US 20050221384A1

(19) **United States**
(12) **Patent Application Publication**
Kolkman et al.
(10) **Pub. No.: US 2005/0221384 A1**
(43) **Pub. Date: Oct. 6, 2005**

(54) COMBINATORIAL LIBRARIES OF MONOMER DOMAINS

(75) Inventors: **Joost A. Kolkman**, Sint-Martens-Latem (BE); **Willem P.C. Stemmer**, Los Gatos, CA (US); **Per-Ola Freskgard**, Norrkoeping (SE)

Correspondence Address:
TOWNSEND AND TOWNSEND AND CREW, LLP
TWO EMBARCADERO CENTER
EIGHTH FLOOR
SAN FRANCISCO, CA 94111-3834 (US)

(73) Assignees: **Avidia Research Institute**, Mountain View, CA (US); **Maxygen, Inc.**, Redwood City, CA (US)

(21) Appl. No.: **10/966,064**

(22) Filed: **Oct. 15, 2004**

Related U.S. Application Data

(63) Continuation of application No. 10/133,128, filed on Apr. 26, 2002.

(60) Provisional application No. 60/374,107, filed on Apr. 18, 2002. Provisional application No. 60/333,359, filed on Nov. 26, 2001. Provisional application No. 60/337,209, filed on Nov. 19, 2001. Provisional application No. 60/286,823, filed on Apr. 26, 2001.

Publication Classification

(51) **Int. Cl.^{7}** **G01N 33/53**; G01N 33/543
(52) **U.S. Cl.** **435/7.1**; 436/518

(57) ABSTRACT

Methods for identifying discrete monomer domains and immuno-domains with a desired property are provided. Methods for generating multimers from two or more selected discrete monomer domains are also provided, along with methods for identifying multimers possessing a desired property. Presentation systems are also provided which present the discrete monomer and/or immuno-domains, selected monomer and/or immuno-domains, multimers and/or selected multimers to allow their selection. Compositions, libraries and cells that express one or more library member, along with kits and integrated systems, are also included in the present invention.

Figure 1

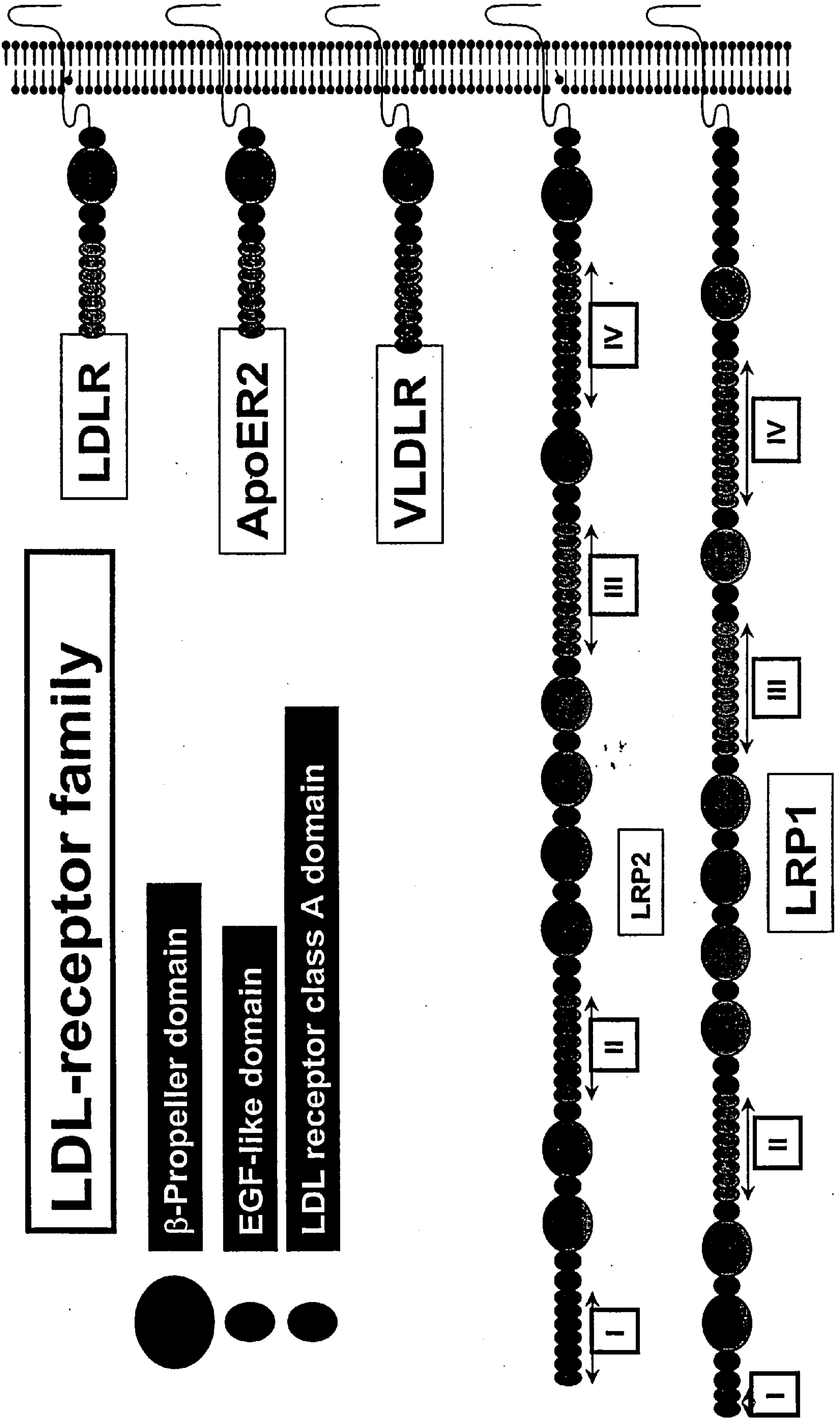

LDL-receptor family
β-Propeller domain
EGF-like domain
LDL receptor class A domain
LDLR
ApoER2
VLDLR
I
II
III
IV
LRP2
I
II
III
IV
LRP1

Figure 2

LDL-receptor class A domain

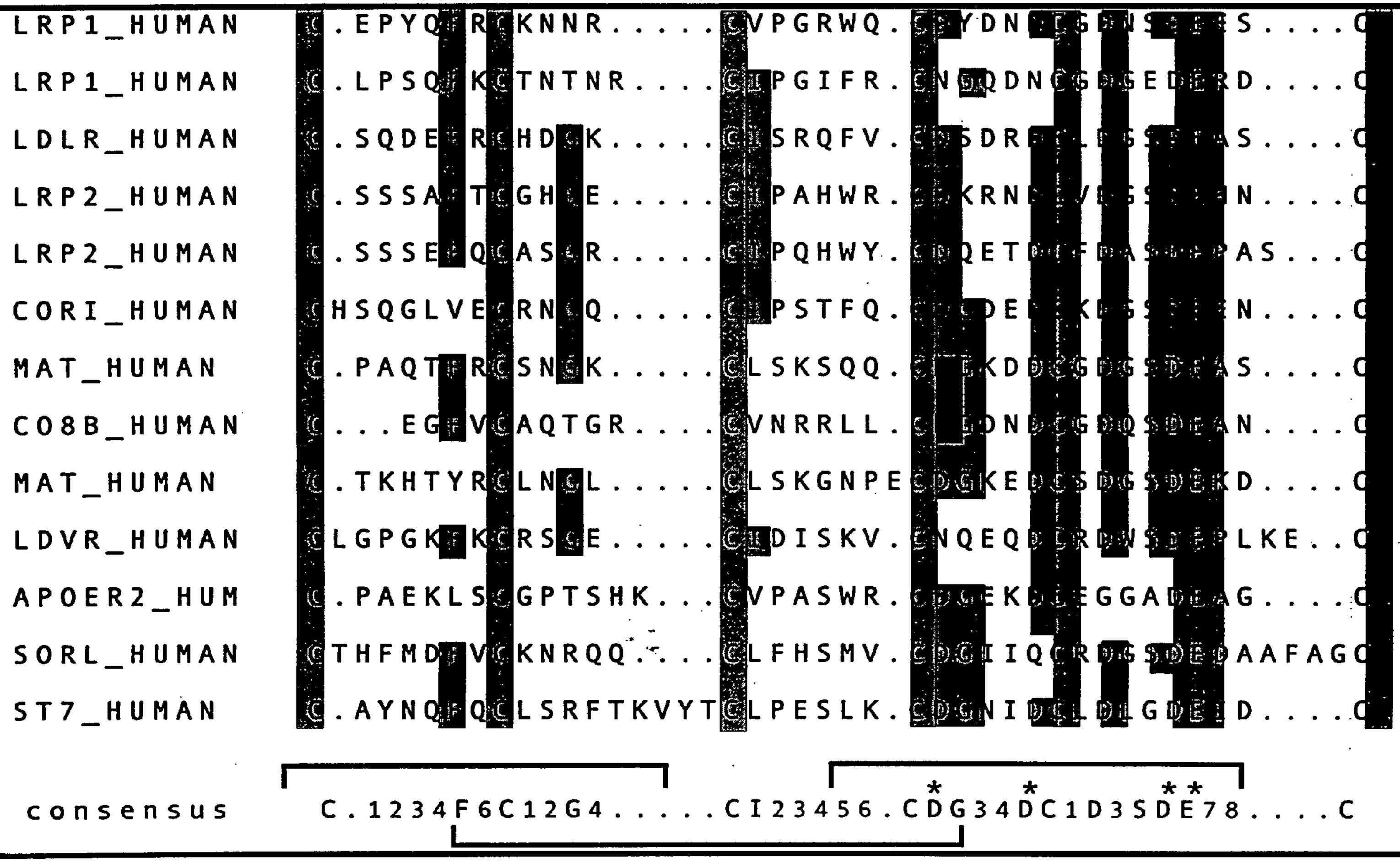


FIGURE 3

A-domains

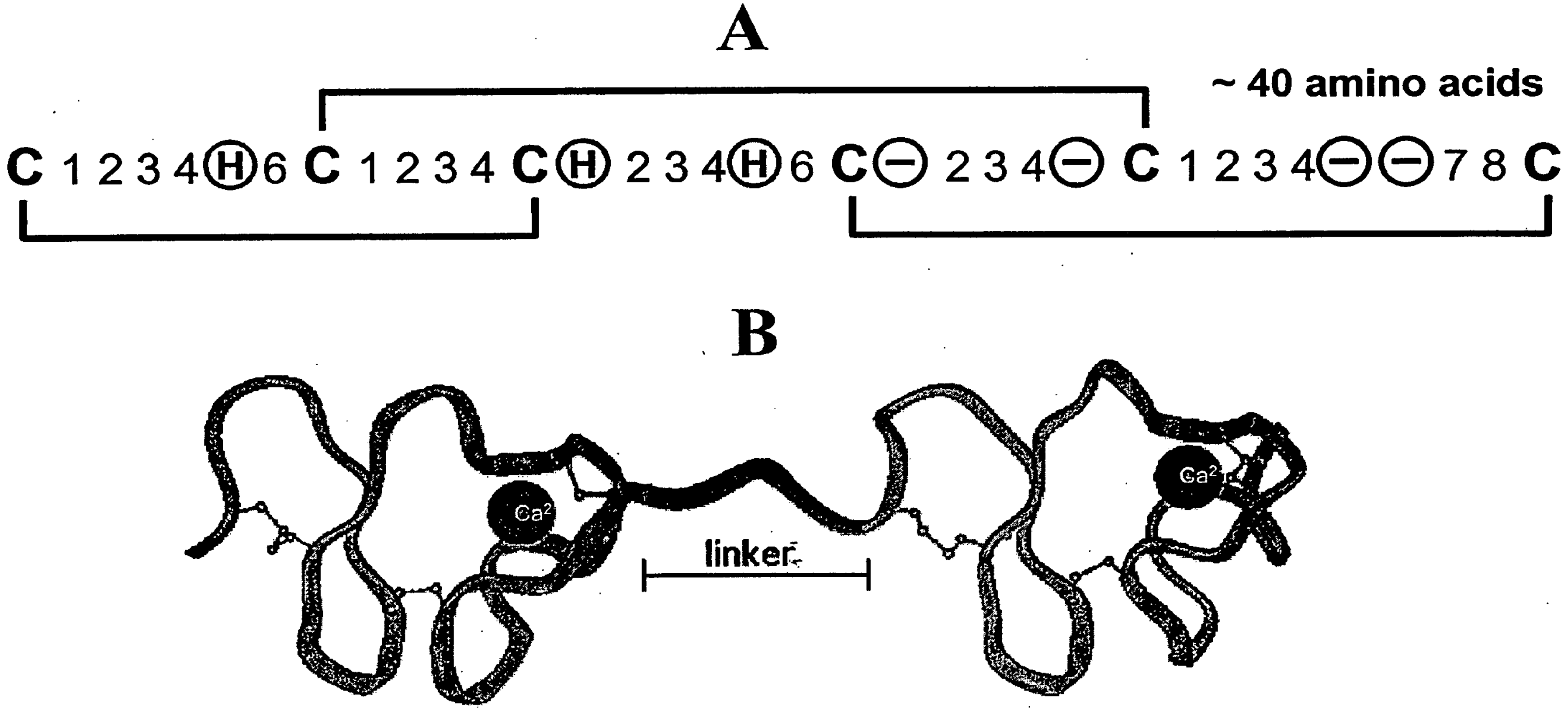


Figure 4

Ligands recognized by LDL-receptor family

- *proteases*
  - factor IXa
  - pro-uPA
  - t-PA
  - plasminogen
  - MMP-9
- *inhibitors*
  - α_2-macroglobulin
  - PAI-1
  - TFPI
  - pancreatic trypsin inhibitor
- *complexes*
  - protease/
    - α_1-antitrypsin
    - protein C inhibitor
    - protease nexin-1
    - antithrombin
    - C1-inhibitor
  - thrombin/heparin cofactor II
  - cathepsin G/α_1-antichymotrypsin
- *vitamin-carrier complexes*
  - vitamin D-bp, vitamin D
  - retinol-bp, vitamin A
  - transcobalamin, vitamin B12
- *proteins involved in lipoprotein metabolism*
  - apoB100
  - apoE
  - apoJ (clusterin)
  - apoH (β_2-glycoprotein I)
  - Lp(a)
  - hepatic lipase
  - lipoprotein lipase
  - IDL
  - VLDL
  - β-VLDL
- *non-human*
  - pseudomonas exotoxin A
  - circumsporozoite protein
  - trichosanthin
  - ricin A
  - saporin
- *antibiotics*
  - gentamicin
  - polymyxin B
- *viruses*
  - HRV2 (Rhino)
  - HCV (Flavi)
  - BVDV (Flavi)
- *miscellaneous*
  - albumin
  - transthyretin
  - β-Amyloid precursor protein
  - RAP
  - complement C3
  - lactoferrin
  - thyroglobulin
  - thrombospondin
  - saposin precursor
  - reelin
  - insulin
  - parathyroid hormone (PTH)
  - aprotinin
  - α-amylase
  - C1q
  - α_1-microglobulin
  - β_2-microglobulin
  - odorant-binding protein
  - epidermal growth factor
  - prolactin
  - lysozyme
  - connective tissue growth factor (CTGF)
  - cytochrome c
  - seminal vesicle secretory protein II
  - clara cell secretory protein (CCSP)
  - cubulin
  - factor VIII

Figure 5

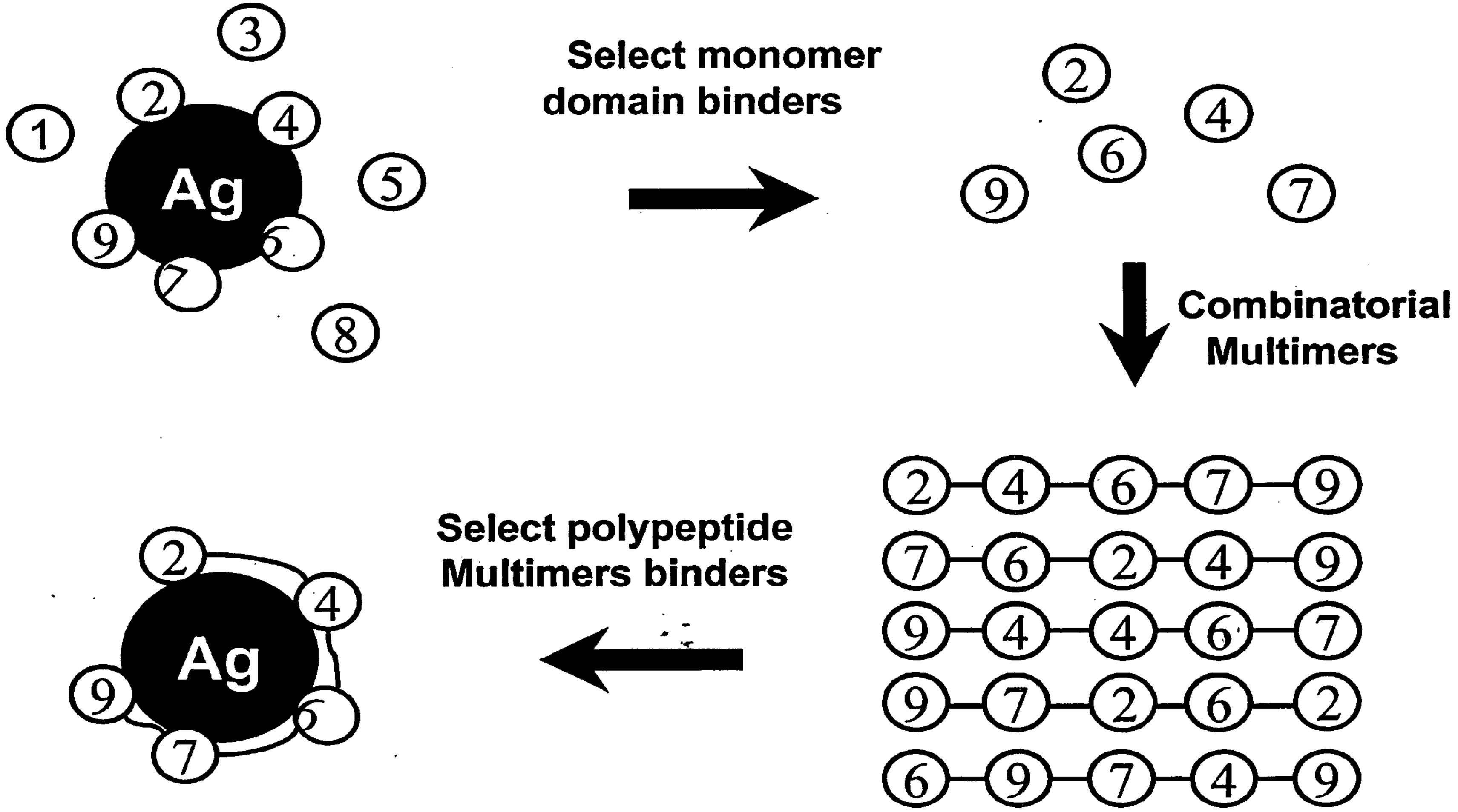

Ag
1
2
3
4
5
6
7
8
9
Select monomer domain binders
2
6
4
9
7
Combinatorial Multimers
2 4 6 7 9
7 6 2 4 9
9 4 4 6 7
9 7 2 6 2
6 9 7 4 9
Select polypeptide Multimers binders
Ag
2
4
6
7
9

Figure 6

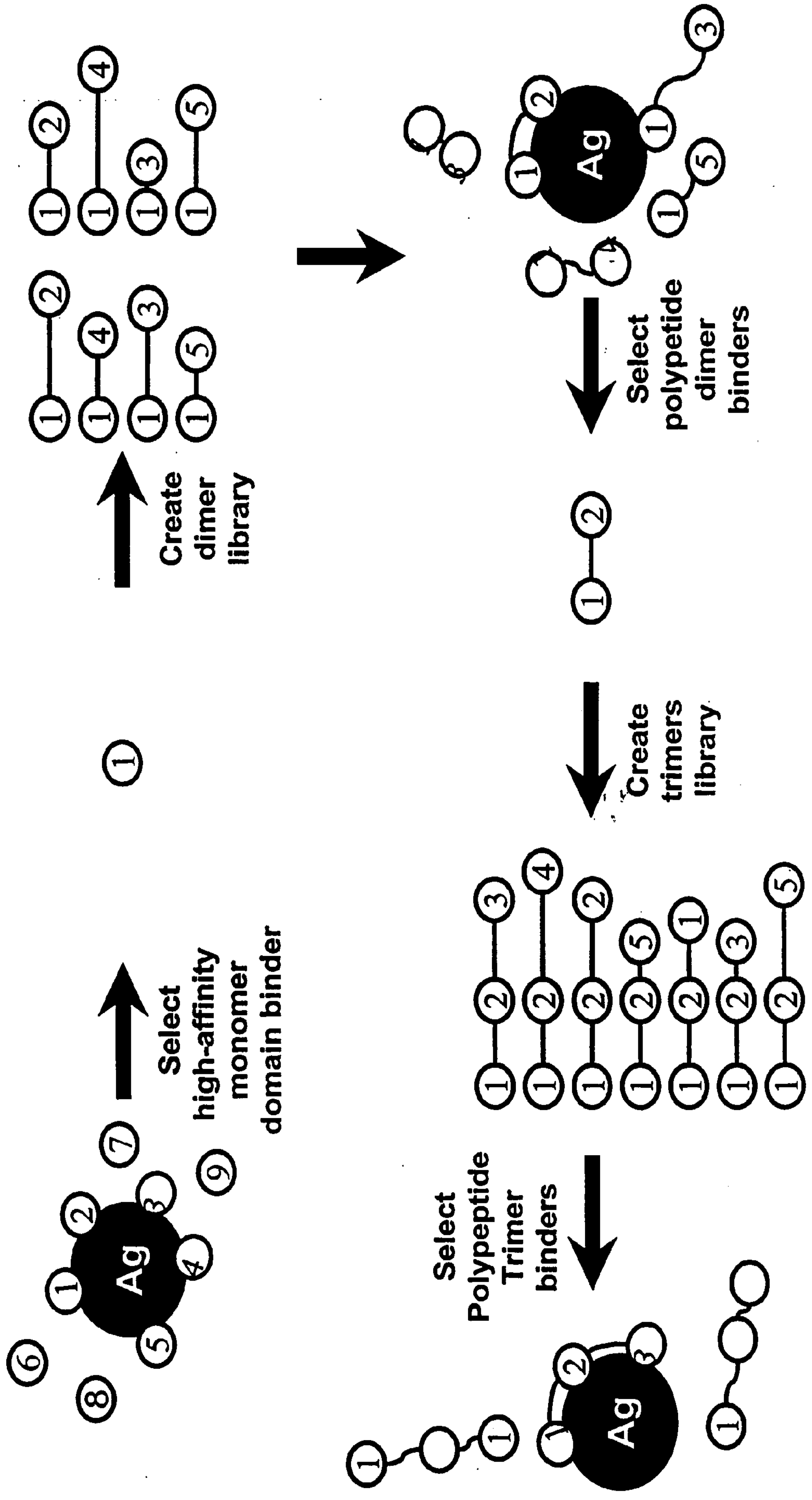

Ag
1
2
3
4
5
6
7
8
9
Select high-affinity monomer domain binder
Create dimer library
Select polypeptide dimer binders
Create trimers library
Select Polypeptide Trimer binders

Figure 7

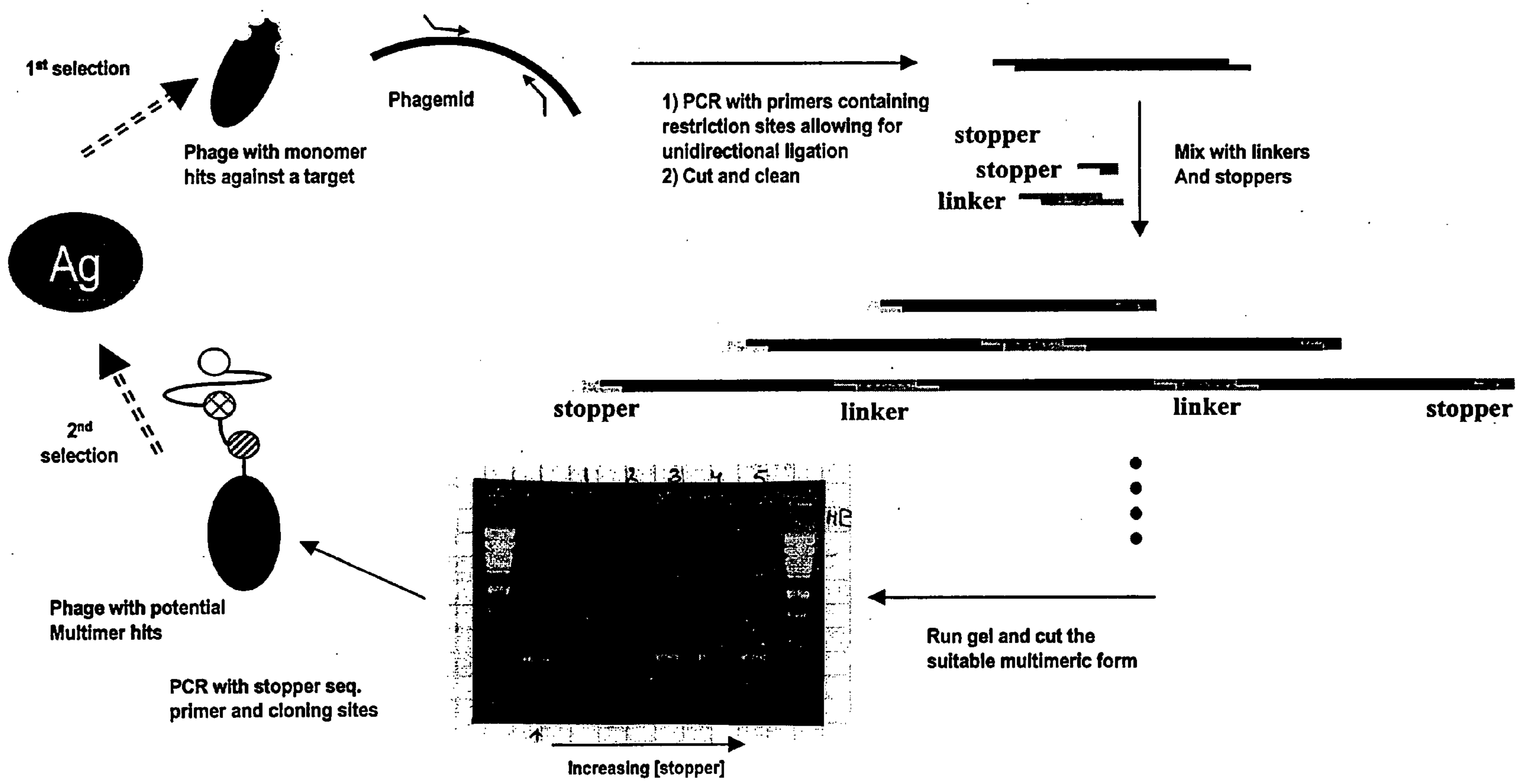

1st selection
Phage with monomer hits against a target
Phagemid
1) PCR with primers containing restriction sites allowing for unidirectional ligation
2) Cut and clean
stopper
stopper
linker
Mix with linkers And stoppers
Ag
2nd selection
stopper
linker
linker
stopper
1
2
3
4
5
Phage with potential Multimer hits
PCR with stopper seq. primer and cloning sites
Increasing [stopper]
Run gel and cut the suitable multimeric form

