Thus, the structure of the monomer domain is often conserved, although the polynucleotide sequence encoding the monomer need not be conserved. For example, the A-domain structure is conserved among the members of the A-domain family, while the A-domain nucleic acid sequence is not. Thus,

for example, a monomer domain is classified as an A-domain by its cysteine residues and its affinity for calcium, not necessarily by its nucleic acid sequence. *See*, Figures 1 and 2.

[121] Specifically, the A-domains (sometimes called "complement-type repeats" or "LDL receptor type or class A domains") contain about 30-50 or 30-65 amino acids. In some embodiments, the domains comprise about 35-45 amino acids and in some cases about 35-40 amino acids. Within the 30-50 amino acids, there are about 6 cysteine residues. Of the six cysteines, disulfide bonds typically are found between the following cysteines: C1 and C3, C2 and C5, C4 and C6. The cysteine residues of the domain are disulfide linked to form a compact, stable, functionally independent moiety. *See*, Figure 3. Clusters of these repeats make up a ligand binding domain, and differential clustering can impart specificity with respect to the ligand binding.

[122] Exemplary A domain sequences and consensus sequences are depicted in Figures 1 and 2. One typical consensus sequence useful to identify A domains is the following: **C**-[VILMA]-$X_{(5)}$-**C**-[DNH]-$X_{(3)}$-[DENQHT]-**C**-$X_{(3,4)}$-[STADE]- [DEH]-[DE]-$X_{(1,5)}$-**C** (SEQ ID NO:340), where the residues in brackets indicate possible residues at one position. "$X_{(\#)}$" indicates number of residues. These residues can be any amino acid residue. Parentheticals containing two numbers refers to the range of amino acids that can occupy that position (e.g., "[DE]-$X_{(1,5)}$-**C**" means that the amino acids DE are followed by 1, 2, 3, 4, or 5 residues, followed by C). This consensus sequence only represents the portion of the A domain beginning at the third cysteine. A second consensus is as follows: **C**-$X_{(3-15)}$-**C**-$X_{(4-15)}$-**C**-$X_{(6-7)}$-**C**-[N,D]-$X_{(3)}$-[D,E,N,Q,H,S,T]-**C**-$X_{(4-6)}$-D-E-$X_{(2-8)}$-**C** (SEQ ID NO:341). The second consensus predicts amino acid residues spanning all six cysteine residues. In some embodiments, A domain variants comprise sequences substantially identical to any of the above-described sequences. Note that reference to "LDL receptor class A" domain, for the purposes of this invention, is not intended to indicate origin or binding properties of the domain.

[123] Additional exemplary A domains include the following sequence:

$C_aX_{3-15}C_bX_{3-15}C_cX_{6-7}C_d(D,N)X_4C_eX_{4-6}DEX_{2-8}C_f$ (SEQ ID NO:342)

wherein C is cysteine, X_{n-m} represents between n and m number of independently selected amino acids, and (D,N) indicates that the position can be either D or N; and wherein C_a-C_c, C_b-C_e and C_d-C_f form disulfide bonds.

[124] To date, at least 190 naturally-occurring human A-domains are identified based on cDNA sequences. *See, e.g.*, Figure 6. Exemplary proteins containing naturally-occurring A-domains include, e.g., complement components (e.g., C6, C7, C8, C9,

and Factor I), serine proteases (e.g., enteropeptidase, matriptase, and corin), transmembrane proteins (e.g., ST7, LRP3, LRP5 and LRP6) and endocytic receptors (e.g., Sortilin-related receptor, LDL-receptor, VLDLR, LRP1, LRP2, and ApoER2). A domains and A domain variants can be readily employed in the practice of the present invention as monomer domains and variants thereof. Further description of A domains can be found in the following publications and references cited therein: Howell and Hertz, *The LDL receptor gene family: signaling functions during development*, (2001) Current Opinion in Neurobiology 11:74-81; Herz (2001), *supra*; Krieger, *The "best" of cholesterols, the "worst" of cholesterols: A tale of two receptors*, (1998) PNAS 95: 4077-4080; Goldstein and Brown, *The Cholesterol Quartet*, (2001) Science, 292: 1310-1312; and, Moestrup and Verroust, *Megalin-and Cubilin-Mediated Endocytosis of Protein-Bound Vitamins, Lipids, and Hormones in Polarized Epithelia*, (2001) Ann. Rev. Nutr. 21:407-28.

[125] A number of other domain types can also be used to generate IL-6-binding monomer domains.

[126] Exemplary EGF monomer domains include the sequence:

$C_aX_{3-14}C_bX_{3-7}C_cX_{4-16}C_dX_{1-2}C_eX_{8-23}C_f$ (SEQ ID NO:343)

wherein C is cysteine, X_{n-m} represents between n and m number of independently selected amino acids; and

wherein C_a-C_c, C_b-C_e and C_d-C_f form disulfide bonds.

[127] Each of the domains described below employ exemplary motifs (*i.e.*, scaffolds). Certain positions are marked x, indicating that any amino acid can occupy the position. These positions can include a number of different amino acid possibilities, thereby allowing for sequence diversity and thus affinity for different target molecules. Use of brackets in motifs indicates alternate possible amino acids within a position (e.g., "[ekq]" indicates that either E, K or Q may be at that position). As used herein, use of parentheses in a sequence motif indicates that that the positions within the parentheses may be present or absent (e.g., "([ekq])" indicates that the position is absent or either E, K, or Q may be at that position). When more than one "x" is used in parentheses (e.g., "(xx)"), each x represents a possible position. Thus "(xx)" indicates that zero, one or two amino acids may be at that position(s), where each amino acid is independently selected from any amino acid.
α represents an aromatic/hydrophobic amino acid such as, *e.g.*, W, Y, F, or L; β represents a hydrophobic amino acid such as, *e.g.*, V, I, L, A, M, or F; χ represents a smaller polar amino acid such as, *e.g.*, G, A, S, or T; δ represents a charged amino acid such as, *e.g.*, K, R, E, Q,

or D; ε represents a small amino acid such as, *e.g.*,; V, A, S, or T; and ϕ represents a negatively charged amino acid such as, *e.g.*, D, E, or N.

[128] Suitable domains include, *e.g.* thrombospondin type 1 domains, trefoil domains, and thyroglobulin domains.

[129] Thrombospondin type 1 ("TSP1") domains contain about 30-50 or 30-65 amino acids. In some embodiments, the domains comprise about 35-55 amino acids and in some cases about 50 amino acids. Within the 35-55 amino acids, there are typically about 4 to about 6 cysteine residues. Of the six cysteines, disulfide bonds typically are found between the following cysteines: C1 and C5, C2 and C6, C3 and C4. The cysteine residues of the domain are disulfide linked to form a compact, stable, functionally independent moiety comprising distorted beta strands. Clusters of these repeats make up a ligand binding domain, and differential clustering can impart specificity with respect to the ligand binding.

[130] Exemplary TSP1 domain sequences and consensus sequences are as follows (SEQ ID NOS:344-349):

(1) (xxxxxx)C_1xxxC_2xxxxx(x)xxxxxC_3xxxx(xxx)xxxxxC_4xxxxxx(x)xxxC_5(x)xxxxC_6;
(2) (wxxWxx)C_1xxxC_2xxGxx(x)xRxxxC_3xxxx(Pxx)xxxxxC_4xxxxxx(x)xxxC_5(x)xxxxC_6
(3) (wxxWxx)C_1sxtC_2xxGxx(x)xRxrxC_3xxxx(Pxx)xxxxxC_4xxxxxx(x)xxxC_5(x)xxxxC_6
(4) (WxxWxx)C_1[Stnd][Vkaq][Tspl]C_2xx[Gq]xx(x)x[Re]x[Rktvm]x[C_3vldr]xxxx([Pq]xx) xxxxx[C_4ldae]xxxxxx(x)xxxC_5(x)xxxxC_6;
(5) (WxxWxx)C_1[Stnd][Vkaq][Tspl]C_2xx[Gq]xx(x)x[Re]x[Rktvm]x[C_3vldr]xxxx([Pq]xx) xxxxx[C_4ldae]xxxxxx(x)xxxC_5(x)xxxxC_6; and
(6) C_1[nst][aegiklqrstv][adenpqrst]C_2[adetgs]xgx[ikqrstv]x[aqrst]x[almrtv]xC_3xxxxxxxxx (xxxxxxx)C_4xxxxxxxxx(xx)C_5xxxxC_6

[131] In some embodiments, thrombospondin type 1 domain variants comprise sequences substantially identical to any of the above-described sequences.

[132] To date, at least 1677 naturally occurring thrombospondin domains have identified based on cDNA sequences. Exemplary proteins containing the naturally occurring thrombospondin domains include, *e.g.*, proteins in the complement pathway (*e.g.*, properdin, C6, C7, C8A, C8B, and C9), extracellular matrix proteins (*e.g.*, mindin, F-spondin, SCO-spondin,), circumsporozoite surface protein 2, and TRAP proteins of *Plasmodium*. Thrombospondin type 1 domains are further described in, *e.g.*, Roszmusz *et al.*, *BBRC* 296:156 (2002); Higgins *et al.*, *J Immunol.* 155:5777-85 (1995); Schultz-Cherry *et al.*, *J. Biol. Chem.* 270:7304-7310 (1995); Schultz-Cherry *et al.*, *J. Biol. Chem.* 269:26783-8 (1994); Bork, *FEBS Lett* 327:125-30 (1993); and Leung-Hagesteijn *et al.*, *Cell* 71:289-99 (1992).

[133] Another exemplary monomer domain suitable for use in the practice of the present invention is the trefoil domain. Trefoil monomer domains are typically about 30-50 or 30-65 amino acids. In some embodiments, the domains comprise about 35-55 amino acids and in some cases about 45 amino acids. Within the 35-55 amino acids, there are typically about 6 cysteine residues. Of the six cysteines, disulfide bonds typically are found between the following cysteines: C1 and C5, C2 and C4, C3 and C6.

[134] To date, at least 149 naturally occurring trefoil domains have identified based on cDNA sequences. Exemplary proteins containing naturally occurring trefoil domains include, *e.g.*, protein pS2 (TFF1), spasmolytic peptide SP (TFF2), intestinal trefoil factor (TFF3), intestinal surcease-isomaltase, and proteins which may be involved in defense against microbial infections by protecting the epithelia (*e.g.*, *Xenopus* xP1, xP4, integumentary mucins A.1 and C.1. Trefoil domains are further described in, *e.g.*, Sands and Podolsky, *Annu. Rev. Physiol.* 58:253-273 (1996); Carr *et al.*, *PNAS USA* 91:2206-2210 (1994); DeA *et al.*, *PNAS USA* 91:1084-1088 (1994); Hoffman *et al.*, *Trends Biochem Sci* 18:239-243 (1993).

[135] Exemplary trefoil domain sequences and consensus sequences are as follows (SEQ ID NOS:350-355):

(1) C_1(xx)xxxxxxxxxC_2xx(x)xxxxxxxC_3xxxxC_4C_5xxxxx(x)xxxxxC_6
(2) C_1(xx)xxxxxxRxxC_2xx(x)xxxxxxxC_3xxxxC_4C_5xxxxx(x)xxxxxC_6
(3) C_1(xx)xxxpxxRxnC_2gx(x)pxitxxxC_3xxxgC_4C_5fdxxx(x)xxxpwC_6f
(4) C_1(xx)xxx[Pvae]xxRx[ndpm]C_2[Gaiy][ypfst]([de]x)[pskq]x[Ivap][Tsa]xx[qedk]C_3xx[krln][Gnk]C_4C_5[Fwy][Dnrs][sdpnte]xx(x)xxx[pki][Weash]C_6[Fy]
(5) C_1(xx)xxx[Pvae]xxRx[ndpm]C_2[Gaiy][ypfst]([de]x)[pskq]x[Ivap][Tsa]xx[keqd]C_3xx[krln][Gnk]C_4C_5[α][Dnrs][sdpnte]xx(x)xxx[pki][Weash]C_6[Fy]
(6) C_1([dnps])[adiklnprstv][dfilmv][adenprst][adelprv][ehklnqrs][adegknsv][kqr][fiklqrtv][dnpqs]C_2[agiy][flpsvy][dknpqs][adfghlp][aipv][st][aegkpqrs][adegkpqs][deiknqt]C_3[adefknqrt][adegknqs][gn]C_4C_5[wyfh][deinrs][adgnpst][aefgqlrstw][giknsvmq]([afmprstv][degklns][afiqstv][iknpv]w)C_6

[136] Another exemplary monomer domain suitable for use in the present invention is the thyroglobulin domain. Thyroglobulin monomer domains are typically about 30-85 or 30-80 amino acids. In some embodiments, the domains comprise about 35-75 amino acids and in some cases about 65 amino acids. Within the 35-75 amino acids, there are typically about 6 cysteine residues. Of the six cysteines, disulfide bonds typically are found between the following cysteines: C1 and C2, C3 and C4, C5 and C6.

[137] To date at least 251 naturally occurring thyroglobulin domains have been identified based on cDNA sequences. The N-terminal section of Tg contains 10 repeats of a domain of about 65 amino acids which is known as the Tg type-1 repeat

PUBMED:3595599, PUBMED:8797845. Exemplary proteins containing naturally occurring thyroglobulin domains include *e.g.*, the HLA class II associated invariant chain, human pancreatic carcinoma marker proteins, nidogen (entactin), insulin-like growth factor binding proteins (IGFBP), saxiphilin, chum salmon egg cysteine proteinase inhibitor, and equistatin. The Thyr-1 and related domains belong to MEROPS proteinase inhibitor family I31, clan IX. Thyroglobulin domains are further described in, *e.g.*, Molina *et al.*, *Eur. J. Biochem.* 240:125-133 (1996); Guncar *et al.*, *EMBO J* 18:793-803 (1999); Chong and Speicher, *DW* 276:5804-5813 (2001).

[138] Exemplary thyroglobulin domain sequences and consensus sequences are as follows (SEQ ID NOS:356-360):

(1) C_1xxxxxxxxxxxxxxx(xxxxxxxxxx)xxxxxxxxxxC$_2$xxxxxxxxxxC$_3$x(x)x(xxx)xxxxC$_4$xC$_5$xxxx(x)xxxxxxxxxxxxxx(xx)xC$_6$

(2) C_1xxxxxxxxxxxxxxx(xxxxxxxxxx)xxxxxxxyxPxC$_2$xxxGxxxxxQC$_3$x(x)x(xxx)xxxxC$_4$WC$_5$Vxxx(x)GxxxxGxxxxxxxx(xx)xC$_6$

(3) C_1xxxxxxxxxxxxxxx(xxxxxxxxxx)xxxxxxxyxPxC$_2$xxxGxyxxxQC$_3$x(x)s(xxx)xxgxC$_4$WC$_5$Vdxx(x)GxxxxGxxxxxgxx(xx)xC$_6$

(4) C_1[qerl]xxxxxxxxxxxxxx(xxxxxxxxxx)xxxxxxx[Yfhp]xPxC$_2$xxxGx[Yf]xx[vkrl]QC$_3$x(x[sa]xxx)xx[Gsa]xC$_4$[Wyf]C$_5$V[Dnyfl]xx(x)Gxxxx[Gdne]xxxxxgxx(xx)xC$_6$

(5) C_1[qerl]xxxxxxxxxxxxxx(xxxxxxxxxx)xxxxxxx[αhp]xPxC2xxxGx[α]xx[vkrl]QC$_3$x(x[sa]xxx)xx[gas]xC$_4$[α]C$_5$V[Dnα]xx(x)Gxxxx[ϕg]xxxxxgxx(xx)xC$_6$

[139] Another exemplary monomer domain that can be used in the present invention is a laminin-EGF domain. Laminin-EGF domains are typically about 30-85 or 30-80 amino acids. In some embodiments, the domains comprise about 45-65 amino acids and in some cases about 50 amino acids. Within the 45-65 amino acids, there are typically about 8 cysteine residues which interact to form 4 disulfide bonds. Laminins are a major noncollagenous component of basement membranes that mediate cell adhesion, growth migration, and differentiation. They are composed of distinct but related alpha, beta, and gamma chains. The three chains form a cross-shaped molecule that consist of a long arm and three short globular arms. The long arm consists of a coiled coil structure contributed by all three chains and cross-linked by interchain disulfide bonds.

[140] Exemplary laminin EGF domain sequences and consensus sequences are as follows (SEQ ID NOS:361-363):

(1) C_1xC$_2$xxxxxx(xxx)xxC$_3$xxx(xxxxxx)xxxxC$_4$xC$_5$xxxxxxxxC$_6$xxC$_7$xxxxxxx(xxxxx)xxxxxC$_8$

(2) C_1xC$_2$xxxxxx(xxx)xxC3xxx(xxxxxx)xxgxC$_4$xC$_5$xxxxxGxxC$_6$xxC$_7$xxxxxxx(xxxxx)xxxxxC$_8$

(3) C_1xC$_2$[ndh]xxxxx(xxx)xxC$_3$xxx(xxxxxx)xxgxC$_4$xC$_5$xxxxxGxxC$_6$[denq]xC$_7$xx[gn][yfht]xxx(xxxxx)xxxxxC$_8$

In some embodiments, the monomer domain is a Notch/LNR monomer domain, a DSL monomer domain, an Anato monomer domain, an integrin beta monomer domain, and a Ca-EGF monomer domain.

[141] In some embodiments, the Ca-EGF monomer domain comprises the following sequence (SEQ ID NO:364):
$DxdEC_1xx(xx)xxxxC_2x(xx)xxxxxC_3xNxxGxfxC_4x(xxx)xC_5xxgxxxxxxx(xxxxx)xxxC_6$.

[142] In some embodiments, the Notch/LNR monomer domain, comprises the following sequence (SEQ ID NO:365):
$C_1xx(xx)xxxC_2xxxxxnGxC_3xxxC_4nxxxC_5xxDGxDC_6$

[143] In some embodiments, the DSL monomer domain comprises the following sequence: $C_1xxxYygxxC_2xxfC_3xxxxdxxxhxxC_4xxxGxxxC_5xxGWxGxxC_6$ (SEQ ID NO:366).

[144] Anato monomer domain comprises the following sequence (SEQ ID NO:366): $C_1C_2xdgxxxxx(x)xxxxC_3exrxxxxxx(xx)xxC_4xxxfxxC_5C_6$ (SEQ ID NO:367).

[145] In some embodiments, the integrin beta monomer domain comprises the following sequence (SEQ ID NO:368):
$C_1xxC_2xxxxpxC_3xwC4xxxxfxxx(gx)xxxxRC_5dxxxxLxxxgC_6$; and "x" is any amino acid.

[146] In some embodiments, C_1-C_5, C_2-C_4 and C_3-C_6 of the Notch/LNR monomer domain form disulfide bonds; and C_1-C_5, C_2-C_4 and C_3-C_6 of the DSL monomer domain form disulfide bonds.

[147] In some embodiments, the Ca-EGF monomer domain comprises the following sequence:
$D[\beta][Dn]EC_1xx(xx)xxxxC_2[pdg](dx)xxxxxC_3xNxxG[sgt][\alpha]xC_4x(xxx)xC_5xx[Gsn][\alpha s]$ xxxxxx(xxxxx)xxxC_6 (SEQ ID NO:369).

[148] In some embodiments, the Notch/LNR monomer domain, comprises the following sequence:
$C_1xx(x[\beta\alpha])xxxC_2x[\phi s]xxx[\phi][Gk]xC_3[nd]x[\phi sa]C_4[\phi s]xx[aeg]C_5x[\alpha]DGxDC_6$ (SEQ ID NO:370).

[149] In some embodiments, the DSL monomer domain comprises the following sequence:
$C_1xxx[\alpha][\alpha h][Gsna]xxC_2xx[\alpha]C_3x[pae]xx[Da]xx[\chi l][Hrgk][\alpha k]xC_4[dnsg]xxGxxxC_5xxG$ $[\alpha]xGxxC_6$ (SEQ ID NO:371).

[150] In some embodiments, the Anato monomer domain comprises the following sequence:
C_1C_2x[Dhtl][Ga]xxxx[plant](xx)xxxxC_3[esqdat]x[Rlps]xxxxxx([gepa]x)xxC_4xx[avfpt][Fqvy]xxC_5C_6 (SEQ ID NO:372).

[151] In some embodiments, the integrin beta monomer domain comprises the following sequence:
C_1xxC_2[β]xx[ghds][Pk]xC_3[χ][α]C_4xxxx[α]xxx([Gr]xx)x[χ]xRC_5[Dnae]xxxxL[βk]xx[Gn]C_6 (SEQ ID NO:373); α is selected from: w, y, f, and l; β is selected from: v, i, l, a, m, and f; χ is selected from: g, a, s, and t; δ is selected from: k, r, e, q, and d; ε is selected from: v, a, s, and t; and ϕ is selected from: d, e, and n.

[152] In some embodiments, the Ca-EGF monomer domain comprises the following sequence:
D[vilf][Dn]EC_1xx(xx)xxxxC_2[pdg](dx)xxxxxC_3xNxxG[sgt][fy]xC_4x(xxx)xC_5xx[Gsn][αs]xxxxxx(xxxxx)xxxC_6 (SEQ ID NO:374).

[153] In some embodiments, the Notch/LNR monomer domain, comprises the following sequence:
C_1xx(x[yiflv])xxxC_2x[dens]xxx[Nde][Gk]xC_3[nd]x[densa]C_4[Nsde]xx[aeg]C_5x[wyf]DGxDC$_6$ (SEQ ID NO:375).

[154] In some embodiments, the DSL monomer domain comprises the following sequence:
C_1xxx[Ywf][Yfh][Gasn]xxC_2xx[Fy]C_3x[pae]xx[Da]xx[glast][Hrgk][ykfw]xC_4[dsgn]xxGxxxC_5xxG[Wlfy]xGxxC_6 (SEQ ID NO:376).

[155] In some embodiments, the Anato monomer domain comprises the following sequence:
C_1C_2x[adehlt]gxxxxxxxx(x)[derst]C_3xxxxxxxxx(xx[aersv])C_4xx[apvt][fmq][eklqrtv][adehqrsk](x)C_5C_6 (SEQ ID NO:377).

[156] In some embodiments, the integrin beta monomer domain comprises the following sequence:
C_1[aegkqrst][kreqd]C_2[il][aelqrv][vilas][dghs][kp]xC_3[gast][wy]C_4xxxx[fl]xxxx(xxxx[vilar]r)C_5[and][dilrt][iklpqrv][adeps][aenq]l[iklqv]x[adknr][gn]C_6 (SEQ ID NO:378).

[157] Polynucleotides encoding the monomer domains are typically employed to make monomer domains via expression. Nucleic acids that encode monomer domains can be derived from a variety of different sources. Libraries of monomer domains

can be prepared by expressing a plurality of different nucleic acids encoding naturally occurring monomer domains, altered monomer domains (i.e., monomer domain variants), or a combinations thereof. For example, libraries may be designed in which a scaffold of amino acids remain constant (e.g., an LDL A receptor domain, EGF domain) while the intervening amino acids in the scaffold comprise randomly generated amino acids.

[158] The invention provides methods of identifying monomer domains that bind to a selected or desired ligand or mixture of ligands. In some embodiments, monomer domains are identified or selected for a desired property (e.g., binding affinity) and then the monomer domains are formed into multimers. *See, e.g.*, Figure 4. For those embodiments, any method resulting in selection of domains with a desired property (e.g., a specific binding property) can be used. For example, the methods can comprise providing a plurality of different nucleic acids, each nucleic acid encoding a monomer domain; translating the plurality of different nucleic acids, thereby providing a plurality of different monomer domains; screening the plurality of different monomer domains for binding of the desired ligand or a mixture of ligands; and, identifying members of the plurality of different monomer domains that bind the desired ligand or mixture of ligands.

[159] Monomer domains can be naturally-occurring or altered (non-natural variants). The term "naturally occurring" is used herein to indicate that an object can be found in nature. For example, natural monomer domains can include human monomer domains or optionally, domains derived from different species or sources, e.g., mammals, primates, rodents, fish, birds, reptiles, plants, etc. The natural occurring monomer domains can be obtained by a number of methods, e.g., by PCR amplification of genomic DNA or cDNA.

[160] Monomer domains of the present invention can be naturally-occurring domains or non-naturally occurring variants. Libraries of monomer domains employed in the practice of the present invention may contain naturally-occurring monomer domain, non-naturally occurring monomer domain variants, or a combination thereof.

[161] Monomer domain variants can include ancestral domains, chimeric domains, randomized domains, mutated domains, and the like. For example, ancestral domains can be based on phylogenetic analysis. Chimeric domains are domains in which one or more regions are replaced by corresponding regions from other domains of the same family. For example, chimeric domains can be constructed by combining loop sequences from multiple related domains of the same family to form novel domains with potentially lowered immunogenicity. Those of skill in the art will recognized the immunologic benefit

of constructing modified binding domain monomers by combining loop regions from various related domains of the same family rather than creating random amino acid sequences. For example, by constructing variant domains by combining loop sequences or even multiple loop sequences that occur naturally in human LDL receptor class A-domains, the resulting domains may contain novel binding properties but may not contain any immunogenic protein sequences because all of the exposed loops are of human origin. The combining of loop amino acid sequences in endogenous context can be applied to all of the monomer constructs of the invention. Thus the present invention provides a method for generating a library of chimeric monomer domains derived from human proteins, the method comprising: providing loop sequences corresponding to at least one loop from each of at least two different naturally occurring variants of a human protein, wherein the loop sequences are polynucleotide or polypeptide sequences; and covalently combining loop sequences to generate a library of at least two different chimeric sequences, wherein each chimeric sequence encodes a chimeric monomer domain having at least two loops. Typically, the chimeric domain has at least four loops, and usually at least six loops. As described above, the present invention provides three types of loops that are identified by specific features, such as, potential for disulfide bonding, bridging between secondary protein structures, and molecular dynamics (i.e., flexibility). The three types of loop sequences are a cysteine-defined loop sequence, a structure-defined loop sequence, and a B-factor-defined loop sequence.

[162] Randomized domains are domains in which one or more regions are randomized. The randomization can be based on full randomization, or optionally, partial randomization based on natural distribution of sequence diversity.

[163] The present invention also provides recombinant nucleic acids encoding one or more polypeptides comprising one or a plurality of monomer domains that bind IL-6. For example, the polypeptide can be selected to comprise a non-naturally occuring domain from the group consisting of: an EGF-like domain, a Kringle-domain, a fibronectin type I domain, a fibronectin type II domain, a fibronectin type III domain, a PAN domain, a Gla domain, a SRCR domain, a Kunitz/Bovine pancreatic trypsin Inhibitor domain, a Kazal-type serine protease inhibitor domain, a Trefoil (P-type) domain, a von Willebrand factor type C domain, an Anaphylatoxin-like domain, a CUB domain, a thyroglobulin type I repeat, LDL-receptor class A domain, a Sushi domain, a Link domain, a Thrombospondin type I domain, an Immunoglobulin-like domain, a C-type lectin domain, a MAM domain, a von Willebrand factor type A domain, a Somatomedin B domain, a WAP-type four disulfide core domain, a F5/8 type C domain, a Hemopexin domain, an SH2 domain, an SH3 domain, a

Laminin-type EGF-like domain, a C2 domain and variants of one or more thereof. In another embodiment, the naturally occuring polypeptide encodes a monomer domain found in the Pfam database and/or the SMART database.

[164] All the compositions of the present invention, including the compositions produced by the methods of the present invention, e.g., monomer domains and/or immuno-domains, as well as multimers and libraries thereof can be optionally bound to a matrix of an affinity material. Examples of affinity material include beads, a column, a solid support, a microarray, other pools of reagent-supports, and the like.

III. MULTIMERS

[165] Methods for generating multimers are a feature of the present invention. Multimers comprise at least two monomer domains. For example, multimers of the invention can comprise from 2 to about 10 monomer domains, from 2 and about 8 monomer domains, from about 3 and about 10 monomer domains, about 7 monomer domains, about 6 monomer domains, about 5 monomer domains, or about 4 monomer domains. In some embodiments, the multimer comprises 3 or at least 3 monomer domains. In some embodiments, the mutimers have no more than 2, 3, 4, 5, 6, 7, or 8 monomer domains. In view of the possible range of monomer domain sizes, the multimers of the invention may be, e.g., less than 100 kD, less than 90kD, less than 80kD, less than 70kD, less than 60kD, less than 50kd, less than 40kD, less than 30kD, less than 25kD, less than 20kD, less than 15kD, less than 10kD or may be smaller or larger. In some cases, the monomer domains have been pre-selected for binding to the target molecule of interest (e.g., IL-6). In some embodiments, the multimers of the invention will bind to IL-6 (or, e.g., a blood factor such as HSA or IgG) with an affinity of less than 10^{-3} M (i.s., binds stronger than an affinity of 10^{-3} M). In some embodiments, the affinity is less than 10^{-4} M, 10^{-5} M, 10^{-6} M, 10^{-7} M, 10^{-8} M, 10^{-9} M, or 10^{-10} M. In some embodiments, the affinity of a multimer will be stronger (less) than the sum of affinities of the respective monomer domains in the multimer.

[166] In some embodiments, each monomer domain specifically binds to one target molecule (e.g., IL-6). In some of these embodiments, each monomer binds to a different position (analogous to an epitope) on a target molecule. Multiple monomer domains that bind to the same target molecule results in an avidity effect resulting in improved affinity of the multimer for the target molecule compared to the affinity of each individual monomer. In some embodiments, the multimer has an avidity of at least about 1.5, 2, 3, 4, 5, 10, 20, 50, 100, 200, 500, or 1000 times the avidity of a monomer domain alone. In some embodiments,

at least one, two, three, four or more (e.g., all) monomers of a multimer bind an ion such as calcium or another ion. Multimers can comprise a variety of combinations of monomer domains. For example, in a single multimer, the selected monomer domains can be identical or different. In addition, the selected monomer domains can comprise various different monomer domains from the same monomer domain family, or various monomer domains from different domain families , or optionally, a combination of both. For example, the monomer domains may be selected from Families 1-7 of IL-6-binding monomer domains. In some embodiments, at least one of the monomer domains is selected from Family 3 and/or 5 of the IL-6 binding monomer domains. Exemplary IL-6-binding dimers and trimers (comprised of two or three, respectively, IL-6-binding monomers) are listed in the examples.

[167] Multimers that are generated in the practice of the present invention may be any of the following:

(1) A homo-multimer (a multimer of the same domain, i.e., A1-A1-A1-A1);

(2) A hetero-multimer of different domains of the same domain class, e.g., A1-A2-A3-A4. For example, hetero-multimer include multimers where A1, A2, A3 and A4 are different non-naturally occurring variants of a particular LDL-receptor class A domains, or where some of A1, A2, A3, and A4 are naturally-occurring variants of a LDL-receptor class A domain.

(3) A hetero-multimer of domains from different monomer domain classes, e.g., A1-B2-A2-B1. For example, where A1 and A2 are two different monomer domains (either naturally occurring or non-naturally-occurring) from LDL-receptor class A, and B1 and B2 are two different monomer domains (either naturally occurring or non-naturally occurring) from class EGF-like domain).

[168] In another embodiment, the multimer comprises monomer domains with specificities for different target molecules (*e.g.*, a blood factor such as serum albumin, immunoglobulin, or erythrocytes). For example, in some embodiments, the multimers of the invention comprises 1, 2, 3, or more monomer domains that bind to IL-6 and at least one monomer domain that binds to a second target molecule. Exemplary target molecules include, e.g., a serum molecule that extends the serum half-life of the multimer (*e.g.*, an immunoglobulin or serum albumin). Exemplary molecules that extend the serum half-life of a multimer include, e.g., red blood cells (*i.e.*, erythrocytes), IgG, and serum albumin such as HSA. An exemplary multimer will include, e.g., a monomer domain from Family 3 and/or Family 5 of the IL-6 binding monomer domains (and optionally 2, 3, 4, or more IL-6 binding

domains) and monomer domain from Family 2 or 3 of the immunoglobulin binding monomer domains.

[169] Multimer libraries employed in the practice of the present invention may contain homo-multimers, hetero-multimers of different monomer domains (natural or non-natural) of the same monomer class, or hetero-multimers of monomer domains (natural or non-natural) from different monomer classes, or combinations thereof.

[170] Monomer domains, as described herein, are also readily employed in a immuno-domain-containing heteromultimer (i.e., a multimer that has at least one immuno-domain variant and one monomer domain variant). Thus, multimers of the present invention may have at least one immuno-domain such as a minibody, a single-domain antibody, a single chain variable fragment (ScFv), or a Fab fragment; and at least one monomer domain, such as, for example, an EGF-like domain, a Kringle-domain, a fibronectin type I domain, a fibronectin type II domain, a fibronectin type III domain, a PAN domain, a Gla domain, a SRCR domain, a Kunitz/Bovine pancreatic trypsin Inhibitor domain, a Kazal-type serine protease inhibitor domain, a Trefoil (P-type) domain, a von Willebrand factor type C domain, an Anaphylatoxin-like domain, a CUB domain, a thyroglobulin type I repeat, LDL-receptor class A domain, a Sushi domain, a Link domain, a Thrombospondin type I domain, an Immunoglobulin-like domain, a C-type lectin domain, a MAM domain, a von Willebrand factor type A domain, a Somatomedin B domain, a WAP-type four disulfide core domain, a F5/8 type C domain, a Hemopexin domain, an SH2 domain, an SH3 domain, a Laminin-type EGF-like domain, a C2 domain, or variants thereof.

[171] Domains need not be selected before the domains are linked to form multimers. On the other hand, the domains can be selected for the ability to bind to a target molecule before being linked into multimers. Thus, for example, a multimer can comprise two domains that bind to one target molecule and a third domain that binds to a second target molecule.