Figure 8

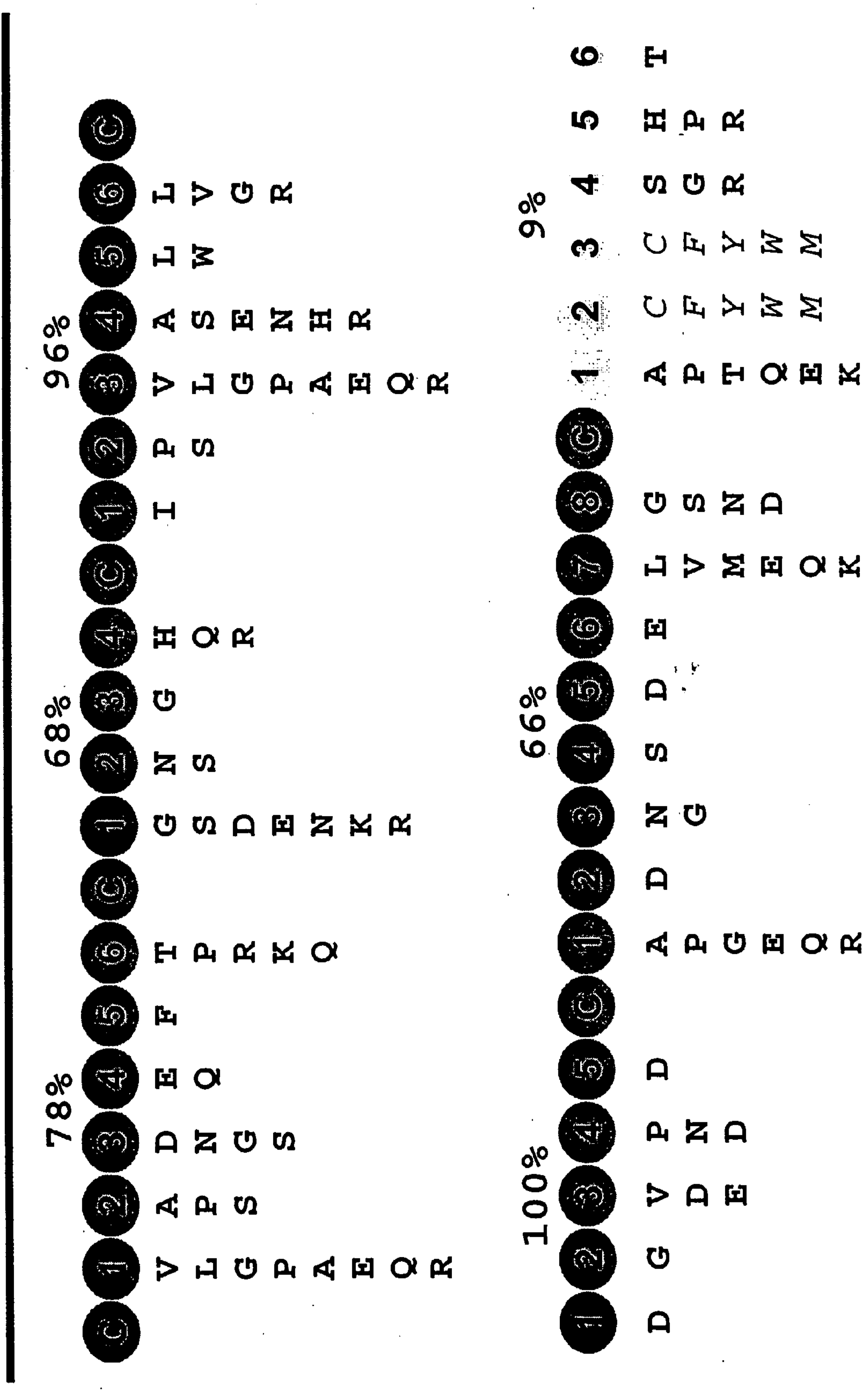
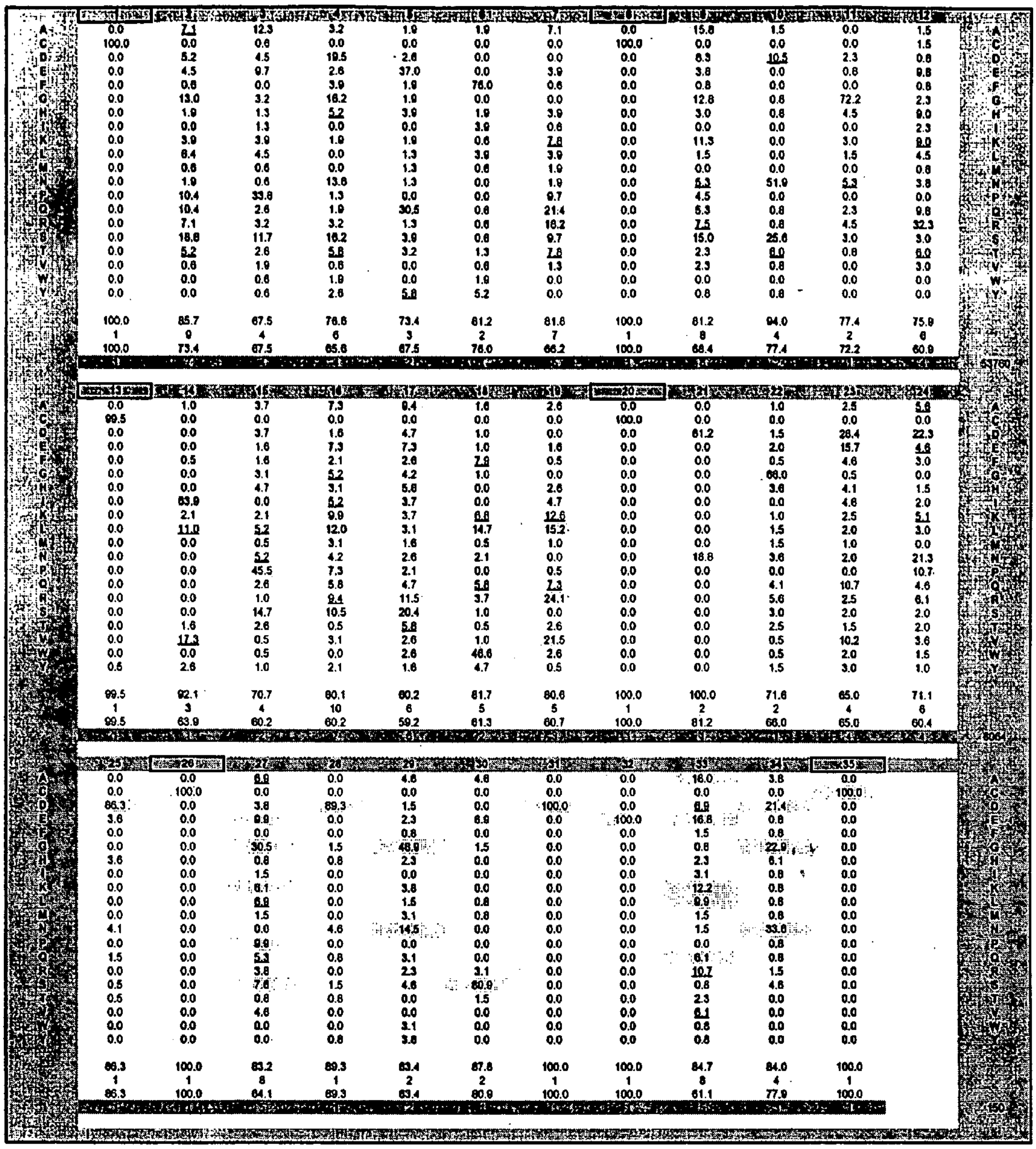

78%
C
1 V L G P A E Q R
2 A P S
3 D N G S
4 E Q
5 F
6 T P R K Q
C
68%
1 G S D E N K R
2 N S
3 G
4 H Q R
C
96%
1 I
2 P S
3 V L G P A E Q R
4 A S E N H R
5 L W
6 L V G R
C
100%
1 D
2 G
3 V D E
4 P N D
5 D
C
66%
1 A P G E Q R
2 D
3 N G
4 S
5 D
6 E
7 L V M E Q K
8 G S N D
C
9%
1 A P T Q E K
2 C F Y W M
3 C F Y W M
4 S G R
5 H P R
6 T

Figure 9



	1	2	3	4	5	6	7	8	9	10	11	12
A	0.0	7.1	12.3	3.2	1.9	1.9	7.1	0.0	15.8	1.5	0.0	1.5
C	100.0	0.0	0.6	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	1.5
D	0.0	5.2	4.5	19.5	2.6	0.0	0.0	0.0	8.3	10.5	2.3	0.8
E	0.0	4.5	9.7	2.6	37.0	0.0	3.9	0.0	3.8	0.0	0.8	9.8
F	0.0	0.6	0.0	3.9	1.9	76.0	0.6	0.0	0.8	0.0	0.0	0.8
G	0.0	13.0	3.2	16.2	1.9	0.0	0.0	0.0	12.8	0.8	72.2	2.3
H	0.0	1.9	1.3	5.2	3.9	1.9	3.9	0.0	3.0	0.8	4.5	9.0
I	0.0	0.0	1.3	0.0	0.0	3.9	0.6	0.0	0.0	0.0	0.0	2.3
K	0.0	3.9	3.9	1.9	1.9	0.6	7.8	0.0	11.3	0.0	3.0	9.0
L	0.0	6.4	4.5	0.0	1.3	3.9	3.9	0.0	1.5	0.0	1.5	4.5
M	0.0	0.6	0.6	0.0	1.3	0.6	1.9	0.0	0.0	0.0	0.0	0.8
N	0.0	1.9	0.6	13.6	1.3	0.0	1.9	0.0	5.3	51.9	5.3	3.8
P	0.0	10.4	33.8	1.3	0.0	0.0	9.7	0.0	4.5	0.0	0.0	0.0
Q	0.0	10.4	2.6	1.9	30.5	0.6	21.4	0.0	5.3	0.8	2.3	9.8
R	0.0	7.1	3.2	3.2	1.3	0.6	16.2	0.0	7.5	0.8	4.5	32.3
S	0.0	18.8	11.7	16.2	3.9	0.6	9.7	0.0	15.0	25.6	3.0	3.0
T	0.0	5.2	2.6	5.8	3.2	1.3	7.8	0.0	2.3	6.0	0.8	6.0
V	0.0	0.6	1.9	0.6	0.0	0.6	1.3	0.0	2.3	0.8	0.0	3.0
W	0.0	0.0	0.6	1.9	0.0	1.9	0.0	0.0	0.0	0.0	0.0	0.0
Y	0.0	0.0	0.6	2.6	5.8	5.2	0.0	0.0	0.8	0.8	0.0	0.0
	100.0	85.7	67.5	76.6	73.4	81.2	81.6	100.0	81.2	94.0	77.4	75.9
	1	9	4	6	3	2	7	1	8	4	2	6
	100.0	73.4	67.5	65.6	67.5	76.0	66.2	100.0	68.4	77.4	72.2	60.9

53760

	13	14	15	16	17	18	19	20	21	22	23	24
A	0.0	1.0	3.7	7.3	9.4	1.6	2.6	0.0	0.0	1.0	2.5	5.6
C	99.5	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0
D	0.0	0.0	3.7	1.6	4.7	1.0	0.0	0.0	81.2	1.5	28.4	22.3
E	0.0	0.0	1.6	7.3	7.3	1.0	1.6	0.0	0.0	2.0	15.7	4.6
F	0.0	0.5	1.6	2.1	2.6	7.8	0.5	0.0	0.0	0.5	4.6	3.0
G	0.0	0.0	3.1	5.2	4.2	1.0	0.0	0.0	0.0	66.0	0.5	0.0
H	0.0	0.0	4.7	3.1	5.8	0.0	2.6	0.0	0.0	3.6	4.1	1.5
I	0.0	63.9	0.0	5.2	3.7	0.0	4.7	0.0	0.0	0.0	4.6	2.0
K	0.0	2.1	2.1	9.9	3.7	6.8	12.6	0.0	0.0	1.0	2.5	5.1
L	0.0	11.0	5.2	12.0	3.1	14.7	15.2	0.0	0.0	1.5	2.0	3.0
M	0.0	0.0	0.5	3.1	1.6	0.5	1.0	0.0	0.0	1.5	1.0	0.0
N	0.0	0.0	5.2	4.2	2.6	2.1	0.0	0.0	16.8	3.6	2.0	21.3
P	0.0	0.0	45.5	7.3	2.1	0.0	0.5	0.0	0.0	0.0	0.0	10.7
Q	0.0	0.0	2.6	5.8	4.7	5.8	7.3	0.0	0.0	4.1	10.7	4.6
R	0.0	0.0	1.0	9.4	11.5	3.7	24.1	0.0	0.0	5.6	2.5	6.1
S	0.0	0.0	14.7	10.5	20.4	1.0	0.0	0.0	0.0	3.0	2.0	2.0
T	0.0	1.6	2.6	0.5	5.8	0.5	2.6	0.0	0.0	2.5	1.5	2.0
V	0.0	17.3	0.5	3.1	2.6	1.0	21.5	0.0	0.0	0.5	10.2	3.6
W	0.0	0.0	0.5	0.0	2.6	46.6	2.6	0.0	0.0	0.5	2.0	1.5
Y	0.5	2.6	1.0	2.1	1.6	4.7	0.5	0.0	0.0	1.5	3.0	1.0
	99.5	92.1	70.7	60.1	60.2	81.7	80.6	100.0	100.0	71.6	65.0	71.1
	1	3	4	10	6	5	5	1	2	2	4	6
	99.5	63.9	60.2	60.2	59.2	61.3	60.7	100.0	81.2	66.0	65.0	60.4

8064

	25	26	27	28	29	30	31	32	33	34	35
A	0.0	0.0	6.9	0.0	4.6	4.8	0.0	0.0	16.0	3.8	0.0
C	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
D	86.3	0.0	3.8	89.3	1.5	0.0	100.0	0.0	6.9	21.4	0.0
E	3.6	0.0	9.9	0.0	2.3	6.9	0.0	100.0	16.8	0.8	0.0
F	0.0	0.0	0.0	0.0	0.8	0.0	0.0	0.0	1.5	0.8	0.0
G	0.0	0.0	30.5	1.5	48.9	1.5	0.0	0.0	0.8	22.9	0.0
H	3.6	0.0	0.8	0.8	2.3	0.0	0.0	0.0	2.3	6.1	0.0
I	0.0	0.0	1.5	0.0	0.0	0.0	0.0	0.0	3.1	0.8	0.0
K	0.0	0.0	6.1	0.0	3.8	0.0	0.0	0.0	12.2	0.8	0.0
L	0.0	0.0	6.9	0.0	1.5	0.8	0.0	0.0	9.9	0.8	0.0
M	0.0	0.0	1.5	0.0	3.1	0.8	0.0	0.0	1.5	0.8	0.0
N	4.1	0.0	0.0	4.6	14.5	0.0	0.0	0.0	1.5	33.6	0.0
P	0.0	0.0	9.9	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.0
Q	1.5	0.0	5.3	0.8	3.1	0.0	0.0	0.0	6.1	0.8	0.0
R	0.0	0.0	3.8	0.0	2.3	3.1	0.0	0.0	10.7	1.5	0.0
S	0.5	0.0	7.6	1.5	4.6	80.9	0.0	0.0	0.8	4.6	0.0
T	0.5	0.0	0.8	0.8	0.0	1.5	0.0	0.0	2.3	0.0	0.0
V	0.0	0.0	4.6	0.0	0.0	0.0	0.0	0.0	6.1	0.0	0.0
W	0.0	0.0	0.0	0.0	3.1	0.0	0.0	0.0	0.8	0.0	0.0
Y	0.0	0.0	0.0	0.8	3.8	0.0	0.0	0.0	0.8	0.0	0.0
	86.3	100.0	83.2	89.3	63.4	87.8	100.0	100.0	84.7	84.0	100.0
	1	1	8	1	2	2	1	1	8	4	1
	86.3	100.0	64.1	89.3	63.4	80.9	100.0	100.0	61.1	77.9	100.0

150

Figure 10

```
                     a             b c               de    f ghi jk lm      nop        q
  1. IDD_HUMAN       C.........NPGQFACRSGTIQ.........CIPLP.WQCDGWATCEDE.....SDEAN......C
  2. LRP3_HUMAN      C.........QADEFRCDNGK...........CLPGP.WQCNTVDECGDG.....SDEGN......C
  3. LRP3_HUMAN      C.........PGGTFPCSGARSTR........CLPVE.RRCDGLQDCGDG.....SDEAG......C
  4. LRP3_HUMAN      C.........LPWEQPCGSSSDSDGGSLGDQGCFSEP.QRCDGWWHCASG.....RDEQG......C
  5. LRP3_HUMAN      C.........PPDQYPCEGGSGL.........CYTPA.DRCNNQKSCPDG.....ADEKN......C
  6. LRP3_HUMAN      C.........QPGTFHCGTNL...........CIFET.WRCDGQEDCQDG.....SDEHG......C
  7. LRP5_HUMAN      C.........SPDQFACATGEID.........CIPGA.WRCDGFPECDDQ.....SDEEG......C
  8. LRP5_HUMAN      C.........SAAQFPCARGQ...........CVDLR.LRCDGEADCQDR.....SDEVD......C
  9. LRP5_HUMAN      C.........LPNQFRCASGQ...........CVLIK.QQCDSFPDCIDG.....SDELM......C
 10. LRP6_HUMAN      C.........SPQQFTCFTGEID.........CIPVA.WRCDGFTECEDH.....SDELN......C
 11. LRP6_HUMAN      C.........SESQFQCASGQ...........CIDGA.LRCNGDANCQDK.....SDEKN......C
 12. LRP6_HUMAN      C.........LIDQFRCANGQ...........CIGKH.KKCDHNVDCSDK.....SDELD......C
 13. ST7_HUMAN       C.........ACDQFRCGNGK...........CIPEA.WKCNNMDECGDS.....SDEEI......C
 14. ST7_HUMAN       C.........AYNQFQCLSRFTKVYT......CLPES.LKCDGNIDCLDL.....GDEID......C
 15. ST7_HUMAN       C.........LPWEIPCGGNWG..........CYTEQ.QRCDGYWHCPNG.....RDETN......C
 16. ST7_HUMAN       C.........QKEEFPCSRNGV..........CYPRS.DRCNYQNHCPNG.....SDEKN......C
 17. ST7_HUMAN       C.........QPGNFHCKNNR...........CVFES.WVCDSQDDCGDG.....SDEEN......C
 18. CORI_HUMAN      C........GRGENFLCASGI...........CIPGK.LQCNGYNDCDDW.....SDEAH......C
 19. CORI_HUMAN      C.........SENLFHCHTGK...........CLNYS.LVCDGYDDCGDL.....SDEQN......C
 20. CORI_HUMAN      C........NPTTEHRCGDGR...........CIAME.WVCDGDHDCVDK.....SDEVN......C
 21. CORI_HUMAN      C........HSQGLVECRNGQ...........CIPST.FQCDGDEDCKDG.....SDEEN......C
 22. CORI_HUMAN      C.........SPSHFKCRSGQ...........CVLAS.RRCDGQADCDDD.....SDEEN......C
 23. CORI_HUMAN      C........KERDLWECPSNKQ..........CLKHT.VICDGFPDCPDY.....MDEKN......C
 24. CORI_HUMAN      C.........QDDELECANHA...........CVSRD.LWCDGEADCSDS.....SDEWD......C
 25. TMS2_HUMAN      C.........SNSGIECDSSGT..........CINPS.NWCDGVSHCPGG.....EDENR......C
 26. TMS3_HUMAN      C..........SGKYRCRSSFK..........CIELI.ARCDGVSDCKDG.....EDEYR......C
 27. MAT_HUMAN       C..........PGQFTCRTGR...........CIRKE.LRCDGWADCTDH.....SDELN......C
 28. MAT_HUMAN       C........DAGHQFTCKNKF...........CKPLF.WVCDSVNDCGDN.....SDEQG......C
 29. MAT_HUMAN       C.........PAQTFRCSNGK...........CLSKS.QQCNGKDDCGDG.....SDEAS......C
 30. MAT_HUMAN       C.........TKHTYRCLNGL...........CLSKGNPECDGKEDCSDG.....SDEKD......C
 31. ENTK_HUMAN      C.........LPGSSPCTDALT..........CIKAD.LFCDGEVNCPDG.....SDEDNKM....C
 32. ENTK_HUMAN      C.........KADHFQCKNGE...........CVPLV.NLCDGHLHCEDG.....SDEAD......C
 33. HAI1_HUMAN      C.........QPTQFRCSNGC...........CIDSF.LECDDTPNCPDA.....SDEAA......C
 34. CPAI_HUMAN      C.YTQKADSPMDDFFQCVNGK...........YISQM.KACDGINDCGDQ.....SDEL.......C
 35. CPAI_HUMAN      C.........QGKGFHCKSGV...........CIPSQ.YQCNGEVDCITG.....EDEVG......C
 36. CO6_HUMAN       C..........KNKFRCDSGR...........CIARK.LECNGENDCGDN.....SDERD......C
 37. CO7_HUMAN       C..........GERFRCFSGQ...........CISKS.LVCNGDSDCDEDS....ADEDR......C
 38. CO8A_HUMAN      C..........GQDFQCKETGR..........CLKRH.LVCNGDQDCLDG.....SDEDD......C
 39. CO8B_HUMAN      C...........EGFVCAQTGR..........CVNRR.LLCNGDNDCGDQ.....SDEAN......C
 40. CO9_HUMAN       C..........GNDFQCSTGR...........CIKMR.LRCNGDNDCGDF.....SDEDD......C
 41. PERL_HUMAN      C.........TEAEFACHSYNE..........CVALE.YRCDRRPDCRDM.....SDELN......C
 42. PERL_HUMAN      C.........GPQEAACRNGH...........CIPRD.YLCDGQEDCEDG.....SDELD......C
 43. PERL_HUMAN      C.........EPNEFPCGNGH...........CALKL.WRCDGDFDCEDR.....TDEAN......C
 44. PERL_HUMAN      C.........GPTQFRCVSTNM..........CIPAS.FHCDEESDCPDR.....SDEFG......C
 45. SORL_HUMAN      C.........LRNQYRCSNGN...........CINSI.WWCDFDNDCGDM.....SDERN......C
 46. SORL_HUMAN      C........DLDTQFRCQESGT..........CIPLS.YKCDLEDDCGDN.....SDESH......C
 47. SORL_HUMAN      C.........RSDEYNCSSGM...........CIRSS.WVCDGDNDCRDW.....SDEAN......C
 48. SORL_HUMAN      C.........EASNFQCRNGH...........CIPQR.WACDGDTDCQDG.....SDEDPVN....C
 49. SORL_HUMAN      C...........NGFRCPNGT...........CIPSS.KHCDGLRDCSDG.....SDEQH......C
 50. SORL_HUMAN      C........THFMDFVCKNRQQ..........CLFHS.MVCDGIIQCRDG.....SDEDAAFAG..C
 51. SORL_HUMAN      C.........DEFGFQCQNGV...........CISLI.WKCDGMDDCGDY.....SDEAN......C
 52. SORL_HUMAN      C........SRYFQFRCENGH...........CIPNR.WKCDRENDCGDW.....SDEKD......C
 53. SORL_HUMAN      C.........LPNYYRCSSGT...........CVMDT.WVCDGYRDCADG.....SDEEA......C
 54. SORL_HUMAN      C.........DRFEFECHQPKT..........CIPNW.KRCDGHQDCQDG.....RDEAN......C
 55. SORL_HUMAN      C.........MSREFQCEDGEA..........CIVLS.ERCDGFLDCSDE.....SDEKA......C
 56. APOER2_HUM      C.........EKDQFQCRNER...........CIPSV.WRCDEDDDCLDH.....SDEDD......C
 57. APOER2_HUM      C.........ADSDFTCDNGH...........CIHER.WKCDGEEECPDG.....SDESEAT....C
 58. APOER2_HUM      C.........PAEKLSCGPTSHK.........CVPAS.WRCDGEKDCEGG.....ADEAG......C
 59. APOER2_HUM      C.........APHEFQCGNRS...........CLAAV.FVCDGDDDCGDG.....SDERG......C
 60. APOER2_HUM      C.........GPREFRCGGDGGGA........CIPER.WVCDRQFDCEDR.....SDEAAEL....C
 61. APOER2_HUM      C........ATVSQFACRSGE...........CVHLG.WRCDGDRDCKDK.....SDEAD......C
 62. APOER2_HUM      C.........RGDEFQCGDGT...........CVLAI.KHCNQEQDCPDG.....SDEAG......C
 63. LDLR_HUMAN      C.........ERNEFQCQDGK...........CISYK.WVCDGSAECQDG.....SDESQET....C
 64. LDLR_HUMAN      C.........KSGDFSCGGRVNR.........CIPQF.WRCDGQVDCDNG.....SDEQG......C
 65. LDLR_HUMAN      C.........SQDEFRCHDGK...........CISRQ.FVCDSDRDCLDG.....SDEAS......C
 66. LDLR_HUMAN      C.........GPASFQCNSST...........CIPQL.WACDNDPDCEDG.....SDEWPQR....C
 67. LDLR_HUMAN      C.........SAFEFHCLSGE...........CIHSS.WRCDGGPDCKDK.....SDEEN......C
 68. LDLR_HUMAN      C.........RPDEFQCSDGN...........CIHGS.RQCDREYDCKDM.....SDEVG......C
 69. LDLR_HUMAN      C........EGPNKFKCHSGE...........CITLD.KVCNMARDCRDW.....SDEPIKE....C
 70. LDVR_HUMAN      C.........EPSQFQCTNGR...........CITLL.WKCDGDEDCVDG.....SDEKN......C
 71. LDVR_HUMAN      C.........AESDFVCNNGQ...........CVPSR.WKCDGDPDCEDG.....SDESPEQ....C
 72. LDVR_HUMAN      C.........RIHEISCGAHSTQ.........CIPVS.WRCDGENDCDSG.....EDEEN......C
 73. LDVR_HUMAN      C.........SPDEFTCSSGR...........CISRN.FVCNGQDDCSDG.....SDELD......C
 74. LDVR_HUMAN      C.........GAHEFQCSTSS...........CIPIS.WVCDDDADCSDQ.....SDESLEQ....C
 75. LDVR_HUMAN      C.........PASEIQCGSGE...........CIHKK.WRCDGDPDCKDG.....SDEVN......C
 76. LDVR_HUMAN      C.........RPDQFECEDGS...........CIHGS.RQCNGIRDCVDG.....SDEVN......C
 77. LDVR_HUMAN      C........LGPGKFKCRSGE...........CIDIS.KVCNQEQDCRDW.....SDEPLKE....C
 78. LRP1_HUMAN      C.........SPKQFACRDQIT..........CISKG.WRCDGERDCPDG.....SDEAPEI....C
 79. LRP1_HUMAN      C.........QPNEHNCLGTEL..........CVPMS.RLCNGVQDCMDG.....SDEGPH.....C
 80. LRP1_HUMAN      C.........QPGEFACANSR...........CIQER.WKCDGDNDCLDN.....SDEAPAL....C
 81. LRP1_HUMAN      C.........PSDRFKCENNR...........CIPNR.WLCDGDNDCGNS.....EDESNAT....C
 82. LRP1_HUMAN      C.........PPNQFSCASGR...........CIPIS.WTCDLDDDCGDR.....SDESAS.....C
```

```
83. LRP1_HUMAN            C........FPLTQFTCNNGR..........CININ.WRCDNDNDCGDN......SDEAG......C
84. LRP1_HUMAN            C.........SSTQFKCNSGR..........CIPEH.WTCDGDNDCGDY......SDETHAN....C
85. LRP1_HUMAN            C.........HTDEFQCRLDGL.........CIPLR.WRCDGDTDCMDS......SDEKS......C
86. LRP1_HUMAN            C........DPSVKFGCKDSAR.........CISKA.WVCDGDNDCEDN......SDEEN......C
87. LRP1_HUMAN            C.........RPPSHPCANNTSV........CLPPD.KLCDGNDDCGDG......SDEGEL.....C
88. LRP1_HUMAN            C........RAQDEFECANGE..........CINFS.LTCDGVPHCKDK......SDEKPSY....C
89. LRP1_HUMAN            C.........KKTFRQCSNGR..........CVSNM.LWCNGADDCGDG......SDEIP......C
90. LRP1_HUMAN            C.........GVGEFRCRDGT..........CIGNS.SRCNQFVDCEDA......SDEMN......C
91. LRP1_HUMAN            CSSYFRLGVKGVLPQPCERTSL.........CYAPS.WVCDGANDCGDY......SDERD......C
92. LRP1_HUMAN            C.........PLNYFACPSGR..........CIPMS.WTCDKEDDCEHG......EDETH......C
93. LRP1_HUMAN            C.........SEAQFECQNHR..........CISKQ.WLCDGSDDCGDG......SDEAAH.....C
94. LRP1_HUMAN            C.........GPSSFSCPGTHV.........CVPER.WLCDGDKDCADG......ADESIAAG...C
95. LRP1_HUMAN            C.........DDREFMCQNRQ..........CIPKH.FVCDHDRDCADG......SDESPE.....C
96. LRP1_HUMAN            C.........GPSEFRCANGR..........CLSSRQWECDGENDCHDQ......SDEAPKNPH..C
97. LRP1_HUMAN            C........NASSQFLCSSGR..........CVAEA.LLCNGQDDCGDS......SDERG......C
98. LRP1_HUMAN            C.........TASQFVCKNDK..........CIPFW.WKCDTEDDCGDH......SDEPPD.....C
99. LRP1_HUMAN            C.........RPGQFQCSTGI..........CTNPA.FICDGDNDCQDN......SDEAN......C
100.   LRP1_HUMAN         C.........LPSQFKCTNTNR.........CIPGI.FRCNGQDNCGDG......EDERD......C
101.   LRP1_HUMAN         C.........APNQFQCSITKR.........CIPRV.WVCDRDNDCVDG......SDEPAN.....C
102.   LRP1_HUMAN         C.........GVDEFRCKDSGR.........CIPAR.WKCDGEDDCGDG......SDEPKEE....C
103.   LRP1_HUMAN         C.........EPYQFRCKNNR..........CVPGR.WQCDYDNDCGDN......SDEES......C
104.   LRP1_HUMAN         C.........SESEFSCANGR..........CIAGR.WKCDGDHDCADG......SDEKD......C
105.   LRP1_HUMAN         C.........DMDQFQCKSGH..........CIPLR.WRCDADADCMDG......SDEEA......C
106.   LRP1_HUMAN         C.........PLDEFQCNNTL..........CKPLA.WKCDGEDDCGDN......SDENPEE....C
107.   LRP1_HUMAN         C........PPNRPFRCKNDRV.........CLWIG.RQCDGTDNCGDG......TDEED......C
108.   LRP1_HUMAN         C........KDKKEFLCRNQR..........CLSSS.LRCNMFDDCGDG......SDEED......C
109.   LRP2_HUMAN         C.........DSAHFRCGSGH..........CIPAD.WRCDGTKDCSDD......ADEIG......C
110.   LRP2_HUMAN         C.........QQGYFKCQSEGQ.........CIPSS.WVCDQDQDCDDG......SDERQD.....C
111.   LRP2_HUMAN         C.........SSHQITCSNGQ..........CIPSE.YRCDHVRDCPDG......ADEND......C
112.   LRP2_HUMAN         C...........EQLTCDNGA..........CYNTS.QKCDWKVDCRDS......SDEIN......C
113.   LRP2_HUMAN         C.........LHNEFSCGNGE..........CIPRA.YVCDHDNDCQDG......SDEHA......C
114.   LRP2_HUMAN         C.........GGYQFTCPSGR..........CIYQN.WVCDGEDDCKDN......GDEDG......C
115.   LRP2_HUMAN         C.........SPREWSCPESGR.........CISIY.KVCDGILDCPGR......EDENNTSTGKYC
116.   LRP2_HUMAN         C.........GLFSFPCKNGR..........CVPNY.YLCDGVDDCHDN......SDEQL......C
117.   LRP2_HUMAN         C.........SSSAFTCGHGE..........CIPAH.WRCDKRNDCVDG......SDEHN......C
118.   LRP2_HUMAN         C.........LDTQYTCDNHQ..........CISKN.WVCDTDNDCGDG......SDEKN......C
119.   LRP2_HUMAN         C.........QPSQFNCPNHR..........CIDLS.FVCDGDKDCVDG......SDEVG......C
120.   LRP2_HUMAN         C.........TASQFKCASGDK.........CIGVT.NRCDGVFDCSDN......SDEAG......C
121.   LRP2_HUMAN         C.........HSDEFQCQEDGI.........CIPNF.WECDGHPDCLYG......SDEHNA.....C
122.   LRP2_HUMAN         C.........PSSYFHCDNGN..........CIHRA.WLCDRDNDCGDM......SDEKD......C
123.   LRP2_HUMAN         C.........PSWQWQCLGHNI.........CVNLS.VVCDGIFDCPNG......TDESPL.....C
124.   LRP2_HUMAN         C.........GASSFTCSNGR..........CISEE.WKCDNDNDCGDG......SDEMESV....C
125.   LRP2_HUMAN         C.........SPTAFTCANGR..........CVQYS.YRCDYYNDCGDG......SDEAG......C
126.   LRP2_HUMAN         C........NATTEFMCNNRR..........CIPRE.FICNGVDNCHDNNT....SDEKN......C
127.   LRP2_HUMAN         C.........QSGYTKCHNSNI.........CIPRV.YLCDGDNDCGDN......SDENPTY....C
128.   LRP2_HUMAN         C.........SSSEFQCASGR..........CIPQH.WYCDQETDCFDA......SDEPAS.....C
129.   LRP2_HUMAN         C.........LADEFKCDGGR..........CIPSE.WICDGDNDCGDM......SDEDKRHQ...C
130.   LRP2_HUMAN         C.........SDSEFLCVNDRPPDRR.....CIPQS.WVCDGDVDCTDG......YDENQN.....C
131.   LRP2_HUMAN         C.........SENEFTCGYGL..........CIPKI.FRCDRHNDCGDY......SDERG......C
132.   LRP2_HUMAN         C.........QQNQFTCQNGR..........CISKT.FVCDEDNDCGDG......SDELMHL....C
133.   LRP2_HUMAN         C.........PPHEFKCDNGR..........CIEMM.KLCNHLDDCLDN......SDEKG......C
134.   LRP2_HUMAN         C.........SSTQFLCANNEK.........CIPIW.WKCDGQKDCSDG......SDELAL.....C
135.   LRP2_HUMAN         C.........RLGQFQCSDGN..........CTSPQ.TLCNAHQNCPDG......SDEDRLL....C
136.   LRP2_HUMAN         C.........DSNEWQCANKR..........CIPES.WQCDTFNDCEDN......SDEDSSH....C
137.   LRP2_HUMAN         C.........RPGQFRCANGR..........CIPQA.WKCDVDNDCGDH......SDEPIEE....C
138.   LRP2_HUMAN         C........DNFTEFSCKTNYR.........CIPKW.AVCNGVDDCRDN......SDEQG......C
139.   LRP2_HUMAN         C........HPVGDFRCKNHH..........CIPLR.WQCDGQNDCGDN......SDEEN......C
140.   LRP2_HUMAN         C.........TESEFRCVNQQ..........CIPSR.WICDHYNDCGDN......SDERD......C
141.   LRP2_HUMAN         C.........HPEYFQCTSGH..........CVHSE.LKCDGSADCLDA......SDEAD......C
142.   LRP2_HUMAN         C.........QATMFECKNHV..........CIPPY.WKCDGDDDCGDG......SDEELHL....C
143.   LRP2_HUMAN         C........NSPNRFRCDNNR..........CIYSH.EVCNGVDDCGDG......TDETEEH....C
144.   LRP2_HUMAN         C.........TEYEYKCGNGH..........CIPHD.NVCDDADDCGDW......SDELG......C
145.   LR1B_HUMAN         C.........DPGEFLCHDHVT.........CVSQS.WLCDGDPDCPDD......SDESLDT....C
146.   LR1B_HUMAN         C.........PLNHIACLGTNK.........CVHLS.QLCNGVLDCPDG......YDEGVH.....C
147.   LR1B_HUMAN         C.........KAGEFRCKNRH..........CIQAR.WKCDGDDDCLDG......SDEDSVN....C
148.   LR1B_HUMAN         C.........PDDQFKCQNNR..........CIPKR.WLCDGANDCGSN......EDESNQT....C
149.   LR1B_HUMAN         C.........QVDQFSCGNGR..........CIPRA.WLCDREDDCGDQ......TDEMAS.....C
150.   LR1B_HUMAN         C........EPLTQFVCKSGR..........CISSK.WHCDSDDDCGDG......SDEVG......C
151.   LR1B_HUMAN         C.........FDNQFRCSSGR..........CIPGH.WACDGDNDCGDP......SDEAQIN....C
152.   LR1B_HUMAN         C.........NGNEFQCHPDGN.........CVPDL.WRCDGEKDCEDG......SDEKG......C
153.   LR1B_HUMAN         C........DHKTKFSCWSTGR.........CINKA.WVCDGDIDCEDQ......SDEDD......C
154.   LR1B_HUMAN         C.........GPPKHPCANDTSV........CLQPE.KLCNGKKDCPDG......SDEGYL.....C
155.   LR1B_HUMAN         C........NAYSEFECGNGE..........CIDYQ.LTCDGIPHCKDK......SDEKLLY....C
156.   LR1B_HUMAN         C.........RRGFKPCYNRR..........CIPHG.KLCDGENDCGDN......SDELD......C
157.   LR1B_HUMAN         C.........ATVEFRCADGT..........CIPRS.ARCNQNIDCADA......SDEKN......C
158.   LR1B_HUMAN         CTHFYKLGVKTTGFIRCNSTSL.........CVLPT.WICDGSNDCGDY......SDELK......C
159.   LR1B_HUMAN         C.........EENYFSCPSGR..........CILNT.WICDGQKDCEDG......RDEFH......C
160.   LR1B_HUMAN         C.........SWNQFACSAQK..........CISKH.WICDGEDDCGDG......LDESDSI....C
161.   LR1B_HUMAN         C.........AADMFSCQGSRA.........CVPRH.WLCDGERDCPDG......SDELSTAG...C
162.   LR1B_HUMAN         C.........DENAFMCHNKV..........CIPKQ.FVCDHDDDCGDG......SDESPQ.....C
163.   LR1B_HUMAN         C.........GTEEFSCADGR..........CLLNTQWQCDGDFDCPDH......SDEAPLNPK..C
164.   LR1B_HUMAN         C.........NSSFFMCKNGR..........CIPSG.GLCDNKDDCGDG......SDERN......C
165.   LR1B_HUMAN         C.........TASQFRCKTDK..........CIPFW.WKCDTVDDCGDG......SDEPDD.....C
166.   LR1B_HUMAN         C.........QPGRFQCGTGL..........CALPA.FICDGENDCGDN......SDELN......C
167.   LR1B_HUMAN         C.........LSGQFKCTKNQK.........CIPVN.LRCNGQDDCGDE......EDERD......C
168.   LR1B_HUMAN         C.........SPDYFQCKTTKH.........CISKL.WVCDEDPDCADA......SDEAN......C
169.   LR1B_HUMAN         C.........GPHEFQCKNNN..........CIPDH.WRCDSQNDCSDN......SDEEN......C
```

```
170.   LR1B_HUMAN          C.........TLKDFLCANGD...........CVSSR.FWCDGDFDCADG......SDERN......C
171.   LR1B_HUMAN          C.........SKDQFRCSNGQ...........CIPAK.WKCDGHEDCKYG......EDEKS......C
172.   LR1B_HUMAN          C.........SSREYICASDG...........CISAS.LKCNGEYDCADG......SDEMD......C
173.   LR1B_HUMAN          C.........KEDQFRCKNKAH..........CIPIR.WLCDGIHDCVDG......SDEEN......C
174.   LR1B_HUMAN          C.........RADEFLCNNSL...........CKLHF.WVCDGEDDCGDN......SDEAPDM....C
175.   LR1B_HUMAN          C........PSTRPHRCRNNRI..........CLQSE.QMCNGIDECGDN......SDEDH......C
176.   LR1B_HUMAN          C.........KKDEPACSNKK...........CIPMD.LQCDRLDDCGDG......SDEQG......C

177.   075851              C.........AEGEALCQENGH..........CVPHG.WLCDNQDDCGDG......SDEEGE.....C
178.   075851              C.........GEGQMTCSSGH...........CLPLA.LLCDRQDDCGDG......TDEPSYP....C
179.   075851              C.........PQGLLACADGR...........CLPPA.LLCDGHPDCLDA......ADEES......C
180.   075851              C.........VPGEVSCVDGT...........CLGAI.QLCDGVWDCPDG......ADEGPGH....C
181.   075851              C..........PGLFPCGVAPGL.........CLTPE.QLCDGIPDCPQG......EDELD......C
182.   075851              C...........PEYTCPNGT...........CIGFQ.LVCDGQPDCGRPGQVGPSPEEQG......C
183.   075851              C.........SPSQLSCGSGE...........CLSAE.RRCDLRPDCQDG......SDEDG......C
184.   075851              C........EPGVGLRCASGE...........CVLRG.GPCDGVLDCEDG......SDEEG......C
185.   075851 =
       ENSP00000262089     C.........GPFEFRCGSGE...........CTPRG.WRCDQEEDCADG......SDERG......C
186.   ENSP00000262089     C.........APHHAPCARGPH..........CVSPE.QLCDGVRQCPDG......SDEGPDA....C
187.   ENSP00000262089     C.........GPGQTPCEVLG...........CVEQA.QVCDGREDCLDG......SDERH......C
188.   C18oRF1             C............KFTCTSGK...........CLYLGSLVCNQQNDCGDN......SDEEN......C
189.   AAH07083            C.........PPTKFQCRTSGL..........CVPLT.WRCDRDLDCSDG......SDEEE......C
190.   AAH07083            C.........LAGELRCTLSDD..........CIPLT.WRCDGHPDCPDS......SDELG......C
191.   Q9HBX9              C.........SLGYFPCGNITK..........CLPQL.LHCNGVDDCGNQ......ADEDN......C
192.   Q9BY79              C.........AHDEFRCDQLI...........CLLPD.SVCDGFANCADG......SDETN......C
193.   Q9BY79              C.........GPSELSCQAGG...........CKGVQ.WMCDMWRDCTDG......SDDN.......C
194.   BAB55257 =
       ENSP00000239367     C.........SRYHFFCDDGC...........CIDIT.LACDGVQQCPDG......SDEDF......C
195.   095518 =           C..........PGEFLCSVNGL..........CVPA....CDGVKDCPNG......LDERN......C
       ENSP00000255793
196.   ENSP00000255793     C..........RATFQCKEDST..........CISLP.KVCDGQPDCLNG......SDEEQ......C
197.   ENSP00000255793     C.........GTFTFQCEDRS...........CVKKPNPQCDGRPDCRDG......SDEEH......C
                           a             b c               de    f ghi  jk lm      nop        q
```

a.	C(197)	100%
b.	F(150)	76%
c.	C(197)	100%
d.	C(196)	99%
e.	I(122); V(35); L(28)	94%
f.	W (91); L(29); F(15); Y(9)	73%
g.	C(197)	100%
h.	D(160); N(37)	100%
i.	G(130);	66%
j.	D(170); N(8)	90%
k.	C(197)	100%
l.	D(177); N(8)	94%
m.	G(101); N(26); A(7); S(6)	71%
n.	S(158)	80%
o.	D(196)	99%
p.	E(196)	99%
q.	C(197)	100%

Figure 11

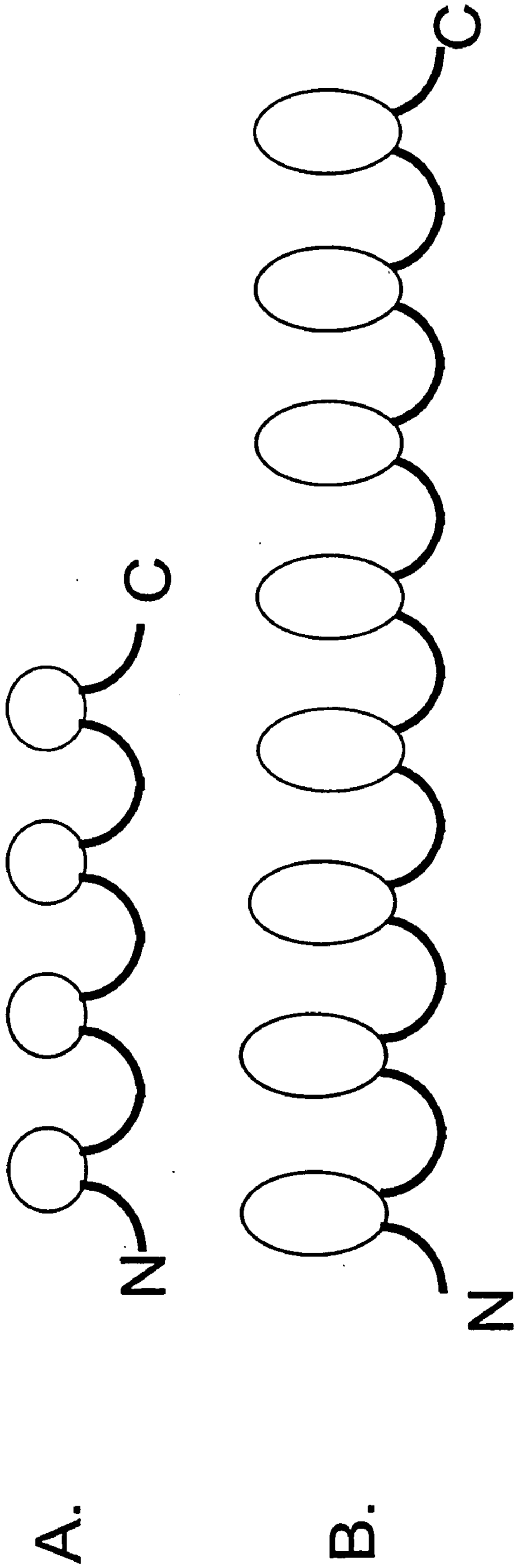
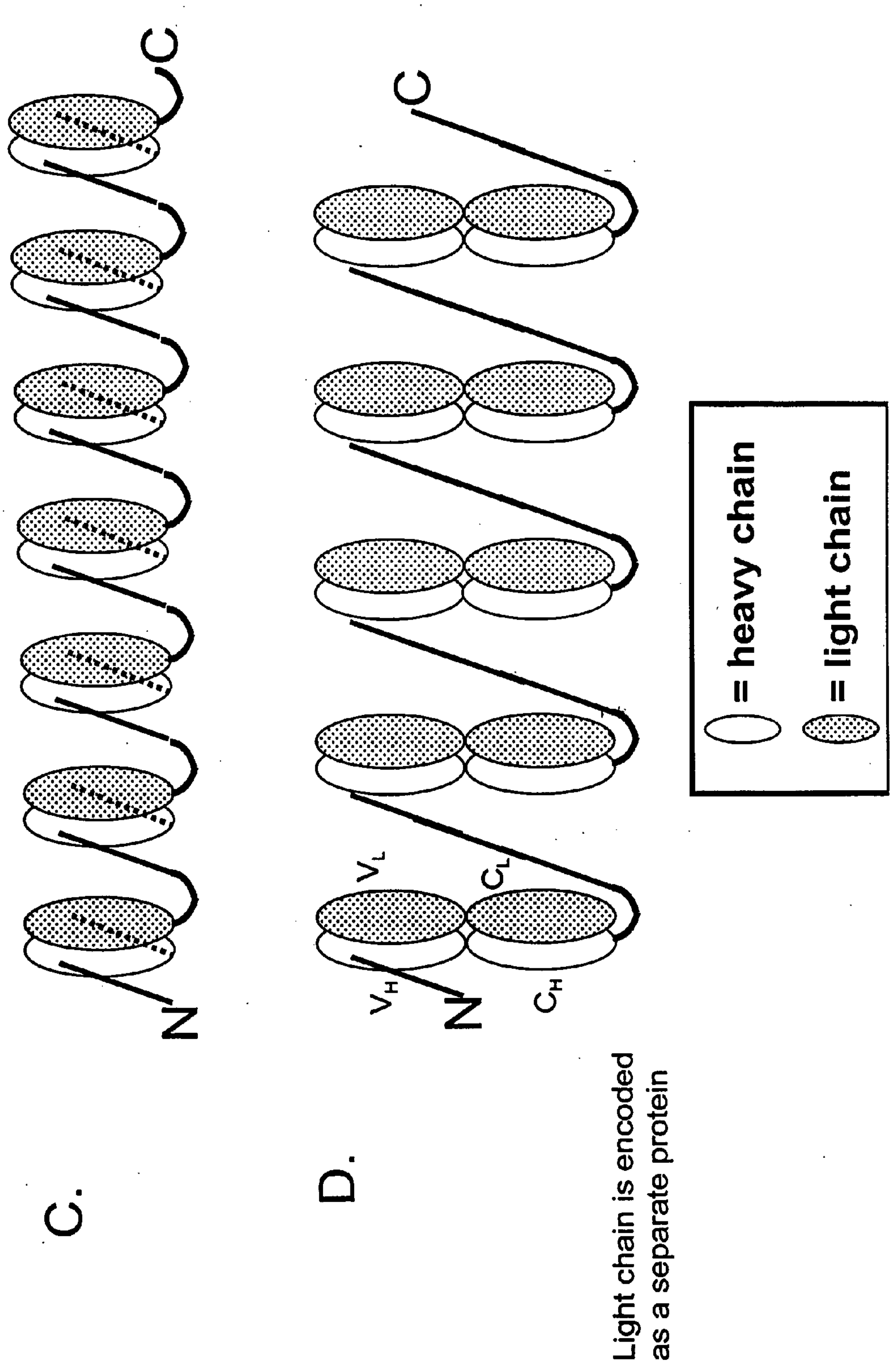


Figure 11



Light chain is encoded as a separate protein

Figure 12

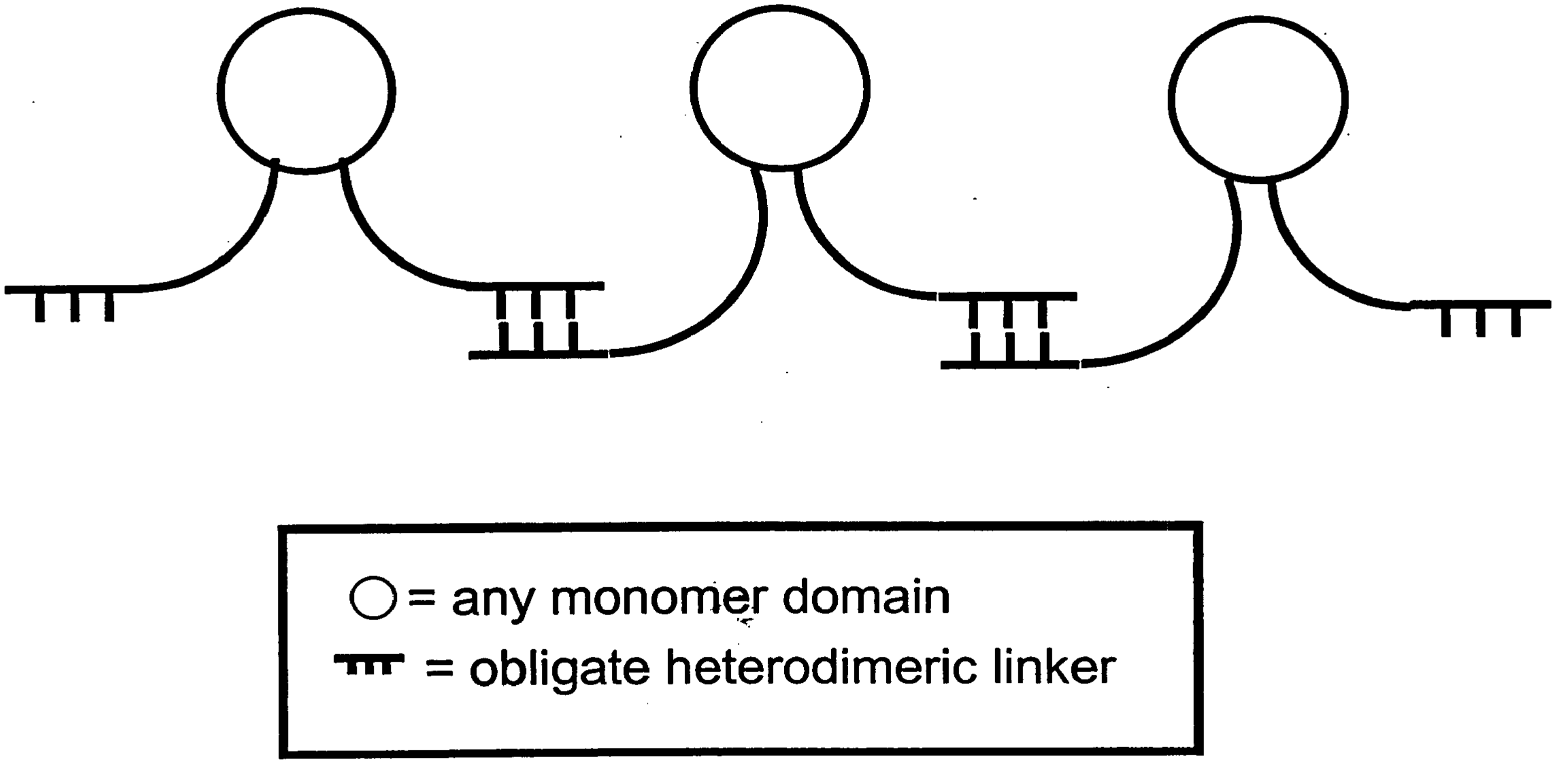
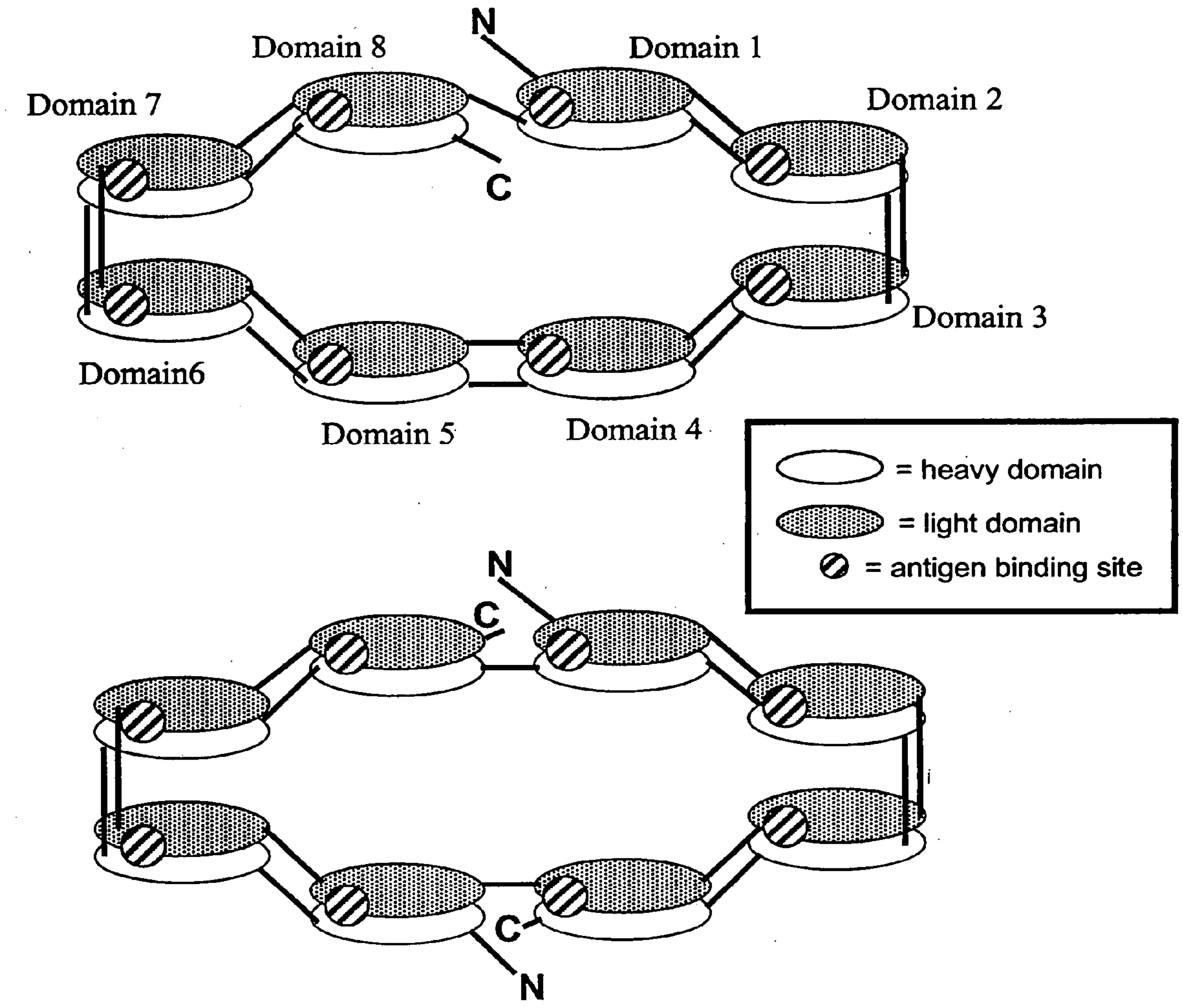


Figure 13



A. One Chain Multimer of Fv's

B. Two Chain Multimer of Fv's

Figure 13C

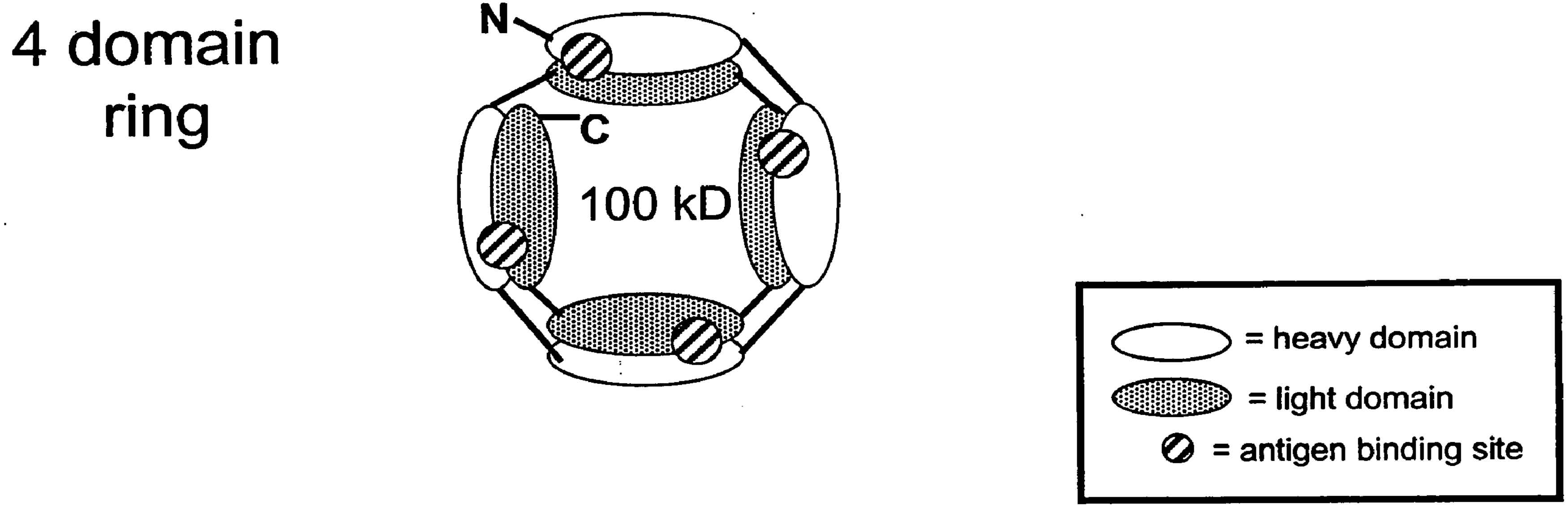
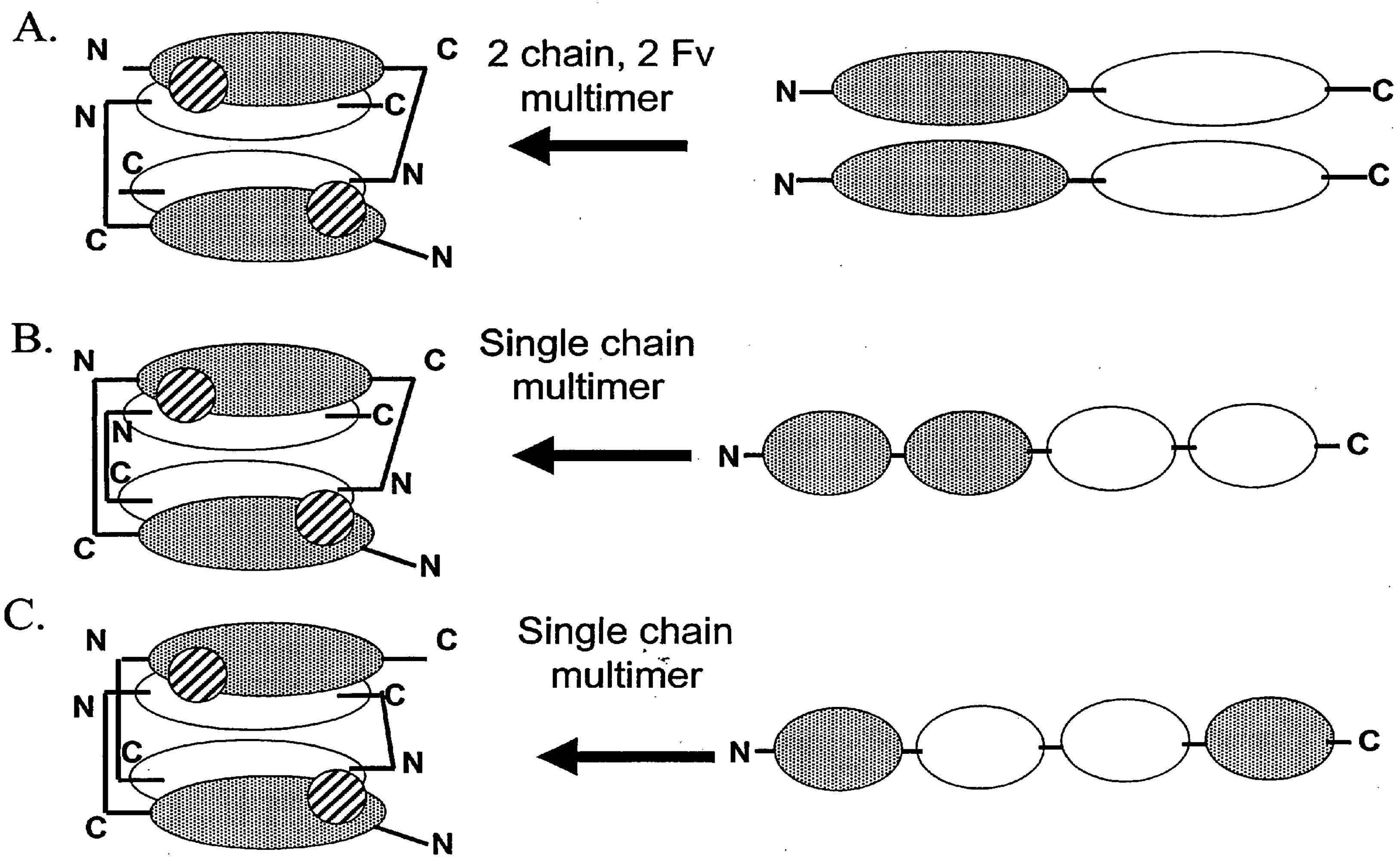

4 domain
ring
N
C
100 kD
= heavy domain
= light domain
= antigen binding site

Figure 14


A.
N
C
N
C
C
N
C
N
2 chain, 2 Fv
multimer
N
C
N
C
B.
N
C
N
C
C
N
C
N
Single chain
multimer
N
C
C.
N
C
N
C
C
N
C
N
Single chain
multimer
N
C

COMBINATORIAL LIBRARIES OF MONOMER DOMAINS

CROSS-REFERENCES TO OTHER APPLICATIONS

[0001] The present application claims benefit of priority and explicitly incorporates by reference the following patent applications: U.S. Provisional Patent Application Ser. No. (USSN) 60/______, filed Apr. 18, 2002 (Attorney Docket No. 18097A-034410US), U.S. Provisional Patent Application Ser. No. (USSN) 60/286,823, filed Apr. 26, 2001, U.S. Provisional Patent Application Ser. No. (USSN) 60/337,209, filed Nov. 19, 2001, and U.S. Provisional Patent Application Ser. No. (USSN) 60/333,359, filed Nov. 26, 2001.