[172] The multimers of the present invention may have the following qualities: multivalent, multispecific, single chain, heat stable, extended serum and/or shelf half-life. Moreover, at least one, more than one or all of the monomer domains may bind an ion (e.g., a metal ion or a calcium ion), at least one, more than one or all monomer domains may be derived from LDL receptor A domains and/or EGF-like domains, at least one, more than one or all of the monomer domains may be non-naturally occurring, and/or at least one, more than one or all of the monomer domains may comprise 1, 2, 3, or 4 disulfide bonds per monomer domain. In some embodiments, the multimers comprise at least two (or at least

three) monomer domains, wherein at least one monomer domain is a non-naturally occurring monomer domain and the monomer domains bind calcium. In some embodiments, the multimers comprise at least 4 monomer domains, wherein at least one monomer domain is non-naturally occurring, and wherein:

a. each monomer domain is between 30-100 amino acids and each of the monomer domains comprise at least one disulfide linkage; or

b. each monomer domain is between 30-100 amino acids and is derived from an extracellular protein; or

c. each monomer domain is between 30-100 amino acids and binds to a protein target.

[173] In some embodiments, the multimers comprise at least 4 monomer domains, wherein at least one monomer domain is non-naturally occurring, and wherein:

a. each monomer domain is between 35-100 amino acids; or

b. each domain comprises at least one disulfide bond and is derived from a human protein and/or an extracellular protein.

[174] In some embodiments, the multimers comprise at least two monomer domains, wherein at least one monomer domain is non-naturally occurring, and wherein each domain is:

a. 25-50 amino acids long and comprises at least one disulfide bond; or

b. 25-50 amino acids long and is derived from an extracellular protein; or

c. 25-50 amino acids and binds to a protein target; or

d. 35-50 amino acids long.

[175] In some embodiments, the multimers comprise at least two monomer domains, wherein at least one monomer domain is non-naturally-occurring and:

a. each monomer domain comprises at least one disulfide bond; or

b. at least one monomer domain is derived from an extracellular protein; or

c. at least one monomer domain binds to a target protein.

[176] The monomer domains and/or multimers identified can have biological activity, which is meant to include at least specific binding affinity for a selected or desired ligand, and, in some instances, will further include the ability to block the binding of other compounds, to stimulate or inhibit metabolic pathways, to act as a signal or messenger, to stimulate or inhibit cellular activity, and the like. Income embodiments, the monomer domains of the invention bind to IL-6R, the IL-6 receptor. In some embodiments, these monomer domains either block Il-6 binding to IL-6R or activate the IL-6 receptor.

[177] A single ligand can be used, or optionally a variety of ligands can be used to select the monomer domains and/or multimers. A monomer domain of the present invention can bind a single ligand or a variety of ligands. A multimer of the present invention can have multiple discrete binding sites for a single ligand, or optionally, can have multiple binding sites for a variety of ligands.

[178] In some embodiments, the multimers of the invention bind to the same or other multimers to form aggregates. Aggregation can be mediated, for example, by the presence of hydrophobic domains on two monomer domains, resulting in the formation of non-covalent interactions between two monomer domains. Alternatively, aggregation may be facilitated by one or more monomer domains in a multimer having binding specificity for a monomer domain in another multimer. Aggregates can also form due to the presence of affinity peptides on the monomer domains or multimers. Aggregates can contain more target molecule binding domains than a single multimer.

[179] Multimers with affinity for both a cell surface target and a second target (e.g., IL-6) may provide for increased avidity effects. In some cases, membrane fluidity can be more flexible than protein linkers in optimizing (by self-assembly) the spacing and valency of the interactions. In some cases, multimers will bind to two different targets, each on a different cell or one on a cell and another on a molecule with multiple binding sites.

[180] In some embodiments, the monomers or multimers of the present invention are linked to another polypeptide to form a fusion protein. Any polypeptide in the art may be used as a fusion partner, though it can be useful if the fusion partner forms multimers. For example, monomers or multimers of the invention may, for example, be fused to the following locations or combinations of locations of an antibody:

1. At the N-terminus of the VH1 and/or VL1 domains, optionally just after the leader peptide and before the domain starts (framework region 1);
2. At the N-terminus of the CH1 or CL1 domain, replacing the VH1 or VL1 domain;
3. At the N-terminus of the heavy chain, optionally after the CH1 domain and before the cysteine residues in the hinge (Fc-fusion);
4. At the N-terminus of the CH3 domain;
5. At the C-terminus of the CH3 domain, optionally attached to the last amino acid residue via a short linker;
6. At the C-terminus of the CH2 domain, replacing the CH3 domain;

7. At the C-terminus of the CL1 or CH1 domain, optionally after the cysteine that forms the interchain disulfide; or

8. At the C-terminus of the VH1 or VL1 domain. *See, e.g.*, Figure 7.

[181] In some embodiments, one or more monomer or multimer domains of the invention is linked to a molecule (e.g., a protein, nucleic acid, organic small molecule, etc.) useful as a pharmaceutical. Exemplary pharmaceutical proteins include, e.g., cytokines, antibodies, chemokines, growth factors, interleukins (e.g., IL-6), cell-surface proteins, extracellular domains, cell surface receptors, cytotoxins, etc. Exemplary small molecule pharmaceuticals include toxins or therapeutic agents. In some embodiments, a metal can be bound to the polypeptides of the invention. This can be useful, e.g., as a contrast agent, e.g., for MRI.

[182] In some embodiments, the monomer or multimers are selected to bind to a tissue- or disease-specific target protein. Tissue-specific proteins are proteins that are expressed exclusively, or at a significantly higher level, in one or several particular tissue(s) compared to other tissues in an animal.

[183] In some embodiments, the monomers or multimers that bind to the target protein are linked to the pharmaceutical protein or small molecule such that the resulting complex or fusion is targeted to the specific tissue or disease-related cell(s) where the target protein (e.g., IL-6) is expressed. Monomers or multimers for use in such complexes or fusions can be initially selected for binding to the target protein and may be subsequently selected by negative selection against other cells or tissue (e.g., to avoid targeting bone marrow or other tissues that set the lower limit of drug toxicity) where it is desired that binding be reduced or eliminated in other non-target cells or tissues. By keeping the pharmaceutical away from sensitive tissues, the therapeutic window is increased so that a higher dose may be administered safely. In another alternative, *in vivo* panning can be performed in animals by injecting a library of monomers or multimers into an animal and then isolating the monomers or multimers that bind to a particular tissue or cell of interest.

[184] The fusion proteins described above may also include a linker peptide between the pharmaceutical protein and the monomer or multimers. A peptide linker sequence may be employed to separate, for example, the polypeptide components by a distance sufficient to ensure that each polypeptide folds into its secondary and tertiary structures. Fusion proteins may generally be prepared using standard techniques, including chemical conjugation. Fusion proteins can also be expressed as recombinant proteins in an expression system by standard techniques.

[185] Multimers or monomer domains of the invention can be produced according to any methods known in the art. In some embodiments, *E. coli* comprising a plasmid encoding the polypeptides under transcriptional control of an inducible promoter are used to express the protein. After harvesting the bacteria, they may be lysed by sonication or homogenization and clarified by centrifugation. The polypeptides may be purified using Ni-NTA agarose elution (if 6xHis (SEQ ID NO:379) tagged) or DEAE sepharose elution (if untagged) and refolded by dialysis. Misfolded proteins may be neutralized by capping free sulfhydrils with iodoacetic acid. Q sepharose elution, butyl sepharose flow-through, SP sepharose elution, DEAE sepharose elution, and/or CM sepharose elution may be used to purify the polypeptides. Equivalent anion and/or cation exchange purification steps may also be employed.

[186] In some embodiments, monomers or multimers are purified using heat lysis, typically followed by a fast cooling to prevent most proteins from renaturing. Due to the heat stability of the proteins of the invention, the desired proteins will not be denatured by the heat and therefore will allow for a purification step resulting in high purity. In some embodiments, a continuous flow heating process to purify the monomers or multimers from bacterial cell cultures is used. For example, a cell suspension can passed through a stainless steel coil submerged in a water bath set to a temperature resulting in lysis of the bacteria (e.g., about 65-100 degrees C). The lysed effluent is routed to a cooling bath to obtain rapid cooling and prevent renaturation of denatured *E. coli* proteins. *E. coli* proteins denature and are prevented from renaturing, but the monomer or multimers do not denature under these conditions due to the exceptional stability of their scaffold. The heating time is controlled by adjusting the flow rate and length of the coil. This approach yields active proteins with high yield and exceptionally high purity (e.g., >80%) compared to alternative approaches and is amenable to large scale production of clinical material.

[187] In some embodiments, following manufacture of the monomers or multimers of the invention, the polypeptides are stored in a solution comprising iodoacetic acid to prevent scrambling of disulfide bind linkages. In some embodiments, 0.1-100 mM (e.g., 1-10 mM) iodoacetic acid is included in the solutions for storage. If desired, the iodoacetic acid can be removed before administered to an individual.

[188] In some embodiments, the polypeptide comprising a monomer or multimer of the invention is linked to itself (C-terminus to N-terminus), e.g., for protein stability.

IV. LINKERS

[189] Monomer domains can be joined by a linker to form a multimer. For example, a linker may be positioned between each separate discrete monomer domain in a multimer.

[190] Joining the selected monomer domains via a linker can be accomplished using a variety of techniques known in the art. For example, combinatorial assembly of polynucleotides encoding selected monomer domains can be achieved by restriction digestion and re-ligation, by PCR-based, self-priming overlap reactions, or other recombinant methods. The linker can be attached to a monomer before the monomer is identified for its ability to bind to a target multimer or after the monomer has been selected for the ability to bind to a target multimer.

[191] The linker can be naturally-occurring, synthetic or a combination of both. For example, the synthetic linker can be a randomized linker, e.g., both in sequence and size. In one aspect, the randomized linker can comprise a fully randomized sequence, or optionally, the randomized linker can be based on natural linker sequences. The linker can comprise, e.g,. a non-polypeptide moiety, a polynucleotide, a polypeptide or the like.

[192] A linker can be rigid, or flexible, or a combination of both. Linker flexibility can be a function of the composition of both the linker and the monomer domains that the linker interacts with. The linker joins two selected monomer domain, and maintains the monomer domains as separate discrete monomer domains. The linker can allow the separate discrete monomer domains to cooperate yet maintain separate properties such as multiple separate binding sites for the same ligand in a multimer, or e.g., multiple separate binding sites for different ligands in a multimer.

[193] Choosing a suitable linker for a specific case where two or more monomer domains (i.e. polypeptide chains) are to be connected may depend on a variety of parameters including, e.g. the nature of the monomer domains, the structure and nature of the target to which the polypeptide multimer should bind and/or the stability of the peptide linker towards proteolysis and oxidation.

[194] The present invention provides methods for optimizing the choice of linker once the desired monomer domains/variants have been identified. Generally, libraries of multimers having a composition that is fixed with regard to monomer domain composition, but variable in linker composition and length, can be readily prepared and screened as described above.

[195] A more detailed discussion of linkers can be found in, e.g., U.S. Patent Publication No. 2005/0048512.

V. IDENTIFYING MONOMERS OR MULTIMERS WITH AFFINITY FOR A TARGET MOLECULE

[196] Those of skill in the art can readily identify monomer domains with a desired property (e.g., binding affinity). For those embodiments, any method resulting in selection of domains with a desired property (e.g., a specific binding property) can be used. For example, the methods can comprise providing a plurality of different nucleic acids, each nucleic acid encoding a monomer domain; translating the plurality of different nucleic acids, thereby providing a plurality of different monomer domains; screening the plurality of different monomer domains for binding of the desired ligand or a mixture of ligands; and, identifying members of the plurality of different monomer domains that bind the desired ligand or mixture of ligands.

[197] In addition, any method of mutagenesis, such as site-directed mutagenesis and random mutagenesis (e.g., chemical mutagenesis) can be used to produce monomer domains, e.g., for a monomer domain library. In some embodiments, error-prone PCR is employed to create variants. Additional methods include aligning a plurality of naturally occurring monomer domains by aligning conserved amino acids in the plurality of naturally occurring monomer domains; and, designing the non-naturally occurring monomer domain by maintaining the conserved amino acids and inserting, deleting or altering amino acids around the conserved amino acids to generate the non-naturally occurring monomer domain. In one embodiment, the conserved amino acids comprise cysteines. In another embodiment, the inserting step uses random amino acids, or optionally, the inserting step uses portions of the naturally occurring monomer domains. The portions could ideally encode loops from domains from the same family. Amino acids are inserted or exchanged using synthetic oligonucleotides, or by shuffling, or by restriction enzyme based recombination. Human chimeric domains of the present invention are useful for therapeutic applications where minimal immunogenicity is desired. The present invention provides methods for generating libraries of human chimeric domains. Human chimeric monomer domain libraries can be constructed by combining loop sequences from different variants of a human monomer domain, as described above. The loop sequences that are combined may be sequence-defined loops, structure-defined loops, B-factor-defined loops, or a combination of any two or more thereof.

[198] Alternatively, a human chimeric domain library can be generated by modifying naturally-occurring human monomer domains at the amino acid level, as compared to the loop level. In some embodiments, to minimize the potential for immunogenicity, only those residues that naturally occur in protein sequences from the same family of human monomer domains are utilized to create the chimeric sequences. This can be achieved by providing a sequence alignment of at least two human monomer domains from the same family of monomer domains, identifying amino acid residues in corresponding positions in the human monomer domain sequences that differ between the human monomer domains, generating two or more human chimeric monomer domains, wherein each human chimeric monomer domain sequence consists of amino acid residues that correspond in type and position to residues from two or more human monomer domains from the same family of monomer domains. Libraries of human chimeric monomer domains can be employed to identify human chimeric monomer domains that bind to a target of interest by: screening the library of human chimeric monomer domains for binding to a target molecule, and identifying a human chimeric monomer domain that binds to the target molecule. Suitable naturally-occurring human monomer domain sequences employed in the initial sequence alignment step include those corresponding to any of the naturally-occurring monomer domains described herein.

[199] Domains of human monomer variant libraries of the present invention (whether generated by varying loops or single amino acid residues) can be prepared by methods known to those having ordinary skill in the art. Methods particularly suitable for generating these libraries are split-pool format and trinucleotide synthesis format as described in WO01/23401.

[200] In some embodiments, monomer domains of the invention are screened for potential immunogenicity by:

providing a candidate protein sequence;

comparing the candidate protein sequence to a database of human protein sequences;

identifying portions of the candidate protein sequence that correspond to portions of human protein sequences from the database; and

determining the extent of correspondence between the candidate protein sequence and the human protein sequences from the database.

[201] In general, the greater the extent of correspondence between the candidate protein sequence and one or more of the human protein sequences from the database, the lower the potential for immunogenicity is predicted as compared to a candidate

protein having little correspondence with any of the human protein sequences from the database. A database of human protein sequences that is suitable for use in the practice of the invention method for screening candidate proteins can be found at ncbi.nlm.nih.gov/blast/Blast.cgi at the World Wide Web (in addition, the following web site can be used to search short, nearly exact matches: cbi.nlm.nih.gov/blast/Blast.cgi?CMD =Web&LAYOUT=TwoWindows&AUTO_FORMAT=Semiauto&ALIGNMENTS=50&ALIGNMENT_VIEW=Pairwise&CLIENT=web&DATABASE=nr&DESCRIPTIONS=100&ENTREZ_QUERY=(none)&EXPECT=1000&FORMAT_OBJECT=Alignment&FORMAT_TYPE=HTML&NCBI_GI=on&PAGE=Nucleotides&PROGRAM=blastn&SERVICE=plain&SET_DEFAULTS.x=29&SET_DEFAULTS.y=6&SHOW_OVERVIEW=on&WORD_SIZE=7&END_OF_HTTPGET=Yes&SHOW_LINKOUT=yes at the World Wide Web). The method is particularly useful in determining whether a crossover sequence in a chimeric protein, such as, for example, a chimeric monomer domain, is likely to cause an immunogenic event. If the crossover sequence corresponds to a portion of a sequence found in the database of human protein sequences, it is believed that the crossover sequence is less likely to cause an immunogenic event.

[202] Information pertaining to portions of human protein sequences from the database can be used to design a protein library of human-like chimeric proteins. Such library can be generated by using information pertaining to "crossover sequences" that exist in naturally occurring human proteins. The term "crossover sequence" refers herein to a sequence that is found in its entirety in at least one naturally occurring human protein, in which portions of the sequence are found in two or more naturally occurring proteins. Thus, recombination of the latter two or more naturally occurring proteins would generate a chimeric protein in which the chimeric portion of the sequence actually corresponds to a sequence found in another naturally occurring protein. The crossover sequence contains a chimeric junction of two consecutive amino acid residue positions in which the first amino acid position is occupied by an amino acid residue identical in type and position found in a first and second naturally occurring human protein sequence, but not a third naturally occurring human protein sequence. The second amino acid position is occupied by an amino acid residue identical in type and position found in a second and third naturally occurring human protein sequence, but not the first naturally occurring human protein sequence. In other words, the "second" naturally occurring human protein sequence corresponds to the naturally occurring human protein in which the crossover sequence appears in its entirety, as described above.

[203] In some embodiments, a library of human-like chimeric proteins is generated by: identifying human protein sequences from a database that correspond to proteins from the same family of proteins; aligning the human protein sequences from the same family of proteins to a reference protein sequence; identifying a set of subsequences derived from different human protein sequences of the same family, wherein each subsequence shares a region of identity with at least one other subsequence derived from a different naturally occurring human protein sequence; identifying a chimeric junction from a first, a second, and a third subsequence, wherein each subsequence is derived from a different naturally occurring human protein sequence, and wherein the chimeric junction comprises two consecutive amino acid residue positions in which the first amino acid position is occupied by an amino acid residue common to the first and second naturally occurring human protein sequence, but not the third naturally occurring human protein sequence, and the second amino acid position is occupied by an amino acid residue common to the second and third naturally occurring human protein sequence, and generating human-like chimeric protein molecules each corresponding in sequence to two or more subsequences from the set of subsequences, and each comprising one of more of the identified chimeric junctions.

[204] Thus, for example, if the first naturally-occurring human protein sequence is, A-B-C, and the second is, B-C-D-E, and the third is, D-E-F, then the chimeric junction is C-D. Alternatively, if the first naturally-occurring human protein sequence is D-E-F-G, and the second is B-C-D-E-F, and the third is A-B-C-D, then the chimeric junction is D-E. Human-like chimeric protein molecules can be generated in a variety of ways. For example, oligonucleotides comprising sequences encoding the chimeric junctions can be recombined with oligonucleotides corresponding in sequence to two or more subsequences from the above-described set of subsequences to generate a human-like chimeric protein, and libraries thereof. The reference sequence used to align the naturally occurring human proteins is a sequence from the same family of naturally occurring human proteins, or a chimera or other variant of proteins in the family.

[205] Nucleic acids encoding fragments of naturally-occurring monomer domains can also be mixed and/or recombined (e.g., by using chemically or enzymatically-produced fragments) to generate full-length, modified monomer domains. The fragments and the monomer domain can also be recombined by manipulating nucleic acids encoding domains or fragments thereof. For example, ligating a nucleic acid construct encoding fragments of the monomer domain can be used to generate an altered monomer domain.

[206] Altered monomer domains can also be generated by providing a collection of synthetic oligonucleotides (e.g., overlapping oligonucleotides) encoding conserved, random, pseudorandom, or a defined sequence of peptide sequences that are then inserted by ligation into a predetermined site in a polynucleotide encoding a monomer domain. Similarly, the sequence diversity of one or more monomer domains can be expanded by mutating the monomer domain(s) with site-directed mutagenesis, random mutation, pseudorandom mutation, defined kernal mutation, codon-based mutation, and the like. The resultant nucleic acid molecules can be propagated in a host for cloning and amplification. In some embodiments, the nucleic acids are shuffled.

[207] The present invention also provides a method for recombining a plurality of nucleic acids encoding monomer domains and screening the resulting library for monomer domains that bind to the desired ligand or mixture of ligands or the like. Selected monomer domain nucleic acids can also be back-crossed by shuffling with polynucleotide sequences encoding neutral sequences (i.e., having insubstantial functional effect on binding), such as for example, by back-crossing with a wild-type or naturally-occurring sequence substantially identical to a selected sequence to produce native-like functional monomer domains. Generally, during back-crossing, subsequent selection is applied to retain the property, e.g., binding to the ligand.

[208] In some embodiments, the monomer library is prepared by shuffling. In such a case, monomer domains are isolated and shuffled to combinatorially recombine the nucleic acid sequences that encode the monomer domains (recombination can occur between or within monomer domains, or both). The first step involves identifying a monomer domain having the desired property, e.g., affinity for a certain ligand. While maintaining the conserved amino acids during the recombination, the nucleic acid sequences encoding the monomer domains can be recombined, or recombined and joined into multimers.

[209] A significant advantage of the present invention is that known ligands, or unknown ligands can be used to select the monomer domains and/or multimers. No prior information regarding ligand structure is required to isolate the monomer domains of interest or the multimers of interest. The monomer domains and/or multimers identified can have biological activity, which is meant to include at least specific binding affinity for a selected or desired ligand, and, in some instances, will further include the ability to block the binding of other compounds, to stimulate or inhibit metabolic pathways, to act as a signal or messenger, to stimulate or inhibit cellular activity, and the like. Monomer domains can be generated to function as ligands for receptors where the natural ligand for the receptor has not yet been

identified (orphan receptors). These orphan ligands can be created to either block or activate the receptor to which they bind.

[210] A single ligand can be used, or optionally a variety of ligands can be used to select the monomer domains and/or multimers. A monomer domain of the present invention can bind a single ligand or a variety of ligands. A multimer of the present invention can have multiple discrete binding sites for a single ligand, or optionally, can have multiple binding sites for a variety of ligands.

[211] In some embodiments, the multimers or monomers are screened for the ability to bind to a protein from multimple (2, 3, or more) different species. For example, human and primate orthologous proteins can be screened, thereby generating polypeptides that bind the proteins (e.g., IL-6, IgG or other serum half-life extender or other protein) from both a human and a primate, thereby allowing for ease of analysis of toxicology in primates for use in humans.

[212] The invention also includes compositions that are produced by methods of the present invention. For example, the present invention includes monomer domains selected or identified from a library and/or libraries comprising monomer domains produced by the methods of the present invention.

[213] The present invention also provides libraries of monomer domains and libraries of nucleic acids that encode monomer domains. The libraries can include, e.g., about 100, 250, 500 or more nucleic acids encoding monomer domains, or the library can include, e.g., about 100, 250, 500 or more polypeptides that encode monomer domains. Libraries can include monomer domains containing the same cysteine frame, e.g., A-domains or EGF-like domains.

[214] In some embodiments, variants are generated by recombining two or more different sequences from the same family of monomer domains (e.g., the LDL receptor class A domain). Alternatively, two or more different monomer domains from different families can be combined to form a multimer. In some embodiments, the multimers are formed from monomers or monomer variants of at least one of the following family classes: an EGF-like domain, a Kringle-domain, a fibronectin type I domain, a fibronectin type II domain, a fibronectin type III domain, a PAN domain, a Gla domain, a SRCR domain, a Kunitz/Bovine pancreatic trypsin Inhibitor domain, a Kazal-type serine protease inhibitor domain, a Trefoil (P-type) domain, a von Willebrand factor type C domain, an Anaphylatoxin-like domain, a CUB domain, a thyroglobulin type I repeat, LDL-receptor class A domain, a Sushi domain, a Link domain, a Thrombospondin type I domain, an

Immunoglobulin-like domain, a C-type lectin domain, a MAM domain, a von Willebrand factor type A domain, a Somatomedin B domain, a WAP-type four disulfide core domain, a F5/8 type C domain, a Hemopexin domain, an SH2 domain, an SH3 domain, a Laminin-type EGF-like domain, a C2 domain and derivatives thereof. In another embodiment, the monomer domain and the different monomer domain can include one or more domains found in the Pfam database and/or the SMART database. Libraries produced by the methods above, one or more cell(s) comprising one or more members of the library, and one or more displays comprising one or more members of the library are also included in the present invention.

[215] Optionally, a data set of nucleic acid character strings encoding monomer domains can be generated e.g., by mixing a first character string encoding a monomer domain, with one or more character string encoding a different monomer domain, thereby producing a data set of nucleic acids character strings encoding monomer domains, including those described herein. In another embodiment, the monomer domain and the different monomer domain can include one or more domains found in the Pfam database and/or the SMART database. The methods can further comprise inserting the first character string encoding the monomer domain and the one or more second character string encoding the different monomer domain in a computer and generating a multimer character string(s) or library(s), thereof in the computer.

[216] The libraries can be screened for a desired property such as binding of a desired ligand or mixture of ligands. For example, members of the library of monomer domains can be displayed and prescreened for binding to a known or unknown ligand or a mixture of ligands. The monomer domain sequences can then be mutagenized (e.g., recombined, chemically altered, etc.) or otherwise altered and the new monomer domains can be screened again for binding to the ligand or the mixture of ligands with an improved affinity. The selected monomer domains can be combined or joined to form multimers, which can then be screened for an improved affinity or avidity or altered specificity for the ligand or the mixture of ligands. Altered specificity can mean that the specificity is broadened, e.g., binding of multiple related viruses, or optionally, altered specificity can mean that the specificity is narrowed, e.g., binding within a specific region of a ligand. Those of skill in the art will recognize that there are a number of methods available to calculate avidity. *See, e.g.*, Mammen *et al.*, *Angew Chem Int. Ed.* 37:2754-2794 (1998); Muller *et al.*, *Anal. Biochem.* 261:149-158 (1998).

VI. SELECTION OF MONOMER DOMAINS THAT BIND IL-6

[217] Preliminary screens can be conducted by screening for agents capable of binding to IL-6, as at least some of the agents so identified are likely IL-6 modulators (e.g., antagonists or agonists). The binding assays usually involve contacting an IL-6 protein (or a fragment thereof) with one or more test agents (i.e., monomers or multimers of the invention) and allowing sufficient time for the protein and test agents to form a binding complex. Any binding complexes formed can be detected using any of a number of established analytical techniques. Protein binding assays include, but are not limited to, immunohistochemical binding assays, flow cytometry or other assays. The IL-6 protein utilized in such assays can be naturally expressed, cloned or synthesized. Similar methods can be used to identify monomer domains or multimers that bind IgG.

[218] The screening methods of the invention can be performed as *in vitro* or cell-based assays. Cell based assays can be performed in any cells in which IL-6 is expressed. Cell-based assays may involve whole cells or cell fractions containing an IL-6 receptor to screen for agent binding or modulation of activity of IL-6 by the agent. Exemplary cell types that can be used according to the methods of the invention include, e.g., any mammalian cells, as well as fungal cells, including yeast, and bacterial cells. Cells can be primary cells or tumor cells or other types of immortal cell lines. Of course, IL-6 can be expressed in cells that do not endogenously contain IL-6. Further, a membrane-bound version of IL-6 may be exogenously expressed on the surface of cells, allowing for flow cytometry or other binding, screening, or selection methods.

[219] IL-6 activity assays may also be used to identify a modulator (antagonist or agonist) of IL-6. In these embodiments, one or more test agents are contacted to a cell expressing IL-6 or to which exogenous IL-6 has been added, and then tested for an activity of IL-6. Exemplary IL-6 activities include cell proliferation, inflammatory responses or differentiation. *See, e.g.*, Naka, T. *et al., Arthritis Res.* 4:S233-S242 (2002). In other embodiments, downstream molecular events can also be monitored to determine signaling activity. For example, IL-6 signal transduction is mediated primarily by two pathways: the JAK-STAT (Janus family tyrosine Kinase - Signal Transducer and Activator of Transcription) pathway and the Ras mitogen-activated protein kinase pathway whose activation can be easily monitored using commercially available kits and methods known to those skilled in the art. *See, e.g., See, e.g.*, Naka, T. *et al., Arthritis Res.* 4:S233-S242 (2002).

[220] In some embodiments, activity assays are also used to confirm that identified antagonist monomers or multimers (i.e., that prevent binding of IL-6 to IL-6R) lack

agonist activity (i.e., that they do not activate IL-6R in the absence of IL-6 or another agonist).

[221] Agents that are initially identified by any of the foregoing screening methods can be further tested to validate the apparent activity. Such studies may be conducted with suitable animal models. The basic format of such methods involves administering a lead compound identified during an initial screen to an animal that serves as a model for humans and then determining if IL-6 is in fact modulated and/or the disease or condition is ameliorated. The animal models utilized in validation studies generally are mammals of any kind. Specific examples of suitable animals include, but are not limited to, primates, mice and rats.

[222] Selection of monomer domains that bind IL-6 from a library of domains can be accomplished by a variety of procedures. For example, one method of identifying monomer domains which have a desired property (e.g., binding IL-6 or IgG) involves translating a plurality of nucleic acids, where each nucleic acid encodes a monomer domain, screening the polypeptides encoded by the plurality of nucleic acids, and identifying those monomer domains that, e.g., bind to a desired ligand or mixture of ligands, thereby producing a selected monomer domain. The monomer domains expressed by each of the nucleic acids can be tested for their ability to bind to the ligand by methods known in the art (i.e. panning, affinity chromatography, FACS analysis).

[223] As mentioned above, selection of monomer domains can be based on binding to a ligand such as IL-6, or a fragment thereof or other target molecule (e.g., lipid, carbohydrate, nucleic acid and the like). Other molecules can optionally be included in the methods along with the target, e.g., ions such as Ca^{+2}.

[224] When a monomer domain of the invention is selected based on its ability to bind to a ligand, the selection basis can include selection based on a slow dissociation rate, which is usually predictive of high affinity. The valency of the ligand can also be varied to control the average binding affinity of selected monomer domains. The ligand can be bound to a surface or substrate at varying densities, such as by including a competitor compound, by dilution, or by other method known to those in the art. High density (valency) of predetermined ligand can be used to enrich for monomer domains that have relatively low affinity, whereas a low density (valency) can preferentially enrich for higher affinity monomer domains.

[225] A variety of reporting display vectors or systems can be used to express nucleic acids encoding the monomer domains and/or multimers of the present

invention and to test for a desired activity. For example, a phage display system is a system in which monomer domains are expressed as fusion proteins on the phage surface (Pharmacia, Milwaukee Wis.). Phage display can involve the presentation of a polypeptide sequence encoding monomer domains on the surface of a filamentous bacteriophage, typically as a fusion with a bacteriophage coat protein.

[226] Generally in these methods, each phage particle or cell serves as an individual library member displaying a single species of displayed polypeptide in addition to the natural phage or cell protein sequences. The nucleic acids are cloned into the phage DNA at a site which results in the transcription of a fusion protein, a portion of which is encoded by the plurality of the nucleic acids. The phage containing a nucleic acid molecule undergoes replication and transcription in the cell. The leader sequence of the fusion protein directs the transport of the fusion protein to the tip of the phage particle. Thus, the fusion protein that is partially encoded by the nucleic acid is displayed on the phage particle for detection and selection by the methods described above and below. For example, the phage library can be incubated with a predetermined ligand such as IL-6 or a fragment thereof, so that phage particles which present a fusion protein sequence that binds to the ligand can be differentially partitioned from those that do not present polypeptide sequences that bind to the predetermined ligand. For example, the separation can be provided by immobilizing the predetermined ligand. The phage particles (i.e., library members) which are bound to the immobilized ligand are then recovered and replicated to amplify the selected phage subpopulation for a subsequent round of affinity enrichment and phage replication. After several rounds of affinity enrichment and phage replication, the phage library members that are thus selected are isolated and the nucleotide sequence encoding the displayed polypeptide sequence is determined, thereby identifying the sequence(s) of polypeptides that bind to the predetermined ligand. Such methods are further described in PCT patent publication Nos. 91/17271, 91/18980, and 91/19818 and 93/08278.

[227] Examples of other display systems include ribosome displays, a nucleotide-linked display (*see, e.g.*, U.S. Patent Nos. 6,281,344; 6,194,550, 6,207,446, 6,214,553, and 6,258,558), polysome display, cell surface displays and the like. The cell surface displays include a variety of cells, e.g., E. coli, yeast and/or mammalian cells. When a cell is used as a display, the nucleic acids, e.g., obtained by PCR amplification followed by digestion, are introduced into the cell and translated. Optionally, polypeptides encoding the monomer domains or the multimers of the present invention can be introduced, e.g., by injection, into the cell.

[228] The monomer and multimer libraries of the invention can be screened for a desired property such as binding of a desired ligand (e.g., IL-6) or mixture of ligands. For example, members of the library of monomer domains can be displayed and prescreened for binding to a known or unknown ligand or a mixture of ligands. The monomer domain sequences can then be mutagenized(e.g., recombined, chemically altered, etc.) or otherwise altered and the new monomer domains can be screened again for binding to the ligand or the mixture of ligands with an improved affinity. The selected monomer domains can be combined or joined to form multimers, which can then be screened for an improved affinity or avidity or altered specificity for the ligand or the mixture of ligands. Altered specificity can mean that the specificity is broadened, e.g., binding of multiple related ligands, or optionally, altered specificity can mean that the specificity is narrowed, e.g., binding within a specific region of a ligand. Those of skill in the art will recognize that there are a number of methods available to calculate avidity. *See, e.g.*, Mammen *et al.*, *Angew Chem Int. Ed.* 37:2754-2794 (1998); Muller *et al.*, *Anal. Biochem.* 261:149-158 (1998).

[229] Those of skill in the art will recognize that the steps of generating variation and screening for a desired property can be repeated (i.e., performed recursively) to optimize results. For example, in a phage display library or other like format, a first screening of a library can be performed at relatively lower stringency, thereby selected as many particles associated with a target molecule as possible. The selected particles can then be isolated and the polynucleotides encoding the monomer or multimer can be isolated from the particles. Additional variations can then be generated from these sequences and subsequently screened at higher affinity.

[230] All the compositions of the present invention, e.g., monomer domains as well as multimers and libraries thereof can be optionally bound to a matrix of an affinity material. Examples of affinity material include beads, a column, a solid support, a microarray, other pools of reagent-supports, and the like.