COPYRIGHT NOTIFICATION

[0002] Pursuant to 37 C.F.R. § 1.7(e), a portion of this patent document contains material that is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure as it appears in the Patent and Trademark Office Patent file or records, but otherwise reserves all copyrights whatsoever.

BACKGROUND OF THE INVENTION

[0003] Analysis of protein sequences and three-dimensional structures have revealed that many proteins are composed of a number of discrete monomer domains. The majority of discrete monomer domain proteins is extracellular or constitutes the extracellular parts of membrane-bound proteins.

[0004] An important characteristic of a discrete monomer domain is its ability to fold independently or with some limited assistance. Limited assistance can include assistance of a chaperonin(s) (e.g., a receptor-associated protein (RAP)). The presence of a metal ion(s) also offers limited assistance. The ability to fold independently prevents misfolding of the domain when it is inserted into a new protein environment. This characteristic has allowed discrete monomer domains to be evolutionarily mobile. As a result, discrete domains have spread during evolution and now occur in otherwise unrelated proteins. Some domains, including the fibronectin type III domains and the immunoglobin-like domain, occur in numerous proteins, while other domains are only found in a limited number of proteins.

[0005] Proteins that contain these domains are involved in a variety of processes, such as cellular transporters, cholesterol movement, signal transduction and signaling functions which are involved in development and neurotransmission. See Herz, *Lipoprotein receptors: beacons to neurons?*, (2001) *Trends in Neurosciences* 24(4):193-195; Goldstein and Brown, *The Cholesterol Quartet*, (2001) *Science* 292: 1310-1312. The function of a discrete monomer domain is often specific but it also contributes to the overall activity of the protein or polypeptide. For example, the LDL-receptor class A domain (also referred to as a class A module, a complement type repeat or an A-domain) is involved in ligand binding while the gamma-carboxyglumatic acid (Gla) domain which is found in the vitamin-K-dependent blood coagulation proteins is involved in high-affinity binding to phospholipid membranes. Other discrete monomer domains include, e.g., the epidermal growth factor (EGF)-like domain in tissue-type plasminogen activator which mediates binding to liver cells and thereby regulates the clearance of this fibrinolytic enzyme from the circulation and the cytoplasmic tail of the LDL-receptor which is involved in receptor-mediated endocytosis.

[0006] Individual proteins can possess one or more discrete monomer domains. These proteins are often called mosaic proteins. For example, members of the LDL-receptor family contain four major structural domains: the cysteine rich A-domain repeats, epidermal growth factor precursor-like repeats, a transmembrane domain and a cytoplasmic domain. The LDL-receptor family includes members that: 1) are cell-surface receptors; 2) recognize extracellular ligands; and 3) internalize them for degradation by lysosomes. See Hussain et al., *The Mammalian Low-Density Lipoprotein Receptor Family*, (1999) *Annu. Rev. Nutr.* 19:141-72. For example, some members include very-low-density lipoprotein receptors (VLDL-R), apolipoprotein E receptor 2, LDLR-related protein (LRP) and megalin. Family members have the following characteristics: 1) cell-surface expression; 2) extracellular ligand binding consisting of A-domain repeats; 3) requirement of calcium for ligand binding; 4) recognition of receptor-associated protein and apolipoprotein (apo) E; 5) epidermal growth factor (EGF) precursor homology domain containing YWTD repeats; 6) single membrane-spanning region; and 7) receptor-mediated endocytosis of various ligands. See Hussain, supra. Yet, the members bind several structurally dissimilar ligands.

[0007] It is advantageous to develop methods for generating and optimizing the desired properties of these discrete monomer domains. However, the discrete monomer domains, while often being structurally conserved, are not conserved at the nucleotide or amino acid level, except for certain amino acids, e.g., the cysteine residues in the A-domain. Thus, existing nucleotide recombination methods fall short in generating and optimizing the desired properties of these discrete monomer domains.

[0008] The present invention addresses these and other problems.

BRIEF SUMMARY OF THE INVENTION

[0009] The present invention provides methods for identifying a multimer that binds to a target molecule. In some embodiments, the method comprises: providing a library of monomer domains; screening the library of monomer domains for affinity to a target molecule; identifying at least one monomer domain that bind to at least one target molecule; linking the identified monomer domains to form a library of multimers; screening the library of multimers for the ability to bind to the target molecule; and identifying a multimer that binds to the target molecule.

[0010] Suitable monomer domains include those that are from 25 and 500 amino acids, 100 and 150 amino acids, or 25 and 50 amino acids in length.

[0011] In some embodiments, each monomer domain of the selected multimer binds to the same target molecule. In some embodiments, the selected multimer comprises at least three monomer domains. In some embodiments, the selected multimer comprises three to ten monomer domains. In some embodiments, at least three monomer domains bind to the same target molecule.

[0012] In some embodiments, the methods comprise identifying a multimer with an improved avidity for the target compared to the avidity of a monomer domain alone for the same target molecule. In some embodiments, the avidity of the multimer is at least two times the avidity of a monomer domain alone.

[0013] In some embodiments, the screening of the library of monomer domains and the identifying of monomer domains occurs simultaneously. In some embodiments, the screening of the library of multimers and the identifying of multimers occurs simultaneously.

[0014] In some embodiments, the polypeptide domain is selected 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, and a C2 domain

[0015] In some embodiments, the methods comprise a further step of mutating at least one monomer domain, thereby providing a library comprising mutated monomer domains. In some embodiments, the mutating step comprises recombining a plurality of polynucleotide fragments of at least one polynucleotide encoding a monomer domain. In some embodiments, the mutating step comprises directed evolution. In some embodiments, the mutating step comprises site-directed mutagenesis.

[0016] In some embodiments, the methods further comprise: screening the library of monomer domains for affinity to a second target molecule; identifying a monomer domain that binds to a second target molecule; linking at least one monomer domain with affinity for the first target molecule with at least one monomer domain with affinity for the second target molecule, thereby forming a library of multimers; screening the library of multimers for the ability to bind to the first and second target molecule; and identifying a multimer that binds to the first and second target molecule, thereby identifying a multimer that specifically binds a first and a second target molecule.

[0017] Certain methods of the present invention further comprise: providing a second library of monomer domains; screening the second library of monomer domains for affinity to at least a second target molecule; identifying a second monomer domain that binds to the second target molecule; linking the identified monomer domains that bind to the first target molecule or the second target molecule, thereby forming a library of multimers; screening the library of multimers for the ability to bind to the first and second target molecule; and identifying a multimer that binds to the first and second target molecules.

[0018] In some embodiments, the target molecule is selected from the group consisting of a viral antigen, a bacterial antigen, a fungal antigen, an enzyme, a cell surface protein, an enzyme inhibitor, a reporter molecule, and a receptor. In some embodiments, the viral antigen is a polypeptide required for viral replication. In some embodiments, the first and at least second target molecules are different components of the same viral replication system. In some embodiments, the selected multimer binds to at least two serotypes of the same virus.

[0019] In some embodiments, the library of multimers is expressed as a phage display, ribosome display or cell surface display. In some embodiments, the library of multimers is presented on a microarray.

[0020] In some embodiments, the monomer domains are linked by a polypeptide linker. In some embodiments, the polypeptide linker is a linker naturally-associated with the monomer domain. In some embodiments, the polypeptide linker is a variant of a linker naturally-associated with the monomer domain. In some embodiments, the linking step comprises linking the monomer domains with a variety of linkers of different lengths and composition.

[0021] In some embodiments, the domains form a secondary structure by the formation of disulfide bonds. In some embodiments, the multimers comprise an A domain connected to a monomer domain by a polypeptide linker. In some embodiments, the linker is from 1-20 amino acids inclusive. In some embodiments, the linker is made up of 5-7 amino acids. In some embodiments, the linker is 6 amino acids in length. In some embodiments, the linker comprises the following sequence, $A_1A_2A_3A_4A_5A_6$, wherein A_1 is selected from the amino acids A, P, T, Q, E and K; A_2 and A_3 are any amino acid except C, F, Y, W, or M; A_4 is selected from the amino acids S, G and R; A_5 is selected from the amino acids H, P, and R; A_6 is the amino acid, T. In some embodiments, the linker comprises a naturally-occurring sequence between the C-terminal cysteine of a first A domain and the N-terminal cysteine of a second A domain.

[0022] In some embodiments, the multimers comprise a C2 domain connected to a monomer domain by a polypeptide linker. In some embodiments, each C2 monomer domain differs from the corresponding wild-type C2 monomer domain in that at least one amino acid residue constituting part of the loop regions has been substituted with another amino acid residue; at least one amino acid residue constituting part of the loop regions has been deleted and/or at least one amino acid residue has been inserted in at least one of the loop regions. In some embodiments, the C2 domain comprises loop regions 1, 2, and 3 and the amino acid sequences outside of the loop regions 1, 2 and 3 are identical for all C2 monomer domains present in the polypeptide multimer. In some of these embodiments, the linker is between 1-20 amino acids. In some embodiments, the linker is between 10-12 amino acids. In some embodiments, the linker is 11 amino acids.

[0023] The present invention also provides polypeptides comprising the multimers selected as described above.

[0024] The present invention also provides polynucleotides encoding the multimers selected as described above.

[0025] The present invention also provides libraries of multimers formed as described above.

[0026] The present invention also provides methods for identifying a multimer that binds to at least one target molecule, comprising the steps of: providing a library of multimers, wherein each multimer comprises at least two monomer domains and wherein each monomer domain exhibits a binding specificity for a target molecule; and screening the library of multimers for target molecule-binding multimers. In some embodiments, the methods further comprise identifying target molecule-binding multimers having an avidity for the target molecule that is greater than the avidity of a single monomer domain for the target molecule. In some embodiments, one or more of the multimers comprises a monomer domain that specifically binds to a second target molecule.

[0027] The present invention also provides libraries of multimers. In some embodiments, each multimer comprises at least two monomer domains connected by a linker; each monomer domain exhibits a binding specificity for a target molecule; and each monomer domain is a non-naturally occurring monomer domain.

[0028] In some embodiments, the linker comprises at least 3 amino acid residues. In some embodiments, the linker comprises at least 6 amino acid residues. In some embodiments, the linker comprises at least 10 amino acid residues.

[0029] The present invention also provides polypeptides comprising at least two monomer domains separated by a heterologous linker sequence. In some embodiments, each monomer domain specifically binds to a target molecule; and each monomer domain is a non-naturally occurring protein monomer domain.

[0030] In some embodiments, polypeptides comprise a first monomer domain that binds a first target molecule and a second monomer domain that binds a second target molecule. In some embodiments, the polypeptides comprise two monomer domains, each monomer domain having a binding specificity that is specific for a different site on the same target molecule. In some embodiments, the polypeptides further comprise a monomer domain having a binding specificity for a second target molecule.

[0031] In some embodiments, the monomer domains of a library, multimer or polypeptide are at least 70% identical.

[0032] The invention also provides polynucleotides encoding the above-described polypeptides.

[0033] The present invention also provides multimers of immuno-domains having binding specificity for a target molecule, as well as methods for generating and screening libraries of such multimers for binding to a desired target molecule. More specifically, the present invention provides a method for identifying a multimer that binds to a target molecule, the method comprising, providing a library of immuno-domains; screening the library of immuno-domains for affinity to a first target molecule; identifying one or more (e.g., two or more) immuno-domains that bind to at least one target molecule; linking the identified monomer domain to form a library of multimers, each multimer comprising at least three immuno-domains (e.g., four or more, five or more, six or more, etc.); screening the library of multimers for the ability to bind to the first target molecule; and identifying a multimer that binds to the first target molecule. Libraries of multimers of at least two immuno-domains that are minibodies, single comain antibodies, Fabs, or combinations thereof are also employed in the practice of the present invention. Such libraries can be readily screened for multimers that bind to desired target molecules in accordance with the invention methods described herein.

[0034] The present invention further provides methods of identifying hetero-immuno multimers that binds to a target molecule. In some embodiments, the methods comprise, providing a library of immuno-domains; screening the library of immuno-domains for affinity to a first target molecule; providing a library of monomer domains; screening the library of monomer domains for affinity to a first target molecule; identifying at least one immuno-domain that binds to at least one target molecule; identifying at least one monomer domain that binds to at least one target molecule; linking the identified immuno-domain with the identified monomer domains to form a library of multimers, each multimer comprising at least two domains; screening the library of multimers for the ability to bind to the first target molecule; and identifying a multimer that binds to the first target molecule.

DEFINITIONS

[0035] Unless otherwise indicated, the following definitions supplant those in the art.

[0036] The term "monomer domain" or "monomer" is used interchangeably 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 specifically 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.

[0037] The term "monomer domain variant" refers to a domain resulting from human-manipulation of a monomer domain sequence. Examples of man-manipulated changes include, e.g., random mutagenesis, site-specific mutagenesis, shuffling, directed evolution, etc. The term "monomer domain variant" does not embrace a mutagenized complementarity determining region (CDR) of an antibody.

[0038] The term "multimer" is used herein to indicate a polypeptide comprising at least two monomer domains and/or immuno-domains (e.g., at least two monomer domains, at least two immuno-domains, or at least one monomer domain and at least one immuno-domain). The separate monomer domains and/or immuno-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.

[0039] 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, fingi, 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).

[0040] As used herein, the term "immuno-domains" refers to protein binding domains that contain at least one complementarity determining region (CDR) of an antibody. Immuno-domains can be naturally occurring immunological domains (i.e. isolated from nature) or can be non-naturally occurring immunological domains that have been altered by human-manipulation (e.g., via mutagenesis methods, such as, for example, random mutagenesis, site-specific mutagenesis, and the like, as well as by directed evolution methods, such as, for example, recursive error-prone PCR, recursive recombination, and the like.). Different types of immuno-domains that are suitable for use in the practice of the present invention include a minibody, a single-domain antibody, a single chain variable fragment (ScFv), and a Fab fragment.

[0041] The term "minibody" refers herein to a polypeptide that encodes only 2 complementarity determining regions (CDRs) of a naturally or non-naturally (e.g., mutagenized) occurring heavy chain variable domain or light chain variable domain, or combination thereof. An example of a minibody is described by Pessi et al., *A designed metal-binding protein with a novel fold*, (1993) Nature 362:367-369. A multimer of minibodies is schematically illustrated in **FIG. 11A**. The circles depict minibodies, and the solid lines depict the linker moieties joining the immuno-domains to each other.

[0042] As used herein, the term "single-domain antibody" refers to the heavy chain variable domain ("V_H") of an antibody, i.e., a heavy chain variable domain without a light chain variable domain. Exemplary single-domain antibodies employed in the practice of the present invention include, for example, the Camelid heavy chain variable domain (about 118 to 136 amino acid residues) as described in Hamers-Casterman, C. et al., *Naturally occurring antibodies devoid of light chains* (1993) *Nature* 363:446-448, and Dumoulin, et al., *Single-domain antibody fragments with high conformational stability* (2002) *Protein Science* 11:500-515. A multimer of single-domain antibodies is depicted in **FIG. 11B**. The ellipses represent the single-domain antibodies, and the solid lines depict the linker moieties joining the single-domain antibodies to each other.

[0043] The terms "single chain variable fragment" or "ScFv" are used interchangeably herein to refer to antibody heavy and light chain variable domains that are joined by a peptide linker having at least 12 amino acid residues. Single chain variable fragments contemplated for use in the practice of the present invention include those described in Bird, et al., *Single-chain antigen-binding proteins* (1988) *Science* 242(4877):423-426 and Huston et al., *Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli* (1988) *Proc Natl Acad Sci USA* 85(16):5879-83. A multimer of single chain variable fragments is illustrated in **FIG. 11C**. The dotted lines represent the peptide linker joining the heavy and light chain variable domains to each other. The solid lines depict the linker moieties joining the heavy chain variable domains to each other.

[0044] As used herein, the term "Fab fragment" refers to an immuno-domain that has two protein chains, one of which is a light chain consisting of two light chain domains (V_L variable domain and C_L constant domain) and a heavy chain consisting of two heavy domains (i.e., a V_H variable and a C_H constant domain). Fab fragments employed in the practice of the present invention include those that have an interchain disulfide bond at the C-terminus of each heavy and light component, as well as those that do not have such a C-terminal disulfide bond. Each fragment is about 47 kD. Fab fragments are described by Pluckthun and Skerra, *Expression of functional antibody Fv and Fab fragments in Escherichia col* (1989) *Methods Enzymol* 178:497-515. A multimer of Fab fragments is depicted in **FIG. 11D**. The white ellipses represent the heavy chain component of the Fab fragment, the filled ellipses represent the light chain component of the Fab.

[0045] 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.

[0046] 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.

[0047] 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 shuffling, chemical mutagenesis and the like.

[0048] 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.

[0049] 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.

[0050] 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.

[0051] The terms "polypeptide,""peptide," and "protein" are used herein interchangeably to refer to an amino acid sequence of two or more amino acids.

[0052] '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. Preferred conservative amino acids substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine.

[0053] 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. It includes chromosomal DNA, self-replicating plasmids and DNA or RNA that performs a primarily structural role.

[0054] 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.

[0055] 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.

[0056] 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.

[0057] 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.

[0058] 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 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., 2×, 5×, 10× or more above background) and does not bind in a significant amount to other molecules present in the sample.

[0059] 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 more preferably over a region that is 100 to 500 or 1000 or more nucleotides or amino acids in length.

[0060] 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 fusion protein).

BRIEF DESCRIPTION OF THE DRAWINGS

[0061] **FIG. 1** schematically illustrates the type, number and order of monomer domains found in members of the

LDL-receptor family. These monomer domains include β-Propeller domains, EGF-like domains and LDL receptor class A-domains. The members shown include low-density lipoprotein receptor (LDLR), ApoE Receptor 2 (ApoER2), very-low-density lipoprotein receptor (VLDLR), LDLR-related protein 2 (LRP2) and LDLR-related protein1 (LRP1).

[0062] **FIG. 2** schematically illustrates the alignment of partial amino acid sequence from a variety of the LDL-receptor class A-domains that include two human LRP1 sequences, two human LRP2 sequences, two human LDLR sequences, two human LDVR sequences, one human LRP3 sequence, one human MAT sequence, a human CO6 sequence, and a human SORL sequence, to demonstrate the conserved cysteines.

[0063] **FIG. 3**, panel A schematically illustrates an example of an A-domain. Panel A schematically illustrates conserved amino acids in an A-domain of about 40 amino acids long. The conserved cysteine residues are indicated by C, and the negatively charged amino acids are indicated by a circle with a minus ("–") sign. Circles with an "H" indicate 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.

[0064] **FIG. 4** indicates some of the ligands recognized by 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.

[0065] **FIG. 5** 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.

[0066] **FIG. 6** 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.

[0067] **FIG. 7** shows the multimerization process of monomer domains. The target-binding monomer hits are amplified from a vector. This mixture of target-binding monomer domains and/or immuno-domains is then cleaved and mixed with an optimal combination of linker and stopper oligonucleotides. The multimers that are generated are then cloned into a suitable vector for the second selection step for identification of target-binding multimers.

[0068] **FIG. 8** depicts common amino acids in each position of the A domain. The percentages above the amino acid positions refer to the percentage of naturally-occurring A domains with the inter-cysteine spacing displayed. Potential amino acid residues in bold depicted under each amino acid position represent common residues at that position. The final six amino acids, depicted as lighter-colored circles, represent linker sequences. The two columns of italicized amino acid residues at positions 2 and 3 of the linker represent amino acid residues that do not occur at that position. Any other amino acid (e.g., A, D, E, G, H, I, K, L, N, P, Q, R, S, T, and V) may be included at these positions.

[0069] **FIG. 9** displays the frequency of occurrence of amino acid residues in naturally-occurring A domains for A domains with the following spacing between cysteines: $CX_6CX_4CX_6CX_5CX_8C$.

[0070] **FIG. 10** depicts an alignment of A domains. At the top and the bottom of the figure, small letters (a-q) indicate conserved residues. The predominant amino acids at these positions and the percent of time they were observed in native A domains is illustrated at the bottom of the figure.

[0071] **FIG. 11** depicts possible multimer conformations comprises of immuno-domains. **FIG. 11A** illustrates a multimer of minibodies. **FIG. 11B** illustrates a multimer of single-domain antibodies. **FIG. 11C** illustrates a immuno-domain multimer of scfvs. **FIG. 11D** illustrates a multimer of Fab fragments.

[0072] **FIG. 12** depicts linkage of domains via partial linkers.

[0073] **FIG. 13** illustrates exemplary multimer ring formations.

[0074] **FIG. 14** illustrates various multimer conformations of heavy and light chains of Fvs.

DETAILED DESCRIPTION OF THE INVENTION

[0075] The invention provides an enhanced approach for selecting and optimizing properties of discrete monomer domains and/or immuno-domains to create multimers. In particular, this disclosure describes methods, compositions and kits for identifying discrete monomer domains and/or immuno-domains that bind to a desired ligand or mixture of ligands and creating multimers (also known as combinatorial mosaic proteins or recombinant mosaic proteins) that comprise two or more monomer domains and/or immuno-domains that are joined via a linker. The multimers can be screened to identify those that have an improved phenotype such as improved avidity or affinity or altered specificity for the ligand or the mixture of ligands, compared to the discrete monomer domain.

[0076] 1. Discrete Monomer Domains

[0077] 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 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 and immuno-domains can typically maintains stable conformation in solution. Sometimes, monomer domains and immuno-domains can fold independently into a stable conformation. In one embodiment, the stable conformation is stabilized by metal 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 (e.g., A, C2) or referenced herein.

[0078] Publications describing monomer domains and mosaic proteins and references cited within include the following: Hegyi, H and Bork, P., *On the classification and evolution of protein modules*, (1997) *J. Protein Chem.,* 16(5):545-551; Baron et al., *Protein modules* (1991) *Trends Biochem. Sci.* 16(1):13-7; Ponting et al., *Evolution of domain families*, (2000), *Adv. Protein Chem.,* 54:185-244; Doolittle, *The multiplicity of domains in proteins*, (1995) *Annu. Rev. Biochem* 64:287-314; Doolitte and Bork, *Evolutionarily mobile modules in proteins* (1993) *Scientific American,* 269 (4):50-6; and Bork, *Shuffled domains in extracellular proteins* (1991), *FEBS letters* 286(1-2):47-54. Monomer domains of the present invention also include those domains found in Pfam database and the SMART database. See Schultz, et al., *SMART: a web-based tool for the study of genetically mobile domains*, (2000) *Nucleic Acid Res.* 28(1):231-34.

[0079] Monomer domains that are particularly suitable for use in the practice of the present invention are (1) β sandwich domains; (2) β-barrel domains; or (3) 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 bands, more typically at least three disulfide bonds.

[0080] 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.

[0081] 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-like 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. For example, **FIG. 1** schematically diagrams various kinds of monomer domains found in molecules in the LDL-receptor family.

[0082] 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).

[0083] 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 Sep. 7, 2001, WO99/16873 entitled "Anticalins" by Beste et al., published Apr. 8, 1999 and WO 00/60070 entitled "A polypeptide structure for use as a scaffold" by Desmet, et al., published on Oct. 12, 2000.

[0084] 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-like domain. The monomer domains can also have a cluster of negatively charged residues. Optionally, the monomer domains contain a repeated sequence, such as YWTD as found in the β-Propeller domain.

[0085] Other features of monomer domains include the ability to bind ligands (e.g., as in the LDL receptor class A domain, or the CUB domain (complement C1r/C1s, Uegf, and bone morphogenic protein-1 domain)), the ability to participate in endocytosis or internalization (e.g., as in the cytoplasmic tail of the LDL receptor or the cytoplasmic tail of Megalin), the ability to bind an ion (e.g., Ca^{2+} binding by the LDL receptor A-domain), and/or the ability to be involved in cell adhesion (e.g., as in the EGF-like domain).

[0086] Characteristics of a monomer domain 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, **FIG. 2**.

[0087] Specifically, the A-domains (sometimes called "complement-type repeats") contain about 30-50 amino acids. In some embodiments, the domains comprise about 35-45 amino acids and in some cases about 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 A domain constitutes a ligand binding moiety. The cysteine residues of the domain are disulfide linked to form a compact, stable, functionally independent moiety. See, **FIG. 3**. Clusters of these repeats make up a ligand binding domain, and differential clustering can impart specificity with respect to the ligand binding.

[0088] Exemplary A domain sequences and consensus sequences are depicted in **FIGS. 2, 3** and **8**. **FIG. 9** displays location and occurrence of residues in A domains with the following spacing between cysteines. In addition, **FIG. 10** depicts a number of A domains and provides a listing of conserved amino acids. 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, 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. 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.

[0089] To date, at least 190 human A-domains are identified based on cDNA sequences. See, e.g., **FIG. 10**. Exemplary proteins containing 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.

[0090] Another exemplary monomer domain suitable for use in the practice of the present invention is the C2 domain. C2 monomer domains are polypeptides containing a compact β-sandwich composed of two, four-stranded β-sheets, where loops at the "top" of the domain and loops at the "bottom" of the domain connect the eight β-strands. C2 monomer domains may be divided into two subclasses, namely C2 monomer domains with topology I (synaptotagmin-like topology) and topology II (cytosolic phospholipase A2-like topology), respectively. C2 monomer domains with topology I contains three loops at the "top" of the molecule (all of which are Ca^{2+} binding loops), whereas C2 monomer domains with topology II contain four loops at the "top" of the molecule (out of which only three are Ca^{2+} binding loops). The structure of C2 monomer domains have been reviewed by Rizo and Südhof, *J. Biol. Chem.* 273;15879-15882 (1998) and by Cho, *J. Biol. Chem.* 276;32407-32410 (2001). The terms "loop region 1", "loop region 2" and "loop region 3" refer to the Ca^{2+} binding loop regions located at the "top" of the molecule. This nomenclature, which is used to distinguish the three Ca^{2+} binding loops located at the "top" of the molecule from the non-Ca^{2+} binding loops (mainly located at the "bottom" of the molecule) is widely used and recognized in the literature. See Rizo and Südhof, *J. Biol. Chem.* 273;15879-15882 (1998). Loop regions 1, 2, and 3 represent target binding regions and thus can be varied to modulate binding specificity and affinity. The remaining portions of the C2 domain can be maintained without alteration if desired. Some exemplary C2 domains are substantially identical to the following sequence:

```
Tyr Ser His Lys Phe Thr Val Val Val Leu Arg Ala Thr Lys Val
1               5                   10                  15

Thr Lys Gly Ala Phe Gly Asp Met Leu Asp Thr Pro Asp Pro Tyr
                20                  25                  30

Val Glu Leu Phe Ile Ser Thr Thr Pro Asp Ser Arg Lys Arg Thr
                35                  40                  45

Arg His Phe Asn Asn Asp Ile Asn Pro Val Trp Asn Glu Thr Phe
                50                  55                  60

Glu Phe Ile Leu Asp Pro Asn Gln Glu Asn Val Leu Glu Ile Thr
                65                  70                  75

Leu Met Asp Ala Asn Tyr Val Met Asp Glu Thr Leu Gly Thr Ala
                80                  85                  90

Thr Phe Thr Val Ser Ser Met Lys Val Gly Glu Lys Lys Glu Val
                95                  100                 105

Pro Phe Ile Phe Asn Gln Val Thr Glu Met Val Leu Glu Met Ser
                110                 115                 120

Leu Glu Val
        123.
```

[0091] Residues 1-16, 29-48, 54-77 and 86-123 constitute positions located outside loop regions 1, 2 and 3 and residues 17-28, 49-53 and 78-85 constitute the loop regions 1, 2 and 3, respectively.

[0092] Other examples of monomer domains can be found in the protein Cubilin, which contains EGF-type repeats and CUB domains. The CUB domains are involved in ligand binding, e.g., some ligands include intrinsic factor (IF)-vitamin B12, receptor associated protein (RAP), Apo A-I, Transferrin, Albumin, Ig light chains and calcium. See, Moestrup and Verroust, supra.

[0093] Megalin also contains multiple monomer domains. Specifically, megalin possesses LDL-receptor type A-domain, EGF-type repeat, a transmembrane segment and a cytoplasmic tail. Megalin binds a diverse set of ligands, e.g., ApoB, ApoE, ApoJ, clusterin, ApopH/Beta2-glycoprotein-I, PTH, Transthyretin, Thyroglobulin, Insulin, Aminoglycosides, Polymyxin B, Aprotinin, Trichosanthin, PAI-1, PAI-1-urokinase, PAI-1-tPA, Pro-urokinase, Lipoprotein lipase, alpha-Amylase, Albumin, RAP, Ig light chains, calcium, C1q, Lactoferrin, beta2-microglobulin, EGF, Prolactin, Lysozyme, Cytochrome c, PAP-1, Odorant-binding protein, seminal vesicle secretory protein II. See, Moestrup & Verroust, supra.