[231] When multimers capable of binding relatively large targets are desired, they can be generated by a "walking" selection method. This method is carried out by providing a library of monomer domains and screening the library of monomer domains for affinity to a first target molecule. Once at least one monomer that binds to the target is identified, that monomer is covalently linked to a new library or each remaining member of the original library of monomer domains. This new library of multimers (dimers) is then screened for multimers that bind to the target with an increased affinity, and a multimer that binds to the target with an increased affinity can be identified. The "walking" monomer

selection method provides a way to assemble a multimer that is composed of monomers that can act additively or even synergistically with each other given the restraints of linker length. This walking technique is very useful when selecting for and assembling multimers that are able to bind large target proteins with high affinity. The walking method can be repeated to add more monomers thereby resulting in a multimer comprising 2, 3, 4, 5, 6, 7, 8 or more monomers linked together.

[232] In some embodiments, the selected multimer comprises more than two domains. Such multimers can be generated in a step fashion, e.g., where the addition of each new domain is tested individually and the effect of the domains is tested in a sequential fashion. *See, e.g.*, Figure 5. In an alternate embodiment, domains are linked to form multimers comprising more than two domains and selected for binding without prior knowledge of how smaller multimers, or alternatively, how each domain, bind.

[233] The methods of the present invention also include methods of evolving monomers or multimers. Intra-domain recombination can be introduced into monomers across the entire monomer or by taking portions of different monomers to form new recombined units. Interdomain recombination (e.g., recombining different monomers into or between multimers) or recombination of modules (e.g., multiple monomers within a multimer) may be achieved. Inter-library recombination is also contemplated.

[234] Methods for evolving monomers or multimers can comprise, e.g., any or all of the following steps: providing a plurality of different nucleic acids, where each nucleic acid encoding a monomer domain; translating the plurality of different nucleic acids, which provides a plurality of different monomer domains; screening the plurality of different monomer domains for binding of the desired ligand (e.g., IL-6) or mixture of ligands; identifying members of the plurality of different monomer domains that bind the desired ligand or mixture of ligands, which provides selected monomer domains; joining the selected monomer domains with at least one linker to generate at least one multimer, wherein the at least one multimer comprises at least two of the selected monomer domains and the at least one linker; and, screening the at least one multimer for an improved affinity or avidity or altered specificity for the desired ligand or mixture of ligands as compared to the selected monomer domains.

[235] Variation can be introduced into either monomers or multimers. An example of improving monomers includes intra-domain recombination in which two or more (e.g., three, four, five, or more) portions of the monomer are amplified separately under conditions to introduce variation (for example by shuffling or other recombination method) in

the resulting amplification products, thereby synthesizing a library of variants for different portions of the monomer. By locating the 5' ends of the middle primers in a "middle" or 'overlap' sequence that both of the PCR fragments have in common, the resulting "left" side and "right" side libraries may be combined by overlap PCR to generate novel variants of the original pool of monomers. These new variants may then be screened for desired properties, e.g., panned against a target or screened for a functional effect. The "middle" primer(s) may be selected to correspond to any segment of the monomer, and will typically be based on the scaffold or one or more consensus amino acids within the monomer (e.g., cysteines such as those found in A domains).

[236] Similarly, multimers may be created by introducing variation at the monomer level and then recombining monomer variant libraries. On a larger scale, multimers (single or pools) with desired properties may be recombined to form longer multimers. In some cases variation is introduced (typically synthetically) into the monomers or into the linkers to form libraries. This may be achieved, e.g., with two different multimers that bind to two different targets, thereby eventually selecting a multimer with a portion that binds to one target and a portion that binds a second target.

[237] Additional variation can be introduced by inserting linkers of different length and composition between domains. This allows for the selection of optimal linkers between domains. In some embodiments, optimal length and composition of linkers will allow for optimal binding of domains. In some embodiments, the domains with a particular binding affinity(s) are linked via different linkers and optimal linkers are selected in a binding assay. For example, domains are selected for desired binding properties and then formed into a library comprising a variety of linkers. The library can then be screened to identify optimal linkers. Alternatively, multimer libraries can be formed where the effect of domain or linker on target molecule binding is not known.

[238] Methods of the present invention also include generating one or more selected multimers by providing a plurality of monomer domains. The plurality of monomer domains is screened for binding of a desired ligand or mixture of ligands. Members of the plurality of domains that bind the desired ligand or mixture of ligands are identified, thereby providing domains with a desired affinity. The identified domains are joined with at least one linker to generate the multimers, wherein each multimer comprises at least two of the selected domains and the at least one linker; and, the multimers are screened for an improved affinity or avidity or altered specificity for the desired ligand or mixture of ligands as compared to the selected domains, thereby identifying the one or more selected multimers.

[239] Multimer libraries may be generated, in some embodiments, by combining two or more libraries or monomers or multimers in a recombinase-based approach, where each library member comprises as recombination site (e.g., a lox site). A larger pool of molecularly diverse library members in principle harbor more variants with desired properties, such as higher target-binding affinities and functional activities. When libraries are constructed in phage vectors, which may be transformed into *E. coli*, library size (10^9 - 10^{10}) is limited by the transformation efficiency of *E. coli*. A recombinase/recombination site system (e.g., the *Cre-lox*P system) and *in vivo* recombination can be exploited to generate libraries that are not limited in size by the transformation efficiency of *E. coli*.

[240] For example, the Cre-loxP system may be used to generate dimer libraries with 10^{10}, 10^{11}, 10^{12}, 10^{13}, or greater diversity. In some embodiments, *E. coli* as a host for one naïve monomer library and a filamentous phage that carries a second naïve monomer library are used. The library size in this case is limited only by the number of infective phage (carrying one library) and the number of infectible *E. coli* cells (carrying the other library). For example, infecting 10^{12} *E. coli* cells (1L at OD600=1) with $>10^{12}$ phage could produce as many as 10^{12} dimer combinations.

[241] Selection of multimers can be accomplished using a variety of techniques including those mentioned above for identifying monomer domains. Other selection methods include, e.g., a selection based on an improved affinity or avidity or altered specificity for the ligand compared to selected monomer domains. For example, a selection can be based on selective binding to specific cell types, or to a set of related cells or protein types (e.g., different virus serotypes). Optimization of the property selected for, e.g., avidity of a ligand, can then be achieved by recombining the domains, as well as manipulating amino acid sequence of the individual monomer domains or the linker domain or the nucleotide sequence encoding such domains, as mentioned in the present invention.

[242] One method for identifying multimers can be accomplished by displaying the multimers. As with the monomer domains, the multimers are optionally expressed or displayed on a variety of display systems, e.g., phage display, ribosome display, polysome display, nucleotide-linked display (*see, e.g.*, U.S. Patent Nos. 6,281,344; 6,194,550, 6,207,446, 6,214,553, and 6,258,558) and/or cell surface display, as described above. Cell surface displays can include but are not limited to *E. coli*, yeast or mammalian cells. In addition, display libraries of multimers with multiple binding sites can be panned for avidity or affinity or altered specificity for a ligand or for multiple ligands.

[243] Monomers or multimers can be screened for target binding activity in yeast cells using a two-hybrid screening assay. In this type of screen the monomer or multimer library to be screened is cloned into a vector that directs the formation of a fusion protein between each monomer or multimer of the library and a yeast transcriptional activator fragment (i.e., Gal4). Sequences encoding the "target" protein are cloned into a vector that results in the production of a fusion protein between the target and the remainder of the Gal4 protein (the DNA binding domain). A third plasmid contains a reporter gene downstream of the DNA sequence of the Gal4 binding site. A monomer that can bind to the target protein brings with it the Gal4 activation domain, thus reconstituting a functional Gal4 protein. This functional Gal4 protein bound to the binding site upstream of the reporter gene results in the expression of the reporter gene and selection of the monomer or multimer as a target binding protein. (see Chien et.al. (1991) *Proc. Natl. Acad. Sci.* (USA) 88:9578; Fields S. and Song O. (1989) *Nature* 340: 245) Using a two-hybrid system for library screening is further described in U.S. Patent No. 5,811,238 (see also Silver S.C. and Hunt S.W. (1993) *Mol. Biol. Rep.* 17:155; Durfee et al. (1993) *Genes Devel.* 7:555; Yang et al. (1992) *Science* 257:680; Luban et al. (1993) *Cell* 73:1067; Hardy et al. (1992) *Genes Devel.* 6:801; Bartel et al. (1993) *Biotechniques* 14:920; and Vojtek et al. (1993) *Cell* 74:205). Another useful screening system for carrying out the present invention is the *E.coli*/BCCP interactive screening system (Germino et al. (1993) *Proc. Nat. Acad. Sci.* (U.S.A.) 90:993; Guarente L. (1993) *Proc. Nat. Acad. Sci.* (U.S.A.) 90:1639).

[244] Other variations include the use of multiple binding compounds, such that monomer domains, multimers or libraries of these molecules can be simultaneously screened for a multiplicity of ligands or compounds that have different binding specificity. Multiple predetermined ligands or compounds can be concomitantly screened in a single library, or sequential screening against a number of monomer domains or multimers. In one variation, multiple ligands or compounds, each encoded on a separate bead (or subset of beads), can be mixed and incubated with monomer domains, multimers or libraries of these molecules under suitable binding conditions. The collection of beads, comprising multiple ligands or compounds, can then be used to isolate, by affinity selection, selected monomer domains, selected multimers or library members. Generally, subsequent affinity screening rounds can include the same mixture of beads, subsets thereof, or beads containing only one or two individual ligands or compounds. This approach affords efficient screening, and is compatible with laboratory automation, batch processing, and high throughput screening methods.

[245] In another embodiment, multimers can be simultaneously screened for the ability to bind multiple ligands, wherein each ligand comprises a different label. For example, each ligand can be labeled with a different fluorescent label, contacted simultaneously with a multimer or multimer library. Multimers with the desired affinity are then identified (e.g., by FACS sorting) based on the presence of the labels linked to the desired labels.

[246] Libraries of either monomer domains or multimers (referred in the following discussion for convenience as "affinity agents") can be screened (i.e., panned) simultaneously against multiple ligands in a number of different formats. For example, multiple ligands can be screened in a simple mixture, in an array, displayed on a cell or tissue (e.g., a cell or tissue provides numerous molecules that can be bound by the monomer domains or multimers of the invention), and/or immobilized. The libraries of affinity agents can optionally be displayed on yeast or phage display systems. Similarly, if desired, the ligands (e.g., encoded in a cDNA library) can be displayed in a yeast or phage display system.

[247] Initially, the affinity agent library is panned against the multiple ligands. Optionally, the resulting "hits" are panned against the ligands one or more times to enrich the resulting population of affinity agents.

[248] If desired, the identity of the individual affinity agents and/or ligands can be determined. In some embodiments, affinity agents are displayed on phage. Affinity agents identified as binding in the initial screen are divided into a first and second portion. The first portion is infected into bacteria, resulting in either plaques or bacterial colonies, depending on the type of phage used. The expressed phage are immobilized and then probed with ligands displayed in phage selected as described below.

[249] The second portion are coupled to beads or otherwise immobilized and a phage display library containing at least some of the ligands in the original mixture is contacted to the immobilized second portion. Phage that bind to the second portion are subsequently eluted and contacted to the immobilized phage described in the paragraph above. Phage-phage interactions are detected (e.g., using a monoclonal antibody specific for the ligand-expressing phage) and the resulting phage polynucleotides can be isolated.

[250] In some embodiments, the identity of an affinity agent-ligand pair is determined. For example, when both the affinity agent and the ligand are displayed on a phage or yeast, the DNA from the pair can be isolated and sequenced. In some embodiments, polynucleotides specific for the ligand and affinity agent are amplified. Amplification

primers for each reaction can include 5' sequences that are complementary such that the resulting amplification products are fused, thereby forming a hybrid polynucleotide comprising a polynucleotide encoding at least a portion of the affinity agent and at least a portion of the ligand. The resulting hybrid can be used to probe affinity agent or ligand (e.g., cDNA-encoded) polynucleotide libraries to identify both affinity agent and ligand.

[251] The above-described methods can be readily combined with "walking" to simultaneously generate and identify multiple multimers, each of which bind to a ligand in a mixture of ligands. In these embodiments, a first library of affinity agents (monomer domains or multimers) are panned against multiple ligands and the eluted affinity agents are linked to the first or a second library of affinity agents to form a library of multimeric affinity agents (e.g., comprising 2, 3, 4, 5, 6, 7, 8, 9, or more monomer), which are subsequently panned against the multiple ligands. This method can be repeated to continue to generate larger multimeric affinity agents. Increasing the number of monomer domains may result in increased affinity and avidity for a particular target. For example, the inventors have found that trimers of monomer domains that bind CD28 have a higher affinity than dimmers, which in turn have a higher affinity than single CD28-binding monomer domains alone. Of course, at each stage, the panning is optionally repeated to enrich for significant binders. In some cases, walking will be facilitated by inserting recombination sites (e.g., lox sites) at the ends of monomers and recombining monomer libraries by a recombinase-mediated event.

[252] The selected multimers of the above methods can be further manipulated, e.g., by recombining or shuffling the selected multimers (recombination can occur between or within multimers or both), mutating the selected multimers, and the like. This results in altered multimers which then can be screened and selected for members that have an enhanced property compared to the selected multimer, thereby producing selected altered multimers.

[253] In view of the description herein, it is clear that the following process may be followed. Naturally or non-naturally occurring monomer domains may be recombined or variants may be formed. Optionally the domains initially or later are selected for those sequences that are less likely to be immunogenic in the host for which they are intended. Optionally, a phage library comprising the recombined domains is panned for a desired affinity. Monomer domains or multimers expressed by the phage may be screened for IC_{50} for a target. Hetero- or homo-meric multimers may be selected. The selected polypeptides may be selected for their affinity to any target, including, e.g., hetero- or homo-multimeric targets.

[254] Linkers, multimers or selected multimers produced by the methods indicated above and below are features of the present invention. Libraries comprising multimers, e.g, a library comprising about 100, 250, 500 or more members produced by the methods of the present invention or selected by the methods of the present invention are provided. In some embodiments, one or more cell comprising members of the libraries, are also included. Libraries of the recombinant polypeptides are also a feature of the present invention, e.g., a library comprising about 100, 250, 500 or more different recombinant polypetides.

[255] Compositions of the present invention can be bound to a matrix of an affinity material, e.g., the recombinant polypeptides. Examples of affinity material include, e.g., beads, a column, a solid support, and/or the like.

VII. THERAPEUTIC AND PROPHYLACTIC TREATMENT METHODS

[256] The present invention also includes methods of therapeutically or prophylactically treating a disease or disorder by administering *in vivo* or *ex vivo* one or more nucleic acids or polypeptides of the invention described above (or compositions comprising a pharmaceutically acceptable excipient and one or more such nucleic acids or polypeptides) to a subject, including, *e.g.,* a mammal, including a human, primate, mouse, pig, cow, goat, rabbit, rat, guinea pig, hamster, horse, sheep; or a non-mammalian vertebrate such as a bird (*e.g.*, a chicken or duck), fish, or invertebrate.

[257] IL-6 antagonists, including polypeptides comprising IL-6-binding monomer domains or multimers of the invention are useful in treatment of a number of human diseases. Exemplary IL-6-binding domains include the specific sequences disclosed herein, including, but not limited to:

Family 1 motif (SEQ ID NO:27):
CX[AS][DN]EFXCXXGXC[ILV]XXAXXCDGV[DN]DCXDNSDEXXC[AST]XX[AST]

Family 2 motif (SEQ ID NO:28):
CXXXQFXC[DN]NGXCVPXXWXCDG[DE][DN]DCXD[DN]SDEXXCXXP[AST]

Family 3 motif (SEQ ID NO:29):
CXXXQFQCXSG[W,Y]C[I,L,F]XXXW[V,L]CDGXNDCEDXSDEXXC

Family 4 motif (SEQ ID NO:30):
CXXX[EQ]FXCXXGXC[ILV]XXXXXCDGX[DN]DCXDXSDEXX(XX)C

Family 5 motif (SEQ ID NO:31):
CX[A,P]X[E,Q]FXCXXGQC[I,L]PXXXXCDGX[D,N]DCXDXSDEXXC

Family 6 motif (SEQ ID NO:32):
CXXX[EQ]FXCX(X)SXXC[ILV]XX[DN]W[ILV]CDGVNDCXDXSDEXGCXXXT

Family 7 motif (SEQ ID NO:33):
CL[AS]X[EQ]FXCXNGXCIP[KQR]XWVCDGDNDC[ILV]DXSDE(P)XXCX[EKQ]XT.

[258] Exemplary human diseases for which antagonism of IL-6 is ameliorative include, e.g.,:

Autoimmune and inflammatory diseases (such as psoriasis, psoriatic arthritis, rheumatoid arthritis, juvenile rheumatoid arthritis, Still's Disease, Crohn's Disease, ulcerative colitis, Systemic Lupus Erythematosus, Lupus Nephritis, Multiple Sclerosis, Osteoarthritis, Castleman's Disease, sepsis and Guillan Barre Syndrome),

Cancer (such as, Multiple Myeloma, Renal Cell Carcinoma, Breast Cancer, Other solid tumors (colon, pancreatic, gastric, etc.), Lymphoma, Leukemia)

Metabolic/endocrine diseases (such as Type II Diabetes (insulin resistance); Osteoporosis; Dyslipidemias (increases high density lipoprotein, HDL, as part of preventing and/or treating artherosclerosis, coronary artery disease, etc.), Anti-Aging, liver disorders (e.g., liver fibrosis));

Cardiovascular disease (such as Coronary Artery Disease);

Pulmonary disease (such as Chronic Obstructive Pulmonary Disease, Interstitial Pulmonary Fibrosis, Asthma); and

Neurological diseases (such as Parkinson's Disease, Alzheimer's Disease, Multiple Sclerosis, Stroke).

[259] In some aspects, a monomer or multimer of the invention is administered to an individual with at least one other molecule used for treatment of one of the above-listed ailments. The other molecule can be prepared in a formulation with the monomer or multimer or can be administered simultaneously or before or after the monomer or multimer is administered.

[260] In some embodiments, a polypeptide comprising a monomer or multimer of the invention that binds IL-6 is used to increase high density lipids (HDLs) in an individual. The individual may be one who has a family history of heart disease or otherwise is at risk from ill health effects of high cholesterol, high LDL levels or low HDL levels. IL-6 is an important mediator of lipid metabolism. This is evidenced by the fact that cholesterol levels go up in the IL-6 knockout mouse, but in IL-6 transgenic mice the cholesterol levels go down. Human anti-IL-6 antibodies cause the total cholesterol level, LDL level, and triglyceride level to go up significantly (P=<0.001). HDL goes up significantly as well (P=<.001). LDL goes up 12-23% in dose-dependent manner with humanized anti-IL-6R antibody (MRA). HDL goes up ~26% in dose-dependent manner with same anti-IL-6R

antibody. *See, e.g.*, Nishimoto *et al, Journal of Rheumatology* (2003); Nishimoto *et al, Arthritis & Rheumatism* (2004).

[261] In some embodiments, a polypeptide comprising a monomer or multimer of the invention that binds IL-6 is administered in combination with a drug that reduces LDL levels. Statins, resins (cholestyramine, colestipol, colesevalam) and nicotinic acid are examples of a drugs that reduces LDL levels. Exemplary statins include, but are not limited to, atorvastatin, fluvastatin, lovastatin, pravastatin, rosuvastatin, and simvastatin.

[262] In one aspect of the invention, in *ex vivo* methods, one or more cells or a population of cells of interest of the subject (*e.g.*, tumor cells, tumor tissue sample, organ cells, blood cells, cells of the skin, lung, heart, muscle, brain, mucosae, liver, intestine, spleen, stomach, lymphatic system, cervix, vagina, prostate, mouth, tongue, *etc.*) are obtained or removed from the subject and contacted with an amount of a selected monomer domain and/or multimer of the invention that is effective in prophylactically or therapeutically treating the disease, disorder, or other condition. The contacted cells are then returned or delivered to the subject to the site from which they were obtained or to another site (*e.g.*, including those defined above) of interest in the subject to be treated. If desired, the contacted cells can be grafted onto a tissue, organ, or system site (including all described above) of interest in the subject using standard and well-known grafting techniques or, *e.g.*, delivered to the blood or lymph system using standard delivery or transfusion techniques.

[263] The invention also provides *in vivo* methods in which one or more cells or a population of cells of interest of the subject are contacted directly or indirectly with an amount of a selected monomer domain and/or multimer of the invention effective in prophylactically or therapeutically treating the disease, disorder, or other condition. In direct contact/administration formats, the selected monomer domain and/or multimer is typically administered or transferred directly to the cells to be treated or to the tissue site of interest (*e.g.*, tumor cells, tumor tissue sample, organ cells, blood cells, cells of the skin, lung, heart, muscle, brain, mucosae, liver, intestine, spleen, stomach, lymphatic system, cervix, vagina, prostate, mouth, tongue, *etc.*) by any of a variety of formats, including topical administration, injection (*e.g.*, by using a needle or syringe), or vaccine or gene gun delivery, pushing into a tissue, organ, or skin site. The selected monomer domain and/or multimer can be delivered, for example, intramuscularly, intradermally, subdermally, subcutaneously, orally, intraperitoneally, intrathecally, intravenously, or placed within a cavity of the body (including, *e.g.*, during surgery), or by inhalation or vaginal or rectal administration. Exemplary delivery include, e.g., oral delivery (buccal or intestinal uptake/delivery), home

injection (e.g., formulated in about a 1.4 ml volume or less), transdermal uptake/delivery, subcutaneously implantable pump, pulmonary uptake/delivery (systemic or non-systemic), blood-brain-barrier uptake/delivery, and/or nasal uptake/delivery.

[264] Design approaches for oral delivery include, e.g., including in the monomer or multimer a binding affinity for mucosal peptide transporters such as PepT1, PepT2, HPT1, ABC-family; addition of a translocation domain (PE38 or PE40 from Pseudomonas exotoxin A) to the monomer or multimer; or including a binding affinity for blood-brain barrier transporters such as Transferrin receptor.

[265] The four primary mechanisms for transepithelial delivery are:

- Transcellular: through the cell by diffusion: mostly used by small molecules.
- Paracellular: around the cells through desmosomes/tight junctions: mostly used by hydrophilic peptides
- Transcytosis: through the cells inside vesicles by pinocytosis
- Carrier- or Transporter-mediated, for molecules that bind to the transporter/carrier

[266] In another approach, the monomer or multimer is designed to bind to receptors on the cell surface or inside vesicles that are actively transported across the cell to the serum, carrying the drug along. Such receptors are called Transporters. One such receptor is the neonatal immunoglobulin Fc receptor (FcRn) which can be used for pulmonary and intestinal delivery to the serum. Another receptor is the Transferrin receptor, which can transport across the blood-brain barrier into the cerebrospinal fluid. Alternatively, the monomer or multimer is designed to bind to the neonatal Fc receptor for pulmonary and intestinal transport. This method can involve engineering the final drug product for pH sensitive release.

[267] In view of the above, several strategies are possible. Monomers or multimers that bind known transporters (e.g., PepT1, HPT-1, PepT2, ABC family, TAP1, TAP2, MDR) can be identified. Other useful binders can be identified by identifying monomers or multimers that bind to Caco2 cells (e.g., using an uptake blocking assay using 14C-Cephalexin uptake by Caco2 cells). See, e.g., Blanchette, J., *et al. Biomedicine and Pharmacotherapy* 58:142-151 (2004); Low, S.C., *et al. Human Reproduction* - advance access April 7 2005; Lee, V. *J. Natl Cancer Institute*, Monographs 29 (2001); Rubas, W., *et al. J. Pharm. Sci.* 85:165-169 (1996); Yang, C.Y., *et al. Pharmaceutical Research* 16:1331-1343 (1999); Dietrich, C.G., *et al. Gut* 52:1788-1799 (2003); Cleland, J.L., *et al. Current Opinion in Biotechnology* 12:212-219 (2001); Nielsen, C.U. and Brodin, B. *Current Drug Targets* 4:373-388 (2004); Kunta, J.R., *et al. Current Drug Metabolism* 5:109-124 (2004).

[268] In another approach, a phage library of multimers can be applied to the oral cavity of a mammal and phage or DNA from the phage can be recovered in the serum of the animal to enrich for sequences that are good delivery agents. Natural domains or peptides that are known to be orally transported can also be fused to monomers or multimers of the invention. In yet another alternative, phage displayed domains are selected for resistance to the appropriate body fluid (nasal, intestinal, etc) or to purified proteases.

[269] In *in vivo* indirect contact/administration formats, the selected monomer domain and/or multimer is typically administered or transferred indirectly to the cells to be treated or to the tissue site of interest, including those described above (such as, *e.g.*, skin cells, organ systems, lymphatic system, or blood cell system, *etc.*), by contacting or administering the polypeptide of the invention directly to one or more cells or population of cells from which treatment can be facilitated. For example, tumor cells within the body of the subject can be treated by contacting cells of the blood or lymphatic system, skin, or an organ with a sufficient amount of the selected monomer domain and/or multimer such that delivery of the selected monomer domain and/or multimer to the site of interest (*e.g.*, tissue, organ, or cells of interest or blood or lymphatic system within the body) occurs and effective prophylactic or therapeutic treatment results. Such contact, administration, or transfer is typically made by using one or more of the routes or modes of administration described above.

[270] In another aspect, the invention provides *ex vivo* methods in which one or more cells of interest or a population of cells of interest of the subject (*e.g.*, tumor cells, tumor tissue sample, organ cells, blood cells, cells of the skin, lung, heart, muscle, brain, mucosae, liver, intestine, spleen, stomach, lymphatic system, cervix, vagina, prostate, mouth, tongue, *etc.*) are obtained or removed from the subject and transformed by contacting said one or more cells or population of cells with a polynucleotide construct comprising a nucleic acid sequence of the invention that encodes a biologically active polypeptide of interest (*e.g.*, a selected monomer domain and/or multimer) that is effective in prophylactically or therapeutically treating the disease, disorder, or other condition. The one or more cells or population of cells is contacted with a sufficient amount of the polynucleotide construct and a promoter controlling expression of said nucleic acid sequence such that uptake of the polynucleotide construct (and promoter) into the cell(s) occurs and sufficient expression of the target nucleic acid sequence of the invention results to produce an amount of the biologically active polypeptide, encoding a selected monomer domain and/or multimer,

effective to prophylactically or therapeutically treat the disease, disorder, or condition. The polynucleotide construct can include a promoter sequence (*e.g.*, CMV promoter sequence) that controls expression of the nucleic acid sequence of the invention and/or, if desired, one or more additional nucleotide sequences encoding at least one or more of another polypeptide of the invention, a cytokine, adjuvant, or co-stimulatory molecule, or other polypeptide of interest.

[271] Following transfection, the transformed cells are returned, delivered, or transferred to the subject to the tissue site or system from which they were obtained or to another site (*e.g.*, tumor cells, tumor tissue sample, organ cells, blood cells, cells of the skin, lung, heart, muscle, brain, mucosae, liver, intestine, spleen, stomach, lymphatic system, cervix, vagina, prostate, mouth, tongue, *etc.*) to be treated in the subject. If desired, the cells can be grafted onto a tissue, skin, organ, or body system of interest in the subject using standard and well-known grafting techniques or delivered to the blood or lymphatic system using standard delivery or transfusion techniques. Such delivery, administration, or transfer of transformed cells is typically made by using one or more of the routes or modes of administration described above. Expression of the target nucleic acid occurs naturally or can be induced (as described in greater detail below) and an amount of the encoded polypeptide is expressed sufficient and effective to treat the disease or condition at the site or tissue system.

[272] In another aspect, the invention provides *in vivo* methods in which one or more cells of interest or a population of cells of the subject (*e.g.*, including those cells and cells systems and subjects described above) are transformed in the body of the subject by contacting the cell(s) or population of cells with (or administering or transferring to the cell(s) or population of cells using one or more of the routes or modes of administration described above) a polynucleotide construct comprising a nucleic acid sequence of the invention that encodes a biologically active polypeptide of interest (*e.g.*, a selected monomer domain and/or multimer) that is effective in prophylactically or therapeutically treating the disease, disorder, or other condition.

[273] The polynucleotide construct can be directly administered or transferred to cell(s) suffering from the disease or disorder (*e.g.*, by direct contact using one or more of the routes or modes of administration described above). Alternatively, the polynucleotide construct can be indirectly administered or transferred to cell(s) suffering from the disease or disorder by first directly contacting non-diseased cell(s) or other diseased cells using one or more of the routes or modes of administration described above with a sufficient amount of the polynucleotide construct comprising the nucleic acid sequence

encoding the biologically active polypeptide, and a promoter controlling expression of the nucleic acid sequence, such that uptake of the polynucleotide construct (and promoter) into the cell(s) occurs and sufficient expression of the nucleic acid sequence of the invention results to produce an amount of the biologically active polypeptide effective to prophylactically or therapeutically treat the disease or disorder, and whereby the polynucleotide construct or the resulting expressed polypeptide is transferred naturally or automatically from the initial delivery site, system, tissue or organ of the subject's body to the diseased site, tissue, organ or system of the subject's body (*e.g.*, via the blood or lymphatic system). Expression of the target nucleic acid occurs naturally or can be induced (as described in greater detail below) such that an amount of expressed polypeptide is sufficient and effective to treat the disease or condition at the site or tissue system. The polynucleotide construct can include a promoter sequence (*e.g.*, CMV promoter sequence) that controls expression of the nucleic acid sequence and/or, if desired, one or more additional nucleotide sequences encoding at least one or more of another polypeptide of the invention, a cytokine, adjuvant, or co-stimulatory molecule, or other polypeptide of interest.

[274] In each of the *in vivo* and *ex vivo* treatment methods as described above, a composition comprising an excipient and the polypeptide or nucleic acid of the invention can be administered or delivered. In one aspect, a composition comprising a pharmaceutically acceptable excipient and a polypeptide or nucleic acid of the invention is administered or delivered to the subject as described above in an amount effective to treat the disease or disorder.

[275] In another aspect, in each *in vivo* and *ex vivo* treatment method described above, the amount of polynucleotide administered to the cell(s) or subject can be an amount such that uptake of said polynucleotide into one or more cells of the subject occurs and sufficient expression of said nucleic acid sequence results to produce an amount of a biologically active polypeptide effective to enhance an immune response in the subject, including an immune response induced by an immunogen (*e.g.*, antigen). In another aspect, for each such method, the amount of polypeptide administered to cell(s) or subject can be an amount sufficient to enhance an immune response in the subject, including that induced by an immunogen (*e.g.*, antigen).

[276] In yet another aspect, in an *in vivo* or *in vivo* treatment method in which a polynucleotide construct (or composition comprising a polynucleotide construct) is used to deliver a physiologically active polypeptide to a subject, the expression of the polynucleotide construct can be induced by using an inducible on- and off-gene expression

system. Examples of such on- and off-gene expression systems include the Tet-On™ Gene Expression System and Tet-Off™ Gene Expression System (*see, e.g.*, Clontech Catalog 2000, pg. 110-111 for a detailed description of each such system), respectively. Other controllable or inducible on- and off-gene expression systems are known to those of ordinary skill in the art. With such system, expression of the target nucleic of the polynucleotide construct can be regulated in a precise, reversible, and quantitative manner. Gene expression of the target nucleic acid can be induced, for example, after the stable transfected cells containing the polynucleotide construct comprising the target nucleic acid are delivered or transferred to or made to contact the tissue site, organ or system of interest. Such systems are of particular benefit in treatment methods and formats in which it is advantageous to delay or precisely control expression of the target nucleic acid (*e.g.*, to allow time for completion of surgery and/or healing following surgery; to allow time for the polynucleotide construct comprising the target nucleic acid to reach the site, cells, system, or tissue to be treated; to allow time for the graft containing cells transformed with the construct to become incorporated into the tissue or organ onto or into which it has been spliced or attached, *etc.*).

VIII. ADDITIONAL MULTIMER USES

[277] The potential applications of multimers of the present invention are diverse and include any use where an affinity agent is desired.

[278] In some cases, a pair of monomers or multimers are selected to bind to the same target (i.e., for use in sandwich-based assays). To select a matched monomer or multimer pair, two different monomers or multimers typically are able to bind the target protein simultaneously. One approach to identify such pairs involves the following:

(1) immobilizing the phage or protein mixture that was previously selected to bind the target protein

(2) contacting the target protein to the immobilized phage or protein and washing;

(3) contacting the phage or protein mixture to the bound target and washing; and

(4) eluting the bound phage or protein without eluting the immobilized phage or protein.

[279] One use of the multimers or monomer domains of the invention is to replace antibodies or other affinity agents in detection or other affinity-based assays. Thus, in some embodiments, monomer domains or multimers are selected against the ability to bind components other than a target in a mixture. The general approach can include performing

the affinity selection under conditions that closely resemble the conditions of the assay, including mimicking the composition of a sample during the assay. Thus, a step of selection could include contacting a monomer domain or multimer to a mixture not including the target ligand and selecting against any monomer domains or multimers that bind to the mixture. Thus, the mixtures (absent the target ligand, which could be depleted using an antibody, monomer domain or multimer) representing the sample in an assay (serum, blood, tissue, cells, urine, semen, etc) can be used as a blocking agent. Such subtraction is useful, e.g., to create pharmaceutical proteins that bind to their target but not to other serum proteins or non-target tissues.

[280] For example, the invention can be used in the application for creating antagonists, where the selected monomer domains or multimers block the interaction between two proteins, e.g., IL-6 and IL-6R. IN some embodiments, monomers can be identified that prevent this interaction either by binding IL-6 or IL-6R. *See, e.g.*, Figure 9. Optionally, the invention can generate agonists. For example, multimers binding two different proteins, e.g., enzyme and substrate, can enhance protein function, including, for example, enzymatic activity and/or substrate conversion.