[0094] Descriptions of some exemplary monomer domains can be found in the following publications and the references cited therein: Yamazaki et al., *Elements of Neural Adhesion Molecules and a Yeast Vacuolar Protein Sorting Receptor are Present in a Novel Mammalian Low Density Lipoprotein Receptor Family Member*, (1996) *Journal of Biological Chemistry* 271(40) 24761-24768; Nakayama et al., *Identification of High-Molecular-Weight Proteins with Multiple EGF-like Motifs by Motif-Trap Screening*, (1998) *Genomics* 51:27-34; Liu et al, *Genomic Organization of New Candidate Tumor Suppressor Gene*, LRP1B, (2000) *Genomics* 69:271-274; Liu et al., *The Putative Tumor Suppressor LRP1B, a Novel Member of the Low Density Lipoprotein* (*LDL*) *Receptor Family, Exhibits Both Overlapping and Distinct Properties with the LDL Receptor-related Protein*, (2001) *Journal of Biological Chemistry* 276(31):28889-28896; Ishii et al, *cDNA of a New Low-Density Lipoprotein Receptor-Related Protein and Mapping of its Gene* (*LRP*3) *to Chromosome Bands* 19*q*12-*q*13.2, (1998) *Genomics* 51:132-135; Orlando et al, *Identification of the second cluster of ligand-binding repeats in megalin as a site for receptor-ligand interactions*, (1997) *PNAS USA* 94:2368-2373; Jeon and Shipley, *Vesicle-reconstituted Low Density Lipoprotein Receptor*, (2000) *Journal of Biological Chemistry* 275(39):30458-30464; Simmons et al., *Human Low Density Lipoprotein Receptor Fragment*, (1997) *Journal of Biological Chemistry* 272(41):25531-25536; Fass et al., *Molecular Basis of familial hypercholesterolaemia from structure of LDL receptor module*, (1997) *Nature* 388:691-93; Daly et al., *Three-dimensional structure of a cysteine-rich repeat from the low-density lipoprotein receptor*, (1995) *PNAS USA* 92:6334-6338; North and Blacklow, *Structural Independence of Ligand-Binding Modules Five and Six of the LDL Receptor*, (1999) *Biochemistry* 38:3926-3935; North and Blacklow, *Solution Structure of the Sixth LDL-A module of the LDL Receptor*, (2000) *Biochemistry* 39:25640-2571; North and Blacklow, *Evidence that Familial Hypercholesterolemia Mutations of the LDL Receptor Cause Limited Local Misfolding in an LDL-A Module Pair*, (2000) *Biochemistry* 39:13127-13135; Beglova et al., *Backbone Dynamics of a Module Pair from the Ligand-Binding Domain of the LDL Receptor*, (2001) *Biochemistry* 40:2808-2815; Bieri et al., *Folding, Calcium binding, and Structural Characterization of a Concatemer of the First and Second Ligand-Binding Modules of the Low-Density Lipoprotein Receptor*, (1998) *Biochemistry* 37:10994-11002; Jeon et al., *Implications for familial hypercholesterolemia from the structure of the LDL receptor YWTD-EGF domain pair*, (2001) *Nature Structural Biology* 8(6):499-504; Kurniawan et al., *NMR structure of a concatemer of the first and second ligand-binding modules of the human low-density lipoprotein receptor*, (2000) *Protein Science* 9:1282-1293; Esser et al., *Mutational Analysis of the Ligand Binding Domain of the Low Density poprotein Receptor*, (1988) *Journal of Biological Chemistry* 263(26):13282-13290; Russell et al., *Different Combinations of Cysteine-rich Repeats Mediate Binding of Low Density Lipoprotein Receptor to Two Different Proteins*, (1989) *Journal of Biological Chemistry* 264(36):21682-21688; Davis et al., *Acid-dependent ligand dissociation and recycling of LDL receptor mediated by growth factor homology region*, (1987) *Nature* 326:760-765; Rong et al., *Conversion of a human low-density lipoprotein receptor ligand-binding repeat to a virus receptor: Identification of residues important for ligand specificity*, (1998) *PNAS USA* 95:8467-8472; Agnello et al., *Hepatitis C virus and other Flaviviridae viruses enter cells via low density lipoprotein receptor*; (1999) *PNAS* 96(22):12766-12771; Esser and Russell, *Transport-deficient Mutations in the Low Density lipoprotein receptor*, (1988) *Journal of Biological Chemistry* 263(26):13276-13281; Davis et al., *The Low Density Lipoprotein Receptor*, (1987) *Journal of Biological Chemistry* 262(9):4075-4082; and, Peacock et al., *Human Low Density Lipoprotein Receptor Expressed in Xenopus Oocytes*, (1988) *Journal of Biological Chemistry* 263(16):7838-7845.

[0095] Others publications that describe the VLDLR, ApoER2 and LRP1 proteins and their monomer domains include the following as well as the references cited therein: Savonen et al., *The Carboxyl-terminal Domain of Receptor-associated Protein Facilitates Proper Folding and Trafficking of the Very Low Density Lipoprotein Receptor by Interaction with the Three Amino-terminal Ligand-binding Repeats of the Receptor*, (1999) *Journal of Biological Chemistry* 274(36):25877-25882; Hewat et al., *The cellular receptor to human rhinovirus* 2 *binds around the* 5-*fold axis and not in the canyon: a structural view*, (2000) *EMBO Journal* 19(23):6317-6325; Okun et al., *VLDL Receptor Fragments of Different Lengths Bind to Human Rhinovirus HRV2 with Different Stoichiometry*, (2001) *Journal of Biological Chemistry* 276(2):1057-1062; Rettenberger et al., *Ligand Binding Properties of the Very Low Density Lipoprotein Receptor*, (1999) *Journal of Biological Chemistry* 274(13):8973-8980; Mikhailenko et al., *Functional Domains of the very low density lipoprotein receptor: molecular analysis of ligand binding and acid-dependent ligand dissociation mechanisms*, (1999) *Journal of Cell Science* 112:3269-3281; Brandes et al., *Alternative Splicing in the Ligand Binding Domain of Mouse ApoE Receptor*-2 *Produces Receptor Variants Binding Reelin but not alpa*2-*macroglobulin*, (2001) *Journal of Biological Chemistry* 276(25):22160-22169; Kim et al., *Exon/Intron Organization, Chromosome Localization, Alternative Splicing, and Transcription Units of the Human Apolipoprotein E Recep-*

tor 2 Gene, (1997) *Journal of Biological Chemistry* 272(13):8498-8504; Obermoeller-McCormick et al., *Dissection of receptor folding and ligand-binding property with functional minireceptors of LDL receptor-related protein*, (2001) *Journal of Cell Science* 114(5):899-908; Horn et al., *Molecular Analysis of Ligand Binding of the Second Cluster of Complement-type Repeats of the Low Density Lipoprotein Receptor-related Protein*, (1997) *Journal of Biological Chemistry* 272(21):13608-13613; Neels et al., *The Second and Fourth Cluster of Class A Cysteine-rich Repeats of the Low Density Lipoprotein Receptor-related Protein Share Ligand-binding Properties*, (1999) *Journal of Biological Chemistry* 274(44):31305-31311; Obermoeller et al., *Differential Functions of the Triplicated Repeats Suggest Two Independent Roles for the Receptor-Associated Protein as a Molecular Chaperone*, (1997) *Journal of Biological Chemistry* 272(16):10761-10768; Andersen et al., *Identification of the Minimal Functional Unit in the Low Density Lipoprotein Receptor-related Protein for Binding the Receptor-associated Protein (RAP)*, (2000) *Journal of Biological Chemistry* 275(28):21017-21024; Andersen et al., *Specific Binding of alpha-Macroglobulin to Complement-Type Repeat CR4 of the Low-Density Lipoprotein Receptor-Related Protein*, (2000) *Biochemistry* 39:10627-10633; Vash et al., *Three Complement-Type Repeats of the Low-Density Lipoprotein Receptor-Related Protein Define a Common Binding Site for RAP, PAI-1, and Lactoferrin*, (1998) *Blood* 92(9):3277-3285; Dolmer et al., *NMR Solution Structure of Complement-like Repeat CR3 from the Low Density Lipoprotein Receptor-related Protein*, (2000) *Journal of Biological Chemistry* 275(5):3264-3269; Huang et al., *NMR Solution Structure of Complement-like Repeat CR8 from the Low Density Lipoprotein Receptor-related Protein*, (1999) *Journal of Biological Chemistry* 274(20):14130-14136; and Liu et al., *Uptake of HIV-1 Tat protein mediated by low-density lipoprotein receptor-related protein disrupts the neuronal metabolic balance of the receptor ligands*, (2000) *Nature Medicine* 6(12):1380-1387.

[0096] Other references regarding monomer domains also include the following publications and references cited therein: FitzGerald et al, *Pseudomonas Exotoxin-mediated Selection Yields Cells with Altered Expression of Low-Density Lipoprotein Receptor-related Protein*, (1995) *Journal of Cell Biology*, 129: 1533-41; Willnow and Herz, *Genetic deficiency in low density lipoprotein receptor-related protein confers cellular resistance to Pseudomonas exotoxin A*, (1994) *Journal of Cell Science*, 107:719-726; Trommsdorf et al., *Interaction of Cytosolic Adaptor Proteins with Neuronal Apolipoprotein E Receptors and the Amyloid Precursor Protein*, (1998) *Journal of Biological Chemistry*, 273(5): 33556-33560; Stockinger et al., *The Low Density Lipoprotein Receptor Gene Family*, (1998) *Journal of Biological Chemistry*, 273(48): 32213-32221; Obermoeller et al., *Ca+2 and Receptor-associated Protein are independently required for proper folding and disulfide bond formation of the low density lipoprotein receptor-related protein*, (1998) *Journal of Biological Chemistry*, 273(35):22374-22381; Sato et al., *39-kDa receptor-associated protein (RAP) facilitates secretion and ligand binding of extracellular region of very-low-density-lipoprotein receptor: implications for a distinct pathway from low-density-lipoprotein receptor*, (1999) *Biochem. J.* 341:377-383; Avromoglu et al, *Functional Expression of the Chicken Low Density Lipoprotein Receptor-related Protein in a mutant Chinese Hamster Ovary Cell Line Restores Toxicity of Pseudomonas Exotoxin A and Degradation of alpha2-Macroglobulin*, (1998) *Journal of Biological Chemistry*, 273(11) 6057-6065; Kingsley and Krieger, *Receptor-mediated endocytosis of low density lipoprotein: Somatic cell mutants define multiple genes required for expression of surface-receptor activity*, (1984) *PNAS USA*, 81:5454-5458; Li et al, *Differential Functions of Members of the Low Density Lipoprotein Receptor Family Suggests by their distinct endocystosis rates*, (2001) *Journal of Biological Chemistry* 276(21):18000-18006; and, Springer, *An Extracellular beta-Propeller Module Predicted in Lipoprotein and Scavenger Receptors, Tyrosine Kinases, Epidermal Growth Factor Precursor, and Extracellular Matrix Components*, (1998) *J. Mol. Biol.* 283:837-862.

[0097] Polynucleotides (also referred to as nucleic acids) 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.

[0098] 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 and/or immuno-domains are identified or selected for a desired property (e.g., binding affinity) and then the monomer domains and/or immuno-domains are formed into multimers. See, e.g., **FIG. 5**. 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.

[0099] As mentioned above, 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.

[0100] 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.

[0101] 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. 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.

[0102] The non-natural monomer domains or altered monomer domains can be produced by a number of methods. Any method of mutagenesis, such as site-directed mutagenesis and random mutatgenesis (e.g., chemical mutagenesis) can be used to produce variants. 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. Amino acids can be inserted synthetically or can be encoded by a nucleic acid.

[0103] Nucleic acids encoding fragments of naturally-occurring monomer domains and/or immuno-domains can also be mixed and/or recombined (e.g., by using chemically or enzymatically-produced fragments) to generate full-length, modified monomer domains and/or immuno-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.

[0104] 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.

[0105] 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.

[0106] 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.

[0107] Selection of monomer domains and/or immuno-domains from a library of domains can be accomplished by a variety of procedures. For example, one method of identifying monomer domains and/or immuno-domains which have a desired property involves translating a plurality of nucleic acids, where each nucleic acid encodes a monomer domain and/or immuno-domain, screening the polypeptides encoded by the plurality of nucleic acids, and identifying those monomer domains and/or immuno-domains that, e.g., bind to a desired ligand or mixture of ligands, thereby producing a selected monomer domain and/or immuno-domain. The monomer domains and/or immuno-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).

[0108] As mentioned above, selection of monomer domains and/or immuno-domains can be based on binding to a ligand such as a target protein 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}. The ligand can be a known ligand, e.g., a ligand known to bind one of the plurality of monomer domains, or e.g., the desired ligand can be an unknown monomer domain ligand. See, e.g., **FIG. 4**, which illustrates some of the ligands that bind to the A-domain. Other selections of monomer domains and/or immuno-domains can be based, e.g., on inhibiting or enhancing a specific function of a target protein or an activity. Target protein activity can include, e.g., endocytosis or internalization, induction of second messenger system, up-regulation or down-regulation of a gene, binding to an extracellular matrix, release of a molecule(s), or a change in conformation. In this case, the ligand does not need to be known. The selection can also include using high-throughput assays.

[0109] When a monomer domain and/or immuno-domain 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 and/or immuno-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.

[0110] A variety of reporting display vectors or systems can be used to express nucleic acids encoding the monomer domains immuno-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 and/or immuno-domains on the surface of a filamentous bacteriophage, typically as a fusion with a bacteriophage coat protein.

[0111] 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 plurality of 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 (desired) ligand, 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.

[0112] Examples of other display systems include ribosome displays, a nucleotide-linked display (see, e.g., U.S. Pat. Nos. 6,281,344; 6,194,550, 6,207,446, 6,214,553, and 6,258,558), 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.

[0113] The invention also includes compositions that are produced by methods of the 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.

[0114] The present invention also provides libraries of monomer domains, immuno-domains and libraries of nucleic acids that encode monomer domains and/or immuno-domains. The libraries can include, e.g., about 100, 250, 500 or more nucleic acids encoding monomer domains and/or immuno-domains, or the library can include, e.g., about 100, 250, 500 or more polypeptides that encode monomer domains and/or immuno-domains. Libraries can include monomer domains containing the same cysteine frame, e.g., A-domains or EGF-like domains.

[0115] In some embodiments, variants are generated by recombining two or more different sequences from the same family of monomer domains and/or immuno-domains (e.g., the LDL receptor class A domain). Alternatively, two or more different monomer domains and/or immuno-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.

[0116] 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.

[0117] 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).

[0118] 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. **FIG. 7** illustrates a generic cycle of selection and generation of variation.

[0119] Compositions of nucleic acids and polypeptides are included in the present invention. For example, the present invention provides a plurality of different nucleic acids wherein each nucleic acid encodes at least one monomer domain or immuno-domain. In some embodiments, at least one monomer domain is selected from the group consisting of: an EGF-like domain, a Kringle-domain, a fibronectin type I domain, a fibronectin type H 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. Suitable monomer domains also include those listed in the Pfam database and/or the SMART database.

[0120] The present invention also provides recombinant nucleic acids encoding one or more polypeptide comprising a plurality of monomer domains and/or immuno-domains, which monomer domains are altered in order or sequence as compared to a naturally occuring polypeptide. For example, the naturally occuring polypeptide can be selected 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.

[0121] 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.

[0122] 2. Multimers (Also Called Recombinant Mosaic Proteins or Combinatorial Mosaic Proteins)

[0123] Methods for generating multimers are a feature of the present invention. Multimers comprise at least two monomer domains and/or immuno-domains. For example, multimers of the invention can comprise from 2 to about 10 monomer domains and/or immuno-domains, from 2 and about 8 monomer domains and/or immuno-domains, from about 3 and about 10 monomer domains and/or immuno-domains, about 7 monomer domains and/or immuno-domains, about 6 monomer domains and/or immuno-domains, about 5 monomer domains and/or immuno-domains, or about 4 monomer domains and/or immuno-domains. In some embodiments, the multimer comprises at least 3 monomer domains and/or immuno-domains. Typically, the monomer domains have been pre-selected for binding to the target molecule of interest.

[0124] In some embodiments, each monomer domain specifically binds to one target molecule. In some of these embodiments, each monomer binds to a different position (analogous to an epitope) on a target molecule. Multiple monomer domains and/or immuno-domains that bind to the same target molecule results in an avidity effect resulting in improved avidity of the multimer for the target molecule compared to each individual monomer. In some embodiments, the multimer has an avidity of at least about 1.5, 2, 3, 4, 5, 10, 20, 50 or 100 times the avidity of a monomer domain alone.

[0125] In another embodiment, the multimer comprises monomer domains with specificities for different target molecules. For example, multimers of such diverse monomer domains can specifically bind different components of a viral replication system or different serotypes of a virus. In some embodiments, at least one monomer domain binds to a toxin and at least one monomer domain binds to a cell surface molecule, thereby acting as a mechanism to target the toxin. In some embodiments, at least two monomer domains and/or immuno-domains of the multimer bind to different target molecules in a target cell or tissue. Similarly, therapeutic molecules can be targeted to the cell or tissue by binding a therapeutic agent to a monomer of the multimer that also contains other monomer domains and/or immuno-domains having cell or tissue binding specificity.

[0126] Multimers can comprise a variety of combinations of monomer domains. For example, in a single multimer, the selected monomer domains can be the same or identical, optionally, different or non-identical. 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.

[0127] Multimers that are generated in the practice of the present invention may be any of the following:

[0128] (1) A homo-multimer (a multimer of the same domain, i.e., A1-A1-A1-A1);

[0129] (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 (see, e.g., **FIG. 10**).

[0130] (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).

[0131] 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. Exemplary heteromultimers comprising immuno-domains include dimers of, e.g., minibodies, single domain antibodies and Fabs, wherein the dimers are linked by a covalent linker. Other exemplary multimers include, e.g., trimers and higher level (e.g., tetramers) multimers of minibodies, single domain antibodies and Fabs. Yet more exemplary multimers include, e.g., dimers, trimers and higher level multimers of single chain antibody fragments, wherein the single chain antibodies are not linked covalently.

[0132] The present invention provides multimers of V_H and V_L domains that associate to form multimers of Fvs as depicted in **FIG. 13** and **FIGS. 14B** and C. As used herein, the term "Fv" refers to a non-covalently associated V_HV_L dimer. Such a dimer is depicted, for example, in **FIG. 13A**, where each pair of overlapping dark and white ellipses represents a single Fv. Fv multimers of the present invention do not comprise a light variable domain covalently linked directly to a heavy variable domain from the same Fv. However, Fv multimers of the present invention can comprise a covalent linkage of the light variable domains and heavy variable domains of the same Fv, that are separated by at least one or more domains. For example, examplary conformations of a multimer are V_{H1}-V_{H2}-V_{L1}-VL2, or V_{H1}-V_{L2}-V_{L1}-V_{H2} (where $V_{L\#}$and $V_{H\#}$represent the heavy and light variable domains, respectively).

[0133] In these and other embodiments, the heavy and light variable domains are aligned such that the corresponding heavy and light variable domains associate to form the corresponding Fv (i.e., Fv1=$V_{H1}V_{L1}$, Fv_2=$V_{H2}V_{L2}$, etc.). **FIGS. 14B** and C illustrate such Fv multimers. Those of ordinary skill in the art will readily appreciate that such Fv multimers can comprise additional heavy or light variable domains of an Fv, to form relatively large multimers of, for example, six, eight of more immuno-domains. See, e.g., **FIG. 13**. The Fvs in an Fv multimer of the present invention are not scFvs (i.e., V_{L1} is not covalently linked to V_{H1}).

[0134] Monomer domain, 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.

[0135] 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 before 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.

[0136] The selected monomer domains are joined by a linker to form a multimer. For example, a linker is positioned between each separate discrete monomer domain in a multimer. Typically, immuno-domains are also linked to each other or to monomer domains via a linker moiety. Linker moieties that can be readily employed to link immuno-domain variants together are the same as those described for multimers of monomer domain variants. Exemplary linker moieties suitable for joining immuno-domain variants to other domains into multimers are described herein.

[0137] 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 DNA ligation, or optionally, by PCR-based, self-priming overlap reactions. 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.

[0138] 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.

[0139] A linker can be rigid, or alternatively, 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.

[0140] 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.

[0141] 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.

[0142] Typically, the linker polypeptide may predominantly include amino acid residues selected from the group consisting of Gly, Ser, Ala and Thr. For example, the peptide linker may contain at least 75% (calculated on the basis of the total number of residues present in the peptide linker), such as at least 80%, e.g. at least 85% or at least 90% of amino acid residues selected from the group consisting of Gly, Ser, Ala and Thr. The peptide linker may also consist of Gly, Ser, Ala and/or Thr residues only. The linker polypeptide should have a length, which is adequate to link two monomer domains in such a way that they assume the correct conformation relative to one another so that they retain the desired activity, for example as antagonists of a given receptor.

[0143] A suitable length for this purpose is a length of at least one and typically fewer than about 50 amino acid residues, such as 2-25 amino acid residues, 5-20 amino acid residues, 5-15 amino acid residues, 8-12 amino acid residues or 11 residues. Similarly, the polypeptide encoding a linker can range in size, e.g., from about 2 to about 15 amino acids, from about 3 to about 15, from about 4 to about 12, about 10, about 8, or about 6 amino acids. In methods and compositions involving nucleic acids, such as DNA, RNA, or combinations of both, the polynucleotide containing the linker sequence can be, e.g., between about 6 nucleotides and about 45 nucleotides, between about 9 nucleotides and about 45 nucleotides, between about 12 nucleotides and about 36 nucleotides, about 30 nucleotides, about 24 nucleotides, or about 18 nucleotides. Likewise, the amino acid residues selected for inclusion in the linker polypeptide should exhibit properties that do not interfere significantly with the activity or function of the polypeptide multimer. Thus, the peptide linker should on the whole not exhibit a charge which would be inconsistent with the activity or function of the polypeptide multimer, or interfere with internal folding, or form bonds or other interactions with amino acid residues in one or more of the monomer domains which would seriously impede the binding of the polypeptide multimer to the target in question.

[0144] In another embodiment of the invention, the peptide linker is selected from a library where the amino acid residues in the peptide linker are randomized for a specific set of monomer domains in a particular polypeptide multimer. A flexible linker could be used to find suitable combinations of monomer domains, which is then optimized using this random library of variable linkers to obtain linkers with optimal length and geometry. The optimal linkers may contain the minimal number of amino acid residues of the right type that participate in the binding to the target and restrict the movement of the monomer domains relative to each other in the polypeptide multimer when not bound to the target.

[0145] The use of naturally occurring as well as artificial peptide linkers to connect polypeptides into novel linked fusion polypeptides is well known in the literature (Hallewell et al. (1989), *J. Biol. Chem.* 264, 5260-5268; Alfthan et al. (1995), *Protein Eng.* 8, 725-731; Robinson & Sauer (1996), *Biochemistry* 35, 109-116; Khandekar et al. (1997), *J. Biol. Chem.* 272, 32190-32197; Fares et al. (1998), *Endocrinology* 139, 2459-2464; Smallshaw et al. (1999), *Protein Eng.* 12, 623-630; U.S. Pat. No. 5,856,456).

[0146] One example where the use of peptide linkers is widespread is for production of single-chain antibodies where the variable regions of a light chain (V_L) and a heavy chain (V_H) are joined through an artificial linker, and a large number of publications exist within this particular field. A widely used peptide linker is a 15mer consisting of three repeats of a Gly-Gly-Gly-Gly-Ser amino acid sequence ($(Gly_4Ser)_3$). Other linkers have been used and phage display technology as well as selective infective phage technology has been used to diversify and select appropriate linker sequences (Tang et al. (1996), *J. Biol. Chem.* 271, 15682-15686; Hennecke et al. (1998), *Protein Eng.* 11, 405-410). Peptide linkers have been used to connect individual chains in hetero- and homo-dimeric proteins such as the T-cell receptor, the lambda Cro repressor, the P22 phage Arc repressor, IL-12, TSH, FSH, IL-5, and interferon-γ. Peptide linkers have also been used to create fusion polypeptides. Various linkers have been used and in the case of the Arc repressor phage display has been used to optimize the linker length and composition for increased stability of the single-chain protein (Robinson and Sauer (1998), *Proc. Natl. Acad. Sci. USA* 95, 5929-5934).

[0147] Another type of linker is an intein, i.e. a peptide stretch which is expressed with the single-chain polypeptide, but removed post-translationally by protein splicing. The use of inteins is reviewed by F. S. Gimble in *Chemistry and Biology,* 1998, Vol 5, No. 10 pp. 251-256.

[0148] Still another way of obtaining a suitable linker is by optimizing a simple linker, e.g. $(Gly_4Ser)_n$, through random mutagenesis.

[0149] As mentioned above, it is generally preferred that the peptide linker possess at least some flexibility. Accordingly, in some embodiments, the peptide linker contains 1-25 glycine residues, 5-20 glycine residues, 5-15 glycine residues or 8-12 glycine residues. The peptide linker will typically contain at least 50% glycine residues, such as at least 75% glycine residues. In some embodiments of the invention, the peptide linker comprises glycine residues only.

[0150] The peptide linker may, in addition to the glycine residues, comprise other residues, in particular residues selected from the group consisting of Ser, Ala and Thr, in particular Ser. Thus, one example of a specific peptide linker includes a peptide linker having the amino acid sequence Gly$_x$-Xaa-Gly$_y$-Xaa-Gly$_z$, wherein each Xaa is independently selected from the group consisting Ala, Val, Leu, Ile, Met, Phe, Trp, Pro, Gly, Ser, Thr, Cys, Tyr, Asn, Gin, Lys, Arg, His, Asp and Glu, and wherein x, y and z are each integers in the range from 1-5. In some embodiments, each Xaa is independently selected from the group consisting of Ser, Ala and Thr, in particular Ser. More particularly, the peptide linker has the amino acid sequence Gly-Gly-Gly-Xaa-Gly-Gly-Gly-Xaa-Gly-Gly-Gly, wherein each Xaa is independently selected from the group consisting Ala, Val, Leu, Ile, Met, Phe, Trp, Pro, Gly, Ser, Thr, Cys, Tyr, Asn, Gin, Lys, Arg, His, Asp and Glu. In some embodiments, each Xaa is independently selected from the group consisting of Ser, Ala and Thr, in particular Ser.

[0151] In some cases it may be desirable or necessary to provide some rigidity into the peptide linker. This may be accomplished by including proline residues in the amino acid sequence of the peptide linker. Thus, in another embodiment of the invention, the peptide linker comprises at least one proline residue in the amino acid sequence of the peptide linker. For example, the peptide linker has an amino acid sequence, wherein at least 25%, such as at least 50%, e.g. at least 75%, of the amino acid residues are proline residues. In one particular embodiment of the invention, the peptide linker comprises proline residues only.

[0152] In some embodiments of the invention, the peptide linker is modified in such a way that an amino acid residue comprising an attachment group for a non-polypeptide moiety is introduced. Examples of such amino acid residues may be a cysteine residue (to which the non-polypeptide moiety is then subsequently attached) or the amino acid sequence may include an in vivo N-glycosylation site (thereby attaching a sugar moiety (in vivo) to the peptide linker).

[0153] In some embodiments of the invention, the peptide linker comprises at least one cysteine residue, such as one cysteine residue. Thus, in some embodiments of the invention the peptide linker comprises amino acid residues selected from the group consisting of Gly, Ser, Ala, Thr and Cys. In some embodiments, such a peptide linker comprises one cysteine residue only.

[0154] In a further embodiment, the peptide linker comprises glycine residues and cysteine residue, such as glycine residues and cysteine residues only. Typically, only one cysteine residue will be included per peptide linker. Thus, one example of a specific peptide linker comprising a cysteine residue, includes a peptide linker having the amino acid sequence Gly$_n$-Cys-Gly$_m$, wherein n and m are each integers from 1-12, e.g., from 3-9, from 4-8, or from 4-7. More particularly, the peptide linker may have the amino acid sequence GGGGG-C-GGGGG.

[0155] This approach (i.e. introduction of an amino acid residue comprising an attachment group for a non-polypeptide moiety) may also be used for the more rigid proline-containing linkers. Accordingly, the peptide linker may comprise proline and cysteine residues, such as proline and cysteine residues only. An example of a specific proline-containing peptide linker comprising a cysteine residue, includes a peptide linker having the amino acid sequence Pro$_n$-Cys-Pro$_m$, wherein n and m are each integers from 1-12, preferably from 3-9, such as from 4-8 or from 4-7. More particularly, the peptide linker may have the amino acid sequence PPPPP-C-PPPPP.

[0156] In some embodiments, the purpose of introducing an amino acid residue, such as a cysteine residue, comprising an attachment group for a non-polypeptide moiety is to subsequently attach a non-polypeptide moiety to said residue. For example, non-polypeptide moieties can improve the serum half-life of the polypeptide multimer. Thus, the cysteine residue can be covalently attached to a non-polypeptide moiety. Preferred examples of non-polypeptide moieties include polymer molecules, such as PEG or MPEG, in particular mPEG as well as non-polypeptide therapeutic agents.

[0157] The skilled person will acknowledge that amino acid residues other than cysteine may be used for attaching a non-polypeptide to the peptide linker. One particular example of such other residue includes coupling the non-polypeptide moiety to a lysine residue.

[0158] Another possibility of introducing a site-specific attachment group for a non-polypeptide moiety in the peptide linker is to introduce an in vivo N-glycosylation site, such as one in vivo N-glycosylation site, in the peptide linker. For example, an in vivo N-glycosylation site may be introduced in a peptide linker comprising amino acid residues selected from the group consisting of Gly, Ser, Ala and Thr. It will be understood that in order to ensure that a sugar moiety is in fact attached to said in vivo N-glycosylation site, the nucleotide sequence encoding the polypeptide multimer must be inserted in a glycosylating, eukaryotic expression host.

[0159] A specific example of a peptide linker comprising an in vivo N-glycosylation site is a peptide linker having the amino acid sequence Gly$_n$-Asn-Xaa-Ser/Thr-Gly$_m$, preferably Gly$_n$-Asn-Xaa-Thr-Gly$_m$, wherein Xaa is any amino acid residue except proline, and wherein n and m are each integers in the range from 1-8, preferably in the range from 2-5.

[0160] Often, the amino acid sequences of all peptide linkers present in the polypeptide multimer will be identical. Nevertheless, in certain embodiments the amino acid sequences of all peptide linkers present in the polypeptide multimer may be different. The latter is believed to be particular relevant in case the polypeptide multimer is a polypeptide tri-mer or tetra-mer and particularly in such cases where an amino acid residue comprising an attachment group for a non-polypeptide moiety is included in the peptide linker.

[0161] Quite often, it will be desirable or necessary to attach only a few, typically only one, non-polypeptide moieties/moiety (such as MPEG, a sugar moiety or a non-polypeptide therapeutic agent) to the polypeptide multimer in order to achieve the desired effect, such as prolonged serum-half life. Evidently, in case of a polypeptide tri-mer, which will contain two peptide linkers, only one peptide linker is typically required to be modified, e.g. by introduction of a cysteine residue, whereas modification of the other peptide linker will typically not be necessary not. In this case all (both) peptide linkers of the polypeptide multimer (tri-mer) are different.

[0162] Accordingly, in a further embodiment of the invention, the amino acid sequences of all peptide linkers present in the polypeptide multimer are identical except for one, two or three peptide linkers, such as except for one or two peptide linkers, in particular except for one peptide linker, which has/have an amino acid sequence comprising an amino acid residue comprising an attachment group for a non-polypeptide moiety. Preferred examples of such amino acid residues include cysteine residues of in vivo N-glycosylation sites.

[0163] A linker can be a native or synthetic linker sequence. An exemplary native linker includes, e.g., the sequence between the last cysteine of a first LDL receptor A domain and the first cysteine of a second LDL receptor A domain can be used as a linker sequence. Analysis of various A domain linkages reveals that native linkers range from at least 3 amino acids to fewer than 20 amino acids, e.g., 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, or 18 amino acids long. However, those of skill in the art will recognize that longer or shorter linker sequences can be used. An exemplary A domain linker sequence is depicted in **FIG. 8**. In some embodiments, the linker is a 6-mer of the following sequence $A_1A_2A_3A_4A_5A_6$, wherein A_1 is selected from the amino acids A, P, T, Q, E and K; A_2 and A_3 are any amino acid except C, F, Y, W, or M; A_4 is selected from the amino acids S, G and R; A_5 is selected from the amino acids H, P, and R; and A_6 is the amino acid, T.