[281] In some embodiments, the monomer domains are used for ligand inhibition, ligand clearance or ligand stimulation. Possible ligands in these methods, include, e.g., IL-6.

[282] If inhibition of ligand binding to a receptor is desired, a monomer domain is selected that binds to the ligand (e.g., IL-6) at a portion of the ligand that contacts the ligand's receptor, or that binds to the receptor at a portion of the receptor that binds contacts the ligand, thereby preventing the ligand-receptor interaction. The monomer domains can optionally be linked to a half-life extender, if desired.

[283] Ligand clearance refers to modulating the half-life of a soluble ligand in bodily fluid. Antibodies against cytokines typically prevent the clearance of the cytokine because the serum half-life of the antibody is much longer than the serum half-life of the cytokine (2-3 weeks versus a few hours). This causes a large increase in the concentration of the cytokine in the serum. In the case of IL-6 this increase was reported to be 700-fold (van Zaanen, *et al. J. Clin. Invest.* 98, 1441-1448 (1996)). Most of the IL-6 is bound to antibody and therefore inactive. However, increased IL-6 levels can in some circumstances create safety issues since under some circumstances (such as in compartments with a low pH or high pH), and in certain tissues (like a filter that prevents Ab penetration but not IL-6) the

IL-6 may separate from the antibody and the high level of active IL-6 could become a serious safety concern.

[284] One alternative approach would be to use an inhibitor with a shorter serum halflife than antibodies, so that the increase in the total IL-6 concentration is less pronounced. The monomers and multimers of the present invention have human serum half lives of 60-180 hours (determined by allometric scaling from monkey data), much less than the 2-3 weeks halflife of many antibodies.

[285] Thus, in some embodiments, IL-6 is inhibited in such a way that the IL-6-inhibitor complex is small enough to be cleared by the kidney. The kidney is a molecular size filter with a cut-off around 45-50kD, letting only smaller molecules pass through. If the size of the cytokine-inhibitor complex can be kept under 35-40kD, efficient kidney clearance is expected. Since most cytokines are around 18-20kD in size, the inhibitor should less than 15-20 kD. The 18kD IL-6 trimer multimers of the invention meet this requirement, but scaffold proteins based on Ig-domains, which are 3-fold larger, generally do not meet this requirement.

[286] In another alternative, multimers of the invention can be designed that, when not bound to its target molecule (e.g., IL-6), have a hydrodynamic Stokes radius of a globular protein that will not be cleared by the kidney (e.g., greater than 45kD, 50kD, 55kD), but, when bound to its target molecule (e.g., IL-6), have a hydrodynamic Stokes radius of a globular protein that will be cleared by the kidney (e.g., less than 55kD, 50kD, 45kD or 40kD). Such multimers will generally comprise at least 3, 4, 5, 6, 7, 8, or more monomer domains such that the multimers, when not bound to the target molecule, form a significantly larger Stokes radius than when bound to the target molecule. Such multimers can be readily selected by, for example, determining the sedimentation coefficient of 1) a globular protein too large to clear the kidney and 2) a globular protein small enough to clear the kidney and then selecting multimers that have a similar sedimentation coefficient to 1) in the absence of target molecule but that have a similar sedimentation coefficient to 2) in the presence of target molecule.

[287] Alternatively, as discussed elsewhere herein, a multimer comprising a first monomer domain that binds to a half-life extender and a second monomer domain that binds to a portion of the ligand that does not bind to the ligand's receptor can be used to increase the half-life of the ligand.

[288] In another embodiment, a multimer comprising a first monomer domain that binds to the ligand and a second monomer domain that binds to the receptor can be used to increase the effective affinity of the ligand for the receptor.

[289] In another embodiment, multimers comprising at least two monomers that bind to receptors are used to bring two receptors into proximity by both binding the multimer, thereby activating the receptors.

[290] Further examples of potential uses of the invention include monomer domains, and multimers thereof, that are capable of drug binding (e.g., binding radionucleotides for targeting, pharmaceutical binding for half-life extension of drugs, controlled substance binding for overdose treatment and addiction therapy), immune function modulating (e.g., immunogenicity blocking by binding such receptors as CTLA-4, immunogenicity enhancing by binding such receptors as CD80,or complement activation by Fc type binding), and specialized delivery (e.g., slow release by linker cleavage, electrotransport domains, dimerization domains, or specific binding to: cell entry domains, clearance receptors such as FcR, oral delivery receptors such as plgR for trans-mucosal transport, and blood-brain transfer receptors such as transferrinR).

[291] In further embodiments, monomers or multimers can be linked to a detectable label (e.g., Cy3, Cy5, etc.) or linked to a reporter gene product (e.g., CAT, luciferase, horseradish peroxidase, alkaline phosphotase, GFP, etc.).

IX. FURTHER MANIPULATING MONOMER DOMAINS AND/OR MULTIMER NUCLEIC ACIDS AND POLYPEPTIDES

[292] As mentioned above, the polypeptide of the present invention can be altered. Descriptions of a variety of diversity generating procedures for generating modified or altered nucleic acid sequences encoding these polypeptides are described herein and the references cited therein.

[293] Another aspect of the present invention includes the cloning and expression of monomer domains, selected monomer domains, multimers and/or selected multimers coding nucleic acids. Thus, multimer domains can be synthesized as a single protein using expression systems well known in the art. General texts which describe molecular biological techniques useful herein, including the use of vectors, promoters and many other topics relevant to expressing nucleic acids such as monomer domains, selected monomer domains, multimers and/or selected multimers, include Berger and Kimmel, Guide

to Molecular Cloning Techniques, Methods in Enzymology volume 152 Academic Press, Inc., San Diego, CA (Berger); Sambrook et al., Molecular Cloning - A Laboratory Manual (2nd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989 ("Sambrook") and Current Protocols in Molecular Biology, F.M. Ausubel et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (supplemented through 1999) ("Ausubel")). Examples of techniques sufficient to direct persons of skill through *in vitro* amplification methods, useful in identifying, isolating and cloning monomer domains and multimers coding nucleic acids, including the polymerase chain reaction (PCR) the ligase chain reaction (LCR), Q-replicase amplification and other RNA polymerase mediated techniques (*e.g.*, NASBA), are found in Berger, Sambrook, and Ausubel, as well as Mullis *et al.*, (1987) U.S. Patent No. 4,683,202; *PCR Protocols A Guide to Methods and Applications* (Innis *et al.* eds) Academic Press Inc. San Diego, CA (1990) (Innis); Arnheim & Levinson (October 1, 1990) *C&EN* 36-47; *The Journal Of NIH Research* (1991) 3, 81-94; (Kwoh *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86, 1173; Guatelli *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87, 1874; Lomell *et al.* (1989) *J. Clin. Chem* 35, 1826; Landegren *et al.*, (1988) *Science* 241, 1077-1080; Van Brunt (1990) *Biotechnology* 8, 291-294; Wu and Wallace, (1989) *Gene* 4, 560; Barringer *et al.* (1990) *Gene* 89, 117, and Sooknanan and Malek (1995) *Biotechnology* 13: 563-564. Improved methods of cloning *in vitro* amplified nucleic acids are described in Wallace *et al.*, U.S. Pat. No. 5,426,039. Improved methods of amplifying large nucleic acids by PCR are summarized in Cheng *et al.* (1994) *Nature* 369: 684-685 and the references therein, in which PCR amplicons of up to 40kb are generated. One of skill will appreciate that essentially any RNA can be converted into a double stranded DNA suitable for restriction digestion, PCR expansion and sequencing using reverse transcriptase and a polymerase. *See*, Ausubel, Sambrook and Berger, *all supra.*

[294] The present invention also relates to the introduction of vectors of the invention into host cells, and the production of monomer domains, selected monomer domains, multimers and/or selected multimers of the invention by recombinant techniques. Host cells are genetically engineered (i.e., transduced, transformed or transfected) with the vectors of this invention, which can be, for example, a cloning vector or an expression vector. The vector can be, for example, in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants, or amplifying the monomer domain, selected monomer domain, multimer and/or selected multimer gene(s) of interest. The culture conditions, such as temperature, pH and the like, are those previously used with the

host cell selected for expression, and will be apparent to those skilled in the art and in the references cited herein, including, e.g., Freshney (1994) *Culture of Animal Cells, a Manual of Basic Technique*, third edition, Wiley- Liss, New York and the references cited therein.

[295] As mentioned above, the polypeptides of the invention can also be produced in non-animal cells such as plants, yeast, fungi, bacteria and the like. Indeed, as noted throughout, phage display is an especially relevant technique for producing such polypeptides. In addition to Sambrook, Berger and Ausubel, details regarding cell culture can be found in Payne *et al.* (1992) *Plant Cell and Tissue Culture in Liquid Systems* John Wiley & Sons, Inc. New York, NY; Gamborg and Phillips (eds) (1995) *Plant Cell, Tissue and Organ Culture*; Fundamental Methods Springer Lab Manual, Springer-Verlag (Berlin Heidelberg New York) and Atlas and Parks (eds) *The Handbook of Microbiological Media* (1993) CRC Press, Boca Raton, FL.

[296] The present invention also includes alterations of monomer domains, immuno-domains and/or multimers to improve pharmacological properties, to reduce immunogenicity, or to facilitate the transport of the multimer and/or monomer domain into a cell or tissue (e.g., through the blood-brain barrier, or through the skin). These types of alterations include a variety of modifications (e.g., the addition of sugar-groups or glycosylation), the addition of PEG, the addition of protein domains that bind a certain protein (e.g., HSA or other serum protein), the addition of proteins fragments or sequences that signal movement or transport into, out of and through a cell. Additional components can also be added to a multimer and/or monomer domain to manipulate the properties of the multimer and/or monomer domain. A variety of components can also be added including, e.g., a domain that binds a known receptor (e.g., a Fc-region protein domain that binds a Fc receptor), a toxin(s) or part of a toxin, a prodomain that can be optionally cleaved off to activate the multimer or monomer domain, a reporter molecule (e.g., green fluorescent protein), a component that bind a reporter molecule (such as a radionuclide for radiotherapy, biotin or avidin) or a combination of modifications.

X. ANIMAL MODELS

[297] Another aspect of the invention is the development of specific non-human animal models in which to test the immunogenicity of the monomer or multimer domains. The method of producing such non-human animal model comprises: introducing into at least some cells of a recipient non-human animal, vectors comprising genes encoding a plurality of human proteins from the same family of proteins, wherein the genes are each

operably linked to a promoter that is functional in at least some of the cells into which the vectors are introduced such that a genetically modified non-human animal is obtained that can express the plurality of human proteins from the same family of proteins.

[298] Suitable non-human animals employed in the practice of the present invention include all vertebrate animals, except humans (e.g., mouse, rat, rabbit, sheep, and the like). Typically, the plurality of members of a family of proteins includes at least two members of that family, and usually at least ten family members. In some embodiments, the plurality includes all known members of the family of proteins. Exemplary genes that can be used include those encoding monomer domains, such as, for example, members of the LDL receptor class A-domain family, the EGF-like domain family, as well as the other domain families described herein.

[299] The non-human animal models of the present invention can be used to screen for immunogenicity of a monomer or multimer domain that is derived from the same family of proteins expressed by the non-human animal model. The present invention includes the non-human animal model made in accordance with the method described above, as well as transgenic non-human animals whose somatic and germ cells contain and express DNA molecules encoding a plurality of human proteins from the same family of proteins (such as the monomer domains described herein), wherein the DNA molecules have been introduced into the transgenic non-human animal at an embryonic stage, and wherein the DNA molecules are each operably linked to a promoter in at least some of the cells in which the DNA molecules have been introduced.

[300] An example of a mouse model useful for screening LDL receptor class A-domain derived binding proteins is described as follows. Gene clusters encoding the wild type human LDL receptor class A-domain monomers are amplified from human cells using PCR. Almost all of the 200 different A-domains can be amplified with only three separate PCR amplification reactions of about 7kb each. These fragments are then used to generate transgenic mice according to the method described above. The transgenic mice will recognize the human A-domains as "self", thus mimicking the "selfness" of a human with regard to A-domains. Individual A-domain-derived monomers or multimers are tested in these mice by injecting the A-domain-derived monomers or multimers into the mice, then analyzing the immune response (or lack of response) generated. The mice are tested to determine if they have developed a mouse anti-human response (MAHR). Monomers and multimers that do not result in the generation of a MAHR are likely to be non-immunogenic when administered to humans.

[301] Historically, MAHR test in transgenic mice is used to test individual proteins in mice that are transgenic for that single protein. In contrast, the above described method provides a non-human animal model that recognizes an entire family of human proteins as "self," and that can be used to evaluate a huge number of variant proteins that each are capable of vastly varied binding activities and uses.

XI. KITS

[302] Kits comprising the components needed in the methods (typically in an unmixed form) and kit components (packaging materials, instructions for using the components and/or the methods, one or more containers (reaction tubes, columns, etc.)) for holding the components are a feature of the present invention. Kits of the present invention may contain a multimer library, or a single type of monomer or multimer, e.g., polypeptides comprising the IL-6 and/or IgG binding domains described herein. Kits can also include reagents suitable for promoting target molecule binding, such as buffers or reagents that facilitate detection, including detectably-labeled molecules. Standards for calibrating a ligand binding to a monomer domain or the like, can also be included in the kits of the invention.

[303] The present invention also provides commercially valuable binding assays and kits to practice the assays. In some of the assays of the invention, one or more ligand is employed to detect binding of a monomer domain, immuno-domains and/or multimer. Such assays are based on any known method in the art, e.g., flow cytometry, fluorescent microscopy, plasmon resonance, and the like, to detect binding of a ligand(s) to the monomer domain and/or multimer.

[304] Kits based on the assay are also provided. The kits typically include a container, and one or more ligand. The kits optionally comprise directions for performing the assays, additional detection reagents, buffers, or instructions for the use of any of these components, or the like. Alternatively, kits can include cells, vectors, (e.g., expression vectors, secretion vectors comprising a polypeptide of the invention), for the expression of a monomer domain and/or a multimer of the invention.

[305] In a further aspect, the present invention provides for the use of any composition, monomer domain, immuno-domain, multimer, cell, cell culture, apparatus, apparatus component or kit herein, for the practice of any method or assay herein, and/or for the use of any apparatus or kit to practice any assay or method herein and/or for the use of

cells, cell cultures, compositions or other features herein as a therapeutic formulation. The manufacture of all components herein as therapeutic formulations for the treatments described herein is also provided.

XII. INTEGRATED SYSTEMS

[306] The present invention provides computers, computer readable media and integrated systems comprising character strings corresponding to monomer domains, selected monomer domains, multimers and/or selected multimers and nucleic acids encoding such polypeptides. These sequences can be manipulated by in silico recombination methods, or by standard sequence alignment or word processing software.

[307] For example, different types of similarity and considerations of various stringency and character string length can be detected and recognized in the integrated systems herein. For example, many homology determination methods have been designed for comparative analysis of sequences of biopolymers, for spell checking in word processing, and for data retrieval from various databases. With an understanding of double-helix pair-wise complement interactions among 4 principal nucleobases in natural polynucleotides, models that simulate annealing of complementary homologous polynucleotide strings can also be used as a foundation of sequence alignment or other operations typically performed on the character strings corresponding to the sequences herein (*e.g.*, word-processing manipulations, construction of figures comprising sequence or subsequence character strings, output tables, etc.). An example of a software package with GOs for calculating sequence similarity is BLAST, which can be adapted to the present invention by inputting character strings corresponding to the sequences herein.

[308] BLAST is described in Altschul *et al.*, (1990) J. Mol. Biol. 215:403-410. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (available on the World Wide Web at ncbi.nlm.nih.gov). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores

are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, a cutoff of 100, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915).

[309] An additional example of a useful sequence alignment algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments. It can also plot a tree showing the clustering relationships used to create the alignment. PILEUP uses a simplification of the progressive alignment method of Feng & Doolittle, (1987) J. Mol. Evol. 35:351-360. The method used is similar to the method described by Higgins & Sharp, (1989) CABIOS 5:151-153. The program can align, e.g., up to 300 sequences of a maximum length of 5,000 letters. The multiple alignment procedure begins with the pairwise alignment of the two most similar sequences, producing a cluster of two aligned sequences. This cluster can then be aligned to the next most related sequence or cluster of aligned sequences. Two clusters of sequences can be aligned by a simple extension of the pairwise alignment of two individual sequences. The final alignment is achieved by a series of progressive, pairwise alignments. The program can also be used to plot a dendogram or tree representation of clustering relationships. The program is run by designating specific sequences and their amino acid or nucleotide coordinates for regions of sequence comparison. For example, in order to determine conserved amino acids in a monomer domain family or to compare the sequences of monomer domains in a family, the sequence of the invention, or coding nucleic acids, are aligned to provide structure-function information.

[310] In one aspect, the computer system is used to perform "in silico" sequence recombination or shuffling of character strings corresponding to the monomer domains. A variety of such methods are set forth in "Methods For Making Character Strings,

Polynucleotides & Polypeptides Having Desired Characteristics" by Selifonov and Stemmer, filed February 5, 1999 (USSN 60/118854) and "Methods For Making Character Strings, Polynucleotides & Polypeptides Having Desired Characteristics" by Selifonov and Stemmer, filed October 12, 1999 (USSN 09/416,375). In brief, genetic operators are used in genetic algorithms to change given sequences, e.g., by mimicking genetic events such as mutation, recombination, death and the like. Multi-dimensional analysis to optimize sequences can be also be performed in the computer system, e.g., as described in the '375 application.

[311] A digital system can also instruct an oligonucleotide synthesizer to synthesize oligonucleotides, e.g., used for gene reconstruction or recombination, or to order oligonucleotides from commercial sources (e.g., by printing appropriate order forms or by linking to an order form on the Internet).

[312] The digital system can also include output elements for controlling nucleic acid synthesis (*e.g.*, based upon a sequence or an alignment of a recombinant, e.g., recombined, monomer domain as herein), i.e., an integrated system of the invention optionally includes an oligonucleotide synthesizer or an oligonucleotide synthesis controller. The system can include other operations that occur downstream from an alignment or other operation performed using a character string corresponding to a sequence herein, e.g., as noted above with reference to assays.

EXAMPLES

[313] The following example is offered to illustrate, but not to limit the claimed invention.

Example 1

[314] This example describes selection of monomer domains and the creation of multimers.

[315] Starting materials for identifying monomer domains and creating multimers from the selected monomer domains and procedures can be derived from any of a variety of human and/or non-human sequences. For example, to produce a selected monomer domain with specific binding for a desired ligand or mixture of ligands, one or more monomer domain gene(s) are selected from a family of monomer domains that bind to a certain ligand. The nucleic acid sequences encoding the one or more monomer domain gene

can be obtained by PCR amplification of genomic DNA or cDNA, or optionally, can be produced synthetically using overlapping oligonucleotides.

[316] Most commonly, these sequences are then cloned into a cell surface display format (*i.e.*, bacterial, yeast, or mammalian (COS) cell surface display; phage display) for expression and screening. The recombinant sequences are transfected (transduced or transformed) into the appropriate host cell where they are expressed and displayed on the cell surface. For example, the cells can be stained with a labeled (e.g., fluorescently labeled), desired ligand. The stained cells are sorted by flow cytometry, and the selected monomer domains encoding genes are recovered (e.g., by plasmid isolation, PCR or expansion and cloning) from the positive cells. The process of staining and sorting can be repeated multiple times (e.g., using progressively decreasing concentrations of the desired ligand until a desired level of enrichment is obtained). Alternatively, any screening or detection method known in the art that can be used to identify cells that bind the desired ligand or mixture of ligands can be employed.

[317] The selected monomer domain encoding genes recovered from the desired ligand or mixture of ligands binding cells can be optionally recombined according to any of the methods described herein or in the cited references. The recombinant sequences produced in this round of diversification are then screened by the same or a different method to identify recombinant genes with improved affinity for the desired or target ligand. The diversification and selection process is optionally repeated until a desired affinity is obtained.

[318] The selected monomer domain nucleic acids selected by the methods can be joined together via a linker sequence to create multimers, e.g., by the combinatorial assembly of nucleic acid sequences encoding selected monomer domains by DNA ligation, or optionally, PCR-based, self-priming overlap reactions. The nucleic acid sequences encoding the multimers are then cloned into a cell surface display format (*i.e.*, bacterial, yeast, or mammalian (COS) cell surface display; phage display) for expression and screening. The recombinant sequences are transfected (transduced or transformed) into the appropriate host cell where they are expressed and displayed on the cell surface. For example, the cells can be stained with a labeled, e.g., fluorescently labeled, desired ligand or mixture of ligands. The stained cells are sorted by flow cytometry, and the selected multimers encoding genes are recovered (e.g., by PCR or expansion and cloning) from the positive cells. Positive cells include multimers with an improved avidity or affinity or altered specificity to the desired ligand or mixture of ligands compared to the selected monomer domain(s). The process of staining and sorting can be repeated multiple times (e.g., using progressively decreasing

concentrations of the desired ligand or mixture of ligands until a desired level of enrichment is obtained). Alternatively, any screening or detection method known in the art that can be used to identify cells that bind the desired ligand or mixture of ligands can be employed.

[319] The selected multimer encoding genes recovered from the desired ligand or mixture of ligands binding cells can be optionally recombined according to any of the methods described herein or in the cited references. The recombinant sequences produced in this round of diversification are then screened by the same or a different method to identify recombinant genes with improved avidity or affinity or altered specificity for the desired or target ligand. The diversification and selection process is optionally repeated until a desired avidity or affinity or altered specificity is obtained.

Example 2

[320] This example describes *in vivo* intra-protein recombination to generate libraries of greater diversity.

[321] A monomer-encoding plasmid vector (pCK-derived vector; see below), flanked by orthologous *loxP* sites, was recombined in a Cre-dependent manner with a phage vector via its compatible *loxP* sites. The recombinant phage vectors were detected by PCR using primers specific for the recombinant construct. DNA sequencing indicated that the correct recombinant product was generated.

Reagents and experimental procedures

[322] **pCK-*cre*-lox-Monomer-loxP**. This vector has two particularly relevant features. First, it carries the *cre* gene, encoding the site-specific DNA recombinase Cre, under the control of P_{lac}. *Cre* was PCR-amplified from p705-*cre* (from *GeneBridges*) with *cre*-specific primers that incorporated *Xba*I (5') and *Sfi*I (3') at the ends of the PCR product. This product was digested with *Xba*I and *Sfi*I and cloned into the identical sites of pCK, a *bla*$^-$, Cm^R derivative of pCK110919-HC-Bla (pACYC ori), yielding pCK-*cre*.

[323] The second feature is the naïve A domain library flanked by two orthologous *loxP* sites, *loxP*(wild-type) and *loxP*(FAS), which are required for the site-specific DNA recombination catalyzed by Cre. *See, e.g.*, Siegel, R.W., *et al.*. FEBS Letters 505:467-473 (2001). These sites rarely recombine with another. *loxP* sites were built into pCK-*cre* sequentially. 5'-phosphorylated oligonucleotides loxP(K) and loxP(K_rc), carrying loxP(WT) and *Eco*RI and *Hin*DIII-compatible overhangs to allow ligation to digested *Eco*RI

and *Hin*DIII-digested pCK, were hybridized together and ligated to pCK-*cre* in a standard ligation reaction (T4 ligase; overnight at 16C).

[324] The resulting plasmid was digested with *EcoR*I and *Sph*I and ligated to the hybridized, 5'-phosphorylated oligos loxP(L) and loxP (L_rc), which carry *loxP*(FAS) and *EcoR*I and *Sph*I-compatible overhangs. To prepare for library construction, a large-scale purification (Qiagen MAXI prep) of pCK-*cre*-lox-P(wt)-loxP(FAS) was performed according to Qiagen's protocol. The Qiagen-purified plasmid was subjected to CsCl gradient centrifugation for further purification. This construct was then digested with *Sph*I and *Bgl*II and ligated to digested naïve A domain library insert, which was obtained via a PCR-amplification of a preexisting A domain library pool. By design, the *loxP* sites and monomer are in-frame, which generates monomers with *loxP*-encoded linkers. This library was utilized in the *in vivo* recombination procedure as detailed below.

[325] fUSE5HA-Monomer-lox-lox vector. The vector is a derivative of fUSE5 from George Smith's laboratory (University of Missouri). It was subsequently modified to carry an HA tag for immunodetection assays. *loxP* sites were built into fUSE5HA sequentially. 5'-phosphorylated oligonucleotides loxP(I) *and* loxP(I)_rc, carrying loxP(WT), a string of stop codons and *Xma*I and *Sfi*I-compatible overhangs, were hybridized together and ligated to *Xma*I- and *Sfi*I-digested fUSE5HA in a standard ligation reaction (New England Biolabs T4 ligase; overnight at 16C).

[326] The resulting phage vector was next digested with *Xma*I and *Sph*I and ligated to the hybridized oligos loxP(J) and loxP(J)_rc, which carry *loxP*(FAS) and overhangs compatible with *Xma*I and *Sph*I. This construct was digested with *Xma*I/*Sfi*I and then ligated to pre-cut (*Xma*I/*Sfi*I) naïve A domain library insert (PCR product). The stop codons are located between the *loxP* sites, preventing expression of *gIII* and consequently, the production of infectious phage.

[327] The ligated vector/library was subsequently transformed into an *E. coli* host bearing a *gIII*-expressing *plasmid* that allows the rescue of the fUSE5HA-Monomer-lox-lox phage, as detailed below.

[328] pCK-*gIII*. This plasmid carries *gIII* under the control of its native promoter. It was constructed by PCR-amplifying *gIII* and its promoter from VCSM13 helper phage (Stratagene) with primers gIIIPromoter_EcoRI and gIIIPromoter_HinDIII. This product was digested with *Eco*RI and *Hin*DIII and cloned into the same sites of pCK110919-

HC-Bla. As *gIII* is under the control of its own promoter, *gIII* expression is presumably constitutive. pCK-*gIII* was transformed into *E. coli* EC100 (Epicentre).

[329] ***In vivo* recombination procedure.** In summary, the procedure involves the following key steps: a) Production of infective (i.e. rescue) of fUSE5HA-Monomer-lox-lox library with an *E. coli* host expressing *gIII* from a plasmid; b) Cloning of 2nd library (pCK) and transformation into F^+ TG1 *E. coli*; c) Infection of the culture carrying the 2nd library with the rescued fUSE5HA-Monomer-lox-lox phage library.

[330] *a. Rescue of phage vector.* Electrocompetent cells carrying pCK-*gIII* were prepared by a standard protocol. These cells had a transformation frequency of 4 x 10^8/μg DNA and were electroporated with large-scale ligations (~5 μg vector DNA) of fUSE5HA-lox-lox vector and the naïve A domain library insert. After individual electroporations (100 ng DNA/electroporation) with ~ 70 μL cells/cuvette, 930 μL warm SOC media were added, and the cells were allowed to recover with shaking at 37C for 1 hour. Next, tetracycline was added to a final concentration of 0.2 μg/mL, and the cells were shaken for ~ 45 minutes at 37C. An aliquot of this culture was removed, 10-fold serially diluted and plated to determine the resulting library size (1.8 x 10^7). The remaining culture was diluted into 2 x 500 mL 2xYT (with 20 μg/mL chloramphenicol and 20 μg/mL tetracycline to select for pCK-*gIII* and the fUSE5HA-based vector, respectively) and grown overnight at 30C.

[331] Rescued phage were harvested using a standard PEG/NaCl precipitation protocol. The titer was approximately 1 x 10^{12} transducing units/mL.

[332] *b. Cloning of the 2nd library and transformation* into *an E. coli host.* The ligated pCK/ naïve A domain library is electroporated into a bacterial F^+ host, with an expected library size of approximately 10^8. After an hour-long recovery period at 37C with shaking, the electroporated cells are diluted to OD_{600}~ 0.05 in 2xYT (plus 20 μg/mL chloramphenicol) and grown to mid-log phase at 37C before infection by fUSEHA-Monomer-lox-lox.

[333] *c. Infection of the culture carrying the 2nd library with the rescued fUSE5HA-Monomer-lox-lox phage library.* To maximize the generation of recombinants, a high infection rate (> 50%) of *E.coli* within a culture is desirable. The infectivity of *E. coli* depends on a number of factors, including the expression of the F pilus and growth conditions. *E. coli* backgrounds TG1 (carrying an F') and K91 (an Hfr strain) were hosts for the recombination system.

Oligonucleotides

loxP(K) (SEQ ID NO:380)
[P-5' agcttataacttcgtatagaaaggtatatacgaagttatagatctcgtgctgcatgcggtgcg]

loxP(K_rc) (SEQ ID NO:381)
[P-5' aattcgcaccgcatgcagcacgagatctataacttcgtatataccttctatacgaagttataagct]

loxP(L) (SEQ ID NO:382)
[P-5' ataacttcgtatagcatacattatacgaagttatcgag]

loxP (L_rc) (SEQ ID NO:383)
[P-5' ctcgataacttcgtataatgtatgctatacgaagttatg]

loxP(I) (SEQ ID NO:384)
[P5'ccgggagcagggcatgctaagtgagtaataagtgagtaaataacttcgtatataccttctatacgaagttatcgtctg]

loxP(I)_rc (SEQ ID NO:385)
[P-5' acgataacttcgtatagaaaggtatatacgaagttatttactcacttattactcacttagcatgccctgctc]

loxP(J) (SEQ ID NO:386)
[5' ccgggaccagtggcctctggggccataacttcgtatagcatacattatacgaagttatg]

loxP(J)_rc (SEQ ID NO:387)
[5' cataacttcgtataatgtatgctatacgaagttatggccccagaggccactggtc]

gIIIPromoter_EcoRI (SEQ ID NO:388)
[5' atggcgaattctcattgtcggcgcaactat

gIIIPromoter_HinDIII (SEQ ID NO:389)
[5' gataagctttcattaagactccttattacgcag]

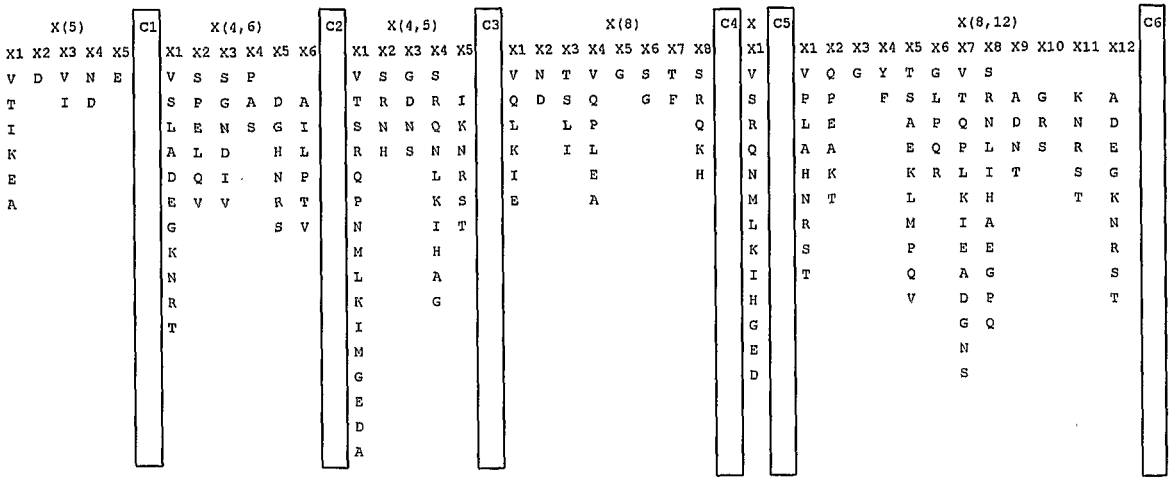
Example 3

[334] This example describes construction of an EGF-based monomer library.

[335] The CaEGF domain library, E3, encodes a protein domain of 36-43 amino acids having the following pattern:

$X_{(5)}C_1$-$X_{(4/6)}$-C_2-$X_{(4,5)}$-C_3-$X_{(8)}$-C_4-$X_{(1)}$-C_5-$X_{(8/12)}$-C_6 (SEQ ID NO:390)

[336] The table below (SEQ ID NOS:391-420) describes for each position which amino acids are encoded in the library based upon the natural diversity of human calcium binding EGF domains:



[337] The library of DNA sequences, E3, encoding monomeric calcium binding EGF domains, was created by assembly PCR as described in Stemmer et al., Gene 164:49-53 (1995). The oligonucleotides used in this PCR reaction are in two groups, 1 and 2. They are:

Group 1 (SEQ ID NOS:421-426):

1. 5'-AAAAGGCCTCGAGGGCCTGGGTGGCAATGGT-3'
2. 5'-CCTGAACCACCACAKHKACCGYKSNBGCACGGAYYCGRCRMACATTCATYAAYATCTDYACCATTGCCACCC-3'
3. 5'-CCTGAACCACCACAKNTGSCGYYGYKMHSGCACGGAYYCGRCRMACATTCATYAAYATCTDYACCATTGCCACCC-3'
4. 5'-CCTGAACCACCACAKHKACCGYKSNBGCAARBAYBCGVAHYCWSKBYACATTCATYAAYATCTDYACCATTGCCACCC-3'
5. 5'-CCTGAACCACCACAKNTGSCGYYGYKMHSGCAARBAYBCGVAHYCWSKBYACATTCATYAAYATCTDYACCATTGCCACCC-3'
6. 5'-TGAATTTTCTGTATGAGGTTTTGCTAAACAACTTTCAACAGTTTCGGCCCCAGAGGCCCTGGAGCCACCTGAACCACCACA-3'

Group 2 (SEQ ID NOS:427-430):

1. 5'-ACGGTGCCTACCCGTATGATGTTCCGGATTATGCCCCGGGTGGCAATGGT-3'
2. 5'-CCTGAACCACCACAGHKTDBACCGGHAWAGCCTKSCRSGCASHBACAKYKAWAGCYACCCDSTRWATYTWBACCATTGCCACCC-3'
3. 5'-CCTGAACCACCACAKBYKBTKCYGKYCBSABYCNGCDBAWAGCCTKBGBKGCASHBACAKYKAWAGCYACCCDSTRWATYTWBACCATTGCCACCC-3'
4. 5'-AAAAGGCCCCAGAGGCCCCTGAACCACCACA-3'

where R=A/G, Y=C/T, M=A/C, K=G/T, S=C/G, W=A/T, B=C/G/T, D=A/G/T, H=A/C/T, V=A/C/G, and N=A/C/G/T.

[338] Following the separate PCRs of the Group 1 and 2 oligonucleotides, the Group 1 PCR fragments were digested with BpmI and group 2 PCR fragments were digested with BsrDI. Digestion products were purified using Qiagen Qiaquick columns and then ligated together. The ligated DNA was then amplified in a PCR using two primers. These are:

5'-AAAAGGCCTCGAGGGCCTGGGTGGCAATGGT-3' (SEQ ID NO:431)
5'-AAAAGGCCCCAGAGGCCCCTGAACCACCACA-3' (SEQ ID NO:432)

[339] The PCR products were purified with Qiagen Qiaquick columns and digested with SfiI. The digested product was purified with Qiagen Qiaquick columns. The DNA fragments were ligated into the SfiI restriction sites of phage display vector fuse5-HA(G4S)4, a derivative of fuse5 carrying an in-frame HA-epitope and a glycine, serine flexible linker. The ligation mixture was electroporated into TransforMax™ EC100™ electrocompetent E. coli cells. Transformed E. coli cells were grown overnight at 37°C in 2xYT medium containing 20 μg/ml tetracycline. The resulting library contained $2X10^9$ independent clones. Phage particles were purified from the culture medium by PEG-precipitation. The titer of the phage was $1.3X10^{12}$/ml. The sequences of 24 individual clones were determined and these were consistent with the library design.

Example 4

[340] This example describes construction of an EGF-based monomer library.

[341] Recombination can be used for intradomain optimization. For example a PCR overlap reaction can be used that recombines two or more segments of a single domain relative to each other. One can use two, three, four, five or more fragment overlap reactions in the same way as illustrated. This recombination process has many applications. One application is to recombine a large pool of hundreds of previously selected clones without sequence information. All that is needed for each overlap to work is one known region of (relatively) constant sequence that exists in the same location in each of the clones (fixed site approach). For A domains, typically these clones would have been derived from a library in which 20-25 amino acids distributed over all five inter-cysteine segments were randomized. The intra-domain recombination method can also be performed on a pool of sequence-related monomer domains by standard DNA recombination (*e.g.*, Stemmer, *Nature* 370:389-391

(1994)) based on random fragmentation and reassembly based on DNA sequence homology, which does not require a fixed overlap site in all of the clones that are to be recombined.

[342] Another application of this process is to create multiple separate, naïve (meaning unpanned) libraries in each of which only one of the intercysteine loops is randomized, to randomize a different loop in each library. After panning of these libraries separately against the target, the selected clones are then recombined. From each panned library only the randomized segment is amplified by PCR and multiple randomized segments are then combined into a single domain, creating a shuffled library which is panned and/or screened for increased potency. This process can also be used to shuffle a small number of clones of known sequence.

[343] Any common sequence may be used as cross-over points. For A domains or other cysteine-containing monomers, the cysteine residues are logical places for the crossover. However, there are other ways to determine optimal crossover sites, such as computer modeling. Alternatively, residues with highest entropy, or the least number of intramolecular contacts, may also be good sites for crossovers.

[344] An exemplary method of generating libraries comprised of proteins with randomized inter-cysteine loops is presented below. In this example, in contrast to the separate loop, separate library approach described above, multiple intercysteine loops are randomized simultaneously in the same library.

[345] An A domain NNK library encoding a protein domain of 39-45 amino acids having the following pattern was constructed:

C_1-$X_{(4,6)}$-E_1-F-R_1-C_2-A-$X_{(2,4)}$-G_1-R_2-C_3-I-P-S_1-S_2-W-V-C_4-D_1-G_2-E_2-D_2-D_3-C_5-G_3-D_4-G_4-S_3-D_5-E_3-$X_{(4,6)}$-C_6 (SEQ ID NOS:433-459);
where,
C_1-C_6: cysteines;
$X_{(n)}$: sequence of n amino acids with any residue at each position;
E_1-E_3: glutamine;
F: phenylalanine;
R_1-R_2: arginine;
A: alanine;
G_1-G_4: glycine;
I: isoleucine;
P: proline;
S_1-S_3: serine;
W: tryptophan;
V: valine;
D_1-D_5: aspartic acid; and
C_1-C_3, C_2-C_5 & C_4-C_6 form disulfides.

[346] The library was constructed by creating a library of DNA sequences, containing tyrosine codons (TAT) or variable non-conserved codons (NNK), by assembly PCR as described in Stemmer *et al.*, *Gene* 164:49-53 (1995). Compared to the native A-domain scaffold and the design that was used to construct library A1 (described previously) this approach: 1) keeps more of the existing residues in place instead of randomizing these potentially critical residues, and 2) inserts a string of amino acids of variable length of all 20 amino acids (NNK codon), such that the average number of inter-cysteine residues is extended beyond that of the natural A domain or the A1 library. The rate of tyrosine residues was increased by including tyrosine codons in the oligonucleotides, because tyrosines were found to be overrepresented in antibody binding sites, presumably because of the large number of different contacts that tyrosine can make. The oligonucleotides used in this PCR reaction are (SEQ ID NOS:460-486):

1. 5′ -ATATCCCGGGTCTGGAGGCGTCTGGTGGTTCGTGTNNKNNKNNKNNKGAATTCCGA- 3′
2. 5′ -ATATCCCGGGTCTGGAGGCGTCTGGTGGTTCGTGTNNKNNKNNKNNKNNKGAATTCCGA- 3′
3. 5′ -ATATCCCGGGTCTGGAGGCGTCTGGTGGTTCGTGTNNKNNKNNKNNKNNKNNKGAATTCCGA-
3′
4. 5′ -ATATCCCGGGTCTGGAGGCGTCTGGTGGTTCGTGTTATNNKNNKNNKGAATTCCGA- 3′
5. 5′ -ATATCCCGGGTCTGGAGGCGTCTGGTGGTTCGTGTNNKTATNNKNNKNNKGAATTCCGA- 3′
6. 5′ -ATATCCCGGGTCTGGAGGCGTCTGGTGGTTCGTGTNNKTATNNKNNKGAATTCCGA- 3′
7. 5′ -ATATCCCGGGTCTGGAGGCGTCTGGTGGTTCGTGTNNKNNKTATNNKGAATTCCGA- 3′
8. 5′ -ATATCCCGGGTCTGGAGGCGTCTGGTGGTTCGTGTNNKNNKNNKTATGAATTCCGA- 3′
9. 5′ -ATATCCCGGGTCTGGAGGCGTCTGGTGGTTCGTGTNNKNNKNNKTATNNKGAATTCCGA- 3′
10. 5′ -ATACCCAAGAAGACGGTATACATCGTCCMNNMNNTGCACATCGGAATTC- 3′
11. 5′ -ATACCCAAGAAGACGGTATACATCGTCCMNNMNNMNNTGCACATCGGAATTC- 3′
12. 5′ -ATACCCAAGAAGACGGTATACATCGTCCMNNMNNMNNMNNTGCACATCGGAATTC- 3′
13. 5′ -ATACCCAAGAAGACGGTATACATCGTCCATAMNNMNNTGCACATCGGAATTC- 3′
14. 5′ -ATACCCAAGAAGACGGTATACATCGTCCMNNATAMNNMNNTGCACATCGGAATTC- 3′
15. 5′ -ATACCCAAGAAGACGGTATACATCGTCCMNNATAMNNTGCACATCGGAATTC- 3′
16. 5′ -ATACCCAAGAAGACGGTATACATCGTCCMNNMNNATATGCACATCGGAATTC- 3′
17. 5′ -ATACCCAAGAAGACGGTATACATCGTCCMNNMNNATAMNNTGCACATCGGAATTC- 3′
18. 5′ -ACCGTCTTCTTGGGTATGTGACGGGGAGGACGATTGTGGTGACGGATCTGACGAG- 3′
19. 5′ -ATATGGCCCCAGAGGCCTGCAATGATCCACCGCCCCCACAMNNMNNMNNMNNCTCGTCAG
ATCCGT- 3′
20. 5′ -ATATGGCCCCAGAGGCCTGCAATGATCCACCGCCCCCACAMNNMNNMNNMNNMNNCTCGTCA
GATCCGT- 3′
21. 5′ -ATATGGCCCCAGAGGCCTGCAATGATCCACCGCCCCCACAMNNMNNMNNMNNMNNMNNC
TCGTCAGATCCGT- 3′
22. 5′ -ATATGGCCCCAGAGGCCTGCAATGATCCACCGCCCCCACAATAMNNMNNMNNCTCGTC
AGATCCGT- 3′
23. 5′ -ATATGGCCCCAGAGGCCTGCAATGATCCACCGCCCCCACAMNNATAMNNMNNMNNCT
CGTCAGATCCGT- 3′
24. 5′ -ATATGGCCCCAGAGGCCTGCAATGATCCACCGCCCCCACAMNNATAMNNMNNCTCGT
CAGATCCGT- 3′
25. 5′ -ATATGGCCCCAGAGGCCTGCAATGATCCACCGCCCCCACAMNNMNNATAMNNCTCG
TCAGATCCGT- 3′
26. 5′ -ATATGGCCCCAGAGGCCTGCAATGATCCACCGCCCCCACAMNNMNNMNNATACTCG
TCAGATCCGT- 3′
27. 5′ -
ATATGGCCCCAGAGGCCTGCAATGATCCACCGCCCCCACAMNNMNNMNNATAMNNCTCGTCAGATCCGT- 3′

where R=A/G, Y=C/T, M=A/C, K=G/T, S=C/G, W=A/T, B=C/G/T, D=A/G/T, H=A/C/T, V=A/C/G, and N=A/C/G/T

[347] The library was constructed though an initial round of 10 cycles of PCR amplification using a mixture of 4 pools of oligonucleotides, each pool containing 400 pmols of DNA. Pool 1 contained oligonucleotides 1-9, pool 2 contained 10-17, pool 3 contained only 18 and pool 4 contained 19-27. The fully assembled library was obtained through an additional 8 cycles of PCR using pool 1 and 4. The library fragments were digested with XmaI and SfiI. The DNA fragments were ligated into the corresponding restriction sites of phage display vector fuse5-HA, a derivative of fuse5 carrying an in-frame HA-epitope. The ligation mixture was electroporated into TransforMax™ EC100™ electrocompetent *E. coli* cells resulting in a library of $2X10^9$ individual clones. Transformed *E. coli* cells were grown overnight at 37°C in 2xYT medium containing 20 μg/ml tetracycline. Phage particles were purified from the culture medium by PEG-precipitation and a titer of $1.1X10^{13}$/ml was determined. Sequences of 24 clones were determined and were consistent with the expectations of the library design.

Example 5

[348] This example describes optimization of multimers by optimizing monomers and/or linkers for binding to a target.

[349] One approach for optimizing multimer binding to targets involves optimization of monomers, multimers and linkers. First a library of monomers is panned for binding to the target (e.g., IL-6). However, some of the monomers may bind at locations on the target that are far away from each other, such that the domains that bind to these sites cannot be connected by a linker peptide. It is therefore useful to create and screen a large library of homo- or heterotrimers from these monomers before optimization of the monomers. These trimer libraries can be screened, e.g., on phage (typical for heterotrimers created from a large pool of monomers) or made and assayed separately (e.g., for homotrimers). By this method, the best trimer is identified. The assays may include binding assays to a target or agonist or antagonist potency determination of the multimer in functional protein- or cell-based assays.

[350] The monomeric domain(s) of the single best trimer are then optimized as a second step. Homomultimers are easiest to optimize, since only one domain sequence exists, though heteromultimers may also be synthesized. For homomultimers, an increase in binding by the multimer compared to the monomer is an avidity effect.

[351] After optimization of the domain sequence itself (e.g., by recombining or NNK randomization) and phage panning, the improved monomers are used to construct a dimer with a linker library. Linker libraries may be formed, e.g., from linkers with an NNK composition and/or variable sequence length.

[352] After panning of this linker library, the best clones (e.g., determined by potency in the inhibition or other functional assay) are converted into multimers composed of multiple (e.g., two, three, four, five, six, seven, eight, etc.) sequence-optimized domains and length- and sequence-optimized linkers.

Example 6

[353] This example describes a structural analysis of A domains.

[354] As with virtually all proteins, only a small fraction of the total surface of an A-domain participates in binding a single target. Based on the solution structure of the domain, adjacent residue positions can be identified which are likely to be able to cooperate in binding to a given target. Herein, such groups of adjacent residues are referred to as structural categories. As an example, four such categories have been identified through examination of the A-domain structure, designated Top, Bottom, Loop 1, and Loop 2. By designing libraries which only allow diversity within a given category, the theoretical sequence space allowed by a library can be significantly reduced, allowing for better coverage of the theoretical space by the physical library. Further, in the case of non-overlapping categories such as the Top and Bottom categories, half-domain sequences selected against different targets can be combined into a single sequence which would be able to bind simultaneously or alternatively to the selected targets. In either case, creating binding sites that occupy only half a domain allows for the creation of molecules that are half as large and would have half the number of immunogenic epitopes, reducing the risk of immunogenicity.

Structural Classification of A-domain Positions

A canonical A-domain sequence (SEQ ID NO:487) is shown below with high-diversity positions represented as an X. Positions that belong to either the Top, Bottom, Loop 1, or Loop 2 categories are designated with a star.

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	C	X	X	X	X	F	X	C	X	X	X	X	C	I	X	X	X	W	X
Top	·	✪	·	·	✪	·	·	·	·	·	·	✪	·	✪	✪	·	✪	✪	·
Bottom	·	·	·	·	·	✪	✪	·	✪	·	·	·	·	·	·	✪	·	·	✪
Loop 1	·	✪	✪	✪	✪	·	·	·	·	·	·	·	·	·	✪	✪	✪	·	·
Loop 2	·	·	·	·	·	·	·	·	✪	✪	✪	✪	·	·	·	·	·	·	·

	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
	C	D	G	X	X	D	C	X	D	X	S	D	E	X	X	C
Top	·	✪	✪	✪	✪	✪	·	✪	·	✪	·	·	·	·	·	·
Bottom	·	·	·	·	·	·	·	·	✪	·	✪	·	·	✪	✪	·
Loop 1	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
Loop 2	·	·	·	·	·	·	·	✪	✪	✪	·	·	·	·	·	·

Example 7

[355] This example describes screening for monomers or multimers that bind human IL-6 .

[356] Phage libraries were panned through several rounds either on solid support (e.g. Nunc Maxisorp plates) or in solution (e.g. Dynal Streptavidin). Output phage pools with (a) the highest frequency of individual phage clones that bind to IL-6 and (b) high sequence diversity among the binding-positive phage clones were chosen for protein screening.

I. Round 1 (Maxisorp plates or Dynal Beads)

1. Coating Target

[357] **A.** ***Coating plates:*** Six wells/library were directly coated with IL-6 (0.5 μg/well) using 100 μL/well of 5 μg/mL human Il-6 (R & D Systems; carrier free) diluted in TBS[pH 7.5]/2 mM $CaCl_2$. When using a biotinylated form of human IL-6, plates were pre-coated with streptavidin for 1 hr at room temperature with shaking. In addition, one negative control well/library was coated with TBS[pH 7.5]/2 mM $CaCl_2$ only. After pre-coating was complete, IL-6 (+ biotin) was added and the plates were incubated for 1.5 hr at room temperature with shaking.

[358] B. ***Coating beads:*** 20 μL Dynal streptavidin (M-280; Dynal ASA) were incubated with 5 μg biotinylated IL-6, in 500 μL TBS[pH 7.5]/2 mM $CaCl_2$ and rotated at room temperature for 1 hr in Eppendorf tubes. As a negative control, 20 μL Dynal streptavidin without target were incubated in 500 μL of TBS[pH 7.5]/2 mM $CaCl_2$ and rotated at room temperature for 1 hr. Note that Dynal beads were washed at least twice with TBS [pH 7.5]/2 mM $CaCl_2$ before adding target, and beads were coated in bulk.

2. Blocking

[359] A. ***Blocking Plates:*** Coating solution was removed and wells were washed one time with 200 μL/well of TBS[pH 7.5]/2 mM $CaCl_2$. 250 μl/well of 1% BSA (protease-free) in TBS[pH 7.5]/2 mM $CaCl_2$ was added and incubated for 1 hr. at room temperature with shaking. Alternative reagents (e.g. casein or milk) can be used for blocking.

[360] B. ***Blocking Beads:*** Coating solution was removed and beads were washed twice with TBS [pH 7.5]/2 mM $CaCl_2$. 500 μl 1% BSA (protease-free) was added in TBS[pH 7.5]/2 mM $CaCl_2$ and rotated for 1 hr at room temperature. As noted above, alternative blocking reagents can be used.

3. Washes

[361] A. ***Wash Plates:*** Wells were washed three times with 200 μL/well of TBS[pH 7.5]/2 mM $CaCl_2$ to remove excess target.

[362] B. ***Wash Beads:*** Beads were washed three times with 1000 μL of TBS[pH 7.5]/2 mM $CaCl_2$ to remove excess target. Beads were allowed to collect on a magnet for a few min after each wash to avoid bead loss.

4. Phage addition

[363] A. ***Phage addition to Plates:*** About 1000 library equivalents (A1 domain naïve phage library) were added in phage addition buffer (1% nonfat dry milk/ 0.2% BSA (protease-free), or other appropriate blocking agent, in TBS [pH 7.5]/2 mM $CaCl_2$) and incubated at room temperature for 2 hr with shaking. In rounds 2-3, 100 μL total of harvested phage was added to 7 wells (6 target + 1 negative control) diluted in phage addition buffer.

[364] **B.** ***Phage addition to Beads:*** About 1000 library equivalents (A1 domain naïve phage library) were added in 500 µl 1% non-fat dry milk + 100 µl 1% BSA (protease-free) in TBS [pH 7.5]/2 mM $CaCl_2$ and incubated with rotation at room temperature for 2 hr. In rounds 2-3, 100 µL total of harvested phage were added to beads.

5. Washes

[365] **A.** ***Washing Plates:*** The plates were washed eight to twelve times with 200 µl/well of TBS [pH 7.5]/2 mM $CaCl_2$/0.1 % Tween-20 over a period of 10 min.

[366] **B.** ***Washing Beads:*** The beads were washed 8-12 times with 800 µl of TBS [pH 7.5]/2 mM $CaCl_2$/0.1 % Tween-20 over a period of 30 - 45 min. Bead resuspension was facilitated by dispensing wash buffer directly onto collected beads or by pipetting up and down (not by vortexing). Alternatively, a KingFisher apparatus (Thermo LabSystems) or equivalent can be used for bead washing.

Conditions for Stringent washes (options)

[367] a. 800 µl of TBS [pH 7.5]/2 mM $CaCl_2$/0.1 % Tween-20 at 37°C;

[368] b. 800 µl of TBS [*450 mM NaCl*, pH 7.5]/2 mM $CaCl_2$/0.1 % Tween-20 at room temperature;

[369] c. Beads were washed normally 6-8 times, then 1 µg of unlabeled IL-6 was added for 1 hr at room temperature or 37°C. Phage that remained bound after this wash were retained for elution/infection;

[370] d. 1% milk/0.2% BSA/with or without 1 M urea/ 37°C (high stringency).

6. Competition (optional):

[371] **A.** ***Competition on Plates:*** Phage were incubated with 100 µL/well of 20 µg/mL (2 µg/well) of sIL-6R plus 7.5 ug gp130 in TBS [pH 7.5]/ 2mM $CaCl_2$ for 1 hr at room temperature with shaking. sIL-6R/gp130 eluates were retained for infection of BlueKan K91 *E. coli*.

[372] **B.** ***Competition on Beads:*** Phage were incubated with 20 μg of sIL-6R and 75 ug of gp130 in 500 μL TBS [pH 7.5]/ 2mM $CaCl_2$ for 1 hr at room temperature with shaking. sIL-6R/gp130 eluates were retained for infection of BlueKan K91 *E. coli*.

7. Phage Elution

[373] **A.** ***Elution off of Plates:*** 100 μL/well of 10 mg/mL trypsin in TBS [pH 7.5]/ 2mM $CaCl_2$ was added, and the plates were incubated at 37°C for 30 min with shaking.

[374] **B.** ***Elution off of Beads:*** 100 μL 10 mg/ml trypsin TBS [pH 7.5]/ 2mM $CaCl_2$ was added to beads, which were then incubated at 37°C (in an Eppendorf rack) for 30 min with shaking.

[375] **C.** ***Alternative elution/infection:*** 200 μL of log-phase BlueKan K91 *E. coli* cells at OD_{600} ~ 0.5 were added to each well (for plates) or to aspirated beads. The infection was allowed to proceed for 30 min at 37°C without shaking. Next, the 200 μL volumes were pooled and added to ~ 3 mL of 2xYT/0.2 μg/mL tetracycline and shaken for 15 min at 37°C.

8. Infection: (same for Plate and Bead Protocol)

[376] An appropriate volume of log-phase BlueKan K91 *E. coli* (in 2xYT/40 μg/mL kanamycin) was grown to OD_{600} ~ 0.5 - 0.6. When the culture reached OD_{600}, it was placed on ice prior to use, although the time on ice was generally minimized.

[377] **A.** In a 50 mL sterile conical tube, eluted phage were mixed with 5 mL log-phase BlueKan K91 *E. coli* culture and incubated at 37°C for 25 min without shaking. The sterile conical tubes were covered with AirPore tape (Qiagen) to facilitate aeration.

[378] **B.** Tetracycline was added to a final concentration of 0.2 μg/mL and shaken for 15 min at 37°C.

[379] **C.** A 10 μL aliquot was sampled for titering and serially diluted 10-fold (10^{-1} to 10^{-6}) in 2xYT, plated in 8 μL/dilution spots on 2xYT/20 μg/ml tetracycline plates and incubated overnight at 30°C or 37°C. The remaining volume of the $10^{-2} - 10^{-4}$ dilutions was plated to obtain single colonies for subsequent phage ELISAs.

[380] **D.** Infected 5 mL cultures were diluted ~10-fold into 50 mL 2xYT/ 20 μg/mL tetracycline and incubated with shaking at 30°C overnight to saturation.

9. Titering input phage was used in the current round of panning (same for Plate and Bead Protocol)

[381] **A.** 100-fold serial dilutions (10^{-4} to 10^{-10}) of harvested phage were made in 2xYT.

[382] **B.** 100 μL/well of a log-phase BlueKan K91 *E. coli* culture at OD_{600} 0.5-0.6 was added to 6 wells of a 96-well polypropylene plate.

[383] **C.** 10 μL of diluted phage was added to the wells containing 100 μL of BlueKan K91 *E. coli.*

[384] **D.** Phage/cell mixtures were incubated at 37°C for 25 min without shaking, and the plates were covered with AirPore tape (Qiagen) to allow for aeration.

[385] **E.** Tetracycline was added to a final concentration of 0.2 μg/mL and the plate was shaken for 15 min at 37°C.

[386] **F.** 8 μL of each dilution (10^{-4} to 10^{-10}) was plated onto a dry 2xYT agar/ 20 μg/mL tetracycline plate.

[387] **G.** Plates were incubated at 30°C or 37°C overnight.

10. *Harvesting phage* (same for Plate and Bead Protocols)

[388] **A.** Overnight cultures were centrifuged at 7000 rpm in disposable 50 mL tubes for 25 min to pellet cells.

[389] **B.** A standard PEG/NaCl phage-precipitation procedure was performed by adding 1/5 volume of a 20% PEG/15% NaCl stock to culture supernatant. It was mixed well by repeatedly inverting and incubating on ice for 45 min to 1 hr.

[390] **C.** The culture was centrifuged at 7000 rpm for 40 min to pellet phage and the supernatant was discarded.

[391] **D.** The phage pellet was resuspended in 1 mL TBS [pH 7.5]/ 2 mM $CaCl_2$, transferred to an Eppendorf tube and centrifuged at 13K rpm for at least 2 min to pellet insoluble material.

[392] **E.** Supernatant was transferred to a fresh tube and 1/5 volume of PEG/NaCl was added, mixed and incubated on ice for ~ 5 min.

[393] **F.** The mixture was then centrifuged at 13000 rpm for at least 2 min, and the supernatant was removed. The pelleted, purified phage were resuspended in up to 1 mL TBS [pH 7.5]/ 2mM $CaCl_2$ and stored at 4°C.

II. Round 2 and round 3 panning

[394] The 2nd and 3rd round panning conditions were generally the same as in Round 1 described above, except the coated target (i.e. IL-6) amount was decreased 2- to 4-fold for each subsequent round, and the plates (or beads) were washed 2 - 4 additional times in each subsequent round of panning.

III. Optional intra-domain recombination

[395] Monomer sequences in phage display-selected phage pools were recombined in the following procedure. This process generated hybrid monomers derived from mixed halves of the starting monomer collection in a given pool(s). For A1-domain-based phage libraries, the primer pairs SHF1 (ATTATGCCCCGGGTCTGGAGGCGTC; SEQ ID NO:488)/ SHBoverlap (CGCCGTCGCAA; SEQ ID NO:489) and SHFoverlap (TTGCGACGGCG; SEQ ID NO:490)/B3 (TCGGCCCCAGAGGCCTGCAATG; SEQ ID NO:491) were used to PCR-amplify the two halves of the monomers. The 2 halves were fused together with LA Taq polymerase (Takara). Next, the fused hybrid coding sequences were amplified by primers SHF2 (CCGGATTATGCCCCGGGTCTGGA; SEQ ID NO:492) and SHB4 (AACAGTTTCGGCCCCAGAGGCCTGC; SEQ ID NO:493). Purified PCR products were digested by SfiI (NEB) and ligated with the SfiI-digested fUSE5HA phage vector to generate recombined monomer libraries. Recombined libraries were panned at least two more rounds against IL-6 and screened as described below. Data from characterization of recombined monomers is in Tables 1 and 2.

IV. Analysis of panning output (same for plate and bead protocols)

[396] ***Phage ELISAs:*** For each output "phage pool" to be analyzed (typically Rounds 2, 3 and 4, if applicable), independent clones were inoculated into 1 mL (2xYT/ 20 μg/mL tetracycline) cultures grown in Costar 96-well polypropylene deep-well plates. Inoculating tips were left in, and plates were shaken overnight at 37°C. Cells were pelleted by centrifugation at 3600 rpm for 15 min. Culture supernatants were retained and ELISAs were performed as described below.

[397] Non-biotinylated IL-6 (0.1 μg/well) was directly coated onto Nunc Maxisorp plates. However, biotinylated IL-6, 96-well Nunc Maxisorp plates should first be coated with 50 μL/well of 50 μg/mL (2.5 μg/well) of streptavidin, diluted in TBS [pH 7.5]/ 2 mM $CaCl_2$. The plate was incubated at 37°C for 1 hr with shaking. Plates were washed three times with 200 μL/well of TBS [pH 7.5]/ 2 mM $CaCl_2$. Wells were blocked with 200 μL/well of 1% BSA (fraction V) and the covered plate was incubated at RT for 1 hr with shaking. The plate was washed three times with TBS [pH 7.5]/2 mM $CaCl_2$. Next, the 96-well Maxisorp plate was coated with 100 μL/well of 1 μg/mL (0.1 μg/well) biotinylated IL-6 diluted in TBS [pH 7.5]/2 mM $CaCl_2$ or 100 μL/well buffer only (negative control). The plate was incubated at RT for 1 hr with shaking. Plates were washed three times with TBS [pH 7.5]/2 mM $CaCl_2$. Next, 30 μL of each phage supernatant is added to wells in the presence of 70 μL of 1% Milk/0.2% BSA/[pH 7.5]/2 mM $CaCl_2$/0.02% Tween-20. Covered plates were incubated at RT for 1.5 hr with shaking.

[398] Plates were washed four times with TBS [pH 7.5]/2 mM $CaCl_2$/0.02% Tween-20. Next, 100 μL/well of α-M13-HRP monoclonal antibody (Amersham Pharmacia), diluted 1:5000 in TBS [pH 7.5]/2 mM $CaCl_2$ + 0.02% Tween-20, was added. Plates were incubated at 4°C for 1 hr with shaking. Plates were washed three times with cold TBS [pH 7.5]/2 mM $CaCl_2$/0.02% Tween-20. 100 μL/well of TMB/H_2O_2 mixture (Pierce), diluted 1:1, was added for ELISA development.

[399] The reactions were allowed to turn blue until the strongest OD_{650} signals reached ~ 1.0. The reaction was stopped with 100 μL/well 2N H_2SO_4, and positive wells changed in color from blue to yellow. Once the reaction was stopped, it was read at OD_{450} on an ELISA plate reader using SoftMaxPro software.

[400] Phage ELISA-positive phage pools were chosen for subcloning into an expression vector if they had (a) a high frequency of individual phage clones that bound to

hIL-6 and (b) high sequence diversity among the binding-positive phage clones. Pools meeting these criteria were chosen for protein screening in the process outlined below. To subclone the monomer or multimer sequences from a given phage pool into the expression vector, pEve, approximately 10^8-10^{10} phage were amplified by 25 cycles of PCR as follows:

PCR recipe
0.5 – 1 μL purified phage
5 μL 10X Buffer
8 μL 2.5mM dNTPs
5 μL 10uM VS-For primer (5'-ATCATCTGGCCGGTCCGGCCTACCCGTATGATGTTCCGGA-3'; SEQ ID NO:494)
5 μL 10uM EveNut primer (5'-AAAAGGCCCCAGAGGCCTTCTGCAATGAC-3'; SEQ ID NO:495)
26 μL H_2O
0.5 μL LA Taq polymerase (1 unit) (Takara)

Cycles: 25 X [94°C/10sec., -45°C/30sec., -72°C/30sec.]

[401] PCR products were run on a 3% agarose gel for analysis. The monomer or multimer product (approximately 200 bp) was purified with a QIAquick spin column (Qiagen), digested with Sfi I (NEB), purified again with a QIAquick column and then ligated using T4 DNA Ligase (NEB) to the Sfi I digested vector, pEve. The ligation was transformed into electrocompetent BL21 (DE3) *E. coli* and plated onto 2xYT plates containing kanamycin at 40 μg/mL. Following overnight growth, approximately 6000 individual clones were inoculated into 2xYT/kanamycin and grown overnight. Positive and negative controls were also included on the plates.

V. Screening of thousands of monomer proteins in 1 mL cell lysates

[402] Protein Production of 1 mL heated lysates (Day 1): Individual clones were inoculated into wells of a 96-well Costar deep-well plate containing 500 μL/well of 2xYT/ 40 μg /mL kanamycin. Cultures were grown overnight (inoculating tips were left in wells) while shaken at 300 rpm at 37°C. This process allowed screening of thousands of individual, partially-purified monomers at the cell-lysate level.

[403] (Day 2) 100 μL of overnight culture was inoculated into new 96-well Costar deep-well plate containing 1 mL/well of 2xYT/ 40 μg/mL kanamycin + 1 mM $CaCl_2$. (The remaining overnight culture was archived by the addition of 25% final glycerol concentration and then stored at -80°C for later use.) Plates were covered with AirPore Tape (Qiagen) and cultures were grown with shaking at 375 rpm at 37°C until an OD_{600} of ~ 0.8 to 1.0 was reached. Once the desired OD_{600} was reached, cultures were induced with 1 mM

IPTG for 3 hr while shaking at 375 rpm at 37°C. Plates containing induced cultures were then centrifuged for 15 min at 3600 rpm at 4°C to pellet cells. Supernatant was removed and discarded, and the remaining cell pellet was resuspended in 100 µL of TBS [pH 7.5]/1 mM $CaCl_2$. Resuspended cells were transferred from the 96-well deep-well plate to a 96-well polypropylene PCR plate and heated for 5 min at 65°C in a PCR machine. Heated/lysed cells were then centrifuged for 15 min at 3600 rpm at 4°C. After centrifugation, protein production was complete, and protein lysates were ready for characterization in a primary screen via binding ELISA and/or competition AlphaScreen assays.

[404] IL-6 Protein ELISA: 96-well Maxisorp plates were coated with 100 µL/well of 1 µg /mL (0.1 µg/well) IL-6(R&D systems) diluted in TBS[pH 7.5]/ 1 mM $CaCl_2$, and the plate was incubated at 4°C overnight or room temperature (RT) for 1.5 hr with shaking. Wells were emptied and then blocked with 200 µL/well of 1% BSA (fraction V)/TBS[pH7.5]/1 mM $CaCl_2$. The covered plate was incubated at RT for 1 hr with shaking. The plate was washed three times with TBS [pH 7.5]/1 mM $CaCl_2$. Protein from 1 mL lysate preparations was added to the wells as a single point concentration diluted 1:10 in TBS [pH 7.5]/1 mM $CaCl_2$/ 0.1% BSA/ 0.02% Tween-20 to a final volume of 100 µL per well. Covered plates were incubated at RT for 1.5 hr with shaking. The plate was washed three times with TBS [pH 7.5]/1 mM $CaCl_2$/ 0.02% Tween-20. 100 µL /well of anti-HA-HRP detection antibody (Roche) diluted 1:2000 in TBS [pH 7.5]/1 mM $CaCl_2$/ 0.1% BSA/ 0.02% Tween-20 was added. Covered plates were incubated at RT for 1 hr with shaking. The plate was washed three times with TBS [pH 7.5]/ 1 mM $CaCl_2$/ 0.02% Tween-20. 100 µL /well of TMB/ H_2SO_4 mixture diluted 1:1 was added. Color was allowed to turn blue until OD_{650} of the strongest signals reached ~ 1.0. The reaction was stopped with 100 µL/well of 2N H_2SO_4. Once stopped, the plate was read on ELISA plate reader at OD_{450}.