[0164] Methods for generating multimers from monomer domains and/or immuno-domains can include joining the selected domains with at least one linker to generate at least one multimer, e.g., the multimer can comprise at least two of the monomer domains and/or immuno-domains and the linker. The multimer(s) is then screened for an improved avidity or affinity or altered specificity for the desired ligand or mixture of ligands as compared to the selected monomer domains. A composition of the multimer produced by the method is included in the present invention.

[0165] In other methods, the selected multimer domains are joined with at least one linker to generate at least two multimers, wherein the two multimers comprise two or more of the selected monomer domains and the linker. The two or more multimers are screened for an improved avidity or affinity or altered specificity for the desired ligand or mixture of ligands as compared to the selected monomer domains. Compositions of two or more multimers produced by the above method are also features of the invention.

[0166] Typically, multimers of the present invention are a single discrete polypeptide. Multimers of partial linker-domain-partial linker moieties are an association of multiple polypeptides, each corresponding to a partial linker-domain-partial linker moiety.

[0167] 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., **FIG. 6**. 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.

[0168] The methods of the present invention also include methods of evolving multimers. The methods 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 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.

[0169] 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 them formed into a library comprising a variety of linkers. The library can then be screened to identify opitmal linkers. Alternatively, multimer libraries can be formed where the effect of domain or linker on target molecule binding is not known.

[0170] 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 and/or immuno-domains are 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.

[0171] 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.

[0172] 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, nucleotide-linked display (see, e.g., U.S. Pat. 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.

[0173] 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.

[0174] 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.

[0175] 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.

[0176] 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.

[0177] 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.

[0178] Suitable linkers employed in the practice of the present invention include an obligate heterodimer of partial linker moieties. The term "obligate heterodimer" refers herein to a dimer of two partial linker moieties that differ from each other in composition, and which associate with each other in a non-covalent, specific manner to join two domains together. The specific association is such that the two partial linkers associate substantially with each other as compared to associating with other partial linkers. Thus, in contrast to multimers of the present invention that are expressed as a single polypeptide, multimers of domains that are linked together via heterodimers are assembled from discrete partial linker-monomer-partial linker units. Assembly of the heterodimers can be achieved by, for example, mixing. Thus, if the partial linkers are polypeptide segments, each partial linker-monomer-partial linker unit may be expressed as a discrete peptide prior to multimer assembly. A disulfide bond can be added to covalently lock the peptides together following the correct non-covalent pairing. A multimer containing such obligate heterodimers is depicted in **FIG. 12**. Partial linker moieties that are appropriate for forming obligate heterodimers include, for example, polynucleotides, polypeptides, and the like. For example, when the partial linker is a polypeptide, binding domains are produced individually along with their unique lining peptide (i.e., a partial linker) and later combined to form multimers. The spacial order of the binding domains in the multimer is thus mandated by the heterodimeric binding specificity of each partial linker. Partial linkers can contain terminal amino acid sequences that specifically bind to a defined heterologous amino acid sequence. An example of such an amino acid sequence is the Hydra neuropeptide head activator as described in Bodenmuller et al., *The neuropeptide head activator loses its biological activity by dimerization*, (1986) *EMBO J.* 5(8):1825-1829. See, e.g., U.S. Pat. No. 5,491,074 and WO 94/28173. These partial linkers allow the multimer to be produced first as monomer-partial linker units or partial linker-monomer-partial linker units that are then mixed together and allowed to assemble into the ideal order based on the binding specificities of each partial linker.

[0179] When the partial linker comprises a DNA binding motiff, each monomer domain has an upstream and a downstream partial linker (i.e., Lp-domain-Lp, where "Lp" is a representation of a partial linker) that contains a DNA binding protein with exclusively unique DNA binding specificity. These domains can be produced individually and then assembled into a specific multimer by the mixing of the domains with DNA fragments containing the proper nucleotide sequences (i.e., the specific recognition sites for the DNA binding proteins of the partial linkers of the two desired domains) so as to join the domains in the desired order. Additionally, the same domains may be assembled into many different multimers by the addition of DNA sequences containing various combinations of DNA binding protein recognition sites. Further randomization of the combinations of DNA binding protein recognition sites in the DNA fragments can allow the assembly of libraries of multimers. The DNA can be synthesized with backbone analogs to prevent degradation in vivo.

[0180] 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, immuno-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.

[0181] A single ligand can be used, or optionally a variety of ligands can be used to select the monomer domains, immuno-domains and/or multimers. A monomer domain and/or immuno-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.

[0182] The potential applications of multimers of the present invention are diverse. 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. 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.

[0183] Other applications include cell targeting. For example, multimers consisting of monomer domains and/or immuno-domains that recognize specific cell surface proteins can bind selectively to certain cell types. Applications involving monomer domains and/or immuno-domains as antiviral agents are also included. For example, multimers binding to different epitopes on the virus particle can be useful as antiviral agents because of the polyvalency. Other applications can include, but are not limited to, protein purification, protein detection, biosensors, ligand-affinity capture experiments and the like. Furthermore, domains or multimers can be synthesized in bulk by conventional means for any suitable use, e.g., as a therapeutic or diagnostic agent.

[0184] In some embodiments, the multimer comprises monomer domains and/or immuno-domains with specificities for different proteins. The different proteins can be related or unrelated. Examples of related proteins including members of a protein family or different serotypes of a virus. Alternatively, the monomer domains and/or immuno-domains of a multimer can target different molecules in a physiological pathway (e.g., different blood coagulation proteins). In yet other embodiments, monomer domains and/or immuno-domains bind to proteins in unrelated pathways (e.g., two domains bind to blood factors, two other domains and/or immuno-domains bind to inflammation-related proteins and a fifth binds to serum albumin).

[0185] The final conformation of the multimers containing immuno-domains can be a ring structure which would offer enhanced stability and other desired characteristics. These cyclic multimers can be expressed as a single polypeptide chain or may be assembled from multiple discrete polypeptide chains. Cyclic multimers assembled from discrete polypeptide chains are typically an assembly of two polypeptide chains. **FIG. 13B** depicts a cyclic multimer of two polypeptide chains. The formation of cyclic multimer structures can be vastly effected by the spatial arrangement (i.e, distance and order) and dimerization specificity of the individual domains. Parameters such as, for example, linker length, linker composition and order of immuno-domains, can be varied to generate a library of cyclic multimers having diverse structures. Libraries of cyclic multimers can be readily screened in accordance with the invention methods described herein to identify cyclic multimers that bind to desired target molecules. After the multimers are generated, optionally a cyclization step can be carried out to generate a library of cyclized multimers that can be further screened for desired binding activity.

[0186] These cyclic ring structures can be, for example, composed of a multimer of ScFv immuno-domains wherein the immuno-domains are split such that a coiling of the polypeptide multimer chain is required for the immuno-domains to form their proper dimeric structures (e.g., N-terminus-V_L1-V_L2-V_L3-V_L4-V_L5-V_L6-V_L7-V_L8-V_H1-V_H2-V_H3-V_H4-V_H5-V_H6-V_H7-V_H8-C-terminus, or N-terminus-V_L1-V_H2-V_L3-V_H4-V_H1-V_L2-V_H3-V_L4-C-terminus, and the like). An example of such a cyclic structure is shown in **FIG. 13A**. The ring could also be formed by the mixing of two polypeptide chains wherein each chain contained half of the immuno-domains. For example, one chain contains the V_L domains and the other chain contains the V_H domains such that the correct pairs of V_L/V_H domains are brought together upon the two strands binding. The circularization of the chains can be mandated by changing the frame of the domain order (i.e., polypeptide one: N-terminus-V_L1-V_L2-V_L3-V_L4-V_L5-V_L6-V_L7-V_L8-C-terminus and polypeptide two: N-terminus-V_H4-V_H5-V_H6-V_H7-V_H8-V_H1-V_H2-V_H3-C-terminus) as depicted in **FIG. 13B**.

[0187] A single polypeptide chain that forms a tetrameric ring structure could be very stable and have strong binding characteristics. An example of such a ring is shown in **FIG. 13C**.

[0188] Cyclic multimers can also be formed by encoding or attaching or linking at least one dimerizing domain at or near the N-terminus of a multimer protein and encoding or attaching or linking at least one second dimerizing domain at or near the C-terminus of the multimer protein wherein the first and second dimerization domain have a strong affinity for each other. As used herein, the term "dimerization domain" refers to a protein binding domain (of either immunological or non-immunological origin) that has the ability to bind to another protein binding domain with great strength and specificity such as to form a dimer. Cyclization of the multimer occurs upon binding of the first and the second dimerization domains to each other. Specifically, dimerization between the two domains will cause the multimer to adopt a cyclical structure. The dimerization domain can form a homodimer in that the domain binds to a protein that is identical to itself. The dimerization domain may form a heterodimer in that the domain binds to a protein binding domain that is different from itself. Some uses for such dimerization domains are described in, e.g., U.S. Pat. No. 5,491,074 and WO 94/28173.

[0189] 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 and/or immuno-domains, resulting in the formation of non-covalent interactions between two monomer domains and/or immuno-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 contain more target molecule binding domains than a single multimer.

[0190] 3. Therapeutic and Prophylactic Treatment Methods

[0191] 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.

[0192] 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.

[0193] 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.

[0194] 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.

[0195] 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.

[0196] 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.

[0197] 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.

[0198] 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.

[0199] 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.

[0200] 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).

[0201] 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.).

[0202] 4. Further Manipulating Monomer Domains and/or Multimer Nucleic Acids and Polypeptides

[0203] 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 above and below in the following publications and the references cited therein: Soong, N. et al., *Molecular breeding of viruses*, (2000) *Nat Genet* 25(4):436-439; Stemmer, et al., *Molecular breeding of viruses for targeting and other clinical properties*, (1999) *Tumor Targeting* 4:1-4; Ness et al., *DNA Shuffling of subgenomic sequences of subtilisin*, (1999) *Nature Biotechnology* 17:893-896; Chang et al., *Evolution of a cytokine using DNA family shuffling*, (1999) *Nature Biotechnology* 17:793-797; Minshull and Stemmer, *Protein evolution by molecular breeding*, (1999) *Current Opinion in Chemical Biology* 3:284-290; Christians et al., *Directed evolution of thymidine kinase for AZT phosphorylation using DNA family shuffling*, (1999) *Nature Biotechnology* 17:259-264; Crameri et al., *DNA shuffling of a family of genes from diverse species accelerates directed evolution*, (1998) *Nature* 391:288-291; Crameri et al., *Molecular evolution of an arsenate detoxification pathway by DNA shuffling*, (1997) *Nature Biotechnology* 15:436-438; Zhang et al., *Directed evolution of an effective fucosidase from a galactosidase by DNA shuffling and screening* (1997) *Proc. Natl. Acad. Sci. USA* 94:4504-4509; Patten et al., *Applications of DNA Shuffling to Pharmaceuticals and Vaccines*, (1997) *Current Opinion in Biotechnology* 8:724-733;

Crameri et al., *Construction and evolution of antibody-phage libraries by DNA shuffling*, (1996) *Nature Medicine* 2:100-103; Crameri et al., *Improved green fluorescent protein by molecular evolution using DNA shuffling*, (1996) *Nature Biotechnology* 14:315-319; Gates et al., *Affinity selective isolation of ligands from peptide libraries through display on a lac repressor 'headpiece dimer'*, (1996) *Journal of Molecular Biology* 255:373-386; Stemmer, *Sexual PCR and Assembly PCR*, (1996) In: *The Encyclopedia of Molecular Biology*. VCH Publishers, New York. pp. 447-457; Crameri and Stemmer, *Combinatorial multiple cassette mutagenesis creates all the permutations of mutant and wildtype cassettes*, (1995) *BioTechniques* 18:194-195; Stemmer et al., *Single-step assembly of a gene and entire plasmid form large numbers of oligodeoxy-ribonucleotides*, (1995) *Gene,* 164:49-53; Stemmer, *The Evolution of Molecular Computation*, (1995) *Science* 270:1510; Stemmer. *Searching Sequence Space*, (1995) *Bio/Technology* 13:549-553; Stemmer, *Rapid evolution of a protein in vitro by DNA shuffling*, (1994) *Nature* 370:389-391; and Stemmer, *DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution*, (1994) *Proc. Natl. Acad. Sci. USA* 91:10747-10751.

[0204] Mutational methods of generating diversity include, for example, site-directed mutagenesis (Ling et al., *Approaches to DNA mutagenesis: an overview*, (1997) *Anal Biochem.* 254(2): 157-178; Dale et al., *Oligonucleotide-directed random mutagenesis using the phosphorothioate method*, (1996) *Methods Mol. Biol.* 57:369-374; Smith, *In vitro mutagenesis*, (1985) *Ann. Rev. Genet.* 19:423-462; Botstein & Shortle, *Strategies and applications of in vitro mutagenesis*, (1985) *Science* 229:1193-1201; Carter, *Site-directed mutagenesis*, (1986) *Biochem. J.* 237:1-7; and Kunkel, *The efficiency of oligonucleotide directed mutagenesis*, (1987) in *Nucleic Acids & Molecular Biology* (Eckstein, F. and Lilley, D. M. J. eds., Springer Verlag, Berlin)); mutagenesis using uracil containing templates (Kunkel, *Rapid and efficient site-specific mutagenesis without phenotypic selection*, (1985) *Proc. Natl. Acad. Sci. USA* 82:488-492; Kunkel et al., *Rapid and efficient site-specific mutagenesis without phenotypic selection*, (1987) *Methods in Enzymol.* 154, 367-382; and Bass et al., *Mutant Trp repressors with new DNA-binding specificities*, (1988) *Science* 242:240-245); oligonucleotide-directed mutagenesis ((1983) *Methods in Enzymol.* 100: 468-500; (1987) *Methods in Enzymol.* 154: 329-350; Zoller & Smith, *Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any DNA fragment*, (1982) *Nucleic Acids Res.* 10:6487-6500; Zoller & Smith, *Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors*, (1983) *Methods in Enzymol.* 100:468-500; and Zoller & Smith, *Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template*, (1987) *Methods in Enzymol.* 154:329-350); phosphorothioate-modified DNA mutagenesis (Taylor et al., *The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA*, (1985) *Nucl. Acids Res.* 13: 8749-8764; Taylor et al., *The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA*, (1985) *Nucl. Acids Res.* 13: 8765-8787; Nakamaye & Eckstein, *Inhibition of restriction endonuclease Nci I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis*, (1986) *Nucl. Acids Res.* 14: 9679-9698; Sayers et al., *Y-T Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis*, (1988) *Nucl. Acids Res.* 16:791-802; and Sayers et al., *Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in the presence of ethidium bromide*, (1988) *Nucl. Acids Res.* 16: 803-814); mutagenesis using gapped duplex DNA (Kramer et al., *The gapped duplex DNA approach to oligonucleotide-directed mutation construction*, (1984) *Nucl. Acids Res.* 12: 9441-9456; Kramer & Fritz *Oligonucleotide-directed construction of mutations via gapped duplex DNA*, (1987) *Methods in Enzymol.* 154:350-367; Kramer et al., *Improved enzymatic in vitro reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations*, (1988) *Nucl. Acids Res.* 16: 7207; and Fritz et al., *Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions in vitro*, (1988) *Nucl. Acids Res.* 16: 6987-6999).

[0205] Additional suitable methods include point mismatch repair (Kramer et al., *Point Mismatch Repair*, (1984) *Cell* 38:879-887), mutagenesis using repair-deficient host strains (Carter et al., *Improved oligonucleotide site-directed mutagenesis using M13 vectors*, (1985) *Nucl. Acids Res.* 13: 4431-4443; and Carter, *Improved oligonucleotide-directed mutagenesis using M13 vectors*, (1987) *Methods in Enzymol.* 154: 382-403), deletion mutagenesis (Eghtedarzadeh & Henikoff, *Use of oligonucleotides to generate large deletions*, (1986) *Nucl. Acids Res.* 14: 5115), restriction-selection and restriction-purification (Wells et al., *Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin*, (1986) *Phil. Trans. R. Soc. Lond.* A 317: 415-423), mutagenesis by total gene synthesis (Nambiar et al., *Total synthesis and cloning of a gene coding for the ribonuclease S protein*, (1984) *Science* 223: 1299-1301; Sakamar and Khorana, *Total synthesis and expression of a gene for the a-subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin)*, (1988) *Nucl. Acids Res.* 14: 6361-6372; Wells et al., *Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites*, (1985) *Gene* 34:315-323; and Grundström et al., *Oligonucleotide-directed mutagenesis by microscale 'shot-gun' gene synthesis*, (1985) *Nucl. Acids Res.* 13: 3305-3316), double-strand break repair (Mandecki, *Oligonucleotide-directed double-strand break repair in plasmids of Escherichia coli: a method for site-specific mutagenesis*, (1986) *Proc. Natl. Acad. Sci. USA,* 83:7177-7181; and Arnold, *Protein engineering for unusual environments*, (1993) *Current Opinion in Biotechnology* 4:450-455). Additional details on many of the above methods can be found in *Methods in Enzymology* Volume 154, which also describes useful controls for trouble-shooting problems with various mutagenesis methods.

[0206] Additional details regarding various diversity generating methods can be found in the following U.S. patents, PCT publications and applications, and EPO publications: U.S. Pat. No. 5,605,793 to Stemmer (Feb. 25, 1997), "Methods for In Vitro Recombination;" U.S. Pat. No. 5,811,238 to Stemmer et al. (Sep. 22, 1998) "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination;" U.S. Pat. No. 5,830,721 to Stemmer et al. (Nov. 3, 1998), "DNA Mutagenesis by Random Fragmentation and Reassembly;" U.S. Pat. No. 5,834,252 to Stemmer, et al. (Nov. 10, 1998) "End-Complementary Polymerase Reaction;" U.S. Pat. No. 5,837,458 to Minshull, et al. (Nov. 17, 1998), "Methods and Compositions for Cellular and Metabolic Engineering;" WO 95/22625, Stemmer and Crameri, "Mutagenesis by Random Fragmentation and Reassembly;" WO 96/33207 by Stemmer and Lipschutz "End Complementary Polymerase Chain Reaction;" WO 97/20078 by Stemmer and Crameri "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination;" WO 97/35966 by Minshull and Stemmer, "Methods and Compositions for Cellular and Metabolic Engineering;" WO 99/41402 by Punnonen et al. "Targeting of Genetic Vaccine Vectors;" WO 99/41383 by Punnonen et al. "Antigen Library Immunization;" WO 99/41369 by Punnonen et al. "Genetic Vaccine Vector Engineering;" WO 99/41368 by Punnonen et al. "Optimization of Immunomodulatory Properties of Genetic Vaccines;" EP 752008 by Stemmer and Crameri, "DNA Mutagenesis by Random Fragmentation and Reassembly;" EP 0932670 by Stemmer "Evolving Cellular DNA Uptake by Recursive Sequence Recombination;" WO 99/23107 by Stemmer et al., "Modification of Virus Tropism and Host Range by Viral Genome Shuffling;" WO 99/21979 by Apt et al., "Human Papillomavirus Vectors;" WO 98/31837 by del Cardayre et al. "Evolution of Whole Cells and Organisms by Recursive Sequence Recombination;" WO 98/27230 by Patten and Stemmer, "Methods and Compositions for Polypeptide Engineering;" WO 98/27230 by Stemmer et al., "Methods for Optimization of Gene Therapy by Recursive Sequence Shuffling and Selection," WO 00/00632, "Methods for Generating Highly Diverse Libraries," WO 00/09679, "Methods for Obtaining in Vitro Recombined Polynucleotide Sequence Banks and Resulting Sequences," WO 98/42832 by Arnold et al., "Recombination of Polynucleotide Sequences Using Random or Defined Primers," WO 99/29902 by Arnold et al., "Method for Creating Polynucleotide and Polypeptide Sequences," WO 98/41653 by Vind, "An in Vitro Method for Construction of a DNA Library," WO 98/41622 by Borchert et al., "Method for Constructing a Library Using DNA Shuffling," and WO 98/42727 by Pati and Zarling, "Sequence Alterations using Homologous Recombination;" WO 00/18906 by Patten et al., "Shuffling of Codon-Altered Genes;" WO 00/04190 by del Cardayre et al. "Evolution of Whole Cells and Organisms by Recursive Recombination;" WO 00/42561 by Crameri et al., "Oligonucleotide Mediated Nucleic Acid Recombination;" WO 00/42559 by Selifonov and Stemmer "Methods of Populating Data Structures for Use in Evolutionary Simulations;" WO 00/42560 by Selifonov et al., "Methods for Making Character Strings, Polynucleotides & Polypeptides Having Desired Characteristics;" WO 01/23401 by Welch et al., "Use of Codon-Varied Oligonucleotide Synthesis for Synthetic Shuffling;" and PCT/US01/06775 "Single-Stranded Nucleic Acid Template-Mediated Recombination and Nucleic Acid Fragment Isolation" by Affholter.

[0207] 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. In addition to the many texts noted above, 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, Calif. (Berger); Sambrook et al., *Molecular Cloning—A Laboratory Manual* (2nd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 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. Pat. No. 4,683,202*; PCR Protocols A Guide to Methods and Applications* (Innis et al. eds) Academic Press Inc. San Diego, Calif. (1990) (Innis); Arnheim & Levinson (Oct. 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 40 kb 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.

[0208] 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 immuno-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.

[0209] 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 through-out, 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, N.Y.; 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, Fla.

[0210] 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., HAS 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.

[0211] 5. Kits

[0212] 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 multimer. 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.

[0213] 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.

[0214] 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.

[0215] 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.

[0216] 6. Integrated Systems

[0217] 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 shuffling methods, or by standard sequence alignment or word processing software.

[0218] 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.

[0219] 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).

[0220] 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.

[0221] 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 Feb. 5, 1999 (U.S. Ser. No. 60/118,854) and "Methods For Making Character Strings, Polynucleotides & Polypeptides Having Desired Characteristics" by Selifonov and Stemmer, filed Oct. 12, 1999 (U.S. Ser. No. 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.

[0222] 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).

[0223] 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., shuffled, 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

[0224] The following example is offered to illustrate, but not to limit the claimed invention.

Example 1

[0225] This example describes selection of monomer domains and the creation of multimers.

[0226] 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.

[0227] 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.

[0228] 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.

[0229] 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.

[0230] 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

[0231] This example describes the development of a library of multimers comprised of C2 domains.

[0232] A library of DNA sequences encoding monomeric C2 domains is created by assembly PCR as described in Stemmer et al., *Gene* 164, 49-53 (1995). The oligonucleotides used in this PCR reaction are:

5'-acactgcaatcgcgccttacggctCCCGGGCGGATCCtcccataagttca

5'-agctaccaaagtgacannknnknnknnknnknnknnknnknnknnknnkccatacgtcgaattgttcat

5'-agctaccaaagtgacaaaaggtgcttttggtgatatgttggatactccagatccatacgtcgaattgttcat

5'-taggaagagaacacgtcattttnnknnknnkattaaccctgtttggaacgagacctttgagt

5'-taggaagagaacacgtcattttaataatgatattaaccctgtttggaacgagacctttgagt

5'-ttggaaatcaccctaatgnnknnknnknnknnknnknnkactctaggtacagcaa

5'-ttggaaatcaccctaatggatgcaaattatgttatggacgaaactctaggtacagcaa

-continued

5'-aagaaggaagtcccatttattttcaatcaagttactgaaatggtcttagagatgtccctt

5'-tgtcactttggtagctcttaacacaactacagtgaacttatgggaGGA

5'-acgtgttctcttcctagaatctggagttgtactgatgaacaattcgacgta

5'-attagggtgatttccaaaacattttcttgattaggatctaatataaactcaaaggtctcgtt

5'-atgggacttccttcttttctcccactttcattgaagatacagtaaacgttgctgtacctagagt

5'-gaccgatagcttgccgattgcagtgtGGCCACAGAGGCCTCGAGaacttcaagggacatctctaaga

[0233] PCR fragments are digested with BamHI and XhoI. Digestion products are separated on 1.5% agarose gel and C2 domain fragments are purified from the gel. The DNA fragments are ligated into the corresponding restriction sites of yeast surface display vector pYD1 (Invitrogen)

[0234] The ligation mixture is used for transformation of yeast strain EBY100. Transformants are selected by growing the cells in glucose-containing selective medium (-Trp) at 30° C.

[0235] Surface display of the C2 domain library is induced by growing the cells in galactose-containing selective medium at 20° C. Cells are rinsed with PBS and then incubated with fluorescently-labeled target protein and rinsed again in PBS.

[0236] Cells are then sorted by FACS and positive cells are regrown in glucose-containing selective medium. The cell culture may be used for a second round of sorting or may be used for isolation of plasmid DNA. Purified plasmid DNA is used as a template to PCR amplify C2 domain encoding DNA sequences.

[0237] The oligonucleotides used in this PCR reaction are:

5'-acactgcaatcgcgccttacggctCAGgtCTGgtggttcccataagttcactgta

5'-gaccgatagcttgccgattgcagtCAGcacCTGaaccaccaccaccagaaccaccaccaccaacttcaagggacatctcta (linker sequence is underlined).

[0238] PCR fragments are then digested with AlwNI, digestion products are separated on 1.5% agarose gel and C2 domain fragments are purified from the gel. Subsequently, PCR fragments are multimerized by DNA ligation in the presence of stop fragments. The stop fragments are listed below:

[0239] Stop1:

5'-gaattcaacgctactaccattagtagaattgatgccaccttttcagc

tcgcgccccaaatgaaaaaatggtcaaactaaatctactcgttcgcagaa

ttgggaatcaactgttacatggaatgaaacttccagacaccgtacttttat

-continued

gaatatttatgacgattccgaggcgcgcccggactacccgtatgatgttc

cggattatgccccgggatcctcaggtgctg-3'

(digested with EcoRI and AlwNI).

[0240] Stop2:

5'-caggtgctgcactcgaggccactgcggccgcatattaacgtagattt

ttcctcccaacgtcctgactggtataatgagccagttcttaaaatcgcat

aaccagtacatggtgattaaagttgaaattaaaccgtctcaagagctttg

ttacgttgatttgggtaatgaagctt-3'

(digested with AlwNI and HindIII).

[0241] The ligation mixture is then digested with EcoRI and HindIII.

[0242] Multimers are separated on 1% agarose gel and DNA fragments corresponding to stop1-C2-C2-stop2 are purified from the gel. Stop1-C2-C2-stop2 fragments are PCR amplified using primers 5' aattcaacgctactaccat-3' and 5'-agcttcattacccaaatcaac-3' and subsequently digested with BamHI and XhoI. Optionally, the polynucleotides encoding the multimers can be put through a further round of affinity screening (e.g., FACS analysis as described above).

[0243] Subsequently, high affinity binders are isolated and sequenced. DNA encoding the high binders is cloned into expression vector and replicated in a suitable host. Expressed proteins are purified and characterized.

Example 3

[0244] This example describes the development of a library of trimers comprised of LDL receptor A domains.

[0245] A library of DNA sequences encoding monomeric A domains is created by assembly PCR as described in Stemmer et al., *Gene* 164, 49-53 (1995). The oligonucleotides used in this PCR reaction are:

5'-CACTATGCATGGACTCAGTGTGTCCGATAAGGGCACACGGTGCCTAC
CCGTATGATGTTCCGGATTATGCCCCGGGCAGTA

5'-CGCCGTCGCATMSCMAGYKCNSAGRAATACAWYGGCCGYTWYYGCAC
BKAAATTSGYYAGVCNSACAGGTACTGCCCGGGGCAT

5'-CGCCGTCGCATMSCMATKCCNSAGRAATACAWYGGCCGYTWYYGCAC
BKAAATTSGYYAGVCNSACAGGTACTGCCCGGGGCAT

5'-ATGCGACGGCGWWRATGATTGTSVAGATGGTAGCGATGAAVWGRRTT
GTVMAVNMVNMVGCCVTACGGGCTCGGCCTCT

5'-ATGCGACGGCGWWCCGGATTGTSVAGATGGTAGCGATGAAVWGRRTT
GTVMAVNMVNMVGCCVTACGGGCTCGGCCTCT

5'-ATGCGACGGCGWWRATGATTGTSVAGATAACAGCGATGAAVWGRRTT
GTVMAVNMVNMVGCCVTACGGGCTCGGCCTCT

5'-ATGCGACGGCGWWCCGGATTGTSVAGATAACAGCGATGAAVWGRRTT
GTVMAVNMVNMVGCCVTACGGGCTCGGCCTCT

5'-TCCTGGTAGTACTTATCTACTACTATTTGTCTGTGTCTGCTCTGGGT
TCCTAACGGTTCGGCCACAGAGGCCGAGCCCGTA

[0246] 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.

[0247] PCR fragments are digested with XmaI and SfiI. Digestion products are separated on 3% agarose gel and A domain fragments are purified from the gel. The DNA fragments are then ligated into the corresponding restriction sites of phage display vector fuse5-HA, a derivative of fuse5. The ligation mixture is electroporated into electrocompetent *E. coli* cells (F-strain e.g. Top10 or MC1061). Transformed *E. coli* cells are grown overnight in 2xYT medium containing 20 μg/ml tetracycline.

[0248] Virions are purified from this culture by PEG-precipitation. Target protein is immobilized on solid surface (e.g. petridish or microtiter plate) directly by incubating in 0.1 M $NaHCO_3$ or indirectly via a biotin-streptavidin linkage. Purified virions are added at a typical number of ~1-3x10^{11} TU. The petridish or microtiter plate is incubated at 4° C., washed several times with washing buffer (TBS/Tween) and bound phages are eluted by adding glycine.HCl buffer. The eluate is neutralized by adding 1 M Tris-HCl (pH 9.1)

[0249] The phages are amplified and subsequently used as input to a second round of affinity selection. ssDNA is extracted from the final eluate using QIAprep M13 kit. ssDNA is used as a template to PCR amplify A domains encoding DNA sequences.

[0250] The oligonucleotides used in this PCR reaction are:

5'-aagcctcagcgaccgaa

5'-agcccaataggaacccat

[0251] PCR fragments are digested with AlwNI and BglI. Digestion products are separated on 3% agarose gel and A domain fragments are purified from the gel. PCR fragments are multimerized by DNA ligation in the presence of the following stop fragments:

[0252] Stop1:

5'-gaattcaacgctactaccattagtagaattgatgccaccttttcagc

tcgcgccccaaatgaaaaaatggtcaaactaaatctactcgttcgcagaa

ttgggaatcaactgttacatggaatgaaacttccagacaccgtactttat

gaatatttatgacgattccgaggcgcgcccggactacccgtatgatgttc

cggattatgccccgggcggatccagtacctg-3'

(digested with EcoRI and ALwNI)

[0253] Stop2:

5'-gccctacgggcctcgaggcacctggtgcggccgcatattaacgtaga

tttttcctcccaacgtcctgactggtataatgagccagttcttaaaatcg

cataaccagtacatggtgattaaagttgaaattaaaccgtctcaagagct

ttgttacgttgatttgggtaatgaagctt-3'

(digested with Bg1I and HindIII)

[0254] The ligation mixture is digested with EcoRI and HindIII.