[405] AlphaScreen IL-6/Fc-biotinylated (bn) IL-6 homogeneous competition assay: All assay components were diluted in AlphaScreen Buffer: 40 mM HEPES [pH 7.4] w/ NaOH,1 mM $CaCl_2$, 0.1 % BSA (w/v), 0.05 % Tween-20, 100 mM NaCl. Three additions were made to a white, 384-well, reduced-volume, Greiner microtiter assay plate with no incubation time in between additions. First, monomers or unlabeled recombinant human IL-6 (rhIL-6, R&D Systems)(as positive control) were added to the plate at 4 µL/well. Monomers from 1 mL heated lysate preparations were added to the wells at a single

concentration (either undiluted [i.e. 1:4 final assay dilution] or up to a 1:100 dilution [1:400 final assay dilution]). As a positive control, instead of monomer protein, 4μL/well of unlabeled rhIL-6 was added to the plate as a twelve-point concentration curve starting with 400 nM (i.e. 100 nM final assay concentration) and then 1:4 serial dilutions thereafter with the last point as buffer only. Secondly, 2 uL/well of bn-rhIL-6 at 1 nM (i.e. 0.25 nM final assay concentration) was added to the plate. Note that the remainder of the assay was done in subdued or green-filtered light as AlphaScreen beads are light sensitive. Thirdly, 2 μL/well of a mixture of IL-6R at 28 nM (7 nM final concentration) and gp130/Fc at 1.6 nM (0.4 nM final concentration) and AlphaScreen Streptavidin "donor beads" and Protein A "acceptor beads" (PerkinElmer) both diluted to 40 μg/mL (i.e., 10 μg/mL final assay concentration) was added to the plate. The assay plate was then covered with topseal and spun down for ~ 30 sec at 800 rpm. The plate was then incubated overnight in the dark at room temperature and read the next day on the Fusion Plate reader (PerkinElmer).

VI. Multimerization and recombination of phage display-selected monomers

[406] Monomers that have been subcloned into pEve (pEve/monomer) were multimerized in the following manner. pEve/monomer plasmids (individually or in pools) were digested with either BsrDI or BpmI (NEB). The ~1.1 kb BsrDI and ~2.9 BpmI fragments were isolated from 1% agarose gels and purified with Qiagen QIAquick spin columns. Pools of each of the two fragments were ligated using T4 DNA ligase (NEB); subsequently, the ligation was purified with a Qiagen QIAquick spin column. Using the primers VS-For and EveNut described in the phage subcloning section above, the multimer coding sequences were PCR-amplified from the ligation. The PCR products were purified and digested with SfiI (NEB), followed by ligation with pEve and transformation of BL21 (DE3) *E. coli*. This method created dimers comprised of different combinations of the starting monomers. This method can also be used to generate other multimers, such as trimers. When making trimers, pools of pEve/dimers (e.g. in above example) and pEve/monomers (the starting collection) are the starting materials. They are processed as above. A molecular biology procedure similar to that described below for making "walking libraries" was also used to generate multimers. In all cases, proteins were expressed, purified and screened as above.

[407] Additional libraries, referred to as "walking libraries," were generated by ligating phage display-selected monomers (i.e. selected monomers) with the full

representation of a naïve monomer library. These libraries were constructed in the following manner. PCR was used to amplify in two separate reactions: a) the coding sequences of the selected monomers with pETF (ACCCGTATGATGTTCCGGATTA; SEQ ID NO:496) /pETB2r (GATGTATTCGGCCCCAGA GGCCTGCAATGAC; SEQ ID NO:497); and b) the coding sequences of naïve monomers in a monomer library with 21new1 (GAAATTCACCTCGAAAGCAA; SEQ ID NO:498)/23 (ATGGGTTCCTATTGGGCT; SEQ ID NO:499). The ~200 bp products were isolated from a 3% agarose gel and purified with Qiagen QIAquick spin columns. Each product from (a) and (b) above was digested with either BsrDI or BpmI (NEB) in separate reactions. BpmI-digested monomers have an overhang which can be ligated to BsrDI-digested monomers. The purified digestion products were ligated to one another using T4 DNA ligase (NEB). Ligation of BsrDI-cut naïve monomers with BpmI-cut selected monomers generates a walking dimer library comprised of N-terminal naïve monomers fused to C-terminal selected monomers. Ligation of BpmI-cut naïve monomers with BsrDI-cut selected monomers generates a walking dimer library comprised of C-terminal naïve monomers fused to N-terminal selected monomers. Primers pETF/pETB2r were used to PCR-amplify the ligated dimer coding sequences from the ligation, and the purified products were digested with SfiI followed by XmaI. The digested products were ligated to the phage vector fUSE5HA for the generation of a phage display dimer "walking library", typically with 10^8-10^9 unique members. A trimer (or larger multimer) "walking library" can be generated in a similar fashion, except that the starting materials are dimers (or larger) and naïve monomers. Walking libraries were panned against IL-6 and screened as described above.

VII. Characterization of purified monomers in binding and competition assays

[408] Once proteins were characterized at the heated protein lysate level, the best monomers were chosen for further characterization. Larger-scale cultures of individual clones were prepared and the monomers, which bear a 6xHis (SEQ ID NO:379) tag, were purified via Ni-NTA resin. These nickel-purified monomers were assayed in binding ELISAs and the AlphaScreen competition assay. Protein sequence data and biochemical data from characterization of purified monomers are in Tables 1 and 2.

[409] Protein Purification 500 mL Cultures for NiNTA: (Day 1) In a 15 mL culture tube containing 3 mL of 2xYT + 40 μg /mL kanamycin, the appropriate "primary hit

well" archived glycerol stock was inoculated. Culture was shaken overnight at 300 rpm at 37°C.

[410] (Day 2) 2 mL of overnight culture was inoculated into 1 L Erlenmeyer shake flask containing 500 mL of 2xYT + 40 μg/mL kanamycin. Cultures were grown with shaking at 375 rpm at 37°C until an OD_{600} of about 0.8-1.0 was reached. Once desired OD_{600} was reached, cultures were induced with 1 mM final concentration of IPTG for 3 hr while shaking at 375 rpm. After 3 hr induction, the 500 mL culture was transferred to clean/autoclaved Sorvall tube and centrifuged for 8 min at 8000 rpm at 4°C to pellet cells.

[411] Once cells were pelleted, supernatant was removed and discarded, and 20 mL of sonication buffer (10% sucrose/20 mM Tris [pH 7.5]/150 mM NaCl/0.2 mM $CaCl_2$) was added to each tube. The pellet was resuspended in sonication buffer with 10 mL serological pipet until there were no visible clumps, and then the resuspended cells (~ 30 mL) were transferred into 35 mL Oakridge Tubes and sonicated for 8 min at ~ 16 power output. After sonication, the warm Oakridge Tubes containing sonicated cells were placed in a ice/water bath for ~ 10 min to cool. Once cooled, tubes were centrifuged for 30 min at 18,000 rpm at 4°C to pellet lysed cells.

[412] While tubes containing lysed cells were being centrifuged, NiNTA resin (Qiagen) was washed with Milli-Q water to remove ethanol. 3 mL of 1:1 diluted NiNTA resin/protein was used (i.e. actually 1.5 mL of resin/protein was used). 3 mL of resin/water mix each was added to appropriately labeled (with protein ID) clean 50 mL screw cap tube. After sonicated cells were pelleted, protein supernatant was removed and added to 50 mL tube containing the 1.5 mL of washed NiNTA resin. Protein was allowed bind to NiNTA resin by rocking gently for 0.5 hr. @ RT. After incubation with NiNTA resin, 50 mL tubes with NiNTA bound to protein was centrifuged for 10 min at ~1500 rpm. Supernatant was gently poured out and discarded.

[413] NiNTA resin + bound protein was transferred to appropriately labeled 15 mL Clontech columns by adding 1 mL of NiNTA Wash Buffer (20 mM Tris [pH 7.5], 200 mM NaCl, 0.1 mM $CaCl_2$, 20 mM imidazole) to 50 mL tube containing resin, swirling to resuspend, then pipetting the mixture into a column which has been mounted on a vacuum manifold. NiNTA resin + bound protein was washed with at least 10 column volumes (15 mL) of NiNTA wash buffer. 15 mL columns containing NiNTA resin + bound and washed protein were transferred to a clean 15 mL screw cap collection tubes and 4 mL of Ni Elution buffer (20 mM Tris [pH 7.5], 200 mM NaCl, 0.1 mM $CaCl_2$, 200 mM imidazole) was added

to each column to elute off protein into the 15 mL collection tube. It was then allowed to elute by gravity.

[414] Eluted protein was transferred to slide-A-lyzer cassette (appropriate MW cutoff—for monomers used 3.5 kDa cutoff and for dimers and trimers used 10 kDa cutoff) using 18.5 gauge needle and 5 mL syringe to load cassette. Slide-A-lyzers containing eluted proteins were placed into overnight dialysis buffer containing redox reagents (20 mM Tris [pH 7.5], 100 mM NaCl, 1 mM $CaCl_2$, 1 mM 2-mercaptoethanol, 0.25 mM 2-hydroxyethyldisulfide).

[415] (Day 3) Slide-A-lyzer cassettes containing overnight dialyzed proteins were transferred into dialysis buffer without redox (20 mM Tris [pH 7.5], 100 mM NaCl, 1 mM $CaCl_2$). After 3 hr dialysis, slide-A-lyzer cassettes were transferred into fresh TBS/ $CaCl_2$ without redox for another 3 hr. After 2nd dialysis change, proteins were removed from slide-A-lyzer cassettes using 18.5 gauge needle and 5 mL syringe, and protein was transferred by filtering using 0.2 micron syringe filter into appropriately labeled 15 mL polypropylene tube.

The anti-IL-6 NiNTA purified proteins, which were selected as the "best inhibitors" in AlphaScreen competition assays or "best binders" in a binding ELISA, were further purified by Q-Sepharose anion exchange to remove contaminants. Q-Sepharose Purification: 1 mL of Q-Sepharose Fast-Flow Resin (Amersham Biosciences) was added to 15 mL Clontech column. Resin with 15 column volumes (or 15 mL) of 20 mM Tris [pH 7.5], 50 mM NaCl, 1 mM $CaCl_2$ was equilibrated. 2 mL (~ 5 mg) of filtered NiNTA-purified protein was added to resin and protein was allowed to bind to resin by gravity. Flow-through into first column of 96-well plate was collected. Columns loaded with protein were transferred to 15 mL collection tube and resin/bound protein were washed with 10 column volumes (or 10 mL) of 20 mM Tris [pH 7.5], 50 mM NaCl, 1 mM $CaCl_2$. Once washed, NaCl gradient elution of protein was started. NaCl concentration was varied in gradient as follows: 100 mM, 150 mM, 200 mM, 250 mM, 300 mM, 350 mM, 400 mM, 500 mM, and finally 1 M NaCl, to a base of 20 mM Tris [pH 7.5], 1 mM $CaCl_2$. Fractions were collected in 96-well deep-well polypropylene plate—2 mL/fraction, in 1 mL increments. Fractions containing protein were tested by Bradford and analyzed by SDS PAGE. Fractions were tested in binding ELISAs and competition assays as described above with the following change. Protein from 500 mL NiNTA purified preparations or NiNTA + Q-sepharose purified preparations was added to the plate as a twelve-point concentration curve starting with a 1:5 to 1:100 first dilution and then 1:4 serial dilutions thereafter with the last point as buffer only.

IL-6 binding monomers & dimers

[416] The following provides a summary of the IL-6 binding monomers identified, grouped by sequence homology. There are 7 families, wherein members of the same family have related sequences.

[417] The information can be summarized as follows. Sequences in brackets ("[]") indicate alternative amino acids at a single position. The consensus sequences below include question marks ("?"). These indicate positions that can be present or absent.

```
Family 1 motif (SEQ ID NO:27):
      CX[AS][DN]EFXCXXGXC[ILV]XXAXXCDGV[DN]DCXDNSDEXXC[AST]XX[AST]

Family 2 motif (SEQ ID NO:28):
      CXXXQFXC[DN]NGXCVPXXWXCDG[DE][DN]DCXD[DN]SDEXXCXXP[AST]

Family 3 motif (SEQ ID NO:29):
      CXXXQFQCXSG[W,Y]C[I,L,F]XXXW[V,L]CDGXNDCEDXSDEXXC

Family 4 motif (SEQ ID NO:30):
      CXXX[EQ]FXCXXGXC[ILV]XXXXXCDGX[DN]DCXDXSDEXX(XX)C

Family 5 motif (SEQ ID NO:31):
      CX[A,P]X[E,Q]FXCXXGQC[I,L]PXXXXCDGX[D,N]DCXDXSDEXXC

Family 6 motif (SEQ ID NO:32):
      CXXX[EQ]FXCX(X)SXXC[ILV]XX[DN]W[ILV]CDGVNDCXDXSDEXGCXXXT

Family 7 motif (SEQ ID NO:33):
      CL[AS]X[EQ]FXCXNGXCIP[KQR]XWVCDGDNDC[ILV]DXSDE(P)XXCX[EKQ]XT
```

[418] Details of each IL-6-binding family follows. Spacing (or dashes ("-") is only for alignment purposes. Each line represents one contiguous monomer domain.

Fam1 (SEQ ID NOS:500-519)

```
CPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT
CPSDEFRCRNGRCVPRAWRCDGVNDCADNSDEEDCTEHT
CPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTENT
CPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTQHT
CPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEGCTPHT
CPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEADCKAHT
CPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEADCAAPT
CPSDEFRCRNGRCIPRAWRCDGVNDCGDNSDETNCAPPT
CPSDEFRCRNGRCIPRAWRCDGVNDCGDSSDETSCSAPASEPPGSL
CPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEANCAGSVPT
CPASEFRCNNGRCIPAQWRCDGVNDCGDDSDEASCQAPT
CQANEFTCSSGNCLSPALLCDGVDDCGDNSDETNCASPAPT
CQANEFTCSSGNCLSPALLCDGVDDCGDNSDETYCASPAPT
CQANEFTCSSGNCLSPALLCDGVDDCGDNSDETYCAGPAPT
CELNEFTCSSGNCLSPALLCDGVDDCGDNSDETNCASPAPT
CGANEFTCDNGRCLSPDLVCDGVDDCADSSDETDCQKPM
CRPDEFKCENGHCIPPGLRCDGVDDCVDNSDETDCETPGRT
CPSDEFRCRNGRCIPPGWLCDGVDDCVDDSDEEGCRTPART
CGSNEFTCSNGRCIPPTLLCDGVNDCGDDSDETGCATPVRT
CASSQFPCDNGKCLPPPLVCDGVDDCGDDSDEPSAPCGSPART
```

score	matches	expected	motif (SEQ ID NOS:520-528)
3738	47	10^{-16}	c...[eq]f.c..g.c[ilv].....cdgv[dn]dc.d.sde.....
3963	46	10^{-18}	c...ef.c..g.c[ilv].....cdgv[dn]dc.d.sde..c.....
4029	45	10^{-19}	c..[dn]ef.c..g.c[ilv].....cdgv[dn]dc.d.sde..c.....
3995	43	10^{-20}	c..[dn]ef.c..g.c[ilv].....cdgv[dn]dc.d[dn]sde..c.....
3922	41	10^{-21}	c.[as][dn]ef.c..g.c[ilv].....cdgv[dn]dc.d[dn]sde..c.....
3919	40	10^{-22}	c.[as][dn]ef.c..g.c[ilv]..[ast]..cdgv[dn]dc.d[dn]sde..c.....
3920	39	10^{-23}	c.[as][dn]ef.c..g.c[ilv]..a..cdgv[dn]dc.dnsde..c.....
3907	38	10^{-23}	c.[as][dn]ef.c..g.c[ilv]..a..cdgv[dn]dc.dnsde..c[ast].....
3890	37	10^{-24}	c.[as][dn]ef.c..g.c[ilv]..a..cdgv[dn]dc.dnsde..c[ast]..[ast]....

Fam2 (SEQ ID NOS:529-532)

CLSSQFQCDNGNCVPQQWVCDGDNDCLDDSDETSCSAPASEPPGSL
CLADQFQCNNGRCVPQSWGCDGDDDCQDNSDETSCAPPT
CAPGEFTCGSGRCVPQTWRCDGVDDCLDSSDETDCSAPASEPPGSL
CPPGQFTCNNGRCVPPTWLCDGEDDCQDDSDEKNCKAPT

score	matches	expected	motif (SEQ ID NOS:533-534)
402	4	10^{-23}	c...[eq]f.c..g.cvp..w.cdg.[dn]dc.d.sde..c..p[ast]
349	3	10^{-28}	c...qf.c[dn]ng.cvp..w.cdg[de][dn]dc.d[dn]sde..c..p[ast]

Fam3 (SEQ ID NOS:535-554)

CRASQFQCGSGYCIPRPWLCDGDNDCEDDSDETNCTPRT
CLPSQFQCRNGRCIPLPWLCDGDNDCEDDSDETNCTPRT
CEADQFQCGSGYCLPRHWVCDGDNDCEDDSDETDCGRPGPGATSAPAV
CLPSQFQCGSGYCISQRWVCDGEHDCEDSSDEANCAGSVPT
CLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPT
CPSSQFQCSSGWCISQRWVCDGENDCEDSSDEANCAGSVPT
CPSSQFQCSSGWCISQRWVCDGENDCEDGSDEANCAGSVPT
.....CSSGWCISQRWVCDGENDCEDSSDEANCAGSVPT
CPSSQFQCGSGWCISQHWVCDGENDCEDGSDEANCAGSVPT
CLPSQFQCGSGWCISQRWVCDGENDCEDSSDEANCAGSVPT
CLPSQFRCGSGCCFSQRWVCDGKNDCEDSSDEANCAGSLPT
CGASQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPT
CLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGSVPT
CAPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGSVPT
CLPDQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPT
CLPGQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGSVPT
CLPDQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGSVPT
CLPDQFQCGSGYCLSQRWVCDGENDCEDGSDEANCAGSVPT
CLPDQFQCGSGYCLPQRWVCDGENDCEDSSDEANCAGSVPT
CLPDQFQCGSGYCLPQRWVCDGENDCEDGSDEANCAGSVPT

score	matches	expected	motif (SEQ ID NOS:555-562)
6106	61	10^{-23}	c...qf[kqr]c..g.c[filmv]...w[ilv]cdg..dced.sde[ast][dn]c.....
6399	60	10^{-25}	c...qf[kqr]c.sg.c[filmv].[kqr].w[ilv]cdg..dced.sde[ast][dn]c.....
6821	59	10^{-27}	c...qfqc.sg[fwy]c[ilv].[kqr].w[ilv]cdg[de].dced.sde[ast][dn]c.....
6968	58	10^{-29}	c...qfqc.sg[fwy]c[ilv].[kqr].w[ilv]cdg[de]ndced.sde[ast][dn]c.....
7842	57	10^{-34}	c...qf[kqr]c.sg.c[filmv].q.wvcdg[ekq].dced.sdeancags[ilv]pt
8205	56	10^{-37}	c...qfqc.sg[fwy]c[ilv].q.wvcdge.dced.sdeancagsvpt
8307	55	10^{-38}	c...qfqc.sg[fwy]c[ilv].q.wvcdgendced.sdeancagsvpt
8387	54	10^{-39}	c...qfqc.sg[fwy]c[ilv].qrwvcdgendced.sdeancagsvpt

Fam4 (SEQ ID NOS:563-576)

CLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETN..CPART
CLPDQFQCGSGYCLPQRWVCDGVDDCEDSSDEKG..CTPHT
CLSDQFQCGSGHCLPASWRCDGEDDCADSSDETS..CTASEPT
CPSSQFQCSSGWCIPEDWRCDGVPDCGDGSDEASDH
CAPSQFQCGSGYCIPLTWVRDGVDDCEDGSDEED..CSQDPEFHKV
CLPNQFTCDSGNCVPQALGCDGDNDCEDNSDESPALCAAPAHT
CLPNEFTCDSGNCVPQALGCDGDNDCEDNSDESPALCAAPAHT
CAPNQFQCGSGYCVPAAWVCDGENDCEDNSDESSALCQERT
CLPDQFPCENGKCLPVNLVCDGDDDCPDGSDESEV.CAGAVRT
CPSSQFRCNNGHCLSVTLVCDGDNDCEDDSDETG..CEGSGHT
CQPGEFTCESGRCLPLRWLCDGEDDCVDSSDEAS..CQPPT
CLSGEFTCENGNCLSLSLVCDGEDDCEDGSDEAS..CPPPT
CPPDEFQCNSGSCLSQPLLCDGEDDCGDGSDEES..CIAAEHT
CQPGQFRCNNGHCLSQPLVCDGVPDCEDGSDEAS..CSAPASEPPGSL

score	matches	expected	motif (SEQ ID NOS:577-579)
994	15	10^{-12}	c...[eq]f.c..g.c[ilv].....dg..dc.d.sde.....
1011	14	10^{-14}	c...[eq]f.c..g.c[ilv].....cdg..dc.d.sde.....
907	12	10^{-15}	c...[eq]f.c..g.c[ilv].....cdg.[dn]dc.d.sde.....

...

Fam5 (SEQ ID NOS:580-591)

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRT
CLPDQFRCGNGQCIPLDWVCDGVNDCGDNSDEAGCPTRT
CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDEAGCPPRT
CLPDQFRCGNGQCIPLDWVCDGVNDCVDGSDEAGCPPRT
CLPDQFRCGNGQCIPLDWVCDGVNDCEDSSDEAGCPPRT
CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPRT
CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPHT
CLPDQFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPPRT
CVPDQFPCGNGQCIPLDWVCDGVDDCEDNSDEKSCPPPR
CLPGQFRCDNGQCIPAHWGCDGVNDCRDNSDEADCPQPT
CQASEFPCDSGQCIPQPLGCDGDNDCADNSDEESCQPRT
CEPGQFPCESGQCLPPPWVCDGVNDCADNSDEEDCGDSHILPFSTPGPST

score	matches	expected	motif (SEQ ID NOS:592-597)
2083	23	10^{-20}	c...[eq]f.c..gqc[ilv]p....cdg.[dn]dc.d.sde..c..... .....
2345	22	10^{-25}	c.p.qf.c..gqc[ilv]p..w.cdgv[dn]dc.d.sde..c..... ....
2510	21	10^{-28}	c[ilv]p.qf.c.ngqcip..w.cdgv[dn]dc.d.sde..cp..... ....
2784	20	10^{-34}	c[ilv]pdqf.cgngqcipldwvcdgv[dn]dc.d.sde..cp..... ....
2850	19	10^{-38}	clpdqfrcgngqcipldwvcdgvndc.d.sde..cp..t.....
2777	18	10^{-39}	clpdqfrcgngqcipldwvcdgvndc.d.sde..cp.rt.....

Fam6 (SEQ ID NOS:598-609)

CLPSEFQCRDSKICLSADWLCDGVNDCADNSDEEGCPPPT
CLPSEFQCRDSKICLSADWLCDGVNDCVDGSDEAGCPPRT
CEANELRCS-SGRCLSVHWLCDGVDDCEDSSDEAGCPPRT
CPASQFPCS-SGHCVPLNWVCDGVNDCVDGSDEAGCPPRT
CESDQFQCD-SGHCVPLNWVCDGVNDCVDGSDEAGCPPRT
CPSDEFTCD-SGRCVSLNWLCDGVNDCGDNSDEAGCPTRT
CLSDEFTCD-SGRCVSLNWLCDGVNDCGDNSDEAGCPTHT
CPSSQFRCS-SGRCVPADWLCDGVNDCADNSDETGCPTRT
CPSSQFRCS-SGRCVPADWLCDGVNDCADNSDETGCATRT
CPSSQFRCS-SGRCVPADWLCDGVNDCGDNSDEAGCPTRT
CRADEFQCG-SGRCVPADWLCDGVNDCADNSDETGCPTRT
CVSSQFQCG-NGHCVPAQWLCDGVNDCGDNSDEASCPPRT

score	matches	expected	motif (SEQ ID NOS:610-615)
2122	23	10^{-20}	c...[eq][fly].c..?...c[ilv]...w[ilv]cdgv[dn]dc.d.sde. .c...t
2199	22	10^{-23}	c...[eq][fly].c..?s..c[ilv]...w[ilv]cdgv[dn]dc.d.sde. gc...t
2236	21	10^{-25}	c...[eq]f.c..?s..c[ilv]..[dn]w[ilv]cdgvndc.d.sde.gc.. .t
2217	20	10^{-26}	c...[eq]f.c..?s..c[ilv]..[dn]w[ilv]cdgvndc.d.sde.gcp. .t
2226	19	10^{-28}	c.[as].[eq]f.c..?sg.cv..[dn]w[ilv]cdgvndc.d.sde[ast]g c...t
2187	18	10^{-29}	c.[as].[eq]f.c..?sg.cv..[dn]w[ilv]cdgvndc.d.sde[ast]g cp..t

Fam7(SEQ ID NOS:616-622)

CLSSEFTCGNGNCIPRPWVCDGDNDCVDDSDE-KGCPEPT
CLANQFKCDNGRCIPQTWVCDGDNDCLDGSDEPVHCAKRT
CLANQFKCDNGRFIPQTWVCDGDNDCLDGSDEPVHCAKRT
CLANQFKCDNGRCIPQTWVCDGDNDCLDGSDEPVHCAKRT
CLANQFKCDNGRCIPQTWVCDGDNDCLDGSDEPVRCAKRT
CLANQFKCDNGRCIPQTWVCDGDNDCLDGSDEPVHCAERT
CLSSEFTCGNGNCIPQTWVCDGDNDCLDGSDEPVHCAKRT

score	matches	expected	motif (SEQ ID NOS:623-626)
1319	11	10^{-29}	cl[as].[eq]f.c.ng..ip[kqr].wvcdgdndc[ilv]d.sde.?..c.[ekq].t
1258	10	10^{-30}	cl[as].[eq]f.c.ng.cip[kqr].wvcdgdndc[ilv]d.sde.?..c.[ekq].t
1286	9	10^{-36}	cl[as].[eq]f.c.ng..ipqtwvcdgdndcldgsde.?v.ca[ekq]rt
1299	8	10^{-41}	clanqfkcdngr.ipqtwvcdgdndcldgsde.?v.ca[ekq]rt

IL-6-binding dimers included the following (SEQ ID NOS:627-781). Dimers are not organized by family.

CLPNQFQCRSGQCIPRPWVCDGDNDCGDDSDEANCKAPVHTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEAN
CGRTEPT

CLSSEFTCGNGNCIPRPWVCDGDNDCVDDSDEKGCPEPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CLSSQFQCSNGNCIPAAWLCDSDNDCGDDSDETSCPAPTCPSSQFQCSSGWCIPEDWRCDGVPDCGDGSDEASDH
CGDSHILPFSTPGPST

CLASEFQCRNGSCIPLQWLCDGDNDCGDDSDEENCATTERTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEAN
CAGSVPT

CLSSEFTCGNGNCIPRPWVCDGDNDCVDDSDEKGCPEPTCPSSQFQCSSGWCIPEDWRCDGVPDCGDGSDEASDH
CGDSHILPFSTPGPST

CGASQFQCHNTSICIPRPFVCDGDNDCGDDSDEADCKTSGHTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEA
NCAGSVPT

CLSSEFTCGNGNCIPRPWVCDGDNDCVDDSDEKGCPEPTCLPDQFQCGSGYCLPQRWVCDGVDDCEDSSDEKGCT
PHT

CPSSQFQCGSGYCISPAWVCDGVNDCEDGSDEKSCAPHTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETNCP
ART

CAPSQFTCGNGNCVPRPWLCDGDNDCEDDSDETSCPPHTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETNCP
ART

CAPSQFTCGNGNCVPRPWLCDGDNDCEDDSDETSCPXHTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETNCP
ART

CLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCRPGQFQCGSGYCIPRPWRCDGDNDCEDDSDEKG
CARSEPT

CLPSQFQCRNGRCIPLPWLCDGDNDCEDDSDETNCTPRTCAPSQFQCGSGYCIPLTWVCDGVDDCEDGSDEEDCE
TAAPT

CLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCRASQFQCGSGYCIPRPWLCDGVNDCEDGSDEAD
CQTRT

CVPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETN
CPART

CLPSQFQCRNGRCIPLPWLCDGDNDCEDDSDETNCTPRTCAPNQFQCGSGYCVPAAWVCDGENDCEDNSDESSAL
CQERT

CAPSQFTCGNGNCIPRPWLCDGDNDCEDDSDETSCPXPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETNCP
ART

CLSSEFTCGNGNCIPRPWVCDGDNDCVDDSDEKGCPEPTCVPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CAPNQFQCGSGYCVPAAWVCDGENDCEDNSDESSALCQERTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEAN
CAGSVPT

CGANEFQCGSGYCLPVGWVCDGENDCEDNSDEKGCPPHTCLPNQFQCRNGRCIPLPWLCDGDNDCEDDSDATNCT
PRT

CAPNQFQCGSGYCVPAAWVCDGENDCEDNSDESSALCQERTCVPSQFQCGSGYCISQRWVCDGENDCEDSSDEAN
CAGSVPT

CLPDQFQCGSGYCLPQRWVCDGVDDCEDSSDEANCAGSVPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETN
CPART

CAPSQFQCGSGYCIPLTWVCDGVDDCEDGSDEEDCETAAPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEAN
CAGSVPT

CLPSQFQCRNGRCIPLPWLCDGDNDCEDDSDETNCTPRTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CAPSQFTCGNGNCVPRPWLCDGDNDCEDDSDETSCPPHTCAPNQFQCGSGYCVPAAWVCDGENDCEDNSDESSAL
CQERT

CLPSEFQCHSTSICVPRQWLCDGDNDCGDDSDEKGCETHTCVPSQFQCGSGYCISQRWVCDGENDCEDSSDEANC
AGSVPT

CPPDQFQCGSGYCLPQAWVCDGENDCEDGSDETNCPARTCLSSEFTCGNGNCIPRPWVCDGDNDCVDDSDEKGCP
EPT

CAPSQFTCGNGNCVPRPWLCDGDNDCEDDSDETSCPPHTCAPSQFQCGSGYCIPLTWVCDGVDDCEDGSDEEDCE
TAAPT

CLPSQFQCRNGRCIPLPWLCDGDNDCEDDSDETNCTPRTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETNCP
ART

CLSSEFTCGNGNCIPRPWVCDGDNDCVDDSDEKGCPEPTCLPDQFQCGSGYCLPQRWVCDGVDDCEDSSDEKGCT
PHT

CAPSQFQCGSGYCIPLTWVCDGVDDCEDGSDEEDCETAAPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETN
CPART

CAPNQFQCGSGYCVPAAWVCDGENDCEDNSDESSALCQERTCLSSQFQCDNGNCVPQQWVCDGDNDCLDDSDETS
CSAPASEPPGSL

CLSSEFTCGHGICIPRPWVCDRDNDCMYHSHEISCPEPTCLPDQFQCGSVYCLPERLVCDCLNDCEDSSDEAXCA
GSXPT

CLPSQFQCRNGRCIPLPWLCDGDNDCEDDSDETNCTPRTCVPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CVPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCGWWGSSEFRCATYPHGRCIPSSXVCDGEDDCGD
GSDEHEVFCGGGG

CAPSQFQCGSGYCIPLTWVCDGVDDCEDGSDEEDCETAAPTCLPSEFQCHSTSICVPRQWLCDGDNDCGDDSDEK
GCETHT

CVPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCLSSEFTCGNGNCIPRPWVCDGDNDCVDDSDEKG
CPEPT

CLPSEFQCHSTSICVPRQWLCDGDNDCGDDSDEKGCETHTCLASQFTCRNGNCIPEPWVCDGDNDCVDDSDEESC
PPHT

CEADQFQCGSGYCLPRHWVCDGDNDCEDDSDETDCGRPGPGATSAPAVSGGSGGSGGSCGS

CRASQFQCGSGYCIPRPWLCDGVNDCEDGSDEADCQTRTCVPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CAPNQFQCGSGYCVPAAWVCDGENDCEDNSDESLAHCQAHTCVSSEFRCSSGQCIPPPWRCDGVDDCADNSDEAG
CTPPT

CQPGQFPCGSGSCVSERLLCDGVDDCADGSDEAGCPPRTCVPSQFQCGSGYCISQRWVCDGENDCEDSSXEANCA
GSVPT

CHATAQFECSSGQCISLDWLCDGVNDCADGSDESQHCPPRTCVPSQFQCGSGYCISQRWVCDGENDCEDSSDEAN
CAXSVPT

CEASQFPCRNSERCVPPHLLCDGVNDCADGSDEIGCPPHTCVPSQFQCGSGYCISQRWVCDGENDCEDSSDEANC
AGSVPT

CPSNQFQCRNNGTCIPLPWLCDGVDDCGDDSDETSCTEPTCVPSQFQCGSGYCISQRWVCDGENDCEDSSDEANC
AGSVPT

CQSSEFTCSSGRCLSAGWLCDGVNDCRDDSDEESCTAPVPTCVPSQFQCGSGYCISQRWVCDGENDCEDSSDEAN
CAGSVPT

CQPGQFQCHDYGRCVSQNWLCDGDPDCLDSSDEANCQPPTCVPSQFQCGSGYCISQRWVCDGENDCEDSSDEANC
AGSVPT

CPAGQFQCRTGRCIPLALLCDGENDCQDGSDEAGCPTPTCVPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CGAGEFPCDNGRCVPAALLCDGDDDCGDNSDEKDCAQPTCVPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CRSSEFRCDSGHCLPLPWVCDGDDDCEDGSDEAQDCEPRTCVPSQFQCGSGYCISQRWVCDGENDCEDSSDEANC
AGSVPT

CRSNQFQCGSGYCIPVKWLCDGENDCEDGSDESLAHCQAHTCVASQFRCDNGRCLSEPLGCDGVNDCPDNSDEED
CTTAEHT

CRADQCQCNSGHCLSRNLRCDGEDDCADNSDEKSCSQDPEFHKVCLPDQFQCGSGYCLPQRWVCDGVDDCEDSSD
EKGCTPHT

CLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCRSSQFQCGNGRCISVHWVCDGDDDCEDSSDESS
AHCKPRT

CPSNEFRCNNGHCIPRAWRCDGDNDCGDSSDETDCTQRTCRSNQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CPADEFPCKNGKCLPLPWVCDGDNDCADNSDETDCTATAPTCRSNQFQCGSGYCISQRWVCDGENDCEDSSDEAN
CAGSVPT

CGSDEFLCNNGKCIPPTWGCDGVNDCQDNSDEAPALCPPPTCRSNQFQCGSGYCIPQRWVCDGVDDCEDSSDEKG
CTPHT

CAPNQFQCGSGYCVPAAWVCDGENDCEDSSDEANCAGSVPTCRSGQFQCGSGYCISQRWXCDGENDCEDSSDEAN
CAGSVPT

CLPGQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCRPSQFPCSNGQCVPQTWVCDGVNDCQDDSDETD
CSAPASEPPGSL

CLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCRPSEYPCGNGRCVPQPLRCDGVNDCEDSSDEAN
CTPHT

CRSNQFQCGSXYCIPVKWLCDGENDCEDGSDEADCAGSVPTCRPDQFPCDNGNCIPGPWGCDGVNDCLDSSDETD
CTPPT

CPSSQFKCNNGSCVPQDWLCDGVNDCVDGSDEPAVCQQRTCRASQFQCGSGYCIPRPWLCDGVNDCEDGSDEADC
QTRA

CASGEFRCGNGRCVPQRWLCDGVPDCEDNSDETDCQPPTCRASQFQCGSGYCIPRPWLCDGVDDCEDSSDEKGCT
PHT

CLPSQFQCGSGYCIPQRWVCDGENDCEDSSDEANCAGSVPTCQSSEFPCSSGHCLPLALGCDGEDDCADNSDEES
CEPRT

CRASQFQCGSGYCIPRPWLCDGVNDCEDSSDEANCAGSVPTCQPSQFRCRSGHCISGALGCDGEPDCGDGSDETG
CRASAHT

CLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCQPDQFRCGSGPLRSTALACDGXKXLSGQLGXKK
AVEPPDRT

CLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCQAGEFRCGNGRCLPPDWRCDGVDDCGDGSXEAS
CKTPT

CQSDEFTCRDSNKCVSRDWLCDGEDDCPDGSDETSCAGSERTCPSSQFQCSSGWCIPEDWRCDGVPDCGDGSDEA
SCPPHT

CVPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCPSNEFPCESGHCVPLTWRCDGEDDCPDSSDEAN
CGRPGPGATSVPAA

CRSNQFQCGSGYCLPQAWVCDGENDCEDGSDETNCPARTCPSGQFRCDNGKCLSEAWVCDGVDDCRDGSXESEHC
APPT

CRSNQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCPSGQFQCRGTGKCIPQPWVCDGDNDCQDDSDEK
DCSAPASEPPGSL

CLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCPSGEFRCGNGSCLPAGWVCDGVDDCVDGSDEKG
CQTPT

CVPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGXNDCADNSDEED
CTEHT

CVPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEED
CTEHT

CRSNQFQCGSGYCIPVKWLCDGENDCEDSSDEESCPXRTCPPSEFRCKGSGKCVPQHWLCDGVNDCEDDSDEKSC
PPPT

CRSNQFQCGSGYCIPVKWLCDGENDCEDSSDEESCPARTCPPSEFRCKGSGKCVPQHWLCDGVNDCEDDSDEKSC
PPPT

CLSSEFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCPADEFRCSSGRCVPATWLCDGVDDCEDSSXESL
ETCTPHT

CRSNQFQCGSGYCIPVKWLCDGENDCEDGSDEADCQTRTCLSSEFQCSNGQCVSQRLLCDXVDDCLDDSDEEDCS
APASEPPGSL

CQANQFRCNNGRCLSPAWLCDGVDDCPNGSDEEGCPPHTCLSSEFQCGSXYCISQRWVCDXENDCEDSSDEANCA
GSVPT

CESGEFTCGSGQCLPATWRCDGDNDCADDSDETNCPQPTCLSSEFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CPPGQFRCRSGNCLPQTWVCDGDDDCEDDSDESPALCSQDPEFHKVCLPSQFQCGSVYCISQRWVRDGXNDCEDS
SDEANCAGSVPT

CQSGQFPCGNGNCIPRPLRCDGVDDCPDGSDEPPECPQPTCLPSQFQCGSGYCISRRWVCDGENDCEDSSDEKGC
TPHT

CQPDEFTCHDKNRCISLALLCDGVNDCEDDSDEEGCAPRTCLPSQFQCGSGYCISQRWVCDXENDCEDSSDEANC
AGSVPT

CVSSQFKCSNGRCLPLGWLCDGVDDCADGSDEKNCPPPTCLPSQFQCGSGYCISQRWVCDXENDCEDSSDEADCA
GSVPX

CVPNQFPCGNGHCLPRDWLCDGVNDCGDSSDESPAHCTEPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSXXAN
CAGSVPT

CRAGQFRCSSGHCLPRRLLCDGEDDCVDGSDETGCTQRTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEKGCT
PHT

CQPNQFRCSNGRCVPLNLLCDGVPDCEDGSDEEGCPPRTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CRSSQFRCSSGQCVSLGWVCDGVNDCPDSSDETGCPKRTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CHAIAQFGCKSGRCLSLGLLCDGVDDCADGSDEAGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANC
AGSVPT

CPPSQFTCSNGNCLSPTLLCDGVNDCVDDSDETSCAASVHTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEAN
CAGSVPT

CLADQFKCHNTNICIPGRLGCDGDPDCEDDSDEKGCPPHTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANC
AGSVPT

CPPGEFPCSSGHCLSQALRCDGVPDCEDDSDETGCPPPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CGANEFTCGNGRCLSPTWLCDGDDDCGDNSDEASCGGTAPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEAN
CAGSVPT

CLPNEFPCGSGHCIPQAWVCDGENDCQDSSDEAPAHCAAPVPACLPSQFQCGSGYCISQRWVCDGENDCEDSSDE
ANCAGSVPT

CPPSQFKCSNGRCLPVQLLCDGVPDCPDGSDETGCPPPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CLASQFRCGNGSCLSQTLLCDGVNDCADNSDEVSCPPPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CLPSQFQCGSGYCISQRWVCDGENDCEDGSDEESCTGPGHTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEAN
CAGSVPT

CPPGEFQCNSGHCLPGTWGCDGEDDCPDGSDESQDCAPPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANC
AGSVPT

CPSSEFKCNNGQCLPVRWLCDGVNDCLDSSDEASCAARTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CHPPSQFRCSNGQCLPLTWRCDGVNDCGDNSDEEDCPERTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANC
AGSVPT

CLANEFRCSSGQCIPARWLCDGVNDCVDDSDEKDCPARTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CLPGQFRCHSTDTCLPESWVCDGVNDCADDSDEASCPPHTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANC
AGSVPT

CLSSQFRCDSGRCISQDWLCDGVNDCQDGSDESLAICPPHTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEAN
CAGSVPT

CPAGQFTCGSGKCVSGGLLCDGDNDCQDGSDEAGCATRTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CAPDQFKCGNGNCLPLPWVCDGDNDCGDNSDEANCPQHTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CAPSQFRCESGQCLPANWRCDGENDCQDGSDESPAICEGTAPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDE
ANCAGSVPT

CPSNQFTCRSGSCLSLAWLCDGVNDCPDNSDEADCAERTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CPSSEFPCHGNSRCIPLTWLCDGDNDCGDDSDEKNCEAAGPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEA
NCAGSVPT

CQPSQFRCGNGHCLPPTLVCDGENDCRDNSDETSCATAVRTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEAN
CAGSVPT

CRADEFRCSSGHCIPAPWGCDGVNDCEDGSDEAGCPARTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CAANEFRCRSTNTCIPQQWLCDGVNDCQDDSDEKSCRTPAPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEA
NCAGSVPT

CAPSQFPCSSNDICVPRRWRCDGVDDCGDSSDEASCGTSVRTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEA
NCAGSVPT

CLPDQFPCSNGRCLPPTWVCDGENDCQDGSDEADCTEPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CLPNQFRCRNGQCLPQNWLCDGVNDCGDSSDETNCEPHTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CQPSQFQCNNGHCLSVRWLCDGVDDCQDGSDETDCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CRSNEFTCSSGSCXPLAWVCDGVNDCLDDSDEAGCPAPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CLSGEFKCNSGHCIPGPWLCDGDDDCPDGSDETDCGSPGPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEAN
CAGSVPT

CPADQFRCSNGRCIPRTWVCDGVDDCGDGSDESPLCESAAHTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEA
NCAGSVPT

CVPSQFQCRNNGICIPQNWHCDGVPDCEDSSDEESCGTTVHTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEA
NCAGSVPT

CLPSEFPCDSGQCLPLAWLCDGVDDCRDSSDESLEICPKHTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEAN
CAGSVPT

CPPGEFKCNSGNCVSPTWGCDGVDDCGDDSDEASAICEAPEHTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDE
ANCAGSVPT

CPPGQFPCSNGRCVPQRLRCDGDPDCEDNSDEAGCATPEPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEAN
CAGSVPT

CVANEFRCGNGHCVPLAWLCDGVPDCVDDSDETSCQARTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPA

CQPGEFRCRNGRCLPLAWLCDGVDDCVDGSDETNCAARTCLPSQFQCGSGYCIPQRWVCDGENDCEDSSDEANCA
GSVPT

CGASEFRCDSGHCISPDWRCDGDNDCVDGSDEKSCEQRTCLPSQFQCGSGYCFSQRWVCDGENDCEDGSDEANCA
GSVPT

CRADEFTCNNGRCVPQTWVCDGDPDCEDSSDEKGCTASVHTCLPSQFQCGSGYXISXXXVCXXENDCEDSSDEAN
CAGSVPT

CLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCLPSEFQCGNGRCLPLRLLCDGDDDCGDNSDEPL
ATCETSAPT

CRSNQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCLPGQFRCNSGQCLPLQWACDGVDDCADNSDEAS
CEPPT

CAAGQFQCHSTSRCVSLRWVCDGEDDCRDGSDEEGCTTAVPTCLPDQFQCGSGYCLPQRWVCDGVDDCEDSSDEK
GCTPHT

CPPSQFPCNSGNCVSAAWVCDGVDDCGDDSDESLAICPTSTCLPDQFQCGSGYCLPQRWVCDGVDDCEDSSDEKG
CTPHT

CRASEFRCRSYERCISPRWVCDGVNDCGDNSDETGCPPRTCLPDQFQCGSGYCLPQRWVCDGVDDCEDSSDEKGC
TPHT

CPLDQFTCGNGSCLPANLLCDGVNDCEDNSDEAGCPTRTCLPDQFQCGSGYCLPQRWVCDGVDDCEDSSDEKGCT
PHM

CPAGQFPCDSGRCVPQPWRCDGVNDCEDGSDEAPATCPEHTCLPDQFQCGSGYCLPQAWVLRXRKXLXGQLGXSK
TVQAPYLT

CRPSQFQCQNSSRCIPLGWRCDGDNDCEDGSDEKDCTATARTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDET
NCPART

CLANQFRCGNGRCVSPSWLCDGVDDCGDDSDEASCAAAEPTCLPDQFQCGSGYCLPQAWVCDGDNDCEDGSDETN
CPART

CRPNQFPCGNGHCLPGRWLCDGDDDCPDGSDEEGWPPPTCLPDQFQCGSGYCLPQRWVCDGVDDCEDSSDEKGCT
PHT

CPANEFTCDNGHCLPPSLLXDGENDCGDDSDETSCAAPGPTCLPSQFQCGSGYCISQRWVCDXENDCEDSSXEAN
CAGSVPT

CRSSQFTCNSGSXLPGHLVCDGVDDCEDNSDEENCAGPGRTCLSSEFQCGSGYCISQRWVCDXENDCEDSSNEAN
CAGSVPT

CQANEFTCSSGNCLSPALLXDGVDDCGDNSDETNCASPAPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEAN
CAGSVPT

CEAGEFRCRNGSCIPEAWGCDGVNDYLDNSDETNCPAHTCLSSEFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CPSGQFRCDNGHCLSPPLLXDGVNDCGDDSDETGCTQRTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CPPSEFRCHSTNTCLSPTLLXDGDNDCGDGSDEASCPQRTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANC
AGSVPT

CVPDEFPCNSGQCVSLAWLCXGVNDCRDGSDEAPYWEPPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANC
AGSVPT

CASNEFPCSNGNCIPGPLRCDGVNDCEDNSDESPLWPLHTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANC
AGSVPT

CQSDQFQCNSGSCIPPPWRXDGVDDCRDSSDEEGCPPRTCLPSQFQCGSGYCISQRWFCDXENDCEDSSXEANCA
GSVPT

CEANQFPCNSGSCIPPRWLXDGVDDCGDNSDEENCTRPTCLPSQFQCGSAYCISQRWVCDGENDGEDSSDEANCA
GSVPT

CLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCGGSEFTCGNGRCLPQHWFCDGENDCLDDSDEAG
CPAPT

CASDQFQCNGSRKCISQSWLCDGVNDCPDSSDEKSCPPHTCGANEFQCGSGYCLPRPWVCDGEDDCEDGSDESLA
HCQAHT

CPANQFRCGNGSCLSANWLCDGVNDCLDNSDEKNCGGAGHTCGANEFQCGSGYCLPRPWVCDGEDDCEDGSDESL
AHCQAHT

CPSGQFRCGSGRCIPQHWRCDGDDDCRDGSDEASCAPPTCAPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPT

CQPFDEFECRNGSCIPLPWRCDGVNDCLDDSDEEDCTGSAPTCAPSQFQCGSGYCIPQRWVCDGVDDCEDSSDEK
GCTPHT

CHPPNKFGCSNGRCVSPALLCDGVDDCQDNSDEKGCPPHTCAPNQFQCGSGYCVPAAWVCDGENDCEDSSDEKGC
TPHT

CLPGQFRCNSGRCVPVVLVCDGVDDCADDSDEAGCPPPTCAPNQFQCGSGYCVPAAWVCDGENDCEDSSDEANCA
GSVPT

CAPGQFPCGNGQCLPLHWVCDGVNDCGDNSDEAGCPPHTCAPNQFQCGSGYCVPAAWVCDGENDCEDNSDESSAL
CQERT

CPPNQFPCGSGQCLPQPLVCDGDDDCGDGSDELQLCEQPTCAPNQFQCGSGYCVPAAWVCDGENDCEDNSDESSA
LCQERT

CRSNQFQCGSGYCIPVKWLCDGENDCVDGSDEESCPARTCAPDEFRCDSGRCVPRTLGXATGDDDCEGRARMKKT
VERPNPT

CRANQFQCGSGYCIPRPWLCDGENDCEDGSDETNCPARTCAANEFTCGNGHCLPPTWACDGVNDCQDGSDEASCQ
TPT

The following lists IL-6-binding trimers, listed by families of related sequences. Note that the family numbers for monomers are not related to family numbers for trimers.

Fam1 (SEQ ID NOS:782-788)

CRASQFQCGSGYCIPRPWLCDGDNDCEDDSDETNCTP--------RTCLPDQFRCG-NGQCIPLDWVCDGVNDCADNSDEAGC--PPRTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPT
CLPDQFPCENGKCLPVNLVCDGDDDCPDGSDESEVCAGAV-----RTCLPDQFRCG-NGQCIPLDWVCDGVNDCADNSDETGC--PTRTCLPSQFQCGSGWCISQRWVCDGENDCEDSSDEANCAGSVPT
CLSSQFQCDNGNCVPQQWVCDGDNDCLDDSDETSCSAPASEPPGSLCPASQFPCS-SGHCVPLNWVCDGVNDCVDGSDEAGC--PPRTCLPDQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPT
CQASEFPCDSGQCIPQPLGCDGDNDCADNSDEESCQPRT--------CPASQFPCS-SGHCVPLNWVCDGVNDCVDGSDEAGC--PPRTCLPDQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPT
CLPNQFTCDSGNCVPQALGCDGDNDCEDNSDESPALCAAP---AHTCQANEFTCS-SGNCLSPALLCDGVDDCGDNSDETYCAGPAPTCLPSQFRCGSGCCFSQRWVCDGKNDCEDSSDEANCAGSLPT
CLPNEFTCDSGNCVPQALGCDGDNDCEDNSDESPALCAAP---AHTCQANEFTCS-SGNCLSPALLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPT
CGASQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVP-----TCLPSEFQCRDSKICLSADWLCDGVNDCADNSDEEGC--PPPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETNCP--ART

score	matches	expected	motif (SEQ ID NOS:789-791)
2137	7	10^{-84}	c...[eq]f.c..g.c[ilv].....cdg[de][dn]dc.d.sde.....?.?.?.?.?.?.?.?.?c...[eq]f.c..?...c[ilv]....[ilv]cdgv[dn]dc.d.sde..c.?.?p..tclp.qf[kqr]cgsg.c[filmv].q.wvcdg[ekq]ndced.sde[ast]nc..?.?..t
2043	6	10^{-95}	c...[eq]f.c..g.c[ilv]p....cdgd[dn]dc.d.sde.....?.?.?.?.?.?.?.?.?c...[eq]f.c..?.g.c[ilv]....[ilv]cdgv[dn]dc.d.sde[ast].c.?.?p..tclp.qf[kqr]cgsg.c[filmv]sqrwvcdg[ekq]ndcedssdeanca.?.?[ilv]pt
1757	5	10^{-98}	c...[eq]f.c..g.c[ilv]p....cdgd[dn]dc.d.sde.....?.?.?.?.?.?.?.?.?c...[eq]f.c..?.g.c[ilv]....[ilv]cdgv[dn]dc.d.sde[ast].c.?.?p..tclp.qfqcgsg[fwy]cisqrwvcdgendcedssdeanca.?.?vpt

Fam2 (SEQ ID NOS:792-813)

CLPDQFRCGNGQCIPLDWVCDGVNDCEDSSDEAGCPP--RTCLPDQFQCGSGYCLPQRWVCDGVDDCEDSSDEKGCTPHT--CLSDQFQCGSGHCLPASWRCDGEDDCADSSDETSCTASEPT
CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPP--RTCLPDQFQCGSGYCLPQRWVCDGVDDCEDSSDEKGCTPHT--CEPGQFPCESGQCLPPPWVCDGVNDCADNSDEEDCGDSHILPFSTPGPST
CLPSQFQCRNGRCIPLPWLCDGDNDCEDDSDETNCTP--RTCPSSQFQCSSGWCIPEDWRCDGVPDCGDGSDEASDHC
CPSSQFRCSSGRCVPADWLCDGVNDCADNSDETGCPT--RTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEH--T
CPSSQFRCSSGRCVPADWLCDGVNDCADNSDETGCAT--RTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEH--T
CPSSQFRCSSGRCVPADWLCDGVNDCGDNSDEAGCPT--RTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEH--T
CPSDEFTCDSGRCVSLNWLCDGVNDCGDNSDEAGCPT--RTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEH--T
CPSDEFTCDSGRCVSLNWLCDGVNDCGDNSDEAGCPT--RTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEH--T
CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPP--RTCLPDQFQCGSGYCLPQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTQH--T
CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPP--RTCLPDQFQCGSGYCLPQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCVPRAWRCDGVNDCADNSDEEDCTEH--T
CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPP--RTCLPDQFQCGSGYCLPQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEADCAAP--T
CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPP--HTCLPDQFQCGSGYCLSQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEH--T
CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPP--RTCLPDQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEH--T
CLPDQFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPP--RTCAPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEH--T
CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPT--RTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEH--T
CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPT--RTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEN--T
CESDQFQCDSGHCVPLNWVCDGVNDCVDGSDEAGCPP--RTCLPGQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEH--T
CEANELRCSSGRCLSVHWLCDGVDDCEDSSDEAGCPP--RTCLPDQFQCGSGYCLPQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEH--T
CGSNEFTCSNGRCIPPTLLCDGVNDCGDDSDETGCATPVRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEH--T
CRADEFQCGSGRCVPADWLCDGVNDCADNSDETGCPT--RTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEH--T
CPSSQFRCSSGRCVPADWLCDGVNDCADNSDETGCPT--RTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEH--T
CPSSQFRCSSGRCVPADWLCDGVNDCADNSDETGCPT--RTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCPPGQFTCNNGRCVPPTWLCDGEDDCQDDSDEEGCRTPART
CPSSQFRCSSGRCVPADWLCDGVNDCADNSDETGCPT--RTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCPPGQFTCNNGRCVPPTWLCDGEDDCQDDSDEKNCKAP-T

score	matches	expected	motif (SEQ ID NOS:814-819)
8844	24	10^{-103}	c...[eq][fly].c..g.c[ilv]....[ilv]cdgv[dn]dc.d.sde..c...?.?.tc.p.qfqcgsgyc[ilv].qrwvcdgendced.sdeancagsvptcp..[eq]f.c.ngrc[iv]p..w.cdg.[dn]dc.d [dn]sde..c....?.?t
8963	23	10^{-110}	c...[eq][fly].c..g.c[ilv]....[ilv]cdgv[dn]dc.d.sde..c...?.?.tc.p.qfqcgsgyc[ilv].qrwvcdgendced.sdeancagsvptcpsdefrcrngrc[iv]p..w.cdgv[dn]dc.d[dn]sde ..c....?.?t
9119	22	10^{-117}	c...[eq][fly].c..g.c[ilv]....[ilv]cdgv[dn]dc.d.sde..c...?.?.tc.p.qfqcgsgyc[ilv].qrwvcdgendced.sdeancagsvptcpsdefrcrngrc[iv]prawrcdgvndcadnsde.dc [ast]...?.?t
8895	21	10^{-120}	c...[eq][fly].c..g.c[ilv]....[ilv]cdgv[dn]dc.d.sde..c...?.?.tc.p.qfqcgsgyc[ilv].qrwvcdgendced.sdeancagsvptcpsdefrcrngrc[iv]prawrcdgvndcadnsdeedct [eq]..?.?t
8599	20	10^{-122}	c...[eq][fly].c..g.c[ilv]...w[ilv]cdgv[dn]dc.d.sde..c...?.?.tc.p.qfqcgsgyc[ilv].qrwvcdgendced.sdeancagsvptcpsdefrcrngrc[iv]prawrcdgvndcadnsdeedct [eq]..?.?t
8309	19	10^{-124}	c...[eq]f.c..g.c[iv]..[dn]w[ilv]cdgvndc.d.sde..c...?.?.tc.p.qfqcgsgyc[ilv].qrwvcdgendced.sdeancagsvptcpsdefrcrngrc[iv]prawrcdgvndcadnsdeedct[eq]..? .?t

Fam3 (SEQ ID NOS:820-829)

CPSDEFTCDSG-RCVSLNWLCDGVNDCGDNSDEAGC--PTRTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPT
CPSSQFRCSSG-RCVPADWLCDGVNDCADNSDETGCPT--RTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCGDSSDETSCSAPASEPPGS-L
CPSDEFTCDSG-RCVSLNWLCDGVNDCGDNSDEAGC--PTRTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCGANEFTCDNGRCLSPDLVCDGVDDCADSSDETDCQ--KPM------
CPSDEFTCDSG-RCVSLNWLCDGVNDCGDNSDEAGC--PTRTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCPPDEFQCNSGSCLSQPLLCDGEDDCGDGSDEESCI--AAEHT----
CLSDEFTCDSG-RCVSLNWLCDGVNDCGDNSDEAGC--PTHTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCPASEFRCNNGRCIPAQWRCDGVNDCGDDSDEASCQ--APT------
CQANEFTCSSG-NCLSPALLCDGVDDCGDNSDETNCASPAPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCT--EHT------
CELNEFTCSSG-NCLSPALLCDGVDDCGDNSDETNCASPAPTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEANCAGSVPT------
CQANEFTCSSG-NCLSPALLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDEASCQ--PPT------
CQANEFTCSSG-NCLSPALLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCLSGEFTCENGNCLSLSLVCDGEDDCEDGSDEASCP--PPT------
CQANEFTCSSG-NCLSPALLCDGVDDCGDNSDETNCASPAPTCLPSQFQCGSGYCISQRWVCDGEHDCEDSSDEANCAGSVPTCAPGEFTCGSGRCVPQTWRCDGVDDCLDSSDETDCS-APASEPPGSL

score	matches	expected	motif (SEQ ID NOS:830-832)
3233	9	10^{-101}	c...[eq]f.c.sg.?.c[ilv]....lcdgv[dn]dc.dnsde[ast].c.?.?.?.?.tclpsqfqcgsgycisqrwvcdge.dced.sdeancagsvptc...ef.c..g.c[ilv].....cdg.[dn]dc.d.sde..c..?.?....?.?.?.?.?.?
3001	8	10^{-105}	c..[dn]eftc.sg.?.c[ilv]s...lcdgv[dn]dcgdnsde[ast].c.?.?.?.?.tclpsqfqcgsgycisqrwvcdge.dced.sdeancagsvptc...ef.c..g.c[ilv].....cdg.[dn]dc.d.sde..c..?.?....?.?.?.?.?
2672	7	10^{-107}	c.[as][dn]eftc.sg.?.c[ilv]s...lcdgv[dn]dcgdnsde[ast].c.?.?.?.?.tclpsqfqcgsgycisqrwvcdge.dcedssdeancagsvptc...ef.c..g.c[ilv].....cdg.[dn]dc.d.sde..c..?.?....?.?.?.?.?

Fam4 (SEQ ID NOS:833-839)

CLANQFKCDNGRFIPQTWVCDGDNDCLDGSDEPVHCAKRTCPSSQFQCSSGWCISQRWVCDGENDCEDSSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDC--TEHT
CLANQFKCDNGRCIPQTWVCDGDNDCLDGSDEPVHCAKRTCPSSQFQCSSGWCISQRWVCDGENDCEDSSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDC--TEHT
CLANQFKCDNGRCIPQTWVCDGDNDCLDGSDEPVHCAERTCPSSQFQCSSGWCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDC--TEHT
CLANQFKCDNGRCIPQTWVCDGDNDCLDGSDEPVRCAKRTCPSSQFQCSSGWCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEGC--TPHT
CLANQFKCDNGRCIPQTWVCDGDNDCLDGSDEPVHCAKRTCPSSQFQCSSGWCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEADC--KAHT
CLANQFKCDNGRCIPQTWVCDGDNDCLDGSDEPVHCAKRTCPSSQFQCGSGWCISQHWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCGDNSDETNC--APPT
CLSSEFTCGNGNCIPQTWVCDGDNDCLDGSDEPVHCAKRTCPSSQFQCSSGWCISQRWVCDGENDCEDSSDEANCAGSVPTCPSSQFRRNNGHCLSVTLVCDGDNDCEDDSDETGCEGSGHT

score	matches	expected	motif (SEQ ID NOS:840-842)
3802	9	10^{-120}	cl[as].[eq]f.c.ng..ipqtwvcdgdndcldgsdepv.ca[ekq]rtcpssqfqc.sgwcisq.wvcdgendced.sdeancagsvptcps.[eq]fr..ng.c[ilv]..[ast]..cdg.ndc.d[dn]sde..c.?.?.. .t
3911	8	10^{-140}	clangfkcdngr.ipqtwvcdgdndcldgsdepv.ca[ekq]rtcpssqfqc.sgwcisq.wvcdgendced.sdeancagsvptcpsdefrcrngrciprawrcdgvndc.dnsde..c.?.?...t
3543	7	10^{-145}	clangfkcdngr.ipqtwvcdgdndcldgsdepv.ca[ekq]rtcpssqfqcssgwcisqrwvcdgendced.sdeancagsvptcpsdefrcrngrciprawrcdgvndcadnsde..c.?.?..ht

Fam5 (SEQ ID NOS:843-851)

CLPDQFRCGNG-QCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDCEDSSDEANCAGS---VPT
CLPDQFRCGNG-QCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDCEDSSDEANCAGS---VPTCQPGQFRCNNGHCLSQPLVCDGVPDCEDGSDEAS--CSAPASEPPGSL
CLPDQFRCGNG-QCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDCEDSSDEANCAGS---VPTCASSQFPCDNGKCLPPPLVCDGVDDCGDDSDEPSAPCGSPART
CLPDQFRCGNG-QCIPLDWVCDGVNDCVDGSDEAGCPPRTCLPDQFQCGSGYCLPQRWVCDGENDCEDGSDEANCAGS---VPTCRPDEFKCENGHCIPPGLRCDGVDDCVDNSDETD--CETPGRT
CLPDQFRCGNG-QCIPLDWVCDGVNDCGDNSDEAGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGS---VPTCLPGQFRCDNGQCIPAHWGCDGVNDCRDNSDEAD--CPQPT
CLSSEFTCGNG-NCIPRPWVCDGDNDCVDDSDEKGCPEPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGS---VPTCEADQFQCGSGYCLPRHWVCDGDNDCEDDSDETD--CGRPGPGATSAPAV
CLSSEFTCGNG-NCIPRPWVCDGDNDCVDDSDEKGCPEPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGS---VPTCAPNQFQCGSGYCVPAAWVCDGENDCEDNSDESSALCQERT
CLPSEFQCRDSKICLSADWLCDGVNDCVDGSDEAGCPPRTCLPDQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGS---VPTCRASQFQCGSGYCIPRPWLCDGDNDCEDDSDETN--CTPRT
CVSSQFQCGNG-HCVPAQWLCDGVNDCGDNSDEASCPPRTCAPSQFQCGSGYCIPLTWVRDGVDDCEDGSDEEDCSQDPEFHKVCLADQFQCNNGRCVPQSWGCDGDDDCQDNSDETS-CAPPT

score	matches	expected	motif (SEQ ID NOS:852-854)
2389	8	10^{-82}	c[ilv]..[eq]f.c.[dn]..?.c[ilv]...w[ilv]cdg.ndc.d.sde..cp..tc.p.qfqcgsgyc[ilv]...wv.dg.[dn]dced.sde.[dn]c[as]...?.?.?...c...[eq]f.c..g.c[ilv].... .cdg..dc.d.sde...?.?c....
2444	7	10^{-98}	cl..[eq]f.c.[dn]..?.c[ilv]...w[ilv]cdg.ndc.d.sde..cp..tclp.qfqcgsgyc[ilv].qrwvcdgendced.sdeancags.?.?.?vptc...[eq]f.c..g.c[ilv].....cdg..dc.d.sd e...?.?c....
2196	6	10^{-103}	cl..[eq]f.cgng.?.cip..wvcdg.ndc.d.sde..cp..tclp.qfqcgsgyc[ilv].qrwvcdgendced.sdeancags.?.?.?vptc...[eq]f.c..g.c[ilv].....cdg..dc.d.sde...?.?c... .

[419] The following (SEQ ID NOS:286-339) represent sequences comprising trimers that bind IL-6 (each monomer in the trimer binds IL-6) linked to an IgG binding monmer domain. Included are 9 trimers with one of two anti-IgG domains at the N-terminus, separated by one of two linkers (9x2x2 = 36 tetramers). Also listed are 9 trimers with one of two anti-IgG domains at the C-terminus, although no linker diversity was investigated for these constructs (9x2 = 18 tetramers).