[0255] Multimers are separated on 1% agarose gel and DNA fragments corresponding to stop1-A-A-A-stop2 are purified from the gel. Stop1-A-A-A-stop2 fragments are subsequently PCR amplified using primers 5'-agcttcattacccaaatcaac-3' and 5' aattcaacgctactaccat-3' and subsequently digested with XmaI and SfiI. Selected polynucleotides are then cloned into a phage expression system and tested for affinity for the target protein.

[0256] High affinity binders are subsequently isolated and sequenced. DNA encoding the high binders is cloned into expression vector and subsequently expressed in a suitable host. The expressed protein is then purified and characterized.

[0257] 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.

```
                              SEQUENCE LISTING


<160> NUMBER OF SEQ ID NOS: 269

<210> SEQ ID NO 1
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 1

Cys Glu Pro Tyr Gln Phe Arg Cys Lys Asn Asn Arg Cys Val Pro Gly
 1               5                  10                  15

Arg Trp Gln Cys Asp Tyr Asp Asn Asp Cys Gly Asp Asn Ser Asp Glu
            20                  25                  30

Glu Ser Cys
        35


<210> SEQ ID NO 2
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 2

Cys Leu Pro Ser Gln Phe Lys Cys Thr Asn Thr Asn Arg Cys Ile Pro
 1               5                  10                  15

Gly Ile Phe Arg Cys Asn Gly Gln Asp Asn Cys Gly Asp Gly Glu Asp
            20                  25                  30

Glu Arg Asp Cys
        35


<210> SEQ ID NO 3
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR) A
      domain

<400> SEQUENCE: 3

Cys Ser Gln Asp Glu Phe Arg Cys His Asp Gly Lys Cys Thr Ser Arg
```

-continued

1 5 10 15

Gln Phe Val Cys Asp Ser Asp Arg Asp Cys Leu Asp Gly Ser Asp Glu
20 25 30

Ala Ser Cys
35

<210> SEQ ID NO 4
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR) related protein 2 (LRP2) A domain

<400> SEQUENCE: 4

Cys Ser Ser Ser Ala Phe Thr Cys Gly His Gly Glu Cys Ile Pro Ala
1 5 10 15

His Trp Arg Cys Asp Lys Arg Asn Asp Cys Val Asp Gly Ser Asp Glu
20 25 30

His Asn Cys
35

<210> SEQ ID NO 5
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR) related protein 2 (LRP2) A domain

<400> SEQUENCE: 5

Cys Ser Ser Ser Glu Phe Gln Cys Ala Ser Gly Arg Cys Ile Pro Gln
1 5 10 15

His Trp Tyr Cys Asp Gln Glu Thr Asp Cys Phe Asp Ala Ser Asp Glu
20 25 30

Pro Ala Ser Cys
35

<210> SEQ ID NO 6
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human CORI A domain

<400> SEQUENCE: 6

Cys His Ser Gln Gly Leu Val Glu Cys Arg Asn Gly Gln Cys Ile Pro
1 5 10 15

Ser Thr Phe Gln Cys Asp Gly Asp Glu Asp Cys Lys Asp Gly Ser Asp
20 25 30

Glu Glu Asn Cys
35

<210> SEQ ID NO 7
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human MAT A domain

<400> SEQUENCE: 7

-continued

```
Cys Pro Ala Gln Thr Phe Arg Cys Ser Asn Gly Lys Cys Leu Ser Lys
 1               5                  10                  15

Ser Gln Gln Cys Asn Gly Lys Asp Asp Cys Gly Asp Gly Ser Asp Glu
            20                  25                  30

Ala Ser Cys
        35


<210> SEQ ID NO 8
<211> LENGTH: 34
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human CO8B A domain

<400> SEQUENCE: 8

Cys Glu Gly Phe Val Cys Ala Gln Thr Gly Arg Cys Val Asn Arg Arg
 1               5                  10                  15

Leu Leu Cys Asn Gly Asp Asn Asp Cys Gly Asp Gln Ser Asp Glu Ala
            20                  25                  30

Asn Cys


<210> SEQ ID NO 9
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human MAT A domain

<400> SEQUENCE: 9

Cys Thr Lys His Thr Tyr Arg Cys Leu Asn Gly Leu Cys Leu Ser Lys
 1               5                  10                  15

Gly Asn Pro Glu Cys Asp Gly Lys Glu Asp Cys Ser Asp Gly Ser Asp
            20                  25                  30

Glu Lys Asp Cys
        35


<210> SEQ ID NO 10
<211> LENGTH: 38
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human LDVR A domain

<400> SEQUENCE: 10

Cys Leu Gly Pro Gly Lys Phe Lys Cys Arg Ser Gly Glu Cys Ile Asp
 1               5                  10                  15

Ile Ser Lys Val Cys Asn Gln Glu Gln Asp Cys Arg Asp Trp Ser Asp
            20                  25                  30

Glu Pro Leu Lys Glu Cys
        35


<210> SEQ ID NO 11
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human ApoE receptor 2 (ApoER2) A domain

<400> SEQUENCE: 11
```

-continued

```
Cys Pro Ala Glu Lys Leu Ser Cys Gly Pro Thr Ser His Lys Cys Val
 1               5                  10                  15

Pro Ala Ser Trp Arg Cys Asp Gly Glu Lys Asp Cys Glu Gly Gly Ala
            20                  25                  30

Asp Glu Ala Gly Cys
        35


<210> SEQ ID NO 12
<211> LENGTH: 41
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human SORL A domain

<400> SEQUENCE: 12

Cys Thr His Phe Met Asp Phe Val Cys Lys Asn Arg Gln Gln Cys Leu
 1               5                  10                  15

Phe His Ser Met Val Cys Asp Gly Ile Ile Gln Cys Arg Asp Gly Ser
            20                  25                  30

Asp Glu Asp Ala Ala Phe Ala Gly Cys
        35                  40


<210> SEQ ID NO 13
<211> LENGTH: 40
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human ST7 A domain

<400> SEQUENCE: 13

Cys Ala Tyr Asn Gln Phe Gln Cys Leu Ser Arg Phe Thr Lys Val Tyr
 1               5                  10                  15

Thr Cys Leu Pro Glu Ser Leu Lys Cys Asp Gly Asn Ile Asp Cys Leu
            20                  25                  30

Asp Leu Gly Asp Glu Ile Asp Cys
        35                  40


<210> SEQ ID NO 14
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: A domain consensus sequence
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (1)...(35)
<223> OTHER INFORMATION: Xaa = any amino acid

<400> SEQUENCE: 14

Cys Xaa Xaa Xaa Xaa Phe Xaa Cys Xaa Xaa Gly Xaa Cys Ile Xaa Xaa
 1               5                  10                  15

Xaa Xaa Xaa Cys Asp Gly Xaa Xaa Asp Cys Xaa Asp Xaa Ser Asp Glu
            20                  25                  30

Xaa Xaa Cys
        35


<210> SEQ ID NO 15
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
```

-continued

```
<223> OTHER INFORMATION: A domain consensus sequence
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (2)...(5)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (6)...(6)
<223> OTHER INFORMATION: Xaa = hydrophobic amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (7)...(7)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (9)...(12)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (14)...(14)
<223> OTHER INFORMATION: Xaa = hydrophobic amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (15)...(17)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (18)...(18)
<223> OTHER INFORMATION: Xaa = hydrophobic amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (19)...(19)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (21)...(21)
<223> OTHER INFORMATION: Xaa = negatively charged amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (22)...(24)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (25)...(25)
<223> OTHER INFORMATION: Xaa = negatively charged amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (27)...(30)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (31)...(32)
<223> OTHER INFORMATION: Xaa = negatively charged amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (33)...(34)
<223> OTHER INFORMATION: Xaa = any amino acid

<400> SEQUENCE: 15

Cys Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa
 1               5                  10                  15

Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa
            20                  25                  30

Xaa Xaa Cys
        35


<210> SEQ ID NO 16
<211> LENGTH: 41
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: A domain consensus sequence
```

-continued

```
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (2)...(2)
<223> OTHER INFORMATION: Xaa = Val, Leu, Gly, Pro, Ala, Glu, Gln or Arg
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (3)...(3)
<223> OTHER INFORMATION: Xaa = Ala, Pro or Ser
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (4)...(4)
<223> OTHER INFORMATION: Xaa = Asp, Asn, Gly or Ser
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (7)...(7)
<223> OTHER INFORMATION: Xaa = Thr, Pro, Arg, Lys or Gln
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (9)...(9)
<223> OTHER INFORMATION: Xaa = Gly, Ser, Asp, Glu, Asn, Lys or Arg
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (10)...(10)
<223> OTHER INFORMATION: Xaa = Asn or Ser
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (12)...(12)
<223> OTHER INFORMATION: Xaa = His, Gln or Arg
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (15)...(15)
<223> OTHER INFORMATION: Xaa = Pro or Ser
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (16)...(16)
<223> OTHER INFORMATION: Xaa = Val, Leu, Gly, Pro, Ala, Glu, Gln or Arg
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (17)...(17)
<223> OTHER INFORMATION: Xaa = Ala, Ser, Glu, Asn, His or Arg
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (18)...(18)
<223> OTHER INFORMATION: Xaa = Leu or Trp
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (19)...(19)
<223> OTHER INFORMATION: Xaa = Leu, Val, Gly or Arg
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (23)...(23)
<223> OTHER INFORMATION: Xaa = Val, Asp or Glu
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (24)...(24)
<223> OTHER INFORMATION: Xaa = Pro, Asn or Asp
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (27)...(27)
<223> OTHER INFORMATION: Xaa = Ala, Pro, Gly, Glu, Gln or Arg
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (29)...(29)
<223> OTHER INFORMATION: Xaa = Asn or Gly
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (33)...(33)
<223> OTHER INFORMATION: Xaa = Leu, Val, Met, Glu, Gln or Lys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (34)...(34)
<223> OTHER INFORMATION: Xaa = Gly, Ser, Asn or Asp
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (36)...(36)
<223> OTHER INFORMATION: Xaa = Ala, Pro, Thr, Gln, Glu or Lys
```

-continued

```
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (37)...(38)
<223> OTHER INFORMATION: Xaa = Ala, Asp, Glu, Gly, His, Ile, Lys, Leu,
      Asn, Pro, Gln, Arg, Ser, Thr or Val
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (39)...(39)
<223> OTHER INFORMATION: Xaa = Ser, Gly or Arg
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (40)...(40)
<223> OTHER INFORMATION: Xaa = His, Pro or Arg

<400> SEQUENCE: 16

Cys Xaa Xaa Xaa Glx Phe Xaa Cys Xaa Xaa Gly Xaa Cys Ile Xaa Xaa
 1               5                  10                  15

Xaa Xaa Xaa Cys Asp Gly Xaa Xaa Asp Cys Xaa Asp Xaa Ser Asp Glu
            20                  25                  30

Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Thr
        35                  40


<210> SEQ ID NO 17
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: A domain consensus sequence
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (1)...(35)
<223> OTHER INFORMATION: Xaa = any amino acid

<400> SEQUENCE: 17

Cys Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa
 1               5                  10                  15

Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa
            20                  25                  30

Xaa Xaa Cys
        35


<210> SEQ ID NO 18
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human IDD A domain

<400> SEQUENCE: 18

Cys Asn Pro Gly Gln Phe Ala Cys Arg Ser Gly Thr Ile Gln Cys Ile
 1               5                  10                  15

Pro Leu Pro Trp Gln Cys Asp Gly Trp Ala Thr Cys Glu Asp Glu Ser
            20                  25                  30

Asp Glu Ala Asn Cys
        35


<210> SEQ ID NO 19
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 3 (LRP3) A domain
```

-continued

```
<400> SEQUENCE: 19

Cys Gln Ala Asp Glu Phe Arg Cys Asp Asn Gly Lys Cys Leu Pro Gly
 1               5                   10                  15

Pro Trp Gln Cys Asn Thr Val Asp Glu Cys Gly Asp Gly Ser Asp Glu
             20                  25                  30

Gly Asn Cys
         35


<210> SEQ ID NO 20
<211> LENGTH: 38
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 3 (LRP3) A domain

<400> SEQUENCE: 20

Cys Pro Gly Gly Thr Phe Pro Cys Ser Gly Ala Arg Ser Thr Arg Cys
 1               5                   10                  15

Leu Pro Val Glu Arg Arg Cys Asp Gly Leu Gln Asp Cys Gly Asp Gly
             20                  25                  30

Ser Asp Glu Ala Gly Cys
         35


<210> SEQ ID NO 21
<211> LENGTH: 46
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 3 (LRP3) A domain

<400> SEQUENCE: 21

Cys Leu Pro Trp Glu Gln Pro Cys Gly Ser Ser Ser Asp Ser Asp Gly
 1               5                   10                  15

Gly Ser Leu Gly Asp Gln Gly Cys Phe Ser Glu Pro Gln Arg Cys Asp
             20                  25                  30

Gly Trp Trp His Cys Ala Ser Gly Arg Asp Glu Gln Gly Cys
         35                  40                  45


<210> SEQ ID NO 22
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 3 (LRP3) A domain

<400> SEQUENCE: 22

Cys Pro Pro Asp Gln Tyr Pro Cys Glu Gly Gly Ser Gly Leu Cys Tyr
 1               5                   10                  15

Thr Pro Ala Asp Arg Cys Asn Asn Gln Lys Ser Cys Pro Asp Gly Ala
             20                  25                  30

Asp Glu Lys Asn Cys
         35


<210> SEQ ID NO 23
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
```

-continued

```
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 3 (LRP3) A domain

<400> SEQUENCE: 23

Cys Gln Pro Gly Thr Phe His Cys Gly Thr Asn Leu Cys Ile Phe Glu
 1               5                  10                  15

Thr Trp Arg Cys Asp Gly Gln Glu Asp Cys Gln Asp Gly Ser Asp Glu
            20                  25                  30

His Gly Cys
        35


<210> SEQ ID NO 24
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 5 (LRP5) A domain

<400> SEQUENCE: 24

Cys Ser Pro Asp Gln Phe Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile
 1               5                  10                  15

Pro Gly Ala Trp Arg Cys Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser
            20                  25                  30

Asp Glu Glu Gly Cys
        35


<210> SEQ ID NO 25
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 5 (LRP5) A domain

<400> SEQUENCE: 25

Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln Cys Val Asp Leu
 1               5                  10                  15

Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp Arg Ser Asp Glu
            20                  25                  30

Val Asp Cys
        35


<210> SEQ ID NO 26
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 5 (LRP5) A domain

<400> SEQUENCE: 26

Cys Leu Pro Asn Gln Phe Arg Cys Ala Ser Gly Gln Cys Val Leu Ile
 1               5                  10                  15

Lys Gln Gln Cys Asp Ser Phe Pro Asp Cys Ile Asp Gly Ser Asp Glu
            20                  25                  30

Leu Met Cys
        35
```

-continued

```
<210> SEQ ID NO 27
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 6 (LRP6) A domain

<400> SEQUENCE: 27

Cys Ser Pro Gln Gln Phe Thr Cys Phe Thr Gly Glu Ile Asp Cys Ile
 1               5                  10                  15

Pro Val Ala Trp Arg Cys Asp Gly Phe Thr Glu Cys Glu Asp His Ser
            20                  25                  30

Asp Glu Leu Asn Cys
        35


<210> SEQ ID NO 28
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 6 (LRP6) A domain

<400> SEQUENCE: 28

Cys Ser Glu Ser Gln Phe Gln Cys Ala Ser Gly Gln Cys Ile Asp Gly
 1               5                  10                  15

Ala Leu Arg Cys Asn Gly Asp Ala Asn Cys Gln Asp Lys Ser Asp Glu
            20                  25                  30

Lys Asn Cys
        35


<210> SEQ ID NO 29
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 6 (LRP6) A domain

<400> SEQUENCE: 29

Cys Leu Ile Asp Gln Phe Arg Cys Ala Asn Gly Gln Cys Ile Gly Lys
 1               5                  10                  15

His Lys Lys Cys Asp His Asn Val Asp Cys Ser Asp Lys Ser Asp Glu
            20                  25                  30

Leu Asp Cys
        35


<210> SEQ ID NO 30
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human ST7 A domain

<400> SEQUENCE: 30

Cys Ala Cys Asp Gln Phe Arg Cys Gly Asn Gly Lys Cys Ile Pro Glu
 1               5                  10                  15

Ala Trp Lys Cys Asn Asn Met Asp Glu Cys Gly Asp Ser Ser Asp Glu
            20                  25                  30

Glu Ile Cys
```

-continued

```
        35


<210> SEQ ID NO 31
<211> LENGTH: 40
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human ST7 A domain

<400> SEQUENCE: 31

Cys Ala Tyr Asn Gln Phe Gln Cys Leu Ser Arg Phe Thr Lys Val Tyr
 1               5                  10                  15

Thr Cys Leu Pro Glu Ser Leu Lys Cys Asp Gly Asn Ile Asp Cys Leu
            20                  25                  30

Asp Leu Gly Asp Glu Ile Asp Cys
        35                  40


<210> SEQ ID NO 32
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human ST7 A domain

<400> SEQUENCE: 32

Cys Leu Pro Trp Glu Ile Pro Cys Gly Gly Asn Trp Gly Cys Tyr Thr
 1               5                  10                  15

Glu Gln Gln Arg Cys Asp Gly Tyr Trp His Cys Pro Asn Gly Arg Asp
            20                  25                  30

Glu Thr Asn Cys
        35


<210> SEQ ID NO 33
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human ST7 A domain

<400> SEQUENCE: 33

Cys Gln Lys Glu Glu Phe Pro Cys Ser Arg Asn Gly Val Cys Tyr Pro
 1               5                  10                  15

Arg Ser Asp Arg Cys Asn Tyr Gln Asn His Cys Pro Asn Gly Ser Asp
            20                  25                  30

Glu Lys Asn Cys
        35


<210> SEQ ID NO 34
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human ST7 A domain

<400> SEQUENCE: 34

Cys Gln Pro Gly Asn Phe His Cys Lys Asn Asn Arg Cys Val Phe Glu
 1               5                  10                  15

Ser Trp Val Cys Asp Ser Gln Asp Asp Cys Gly Asp Gly Ser Asp Glu
            20                  25                  30

Glu Asn Cys
```

-continued

```
        35


<210> SEQ ID NO 35
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human CORI A domain

<400> SEQUENCE: 35

Cys Gly Arg Gly Glu Asn Phe Leu Cys Ala Ser Gly Ile Cys Ile Pro
 1               5                  10                  15

Gly Lys Leu Gln Cys Asn Gly Tyr Asn Asp Cys Asp Asp Trp Ser Asp
            20                  25                  30

Glu Ala His Cys
        35


<210> SEQ ID NO 36
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human CORI A domain

<400> SEQUENCE: 36

Cys Ser Glu Asn Leu Phe His Cys His Thr Gly Lys Cys Leu Asn Tyr
 1               5                  10                  15

Ser Leu Val Cys Asp Gly Tyr Asp Asp Cys Gly Asp Leu Ser Asp Glu
            20                  25                  30

Gln Asn Cys
        35


<210> SEQ ID NO 37
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human CORI A domain

<400> SEQUENCE: 37

Cys Asn Pro Thr Thr Glu His Arg Cys Gly Asp Gly Arg Cys Ile Ala
 1               5                  10                  15

Met Glu Trp Val Cys Asp Gly Asp His Asp Cys Val Asp Lys Ser Asp
            20                  25                  30

Glu Val Asn Cys
        35


<210> SEQ ID NO 38
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human CORI A domain

<400> SEQUENCE: 38

Cys His Ser Gln Gly Leu Val Glu Cys Arg Asn Gly Gln Cys Ile Pro
 1               5                  10                  15

Ser Thr Phe Gln Cys Asp Gly Asp Glu Asp Cys Lys Asp Gly Ser Asp

            20                  25                  30

Glu Glu Asn Cys
```

```
        35


<210> SEQ ID NO 39
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human CORI A domain

<400> SEQUENCE: 39

Cys Ser Pro Ser His Phe Lys Cys Arg Ser Gly Gln Cys Val Leu Ala
 1              5                   10                  15

Ser Arg Arg Cys Asp Gly Gln Ala Asp Cys Asp Asp Asp Ser Asp Glu
            20                  25                  30

Glu Asn Cys
        35


<210> SEQ ID NO 40
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human CORI A domain

<400> SEQUENCE: 40

Cys Lys Glu Arg Asp Leu Trp Glu Cys Pro Ser Asn Lys Gln Cys Leu
 1              5                   10                  15

Lys His Thr Val Ile Cys Asp Gly Phe Pro Asp Cys Pro Asp Tyr Met
            20                  25                  30

Asp Glu Lys Asn Cys
        35


<210> SEQ ID NO 41
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human CORI A domain

<400> SEQUENCE: 41

Cys Gln Asp Asp Glu Leu Glu Cys Ala Asn His Ala Cys Val Ser Arg
 1              5                   10                  15

Asp Leu Trp Cys Asp Gly Glu Ala Asp Cys Ser Asp Ser Ser Asp Glu
            20                  25                  30

Trp Asp Cys
        35


<210> SEQ ID NO 42
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human TMS2 A domain

<400> SEQUENCE: 42

Cys Ser Asn Ser Gly Ile Glu Cys Asp Ser Ser Gly Thr Cys Ile Asn
 1              5                   10                  15

Pro Ser Asn Trp Cys Asp Gly Val Ser His Cys Pro Gly Gly Glu Asp
            20                  25                  30

Glu Asn Arg Cys
```

```
        35


<210> SEQ ID NO 43
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human TMS3 A domain

<400> SEQUENCE: 43

Cys Ser Gly Lys Tyr Arg Cys Arg Ser Ser Phe Lys Cys Ile Glu Leu
 1               5                  10                  15

Ile Ala Arg Cys Asp Gly Val Ser Asp Cys Lys Asp Gly Glu Asp Glu
            20                  25                  30

Tyr Arg Cys
        35


<210> SEQ ID NO 44
<211> LENGTH: 34
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human MAT A domain

<400> SEQUENCE: 44

Cys Pro Gly Gln Phe Thr Cys Arg Thr Gly Arg Cys Ile Arg Lys Glu
 1               5                  10                  15

Leu Arg Cys Asp Gly Trp Ala Asp Cys Thr Asp His Ser Asp Glu Leu
            20                  25                  30

Asn Cys


<210> SEQ ID NO 45
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human MAT A domain

<400> SEQUENCE: 45

Cys Asp Ala Gly His Gln Phe Thr Cys Lys Asn Lys Phe Cys Lys Pro
 1               5                  10                  15

Leu Phe Trp Val Cys Asp Ser Val Asn Asp Cys Gly Asp Asn Ser Asp
            20                  25                  30

Glu Gln Gly Cys
        35


<210> SEQ ID NO 46
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human MAT A domain

<400> SEQUENCE: 46

Cys Pro Ala Gln Thr Phe Arg Cys Ser Asn Gly Lys Cys Leu Ser Lys
 1               5                  10                  15

Ser Gln Gln Cys Asn Gly Lys Asp Asp Cys Gly Asp Gly Ser Asp Glu
            20                  25                  30

Ala Ser Cys
        35
```

-continued

```
<210> SEQ ID NO 47
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human MAT A domain

<400> SEQUENCE: 47

Cys Thr Lys His Thr Tyr Arg Cys Leu Asn Gly Leu Cys Leu Ser Lys
 1               5                  10                  15

Gly Asn Pro Glu Cys Asp Gly Lys Glu Asp Cys Ser Asp Gly Ser Asp
            20                  25                  30

Glu Lys Asp Cys
        35


<210> SEQ ID NO 48
<211> LENGTH: 38
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human ENTK A domain

<400> SEQUENCE: 48

Cys Leu Pro Gly Ser Ser Pro Cys Thr Asp Ala Leu Thr Cys Ile Lys
 1               5                  10                  15

Ala Asp Leu Phe Cys Asp Gly Glu Val Asn Cys Pro Asp Gly Ser Asp
            20                  25                  30

Glu Asp Asn Lys Met Cys
        35


<210> SEQ ID NO 49
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human ENTK A domain

<400> SEQUENCE: 49

Cys Lys Ala Asp His Phe Gln Cys Lys Asn Gly Glu Cys Val Pro Leu
 1               5                  10                  15

Val Asn Leu Cys Asp Gly His Leu His Cys Glu Asp Gly Ser Asp Glu
            20                  25                  30

Ala Asp Cys
        35


<210> SEQ ID NO 50
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human HAI1 A domain

<400> SEQUENCE: 50

Cys Gln Pro Thr Gln Phe Arg Cys Ser Asn Gly Cys Cys Ile Asp Ser
 1               5                  10                  15

Phe Leu Glu Cys Asp Asp Thr Pro Asn Cys Pro Asp Ala Ser Asp Glu
            20                  25                  30

Ala Ala Cys
        35
```

-continued

```
<210> SEQ ID NO 51
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human CFAI A domain

<400> SEQUENCE: 51

Cys Tyr Thr Gln Lys Ala Asp Ser Pro Met Asp Asp Phe Phe Gln Cys
 1               5                  10                  15

Val Asn Gly Lys Tyr Ile Ser Gln Met Lys Ala Cys Asp Gly Ile Asn
            20                  25                  30

Asp Cys Gly Asp Gln Ser Asp Glu Leu Cys
        35                  40


<210> SEQ ID NO 52
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human CFAI A domain

<400> SEQUENCE: 52

Cys Gln Gly Lys Gly Phe His Cys Lys Ser Gly Val Cys Ile Pro Ser
 1               5                  10                  15

Gln Tyr Gln Cys Asn Gly Glu Val Asp Cys Ile Thr Gly Glu Asp Glu
            20                  25                  30

Val Gly Cys
        35


<210> SEQ ID NO 53
<211> LENGTH: 34
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human CO6 A domain

<400> SEQUENCE: 53

Cys Lys Asn Lys Phe Arg Cys Asp Ser Gly Arg Cys Ile Ala Arg Lys
 1               5                  10                  15

Leu Glu Cys Asn Gly Glu Asn Asp Cys Gly Asp Asn Ser Asp Glu Arg
            20                  25                  30

Asp Cys


<210> SEQ ID NO 54
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human CO7 A domain

<400> SEQUENCE: 54

Cys Gly Glu Arg Phe Arg Cys Phe Ser Gly Gln Cys Ile Ser Lys Ser
 1               5                  10                  15

Leu Val Cys Asn Gly Asp Ser Asp Cys Asp Glu Asp Ser Ala Asp Glu
            20                  25                  30

Asp Arg Cys
        35
```

-continued

<210> SEQ ID NO 55
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human CO8A A domain

<400> SEQUENCE: 55

```
Cys Gly Gln Asp Phe Gln Cys Lys Glu Thr Gly Arg Cys Leu Lys Arg
 1               5                  10                  15

His Leu Val Cys Asn Gly Asp Gln Asp Cys Leu Asp Gly Ser Asp Glu
            20                  25                  30

Asp Asp Cys
        35
```

<210> SEQ ID NO 56
<211> LENGTH: 34
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human CO8B A domain

<400> SEQUENCE: 56

```
Cys Glu Gly Phe Val Cys Ala Gln Thr Gly Arg Cys Val Asn Arg Arg
 1               5                  10                  15

Leu Leu Cys Asn Gly Asp Asn Asp Cys Gly Asp Gln Ser Asp Glu Ala
            20                  25                  30

Asn Cys
```

<210> SEQ ID NO 57
<211> LENGTH: 34
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human CO9 A domain

<400> SEQUENCE: 57

```
       Cys Gly Asn Asp Phe Gln Cys Ser Thr Gly Arg Cys Ile Lys Met Arg
 1               5                  10                  15
Leu Arg Cys Asn Gly Asp Asn Asp Cys Gly Asp Phe Ser Asp Glu Asp

            20                  25                  30
Asp Cys
```

<210> SEQ ID NO 58
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human PERL A domain

<400> SEQUENCE: 58

```
Cys Thr Glu Ala Glu Phe Ala Cys His Ser Tyr Asn Glu Cys Val Ala
 1               5                  10                  15

Leu Glu Tyr Arg Cys Asp Arg Arg Pro Asp Cys Arg Asp Met Ser Asp
            20                  25                  30

Glu Leu Asn Cys
        35
```

<210> SEQ ID NO 59

```
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human PERL A domain

<400> SEQUENCE: 59

Cys Gly Pro Gln Glu Ala Ala Cys Arg Asn Gly His Cys Ile Pro Arg
 1               5                  10                  15

Asp Tyr Leu Cys Asp Gly Gln Glu Asp Cys Glu Asp Gly Ser Asp Glu
            20                  25                  30

Leu Asp Cys
        35


<210> SEQ ID NO 60
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human PERL A domain

<400> SEQUENCE: 60

Cys Glu Pro Asn Glu Phe Pro Cys Gly Asn Gly His Cys Ala Leu Lys
 1               5                  10                  15

Leu Trp Arg Cys Asp Gly Asp Phe Asp Cys Glu Asp Arg Thr Asp Glu
            20                  25                  30

Ala Asn Cys
        35


<210> SEQ ID NO 61
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human PERL A domain

<400> SEQUENCE: 61

Cys Gly Pro Thr Gln Phe Arg Cys Val Ser Thr Asn Met Cys Ile Pro
 1               5                  10                  15

Ala Ser Phe His Cys Asp Glu Glu Ser Asp Cys Pro Asp Arg Ser Asp
            20                  25                  30

Glu Phe Gly Cys
        35


<210> SEQ ID NO 62
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human SORL A domain

<400> SEQUENCE: 62

Cys Leu Arg Asn Gln Tyr Arg Cys Ser Asn Gly Asn Cys Ile Asn Ser
 1               5                  10                  15

Ile Trp Trp Cys Asp Phe Asp Asn Asp Cys Gly Asp Met Ser Asp Glu
            20                  25                  30

Arg Asn Cys
        35


<210> SEQ ID NO 63
```