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRTCLPDQFRCGNGQCIPLD
WVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRWVCDGENDCEDGSDEAN
CAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRTCLPDQFRCGNGQCIPLD
WVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEA
NCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRTCRADEFQCGSGRCVPAD
WLCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEAN
CAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRTCLPDQFRCGNGQCIPLD
WVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDCEDGSDEA
NCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTQHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRTCQANEFTCSSGNCLSPA
LLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDE
ANCAGSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDEASCQPPT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRTCLPDQFRCGNGQCIPLD
WVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDCEDGSDEA
NCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEADCAAPT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRTCLPDQFRCGNGQCIPLD
WVCDGVNDCEDDSDEESCPPHTCLPDQFQCGSGYCLSQRWVCDGENDCEDGSDEA
NCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRTCESDQFQCDSGHCVPLN
WVCDGVNDCVDGSDEAGCPPRTCLPGQFQCGSGYCISQRWVCDGENDCEDGSDEA
NCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRTCGASQFQCGSGYCISQR
WVCDGENDCEDSSDEANCAGSVPTCLPSEFQCRDSKICLSADWLCDGVNDCADNSD
EEGCPPPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETNCPART

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCKPRTCLPDQFRCGNGQCIPL
DWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRWVCDGENDCEDGSDEA
NCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCKPRTCLPDQFRCGNGQCIPL
DWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDE
ANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCKPRTCRADEFQCGSGRCVP
ADWLCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDE
ANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCKPRTCLPDQFRCGNGQCIPL
DWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDCEDGSDE
ANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTQHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCKPRTCQANEFTCSSGNCLSP
ALLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSD
EANCAGSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDEASCQPPT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCKPRTCLPDQFRCGNGQCIPL
DWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDCEDGSDE
ANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEADCAAPT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCKPRTCLPDQFRCGNGQCIPL
DWVCDGVNDCEDDSDEESCPPHTCLPDQFQCGSGYCLSQRWVCDGENDCEDGSDE
ANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCKPRTCESDQFQCDSGHCVP
LNWVCDGVNDCVDGSDEAGCPPRTCLPGQFQCGSGYCISQRWVCDGENDCEDGSD
EANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCKPRTCGASQFQCGSGYCIS
QRWVCDGENDCEDSSDEANCAGSVPTCLPSEFQCRDSKICLSADWLCDGVNDCADN
SDEEGCPPPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETNCPART

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCSAPASEPPGSLCLPDQFRCGN
GQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRWVCDGENDCE
DGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCSAPASEPPGSLCLPDQFRCGN
GQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCE
DGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCSAPASEPPGSLCRADEFQCGS
GRCVPADWLCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCE
DGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCSAPASEPPGSLCLPDQFRCGN
GQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDCE
DGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTQHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCSAPASEPPGSLCQANEFTCSSG
NCLSPALLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQRWVCDGENDCE
DSSDEANCAGSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDEASCQPPT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCSAPASEPPGSLCLPDQFRCGN
GQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDCE
DGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEADCAAPT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCSAPASEPPGSLCLPDQFRCGN
GQCIPLDWVCDGVNDCEDDSDEESCPPHTCLPDQFQCGSGYCLSQRWVCDGENDCE
DGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCSAPASEPPGSLCESDQFQCDSG
HCVPLNWVCDGVNDCVDGSDEAGCPPRTCLPGQFQCGSGYCISQRWVCDGENDCE
DGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCSAPASEPPGSLCGASQFQCGS
GYCISQRWVCDGENDCEDSSDEANCAGSVPTCLPSEFQCRDSKICLSADWLCDGVN
DCADNSDEEGCPPPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETNCPART

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCRADEFQCG
SGRCVPADWLCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTQHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCQANEFTCS
SGNCLSPALLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQRWVCDGEND
CEDSSDEANCAGSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDEASCQPPT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEADCAAPT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCEDDSDEESCPPHTCLPDQFQCGSGYCLSQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCESDQFQCD
SGHCVPLNWVCDGVNDCVDGSDEAGCPPRTCLPGQFQCGSGYCISQRWVCDGEND
CEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCGASQFQCG
SGYCISQRWVCDGENDCEDSSDEANCAGSVPTCLPSEFQCRDSKICLSADWLCDGVN
DCADNSDEEGCPPPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETNCPART

CLPDQFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRW
VCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEE
DCTEHTCLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRT

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHTCLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRT

CRADEFQCGSGRCVPADWLCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHTCLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRT

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTQHTCLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRT

CQANEFTCSSGNCLSPALLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQR
WVCDGENDCEDSSDEANCAGSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDE
ASCQPPTCLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRT

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EADCAAPTCLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRT

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPHTCLPDQFQCGSGYCLSQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHTCLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRT

CESDQFQCDSGHCVPLNWVCDGVNDCVDGSDEAGCPPRTCLPGQFQCGSGYCISQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHTCLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRT

CGASQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCLPSEFQCRDSKICLS
ADWLCDGVNDCADNSDEEGCPPPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSD
ETNCPARTCLSSEFQCQSSGRCIPLAWVCDGDNDCRDDSDEKSCKPRT

CLPDQFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRW
VCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEE
DCTEHTCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSL

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHTCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSL

CRADEFQCGSGRCVPADWLCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHTCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSL

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTQHTCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSL

CQANEFTCSSGNCLSPALLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQR
WVCDGENDCEDSSDEANCAGSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDE
ASCQPPTCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSL

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EADCAAPTCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSL

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPHTCLPDQFQCGSGYCLSQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHTCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSL

CESDQFQCDSGHCVPLNWVCDGVNDCVDGSDEAGCPPRTCLPGQFQCGSGYCISQR
WVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSD
EEDCTEHTCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSL

CGASQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCLPSEFQCRDSKICLS
ADWLCDGVNDCADNSDEEGCPPPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSD
ETNCPARTCHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSL

Example 8

[420] This example describes experiments demonstrating inhibition of IL-6-induced cell proliferation by IL-6-binding trimers.

[421] TF-1, a GM-CSF dependent human erythroleukemia cell line, was purchased from ATCC (CRL-2003). IL-6 is a potent stimulator of TF-1 cell proliferation. The use of TF-1 cells for determining efficacy of IL-6, IL-6R, and/or gp130 inhibitors is well-established in the art.

[422] TF-1 cells were maintained in growth medium (RPMI 1640 supplemented with 50 uM 2-ME, antibiotic-antimycotic), 10% FBS and 5 ng/ml hGM-CSF (Leukine). Prior to assay, TF-1 cells were washed and cultured overnight in growth medium plus 1% FBS and Leukine. TF-1 cells were then washed three times with growth medium plus 1% FBS, without Leukine prior to the assay.

[423] Serial dilutions of anti-IL-6 monoclonal antibodies (Clone 6708-111, R&D Systems, and Clone MQ2-13A5, eBioscience) or an anti-IL-6 trimer further comprising an IgG binding domain at the N-terminus (which in other experiments was shown not to affect Il-6 binding) was prepared in growth medium plus 1% FBS and incubated with a fixed concentration of rhIL-6 (0.2 ng/ml from R&D Systems) for 45 min in a humidified incubator at 37°C containing 5% CO_2.

[424] TF-1 cells ($1x10^4$/well) were cultured in the presence or absence of the anti-IL-6 inhibitors-hIL-6 mixture in a 96-well flat bottom plate (Becton Dickinson) in a final volume of 200 ul/well for 40 hr in a humidified incubator at 37°C, 5% CO2. Cells were pulse-labeled with 1 uCi/well of [^{3}H]-Thymidine (Amersham Biosciences) during the final 16 hr of culture. Plates were harvested with a semiautomatic cell harvester (Tomtec, Inc.) and [^{3}H]-Thymidine incorporation was measured using a Microbeta-counter (Wallac-Perkin Elmer). Results are expressed as average cpm from triplicate wells.

[425] Control experiments determined 0.2 ng/mL as the EC50 of rhIL-6 for stimulation of TF-1 cell proliferation.

[426] The anti-IL-6 trimer inhibited TF-1 cell proliferation induced by 0.2 ng/mL IL-6 to 50% of maximum at a concentration of 0.8 picomolar.

[427] This assay provides a means to screen monomers or multimers for anti-IL-6 activity in an in vitro bioassay using human cells. By determining IC50 values for tested multimers or monomers, optimal molecules can be identified and ranked on the basis of their biological activity.

Example 9

[428] This example describes experiments demonstrating monomer binding to a IL-6-expressing human cell line.

[429] Monomers were constructed to include an influenza hemagluttinin (HA) epitope tag. This enables the monomer to be used as a primary flow cytometry detection agent, with a fluorescent-tagged anti-HA secondary antibody being used as the secondary detection agent.

[430] 293FT human embryonic kidney cells (Invitrogen) were transiently transfected with a CD23-IL-6 fusion in the pDisplay vector (Invitrogen), resulting in cell surface expression of IL-6.

[431] 9 trimers were tested for the ability to bind 293FT/CD23-IL-6 cells. Untransfected 293FT cells, as well as Jurkat, HuT78, ArentAR, U266-BL, Raji, RPMI8226, and Daudi cell lines, were used as negative controls.

[432] To determine monomer binding, $2.5X10^5$ cells were stained with 10 uM IL-6 trimer in 100 ul FACS staining buffer (PBS pH 7.4, 5% fetal calf serum, 0.01% sodium azide) on ice for 30 min. Cells were washed once with 4 ml ice-cold FACS staining buffer, then resuspended in 100 ul FACS staining buffer plus 0.2 ug FITC-conjugated anti-HA monoclonal antibody (Santa Cruz Biotechnology) and incubated on ice for 30 min. Cells were washed once with 4 ml ice-cold FACS staining buffer, then resuspended in 200 ul FACS staining buffer and analyzed using a FACSCalibur Flow Cytometer (BD Biosciences). Data were collected and analyzed using CellQuest Pro (BD Biosciences). The geometric mean fluorescence was determined for each cell line plus IL-6 trimer, and normalized against the geometric mean fluorescence for that cell line plus FITC-conjugated anti-HA monoclonal antibody alone.

[433] The anti-IL-6 trimer showed significant binding only to 293FT/CD23-IL-6 cells, with no binding detected for any of the other cell lines tested.

[434] This assay provides a means to screen monomers or multimers for IL-6 binding specificity. Specificity for IL-6 over other cell-surface proteins can be an important factor in the development of a therapeutic, and can be used to screen and select for optimal molecules.

Example 10

[435] This example describes experiments designed to measure the in vitro binding affinity of anti-IL-6 monomers.

[436] Experiments were performed on a Biacore 1000 SPR machine. The running buffer for all experiments was HBS-P (Biacore) with the addition of 1 mM $CaCl_2$.

[437] ~200 RUs of human IL-6 were conjugated to a flowcell of a CM5 chip using NHS/EDC coupling and ethanolamine blocking according to the manufacturer's protocols.

[438] Serial dilutions of the Il-6 binding trimer used in Example 9 were prepared in the running buffer, and injected over the IL-6-conjugated flowcell. Signal resulting from flowing over a mock coupled flowcell was subtracted from the data to

minimize bulk shift effects. Data were analyzed using the BiaEvaluation software (Biacore), using a 1:1 Langmuir binding model.

[439] Independent fits of the data to a 1:1 binding model yielded values of $k_{on} = 1.95x10^6\ s^{-1}$, $k_{off} = 4.93x10^{-4}\ M^{-1}s^{-1}$, $K_D = 2.53x10^{-10}$ M.

[440] This assay provides a means to screen monomers or multimers for IL-6 binding affinity. Affinity to IL-6 may be used as an indicator of potential biological activity, and can be used to screen and select for optimal molecules.

Example 11

[441] This example describes experiments designed to show in vivo activity of anti-IL-6 trimers in mice.

[442] 8 week old female CD-1 mice were used for all experiments.

[443] Optimal IL-6 dose (600 ng/mouse in 100 ul PBS, 1% mouse serum, intraperitoneal injection) and time to sacrifice following IL-6 injection (6 hours) were predetermined by dose titration and measurement of serum amyloid A (SAA) in serum harvested from sacrificed mice by intracardiac puncture.

[444] SAA was measured by ELISA in accordance with the manufacturer's instructions (Biosource). All experiments included three mice per treatment group, with duplicate SAA measurements being made for each mouse.

[445] Trimer doses were prepared in 20 mM Tris, pH7.6, 260 mM NaCl, 1 mM $CaCl_2$, with each dose having a volume of 200 ul per mouse. All monomer preparations contained less than 4 EU/mg endotoxin as measured by the LAL endpoint method (Associates of Cape Cod).

[446] Trimer doses, control monoclonal antibodies (anti-IL-6 or mouse IgG1, 500 ug per mouse in 200 ul) or vehicle alone were injected via the tail vein. One hour later, 600 ng IL-6, 200 ng IL-1, or vehicle alone (PBS, 1% mouse serum) was injected intraperitoneally. Six hours after this second injection the mice were sacrificed, and serum harvested. Serum was frozen for shipment and storage prior to analysis.

[447] The anti-IL-6 trimer completely abrogated acute phase protein induction by human IL-6. In contrast, the anti-IL-6 trimer showed no affect on acute phase protein induction by human IL-1, indicating that the inhibitory effect is specific. The inhibitory effect of anti-IL-6 trimer was also found to be dose-dependent. These results

indicate that anti-IL-6 trimer can inhibit specifically the biological functions of IL-6 in an *in vivo* setting.

Example 12

[448] This example describes generation of additional IL-6 binding sequences based on the A domain scaffold.

The following sequences were identified as binding IL-6 as measured by ELISA binding to IL-6 and also by biological activity in the TF1 cell assay (SEQ ID NOS:855-928):

CESDQFQCDSGHCVPLNWVCDGVNDCVDGSDEAGCPPRTCLPGQFQCGSGYCISQRWVCDGENDCEDGSDEANCAG
SVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLPDQFRCGNGQCMPLDWVCDGVYDCPDDSDEEGCPPRTCAQSKFQCGSGYCISQRWVCDGEYDCEDGSDEANCAD
SVPTCPSDEFRCRNGRCIPRGWRCDGVNDCADSSDEEDCTENP

CLSVQFRCGNGQCIPLDWVCDGVKDCPDDSDEEGCPPWTCAPSQFQCGSGYCISQRWVCDGESDCEDGSDEANCAG
SVPTCPSDEFRCRNGRCIPRGWLCDGVNDCADNSDEEDCTDNT

CLPDQFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGWCISTRWVCDGENDCEDGSDEANCPG
TVPTCPSDEFRCKNGRCVPRSWCCDGVNDCADNSDEEDCTDPT

CLPDQFRCGNGQCIPLDWVCDGVSDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGS
VPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGS
VPTCPSGEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGS
VPTCPSDEFRCRNGCCIPRAWRCDGVNDCTDNSDEGDCTEHT

CLSDQFRCGNGDCIRLDWVCDGVNDCTDDSDEEGCPQRSCAPSQFQCGSGYCISQGWVCDGENDCEDGSDEANCAG
CVPTCRSDEFRCRNGRCIPLAWRCDGVNDCVDNSDEEDCPEHT

CLPDQFRCGNGQCIPLDWVCDGVNDCADXSDEXXCPTRTCXPSQFQCGXXYCISQRWVCDGEXDCEDGSDEANCAGS
VPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CPSDKFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAG
SVPTCPSDEFRCRNGQCIPLDWVCDGVNDCPDDSDEEDCTEHT

CPSDEFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAG
SVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CPSDEFRCGNGQCIPLDWVCDGVNDCPDDSDEXGCPPRTCAPSQFQCGSGYCISQRWVCDGXNDCEDGSDEANCAG
SVPTCPSDEFRCXNGRCIPRAWRCDGVNDCADNSDEXXCTEHT

CPPDQFRCGNGQCIPLDWVCDGVHDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAG
SVPTCPWNEFRCRNGRCIPWSWRCDGVNDCADNSDEEECTEHT

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGS
VPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETSCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGS
VPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLPDQFRCGNGQCTPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCIPQRWVCDGENDCEDGSDEANCAG
SVPTCPSGEFQCSNGRCIPWAWRCDGVNDCADNSDEEDCTEPP

CLPDQFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCPPSQFQCGSGYCITERWVCDGENDCEDGSDEANCAG
SVPTCSSDEFRCRAGTCIPRAWCCDGVKDCADNSDEEDCTENT

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGS
VPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLPDEFQCGNGQCISLDWVCDGVNDCPDDSDEXGCPPSSCPPNQFQCGSGYCISQRWVCDGXNDCEDGSXEANCAG
SVPACPSDEFRCRDGRCIPRAWLCDGINDCADNSDEEDCTEHT

CLAHQFLCGNGQCLPLDWVCDGVTDCPDDSDEVGCPPRTCAPSHFQCGSGYCISQSWVCDGEYDCEDGSDEASCAGS

VPSCPSDEFQCRNGRCIPRPWRCDGVNDCADNSDEEDCTEHT

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGS
VPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGS
VPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CQANEFTCSSGNCLSPALLCDGVDDCGDNSDEAYCASPAPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDEASCQPPT

CQANEFTCSSGNCLSPALLCDGVDDCGDNSDEAYCASPAPTCLPSQFQCGSGYCISQRWVCDGENDCEDSSDEANCA
GSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDEASCQPPT

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGS
VPTCPSDEFRCRNGRCIPRGWRCDGVNDCADTSDEEDCKEHC

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGS
VPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGS
VPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLPDQFRCGNGKCIPLDWVCDGVNDCPDASDEEGCPPRTCGPSQFQCGSGYCISQGWVCDGEIDCEDSSDEANCAGS
VPTCPSDQFQCSNGRCIPRAWHCDGINDCEDNSDEEDCTEHT

CLPGQFRCANGQCIPLEWVCDGDNDCPDDSDEDGCPPRTCGPSQFKCGSGYCISESWVCDGEYDCEDGSDEATCVGS
VPTCPWDEFRCRSGRCISRAWRCDGVNDCADNSDEEACTDHT

CLSDQFRCGEGQCIPLDWVCDGVSDCGDDSDEEGCPPSTCAPSQFQCGSGYCISQNWVCDGENDCEDGSDEANCAG
CVPTCPADEFRCRDGRCIPRAWRCDGVNDCGDKSDEEDCTENT

CRPDQFRCGNGHCIPLDWVCDGVNDCTDDSDEGGCPQRSCAPRQFQCGSGYCISQRWVCDGEYDCEDGSDEANCTG
SVAACPYDEFRCINGRCIPRSWCCDGVNDCADNSDEADCPEHT

CLPDEFRCGNGHCIPLDWVCDGVNDCPDDSDEEGCPPQTCAPSQFQCGSGFCISQRWVCDGEQDCVDGSDESNCAD
SVPTCPSDEFRCKNGRCIPLVWRCDGVNDCSDNSDEEDCKVHS

CRRDQFRCGNGQCISVDWVCDGVNDCPDESDEEGCPPRTCTPSQFMCGSXYCISQRWVCDGEYDCEDGSDXXNCAG
SVPTCPKDXFQCRNGRCIPRAWGCDXX

XXXXXFPGXXRXXXXXXXXXXXXYDCLDXSDEXXCPARTCXXSQFQCGSGYCISQRWVCDGEYDCEDGSDXXNXXXXX
XTCXSXEFRCXXGRCIPRSWRCDGVNDCADNSDEXDCTGHS

CLPGQFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPNQFQCGSGYCISQRWVCDGENDCEDGSDEANCVG
SVTTCPSDEFRCRNGRCIPRVWRCDGEDDCADNSDEEDCTECT

CLXYXFLCVNGQCIXWXCXXXXXLXXXXXXKXLXXXYLCXXXXXXRXRLLHXXXLGLXXXKXLXXXLGXNKLCWPXTXXS
XXXVXVQXXXLXPXCXAXRXXXXLXGXLXXXXXXRTHFXAXXXXXPXXXXPXXXPL

CXPDQFRCXXGQCIPLDWVCDGVNDCPDNSDEKXCPPRTCXPSQFQCXSGYCISQXWVCDGENDCEDGSDEANCAGS
VPXCPSDEFXCRNGLCIPRAWRCDGVNDCXDNSDEEXCTEHT

CLPNQFRCRNGQCIPLDWVCDGINDCPDESDEEGCPPRTCAPSQFQCGSGYCISRRWVCDGEIDCDDGSDEANCAGSV
PACPSDEFRCSNCRCIPRAWRCDGVNDCSDNSDEEDCTGHT
CLPAQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCXDGSDXXNCAGS
VXTCPSDXFRCXNGRCIPRAWRCXGXNXCGGQL

XXSDQFLCXXGXCIPLXWVCDGIHDXXXXXEXXRXXXXXXXXXXXXXRXRXXLXXXXXQRQXXXXXRXXXXKXXSFXTX
XPXXXFXXQXXXLXXXXXXXXRXXXXCEQLEXKQXXSAXXFAKGXXGPQAPPPPPPPX

CLQDQFRCGDGQCIPLEWVCDGVDDCADHSDEEGCPPRTCSPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCPG
SVPTCPSDEFRCRNGRCIPRAWRCDGVSDCADNSDEEDCTEET

CLPDQFRCGNGECIPLDWVCDGVKDCPDDSDEEGCPPPTCAPGQFQCGSGYCISKRWVCDGENDCEDGSDEANCEGS
VPACPSDEFRCGNGRCIPRDWRCDGVHDCEDNSDEEXCTEHN

CLPDQFRCSNGHCIPLDWVCDGVNDCPDDSDEGGCPPRTCAPSQFQCGSGFCISQRWVCDGEHDCEDGSDEANCAVS
IPTCPPDDFRCRNGRCIPRAWCCDGVNDCADNSDEEGCREHT

CPPKQFRCGNGQCIPLDWICDGVNDCTDDSDEEGCPPRSCAPGQFQCGSGYCISQRWVCDGENDCEDGSDEANCPCS
VPTCPSDEFRCRNGPCIPRAWRCDGVNDCADNSDEDDCTGHT

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGS
VPTCPSDEFRCRNGRCIPRHWCCDGVSDCPDNSDEEDCKDHT

CESDQFQCDSGHCVPLNWVCDGVNDCVDGSDEAGCPPRTCLPGQFQCGSGYCISQRWVCDGENDCEDGSDEANCAG
SVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLPYQFRCRNGRCIPLDWLCDGVNDCPDDSDEEGCPALTCAPSQFQCGSGYCISQRWVCDGEIDCEDGSDEANCAGSE
PTCPSDEFRCXNGRCIPRAWCCDGVNDCVDNSDEEDCAPHT

CRPDQFRCDNGHCIQLDWVCDGVNDCQDDSDEDGCPPRTCAQSQFQCGSGYCISQSWVCDGENDCEDGSDEANCAG
SVPTCPSDEFRCRNGRCIPRAWRCDGINDCEDHSDEEDCKAHT

CLPDQFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAG
SVPTCPSEEFRCRNGXCXXRAWXCDGVNDCXDNSDEEXCTEHP

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLSSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGS
VPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLPDQFRCGNGQCIPLDWACDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGS
VPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGS
VPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLPDQFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAG
SVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLPDQFRCGNGRCIPLDWVCDGVNDCQDDSDEEGCPPRTCAPSQFQCGSGYCISQGWVCDGENDCEDGSDEANCAG
PVPTCAADEFRCRNGRCIPRAWRCDGVNDCADNSDEEACTEHT

CPRTSSAAATANAIHGPGFATAILICPDDSDEDGCPPRTCSPNQFQCGSGYCISQRWVCDGEIDCEDGSDEANCAGSVP
KCPSDAFRCRNGRCIPRAWRCDGINDCVDNSDEEECREHT

CPRTSSAAATANAIHGPGFATAILICPDDSDEDGCPPRTCSPNQFQCGSGYCISQRWVCDGEIDCEDGSDEANCAGSVP
KCPSDAFRCRNGRCIPRAWRCDGINDCVDNSDEEECREHT

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGS
VPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CPPEQFRCGTGKCIPMDWVCDGVKDCPDDSDEEGCPPRTCAPHQFQCGSGFCISQRWVCDGENDCEDGSDEANCAG
SVPTCPSNEFRCRNGRCIPCSWHCDGVNDCADNSDEEDCTPHP

CVPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEAHCAG
SVPTCPADEFRCRNGGCIPLAWRCDGVNDCADNSDEEDCTDDT

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGS
VPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGS
VPTCPSDEFRCRNGRCITRAWRCDGVNDCADNSDEEDCTEHT

CLPGQFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPNQFQCGSGYCISQRWVCDGENDCEDGSDEANCVG
SVTTCPSDEFRCRNGRCIPRVWRCDGEDDCADNSDEEDCTECT

CPSPVPVRQRPLRSTGLGLRRRQCLSDDSDEEGCPARTCGPGQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGSV
PTCTSDEFRCRNGRCIPRAWRCDGVKDCADNSDEEPCKEHQ

CLPDQFRCDNGQCIPLDWVCDGVNDCPDDSDEQGCPPRSCAPSQFQCGSGYCISQRWVCDGEYDCEDGSDEANCAG
SAPTCPSEEFRCRNGRCIPNAWCCDGLNDCSDNSDEEDCAEHP

CLPDQFRCGNGQCIPLDWICDGLIDCPDESDEEGCPPRTCEPSQFQCGSGFCISQGWVCDGEIDCEDGSDEANCAGSV
PTCPSDEFRCRSGRCIPRAWRCDGISDCADGSDEEDCTEDT

CAPSQFQCGSGYCISQRWVCDGENDCEDGSDEVGCPPRTCAPSQFQCGSGWCISQRWVCDGENDCEDGSDEANCA
GSVPTCPSDEFRCRSGRCIPRAWRCDGVSDCADGSDEEDCTEDT

CLRDQFRCGNGQCIPLDWACDGINDCPDDSDEGGCPPRMCAPSQFQCGSGYCIPLGWVCDGENDCEDGSDEANCAG
SEPTCPSGEFRCRNGRCIPRAWLCDGVNDCADNSDEEDCTGHT

CLPDMFRCGTRQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCAGS
VPTCPSDEFRCRNGRCIPRSWCCDGVNDCADNSDEEDCTEHT

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDCEDGSDEANCADS
VPTCPSYEFRCRNGRCIPRTWRCDGVNDCADNSDEEDCTEHP

CLANQFKCDNGRCIPQTWVCDGDNDCLDGSDEPVHCAKRTCPSSQFQCSSGWCISQRWVCDGENDCEDNSDEANCA
GSVPTCGPGQFPCNDSDRCLSVPRLCDGDDDCGDDSDEASCQPPT

CLPDQFRCGNGQCMPLDWVCDGVNDCPDDSDEEGCPPRTCEPGQFQCGSGYCISQRWVCDGENDCEDGSDEATCA
GSVPQCPSDEFRCRTGRCIPQAWRCDGVNDCVDNSDEECCKEHT

CLANQFKCDNGRCIPQTWVCDGDNDCLDGSDEPVHCAKRTCPSSQFQCSSGWCISQRWVCDGENDCEDSSDEANCA
GSVPT

CVSDQFRCGNGQCLPLDWVCDGVNDCPDDSDEEGCPPLTCAPSQFQCGSGYCISQRWVCDGEYDCEDGSDEANCAG
SVPTCPSDEFRCRNGRCIPRAWCCDGINDCTDNSDEEDCTEYT

CPRTSSAAATANAIHGPGFATAILICPDDSDEDGCPPRTCSPNQFQCGSGYCISQRWVCDGEIDCEDGSDEANCAGSVP
KCPSDAFRCRNGRCIPRAWRCDGINDCVDNSDEEECREHT

CLPDRFRCGNGQCMPMDWVCDGVNDCPDDSDERGCPPRTCAQSQFQCGSGYCISESWVCDGENDCEDGSDEANCA
GSVPTCPSDEFRCSNGRCIPRSWLCDGVNDCSDNSDEEDCAEHP

[449] While the foregoing invention has been described in some detail for purposes of clarity and understanding, it will be clear to one skilled in the art from a reading of this disclosure that various changes in form and detail can be made without departing from the true scope of the invention. For example, all the techniques, methods, compositions, apparatus and systems described above can be used in various combinations. All publications, patents, patent applications, or other documents cited in this application are incorporated by reference in their entirety for all purposes to the same extent as if each individual publication, patent, patent application, or other document were individually indicated to be incorporated by reference for all purposes. PCT applications PCT/US02/13257 and PCT/US03/35664 and U.S. Patent Publication No. 2005/0089932 are hereby incorporated by reference in their entirety.

SEQUENCE LISTING

SEQ ID NO:1 - IL-6 processed amino acid sequence

APVPPGEDSKDVAAPHRQPLTSSERIDKQIRYILDGISALRKETCNKSNMCESSKEAL
AENNLNLPKMAEKDGCFQSGFNEETCLVKIITGLLEFEVYLEYLQNRFESSEEQARAV
QMSTKVLIQFLQKKAKNLDAITTPDPTTNASLLTKLQAQNQWLQDMTTHLILRSFKE
FLQSSLRALRQM

WHAT IS CLAIMED IS:

1. A polypeptide comprising a monomer domain that selectively binds to IL-6, wherein the monomer domain is an LDL receptor class A monomer domain comprising either of the following sequences (SEQ ID NOS:1 and 2):

CXXXQFQCXSG[W,Y]C[I,L,F]XXXW[V,L]CDGXNDCEDXSDEXXC

or

CX[A,P]X[E,Q]FXCXXGQC[I,L]PXXXXCDGX[D,N]DCXDXSDEXXC

wherein X is any amino acid and amino acids in brackets indicate alternative amino acids at a specified position.

2. The polypeptide of claim 1, wherein the polypeptide comprises CXXXQFQCXSG[W,Y]C[I,L,F]XXXW[V,L]CDGXNDCEDXSDEXXC (SEQ ID NO:1).

3. The polypeptide of claim 1, wherein the polypeptide comprises CX[A,P]X[E,Q]FXCXXGQC[I,L]PXXXXCDGX[D,N]DCXDXSDEXXC (SEQ ID NO:2).

4. The polypeptide of claim 1, wherein the polypeptide comprises at least one and no more than six monomer domains that selectively bind IL-6.

5. The polypeptide of claim 1, wherein the polypeptide comprises at least two monomer domains that selectively bind IL-6.

6. The polypeptide of claim 5, wherein:

the polypeptide comprises a first monomer domain with selective binding specificity for IL-6 and a second monomer domain with selective binding specificity for IL-6;

the first monomer domain comprises

CXXXQFQCXSG[W,Y]C[I,L,F]XXXW[V,L]CDGXNDCEDXSDEXXC (SEQ ID NO:1); and

the second monomer domain comprises

CX[A,P]X[E,Q]FXCXXGQC[I,L]PXXXXCDGX[D,N]DCXDXSDEXXC (SEQ ID NO:2).

7. The polypeptide of claim 6, wherein the polypeptide comprises:

CX[A,P]X[E,Q]FXCXXGQC[I,L]PXXXXCDGX[D,N]DCXDXSDEXXCX$_n$
CXXXQFQCXSG[W,Y]C[I,L,F]XXXW[V,L]CDGXNDCEDXSDEXXC,
wherein n is between 0 and 100 (SEQ ID NO:3).

8. The polypeptide of claim 1, wherein the polypeptide comprises any of the following sequences (SEQ ID NOS:5-13):

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQCGSGYCISQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCRADEFQCG
SGRCVPADWLCDGVNDCADNSDETGCPTRTCLPSQFQCGSGYCISQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTQHT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCQANEFTCS
SGNCLSPALLCDGVDDCGDNSDETYCASPAPTCLPSQFQCGSGYCISQRWVCDGEND
CEDSSDEANCAGSVPTCQPGEFTCESGRCLPLRWLCDGEDDCVDSSDEASCQPPT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQCGSGYCLPQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEADCAAPT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCLPDQFRCG
NGQCIPLDWVCDGVNDCEDDSDEESCPPHTCLPDQFQCGSGYCLSQRWVCDGENDC
EDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT; or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCESDQFQCD
SGHCVPLNWVCDGVNDCVDGSDEAGCPPRTCLPGQFQCGSGYCISQRWVCDGEND
CEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGVNDCADNSDEEDCTEHT;
or

CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENCSAPASEPPGSLCGASQFQCG
SGYCISQRWVCDGENDCEDSSDEANCAGSVPTCLPSEFQCRDSKICLSADWLCDGVN
DCADNSDEEGCPPPTCLPDQFQCGSGYCLPQAWVCDGENDCEDGSDETNCPART.

9. The polypeptide of claim 1, wherein the polypeptide comprises any of the following sequences (SEQ ID NOS:14-22):

3 CLPDQFRCGNGQCIPLDWVCDGVNDCPDDSDEEGCPPRTCAPSQFQC
4 GSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGV
5 NDCADNSDEEDCTEHT;

CLPDQFRCGNGQCIPLDWVCDGVNDCADNSDETGCPTRTCLPSQFQC
GSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGV
NDCADNSDEEDCTEHT;

CRADEFQCGSGRCVPADWLCDGVNDCADNSDETGCPTRTCLPSQFQC
GSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGV
NDCADNSDEEDCTEHT;

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQC
GSGYCLPQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGV
NDCADNSDEEDCTQHT;

CQANEFTCSSGNCLSPALLCDGVDDCGDNSDETYCASPAPTCLPSQFQ
CGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCQPGEFTCESGRCLPLRWLCDG
EDDCVDSSDEASCQPPT;

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPRTCLPDQFQC
GSGYCLPQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGV
NDCADNSDEADCAAPT;

CLPDQFRCGNGQCIPLDWVCDGVNDCEDDSDEESCPPHTCLPDQFQC
GSGYCLSQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGV
NDCADNSDEEDCTEHT;

CESDQFQCDSGHCVPLNWVCDGVNDCVDGSDEAGCPPRTCLPGQFQC
GSGYCISQRWVCDGENDCEDGSDEANCAGSVPTCPSDEFRCRNGRCIPRAWRCDGV
NDCADNSDEEDCTEHT; or

CGASQFQCGSGYCISQRWVCDGENDCEDSSDEANCAGSVPTCLPSEFQ
CRDSKICLSADWLCDGVNDCADNSDEEGCPPPTCLPDQFQCGSGYCLPQAWVCDGE
NDCEDGSDETNCPART.

10. The polypeptide of claim 1, wherein the polypeptide comprises three monomer domains that selectively bind IL-6.

11. The polypeptide of claim 1, wherein the polypeptide further comprises a monomer domain that has a selective binding specificity for a blood factor, thereby increasing the serum half-life of the polypeptide when the polypeptide is injected into an

4 animal compared to the serum half-life of a polypeptide lacking the blood factor-binding
5 monomer domain.

12. The polypeptide of claim 11, wherein the blood factor is serum albumin, an immunoglobulin or an erythrocyte.

13. The polypeptide of claim 11, wherein the monomer domain that has a selective binding specificity for a blood factor binds to immunoglobulin (IgG) and the monomer domain that has a binding specificity for a blood factor is an LDL receptor class A monomer domain comprising a sequence selected from the following (SEQ ID NOS:23 and 24):

CXSSGRCIPXXWVCDGXXDCRDXSDE, and

CXSSGRCIPXXWLCDGXXDCRDXSDE,

wherein X is any amino acid.

14. The polypeptide of claim 11, wherein the monomer domain that has a selective binding specificity for a blood factor binds to immunoglobulin (IgG) and the monomer domain that has a binding specificity for a blood factor is an LDL receptor class A monomer domain comprising the following sequence (SEQ ID NO:25):

[EQ]FXCRX[ST]XRC[IV]XXXW[ILV]CDGXXDCXD[DN]SDE,

wherein X is any amino acid and amino acids in brackets are alternative amino acids at a single position.

15. The polypeptide of claim 11, wherein the monomer domain that has a selective binding specificity for a blood factor comprises CHPTGQFRCRSSGRCVSPTWVCDGDNDCGDNSDEENC (SEQ ID NO:26).

16. The polypeptide of claim 1, wherein the monomer domain is between 35 to 45 amino acids.

17. The polypeptide of claim 1, wherein the monomer comprises an amino acid sequence in which:

at least 10% of the amino acids in the sequence are cysteine; and/or

at least 25% of the amino acids are non-naturally-occurring amino acids.

18. A polynucleotide encoding the polypeptide of claim 1.