-continued

```
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human SORL A domain

<400> SEQUENCE: 63

Cys Asp Leu Asp Thr Gln Phe Arg Cys Gln Glu Ser Gly Thr Cys Ile
 1              5                   10                  15

Pro Leu Ser Tyr Lys Cys Asp Leu Glu Asp Asp Cys Gly Asp Asn Ser
            20                  25                  30

Asp Glu Ser His Cys
        35


<210> SEQ ID NO 64
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human SORL A domain

<400> SEQUENCE: 64

Cys Arg Ser Asp Glu Tyr Asn Cys Ser Ser Gly Met Cys Ile Arg Ser
 1              5                   10                  15

Ser Trp Val Cys Asp Gly Asp Asn Asp Cys Arg Asp Trp Ser Asp Glu
            20                  25                  30

Ala Asn Cys
        35


<210> SEQ ID NO 65
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human SORL A domain

<400> SEQUENCE: 65

Cys Glu Ala Ser Asn Phe Gln Cys Arg Asn Gly His Cys Ile Pro Gln
 1              5                   10                  15

Arg Trp Ala Cys Asp Gly Asp Thr Asp Cys Gln Asp Gly Ser Asp Glu
            20                  25                  30

Asp Pro Val Asn Cys
        35


<210> SEQ ID NO 66
<211> LENGTH: 33
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human SORL A domain

<400> SEQUENCE: 66

Cys Asn Gly Phe Arg Cys Pro Asn Gly Thr Cys Ile Pro Ser Ser Lys
 1              5                   10                  15

His Cys Asp Gly Leu Arg Asp Cys Ser Asp Gly Ser Asp Glu Gln His
            20                  25                  30

Cys


<210> SEQ ID NO 67
<211> LENGTH: 41
```

-continued

```
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human SORL A domain

<400> SEQUENCE: 67

Cys Thr His Phe Met Asp Phe Val Cys Lys Asn Arg Gln Gln Cys Leu
 1               5                  10                  15

Phe His Ser Met Val Cys Asp Gly Ile Ile Gln Cys Arg Asp Gly Ser
            20                  25                  30

Asp Glu Asp Ala Ala Phe Ala Gly Cys
        35                  40


<210> SEQ ID NO 68
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human SORL A domain

<400> SEQUENCE: 68

Cys Asp Glu Phe Gly Phe Gln Cys Gln Asn Gly Val Cys Ile Ser Leu
 1               5                  10                  15

Ile Trp Lys Cys Asp Gly Met Asp Asp Cys Gly Asp Tyr Ser Asp Glu
            20                  25                  30

Ala Asn Cys
        35


<210> SEQ ID NO 69
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human SORL A domain

<400> SEQUENCE: 69

Cys Ser Arg Tyr Phe Gln Phe Arg Cys Glu Asn Gly His Cys Ile Pro
 1               5                  10                  15

Asn Arg Trp Lys Cys Asp Arg Glu Asn Asp Cys Gly Asp Trp Ser Asp
            20                  25                  30

Glu Lys Asp Cys
        35


<210> SEQ ID NO 70
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human SORL A domain

<400> SEQUENCE: 70

Cys Leu Pro Asn Tyr Tyr Arg Cys Ser Ser Gly Thr Cys Val Met Asp
 1               5                  10                  15

Thr Trp Val Cys Asp Gly Tyr Arg Asp Cys Ala Asp Gly Ser Asp Glu
            20                  25                  30

Glu Ala Cys
        35


<210> SEQ ID NO 71
<211> LENGTH: 36
```

-continued

```
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human SORL A domain

<400> SEQUENCE: 71

Cys Asp Arg Phe Glu Phe Glu Cys His Gln Pro Lys Thr Cys Ile Pro
 1               5                  10                  15

Asn Trp Lys Arg Cys Asp Gly His Gln Asp Cys Gln Asp Gly Arg Asp
            20                  25                  30

Glu Ala Asn Cys
        35


<210> SEQ ID NO 72
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human SORL A domain

<400> SEQUENCE: 72

Cys Met Ser Arg Glu Phe Gln Cys Glu Asp Gly Glu Ala Cys Ile Val
 1               5                  10                  15

Leu Ser Glu Arg Cys Asp Gly Phe Leu Asp Cys Ser Asp Glu Ser Asp
            20                  25                  30

Glu Lys Ala Cys
        35


<210> SEQ ID NO 73
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human SORL A domain

<400> SEQUENCE: 73

Cys Glu Lys Asp Gln Phe Gln Cys Arg Asn Glu Arg Cys Ile Pro Ser
 1               5                  10                  15

Val Trp Arg Cys Asp Glu Asp Asp Asp Cys Leu Asp His Ser Asp Glu
            20                  25                  30

Asp Asp Cys
        35


<210> SEQ ID NO 74
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human SORL A domain

<400> SEQUENCE: 74

Cys Ala Asp Ser Asp Phe Thr Cys Asp Asn Gly His Cys Ile His Glu
 1               5                  10                  15

Arg Trp Lys Cys Asp Gly Glu Glu Glu Cys Pro Asp Gly Ser Asp Glu
            20                  25                  30

Ser Glu Ala Thr Cys
        35


<210> SEQ ID NO 75
<211> LENGTH: 37
```

-continued

```
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human ApoER2 A domain

<400> SEQUENCE: 75

Cys Pro Ala Glu Lys Leu Ser Cys Gly Pro Thr Ser His Lys Cys Val
 1               5                   10                  15

Pro Ala Ser Trp Arg Cys Asp Gly Glu Lys Asp Cys Glu Gly Gly Ala
            20                  25                  30

Asp Glu Ala Gly Cys
        35


<210> SEQ ID NO 76
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human ApoER2 A domain

<400> SEQUENCE: 76

Cys Ala Pro His Glu Phe Gln Cys Gly Asn Arg Ser Cys Leu Ala Ala
 1               5                   10                  15

Val Phe Val Cys Asp Gly Asp Asp Asp Cys Gly Asp Gly Ser Asp Glu
            20                  25                  30

Arg Gly Cys
        35


<210> SEQ ID NO 77
<211> LENGTH: 40
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human ApoER2 A domain

<400> SEQUENCE: 77

Cys Gly Pro Arg Glu Phe Arg Cys Gly Gly Asp Gly Gly Gly Ala Cys
 1               5                   10                  15

Ile Pro Glu Arg Trp Val Cys Asp Arg Gln Phe Asp Cys Glu Asp Arg
            20                  25                  30

Ser Asp Glu Ala Ala Glu Leu Cys
        35                  40


<210> SEQ ID NO 78
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human ApoER2 A domain

<400> SEQUENCE: 78

Cys Ala Thr Val Ser Gln Phe Ala Cys Arg Ser Gly Glu Cys Val His
 1               5                   10                  15

Leu Gly Trp Arg Cys Asp Gly Asp Arg Asp Cys Lys Asp Lys Ser Asp
            20                  25                  30

Glu Ala Asp Cys
        35


<210> SEQ ID NO 79
<211> LENGTH: 35
```

-continued

```
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human ApoER2 A domain

<400> SEQUENCE: 79

Cys Arg Gly Asp Glu Phe Gln Cys Gly Asp Gly Thr Cys Val Leu Ala
 1               5                  10                  15

Ile Lys His Cys Asn Gln Glu Gln Asp Cys Pro Asp Gly Ser Asp Glu
            20                  25                  30

Ala Gly Cys
        35


<210> SEQ ID NO 80
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR) A
      domain

<400> SEQUENCE: 80

Cys Glu Arg Asn Glu Phe Gln Cys Gln Asp Gly Lys Cys Ile Ser Tyr

 1               5                  10                  15
Lys Trp Val Cys Asp Gly Ser Ala Glu Cys Gln Asp Gly Ser Asp Glu

            20                  25                  30
Ser Gln Glu Thr Cys
               35


<210> SEQ ID NO 81
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR) A
      domain

<400> SEQUENCE: 81

Cys Lys Ser Gly Asp Phe Ser Cys Gly Gly Arg Val Asn Arg Cys Ile
 1               5                  10                  15

Pro Gln Phe Trp Arg Cys Asp Gly Gln Val Asp Cys Asp Asn Gly Ser
            20                  25                  30

Asp Glu Gln Gly Cys
        35


<210> SEQ ID NO 82
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR) A
      domain

<400> SEQUENCE: 82

Cys Ser Gln Asp Glu Phe Arg Cys His Asp Gly Lys Cys Ile Ser Arg
 1               5                  10                  15

Gln Phe Val Cys Asp Ser Asp Arg Asp Cys Leu Asp Gly Ser Asp Glu
            20                  25                  30

Ala Ser Cys
        35
```

-continued

```
<210> SEQ ID NO 83
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR) A
      domain

<400> SEQUENCE: 83

Cys Gly Pro Ala Ser Phe Gln Cys Asn Ser Ser Thr Cys Ile Pro Gln
 1               5                  10                  15

Leu Trp Ala Cys Asp Asn Asp Pro Asp Cys Glu Asp Gly Ser Asp Glu
            20                  25                  30

Trp Pro Gln Arg Cys
        35


<210> SEQ ID NO 84
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR) A
      domain

<400> SEQUENCE: 84

Cys Ser Ala Phe Glu Phe His Cys Leu Ser Gly Glu Cys Ile His Ser
 1               5                  10                  15

Ser Trp Arg Cys Asp Gly Gly Pro Asp Cys Lys Asp Lys Ser Asp Glu
            20                  25                  30

Glu Asn Cys
        35


<210> SEQ ID NO 85
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR) A
      domain

<400> SEQUENCE: 85

Cys Arg Pro Asp Glu Phe Gln Cys Ser Asp Gly Asn Cys Ile His Gly
 1               5                  10                  15

Ser Arg Gln Cys Asp Arg Glu Tyr Asp Cys Lys Asp Met Ser Asp Glu
            20                  25                  30

Val Gly Cys
        35


<210> SEQ ID NO 86
<211> LENGTH: 38
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR) A
      domain

<400> SEQUENCE: 86

Cys Glu Gly Pro Asn Lys Phe Lys Cys His Ser Gly Glu Cys Ile Thr
 1               5                  10                  15

Leu Asp Lys Val Cys Asn Met Ala Arg Asp Cys Arg Asp Trp Ser Asp
            20                  25                  30
```

-continued

```
Glu Pro Ile Lys Glu Cys
        35


<210> SEQ ID NO 87
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human LDVR A domain

<400> SEQUENCE: 87

Cys Glu Pro Ser Gln Phe Gln Cys Thr Asn Gly Arg Cys Ile Thr Leu
 1               5                  10                  15

Leu Trp Lys Cys Asp Gly Asp Glu Asp Cys Val Asp Gly Ser Asp Glu
            20                  25                  30

Lys Asn Cys
        35


<210> SEQ ID NO 88
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human LDVR A domain

<400> SEQUENCE: 88

Cys Ala Glu Ser Asp Phe Val Cys Asn Asn Gly Gln Cys Val Pro Ser
 1               5                  10                  15

Arg Trp Lys Cys Asp Gly Asp Pro Asp Cys Glu Asp Gly Ser Asp Glu
            20                  25                  30

Ser Pro Glu Gln Cys
        35


<210> SEQ ID NO 89
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human LDVR A domain

<400> SEQUENCE: 89

Cys Arg Ile His Glu Ile Ser Cys Gly Ala His Ser Thr Gln Cys Ile
 1               5                  10                  15

Pro Val Ser Trp Arg Cys Asp Gly Glu Asn Asp Cys Asp Ser Gly Glu
            20                  25                  30

Asp Glu Glu Asn Cys
        35


<210> SEQ ID NO 90
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human LDVR A domain

<400> SEQUENCE: 90

Cys Ser Pro Asp Glu Phe Thr Cys Ser Ser Gly Arg Cys Ile Ser Arg
 1               5                  10                  15

Asn Phe Val Cys Asn Gly Gln Asp Asp Cys Ser Asp Gly Ser Asp Glu
            20                  25                  30
```

-continued

```
Leu Asp Cys
        35


<210> SEQ ID NO 91
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human LDVR A domain

<400> SEQUENCE: 91

Cys Gly Ala His Glu Phe Gln Cys Ser Thr Ser Ser Cys Ile Pro Ile
 1               5                  10                  15

Ser Trp Val Cys Asp Asp Asp Ala Asp Cys Ser Asp Gln Ser Asp Glu
            20                  25                  30

Ser Leu Glu Gln Cys
        35


<210> SEQ ID NO 92
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human LDVR A domain

<400> SEQUENCE: 92

Cys Pro Ala Ser Glu Ile Gln Cys Gly Ser Gly Glu Cys Ile His Lys
 1               5                  10                  15

Lys Trp Arg Cys Asp Gly Asp Pro Asp Cys Lys Asp Gly Ser Asp Glu
            20                  25                  30

Val Asn Cys
        35


<210> SEQ ID NO 93
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human LDVR A domain

<400> SEQUENCE: 93

Cys Arg Pro Asp Gln Phe Glu Cys Glu Asp Gly Ser Cys Ile His Gly
 1               5                  10                  15

Ser Arg Gln Cys Asn Gly Ile Arg Asp Cys Val Asp Gly Ser Asp Glu
            20                  25                  30

Val Asn Cys
        35


<210> SEQ ID NO 94
<211> LENGTH: 38
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human LDVR A domain

<400> SEQUENCE: 94

Cys Leu Gly Pro Gly Lys Phe Lys Cys Arg Ser Gly Glu Cys Ile Asp
 1               5                  10                  15

Ile Ser Lys Val Cys Asn Gln Glu Gln Asp Cys Arg Asp Trp Ser Asp
            20                  25                  30
```

-continued

```
Glu Pro Leu Lys Glu Cys
        35


<210> SEQ ID NO 95
<211> LENGTH: 38
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 95

Cys Ser Pro Lys Gln Phe Ala Cys Arg Asp Gln Ile Thr Cys Ile Ser
 1               5                  10                  15

Lys Gly Trp Arg Cys Asp Gly Glu Arg Asp Cys Pro Asp Gly Ser Asp
            20                  25                  30

Glu Ala Pro Glu Ile Cys
        35


<210> SEQ ID NO 96
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 96

Cys Gln Pro Asn Glu His Asn Cys Leu Gly Thr Glu Leu Cys Val Pro
 1               5                  10                  15

Met Ser Arg Leu Cys Asn Gly Val Gln Asp Cys Met Asp Gly Ser Asp
            20                  25                  30

Glu Gly Pro His Cys
        35


<210> SEQ ID NO 97
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 97

Cys Gln Pro Gly Glu Phe Ala Cys Ala Asn Ser Arg Cys Ile Gln Glu
 1               5                  10                  15

Arg Trp Lys Cys Asp Gly Asp Asn Asp Cys Leu Asp Asn Ser Asp Glu
            20                  25                  30

Ala Pro Ala Leu Cys
        35


<210> SEQ ID NO 98
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 98

Cys Pro Ser Asp Arg Phe Lys Cys Glu Asn Asn Arg Cys Ile Pro Asn
```

```
                              -continued
 1               5                   10                  15

Arg Trp Leu Cys Asp Gly Asp Asn Asp Cys Gly Asn Ser Glu Asp Glu
            20                  25                  30

Ser Asn Ala Thr Cys
        35


<210> SEQ ID NO 99
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 99

Cys Pro Pro Asn Gln Phe Ser Cys Ala Ser Gly Arg Cys Ile Pro Ile
 1               5                   10                  15

Ser Trp Thr Cys Asp Leu Asp Asp Asp Cys Gly Asp Arg Ser Asp Glu
            20                  25                  30

Ser Ala Ser Cys
        35


<210> SEQ ID NO 100
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 100

Cys Phe Pro Leu Thr Gln Phe Thr Cys Asn Asn Gly Arg Cys Ile Asn
 1               5                   10                  15

Ile Asn Trp Arg Cys Asp Asn Asp Asn Asp Cys Gly Asp Asn Ser Asp
            20                  25                  30

Glu Ala Gly Cys
        35


<210> SEQ ID NO 101
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 101

Cys Ser Ser Thr Gln Phe Lys Cys Asn Ser Gly Arg Cys Ile Pro Glu
 1               5                   10                  15

His Trp Thr Cys Asp Gly Asp Asn Asp Cys Gly Asp Tyr Ser Asp Glu
            20                  25                  30

Thr His Ala Asn Cys
        35


<210> SEQ ID NO 102
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain
```

-continued

```
<400> SEQUENCE: 102

Cys His Thr Asp Glu Phe Gln Cys Arg Leu Asp Gly Leu Cys Ile Pro
 1               5                   10                  15

Leu Arg Trp Arg Cys Asp Gly Asp Thr Asp Cys Met Asp Ser Ser Asp
            20                  25                  30

Glu Lys Ser Cys
        35


<210> SEQ ID NO 103
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 103

Cys Asp Pro Ser Val Lys Phe Gly Cys Lys Asp Ser Ala Arg Cys Ile
 1               5                   10                  15

Ser Lys Ala Trp Val Cys Asp Gly Asp Asn Asp Cys Glu Asp Asn Ser
            20                  25                  30

Asp Glu Glu Asn Cys
        35


<210> SEQ ID NO 104
<211> LENGTH: 38
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 104

Cys Arg Pro Pro Ser His Pro Cys Ala Asn Asn Thr Ser Val Cys Leu
 1               5                   10                  15

Pro Pro Asp Lys Leu Cys Asp Gly Asn Asp Asp Cys Gly Asp Gly Ser
            20                  25                  30

Asp Glu Gly Glu Leu Cys
        35


<210> SEQ ID NO 105
<211> LENGTH: 38
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 105

Cys Arg Ala Gln Asp Glu Phe Glu Cys Ala Asn Gly Glu Cys Ile Asn
 1               5                   10                  15

Phe Ser Leu Thr Cys Asp Gly Val Pro His Cys Lys Asp Lys Ser Asp
            20                  25                  30

Glu Lys Pro Ser Tyr Cys
        35


<210> SEQ ID NO 106
<211> LENGTH: 35
<212> TYPE: PRT
```

-continued

```
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 106

Cys Lys Lys Thr Phe Arg Gln Cys Ser Asn Gly Arg Cys Val Ser Asn
 1              5                   10                  15

Met Leu Trp Cys Asn Gly Ala Asp Asp Cys Gly Asp Gly Ser Asp Glu
            20                  25                  30

Ile Pro Cys
        35


<210> SEQ ID NO 107
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 107

Cys Gly Val Gly Glu Phe Arg Cys Arg Asp Gly Thr Cys Ile Gly Asn
 1              5                   10                  15

Ser Ser Arg Cys Asn Gln Phe Val Asp Cys Glu Asp Ala Ser Asp Glu
            20                  25                  30

Met Asn Cys
        35


<210> SEQ ID NO 108
<211> LENGTH: 45
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 108

Cys Ser Ser Tyr Phe Arg Leu Gly Val Lys Gly Val Leu Phe Gln Pro
 1              5                   10                  15

Cys Glu Arg Thr Ser Leu Cys Tyr Ala Pro Ser Trp Val Cys Asp Gly
            20                  25                  30

Ala Asn Asp Cys Gly Asp Tyr Ser Asp Glu Arg Asp Cys
        35                  40                  45


<210> SEQ ID NO 109
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 109

Cys Pro Leu Asn Tyr Phe Ala Cys Pro Ser Gly Arg Cys Ile Pro Met
 1              5                   10                  15

Ser Trp Thr Cys Asp Lys Glu Asp Asp Cys Glu His Gly Glu Asp Glu
            20                  25                  30

Thr His Cys
        35
```

-continued

```
<210> SEQ ID NO 110
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 110

Cys Ser Glu Ala Gln Phe Glu Cys Gln Asn His Arg Cys Ile Ser Lys
 1               5                  10                  15

Gln Trp Leu Cys Asp Gly Ser Asp Asp Cys Gly Asp Gly Ser Asp Glu
            20                  25                  30

Ala Ala His Cys
        35


<210> SEQ ID NO 111
<211> LENGTH: 39
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 111

Cys Gly Pro Ser Ser Phe Ser Cys Pro Gly Thr His Val Cys Val Pro
 1               5                  10                  15

Glu Arg Trp Leu Cys Asp Gly Asp Lys Asp Cys Ala Asp Gly Ala Asp
            20                  25                  30

Glu Ser Ile Ala Ala Gly Cys
        35


<210> SEQ ID NO 112
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 112

Cys Asp Asp Arg Glu Phe Met Cys Gln Asn Arg Gln Cys Ile Pro Lys
 1               5                  10                  15

His Phe Val Cys Asp His Asp Arg Asp Cys Ala Asp Gly Ser Asp Glu
            20                  25                  30

Ser Pro Glu Cys
        35


<210> SEQ ID NO 113
<211> LENGTH: 40
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 113

Cys Gly Pro Ser Glu Phe Arg Cys Ala Asn Gly Arg Cys Leu Ser Ser
 1               5                  10                  15

Arg Gln Trp Glu Cys Asp Gly Glu Asn Asp Cys His Asp Gln Ser Asp
            20                  25                  30
```

-continued

```
Glu Ala Pro Lys Asn Pro His Cys
        35                  40


<210> SEQ ID NO 114
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 114

Cys Asn Ala Ser Ser Gln Phe Leu Cys Ser Ser Gly Arg Cys Val Ala
 1               5                  10                  15

Glu Ala Leu Leu Cys Asn Gly Gln Asp Asp Cys Gly Asp Ser Ser Asp
            20                  25                  30

Glu Arg Gly Cys
        35


<210> SEQ ID NO 115
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 115

Cys Thr Ala Ser Gln Phe Val Cys Lys Asn Asp Lys Cys Ile Pro Phe
 1               5                  10                  15

Trp Trp Lys Cys Asp Thr Glu Asp Asp Cys Gly Asp His Ser Asp Glu
            20                  25                  30

Pro Pro Asp Cys
        35


<210> SEQ ID NO 116
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 116

Cys Arg Pro Gly Gln Phe Gln Cys Ser Thr Gly Ile Cys Thr Asn Pro
 1               5                  10                  15

Ala Phe Ile Cys Asp Gly Asp Asn Asp Cys Gln Asp Asn Ser Asp Glu
            20                  25                  30

Ala Asn Cys
        35


<210> SEQ ID NO 117
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 117

Cys Leu Pro Ser Gln Phe Lys Cys Thr Asn Thr Asn Arg Cys Ile Pro
```

```
 1               5                   10                  15

Gly Ile Phe Arg Cys Asn Gly Gln Asp Asn Cys Gly Asp Gly Glu Asp
            20                  25                  30

Glu Arg Asp Cys
        35


<210> SEQ ID NO 118
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 118

Cys Ala Pro Asn Gln Phe Gln Cys Ser Ile Thr Lys Arg Cys Ile Pro
 1               5                   10                  15

Arg Val Trp Val Cys Asp Arg Asp Asn Asp Cys Val Asp Gly Ser Asp
            20                  25                  30

Glu Pro Ala Asn Cys
        35


<210> SEQ ID NO 119
<211> LENGTH: 38
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 119

Cys Gly Val Asp Glu Phe Arg Cys Lys Asp Ser Gly Arg Cys Ile Pro
 1               5                   10                  15

Ala Arg Trp Lys Cys Asp Gly Glu Asp Asp Cys Gly Asp Gly Ser Asp
            20                  25                  30

Glu Pro Lys Glu Glu Cys
        35


<210> SEQ ID NO 120
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 120

Cys Glu Pro Tyr Gln Phe Arg Cys Lys Asn Asn Arg Cys Val Pro Gly
 1               5                   10                  15

Arg Trp Gln Cys Asp Tyr Asp Asn Asp Cys Gly Asp Asn Ser Asp Glu
            20                  25                  30

Glu Ser Cys
        35


<210> SEQ ID NO 121
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain
```

-continued

```
<400> SEQUENCE: 121

Cys Ser Glu Ser Glu Phe Ser Cys Ala Asn Gly Arg Cys Ile Ala Gly
 1               5                  10                  15

Arg Trp Lys Cys Asp Gly Asp His Asp Cys Ala Asp Gly Ser Asp Glu
            20                  25                  30

Lys Asp Cys
        35


<210> SEQ ID NO 122
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 122

Cys Asp Met Asp Gln Phe Gln Cys Lys Ser Gly His Cys Ile Pro Leu
 1               5                  10                  15

Arg Trp Arg Cys Asp Ala Asp Ala Asp Cys Met Asp Gly Ser Asp Glu
            20                  25                  30

Glu Ala Cys
        35


<210> SEQ ID NO 123
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 123

Cys Pro Leu Asp Glu Phe Gln Cys Asn Asn Thr Leu Cys Lys Pro Leu
 1               5                  10                  15

Ala Trp Lys Cys Asp Gly Glu Asp Asp Cys Gly Asp Asn Ser Asp Glu
            20                  25                  30

Asn Pro Glu Glu Cys
        35


<210> SEQ ID NO 124
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 124

Cys Pro Pro Asn Arg Pro Phe Arg Cys Lys Asn Asp Arg Val Cys Leu
 1               5                  10                  15

Trp Ile Gly Arg Gln Cys Asp Gly Thr Asp Asn Cys Gly Asp Gly Thr
            20                  25                  30

Asp Glu Glu Asp Cys
        35


<210> SEQ ID NO 125
<211> LENGTH: 36
<212> TYPE: PRT
```

-continued

```
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1 (LRP1) A domain

<400> SEQUENCE: 125

Cys Lys Asp Lys Lys Glu Phe Leu Cys Arg Asn Gln Arg Cys Leu Ser
 1               5                  10                  15

Ser Ser Leu Arg Cys Asn Met Phe Asp Asp Cys Gly Asp Gly Ser Asp
            20                  25                  30

Glu Glu Asp Cys
        35


<210> SEQ ID NO 126
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 126

Cys Asp Ser Ala His Phe Arg Cys Gly Ser Gly His Cys Ile Pro Ala
 1               5                  10                  15

Asp Trp Arg Cys Asp Gly Thr Lys Asp Cys Ser Asp Asp Ala Asp Glu
            20                  25                  30

Ile Gly Cys
        35


<210> SEQ ID NO 127
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 127

Cys Gln Gln Gly Tyr Phe Lys Cys Gln Ser Glu Gly Gln Cys Ile Pro
 1               5                  10                  15

Ser Ser Trp Val Cys Asp Gln Asp Gln Asp Cys Asp Asp Gly Ser Asp
            20                  25                  30

Glu Arg Gln Asp Cys
        35


<210> SEQ ID NO 128
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 128

Cys Ser Ser His Gln Ile Thr Cys Ser Asn Gly Gln Cys Ile Pro Ser
 1               5                  10                  15

Glu Tyr Arg Cys Asp His Val Arg Asp Cys Pro Asp Gly Ala Asp Glu
            20                  25                  30

Asn Asp Cys
        35
```

-continued

```
<210> SEQ ID NO 129
<211> LENGTH: 33
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 129

Cys Glu Gln Leu Thr Cys Asp Asn Gly Ala Cys Tyr Asn Thr Ser Gln
 1               5                   10                  15

Lys Cys Asp Trp Lys Val Asp Cys Arg Asp Ser Ser Asp Glu Ile Asn
             20                  25                  30

Cys


<210> SEQ ID NO 130
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 130

Cys Leu His Asn Glu Phe Ser Cys Gly Asn Gly Glu Cys Ile Pro Arg
 1               5                   10                  15

Ala Tyr Val Cys Asp His Asp Asn Asp Cys Gln Asp Gly Ser Asp Glu
             20                  25                  30

His Ala Cys
         35


<210> SEQ ID NO 131
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 131

Cys Gly Gly Tyr Gln Phe Thr Cys Pro Ser Gly Arg Cys Ile Tyr Gln
 1               5                   10                  15

Asn Trp Val Cys Asp Gly Glu Asp Asp Cys Lys Asp Asn Gly Asp Glu
             20                  25                  30

Asp Gly Cys
         35


<210> SEQ ID NO 132
<211> LENGTH: 42
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 132

Cys Ser Pro Arg Glu Trp Ser Cys Pro Glu Ser Gly Arg Cys Ile Ser
 1               5                   10                  15

Ile Tyr Lys Val Cys Asp Gly Ile Leu Asp Cys Pro Gly Arg Glu Asp
             20                  25                  30
```

-continued

```
Glu Asn Asn Thr Ser Thr Gly Lys Tyr Cys
        35                  40


<210> SEQ ID NO 133
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 133

Cys Gly Leu Phe Ser Phe Pro Cys Lys Asn Gly Arg Cys Val Pro Asn
 1               5                  10                  15

Tyr Tyr Leu Cys Asp Gly Val Asp Asp Cys His Asp Asn Ser Asp Glu
            20                  25                  30

Gln Leu Cys
        35


<210> SEQ ID NO 134
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 134

Cys Ser Ser Ser Ala Phe Thr Cys Gly His Gly Glu Cys Ile Pro Ala
 1               5                  10                  15

His Trp Arg Cys Asp Lys Arg Asn Asp Cys Val Asp Gly Ser Asp Glu
            20                  25                  30

His Asn Cys
        35


<210> SEQ ID NO 135
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 135

Cys Leu Asp Thr Gln Tyr Thr Cys Asp Asn His Gln Cys Ile Ser Lys
 1               5                  10                  15

Asn Trp Val Cys Asp Thr Asp Asn Asp Cys Gly Asp Gly Ser Asp Glu
            20                  25                  30

Lys Asn Cys
        35


<210> SEQ ID NO 136
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 136

Cys Gln Pro Ser Gln Phe Asn Cys Pro Asn His Arg Cys Ile Asp Leu
 1               5                  10                  15
```

-continued

```
Ser Phe Val Cys Asp Gly Asp Lys Asp Cys Val Asp Gly Ser Asp Glu
            20                  25                  30

Val Gly Cys
        35


<210> SEQ ID NO 137
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 137

Cys Thr Ala Ser Gln Phe Lys Cys Ala Ser Gly Asp Lys Cys Ile Gly
 1               5                  10                  15

Val Thr Asn Arg Cys Asp Gly Val Phe Asp Cys Ser Asp Asn Ser Asp
            20                  25                  30

Glu Ala Gly Cys
        35


<210> SEQ ID NO 138
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 138

Cys His Ser Asp Glu Phe Gln Cys Gln Glu Asp Gly Ile Cys Ile Pro
 1               5                  10                  15

Asn Phe Trp Glu Cys Asp Gly His Pro Asp Cys Leu Tyr Gly Ser Asp
            20                  25                  30

Glu His Asn Ala Cys
        35


<210> SEQ ID NO 139
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 139

Cys Pro Ser Ser Tyr Phe His Cys Asp Asn Gly Asn Cys Ile His Arg
 1               5                  10                  15

Ala Trp Leu Cys Asp Arg Asp Asn Asp Cys Gly Asp Met Ser Asp Glu
            20                  25                  30

Lys Asp Cys
        35


<210> SEQ ID NO 140
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain
```

```
<400> SEQUENCE: 140

Cys Pro Ser Trp Gln Trp Gln Cys Leu Gly His Asn Ile Cys Val Asn
 1              5                   10                  15

Leu Ser Val Val Cys Asp Gly Ile Phe Asp Cys Pro Asn Gly Thr Asp
            20                  25                  30

Glu Ser Pro Leu Cys
        35


<210> SEQ ID NO 141
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 141

Cys Gly Ala Ser Ser Phe Thr Cys Ser Asn Gly Arg Cys Ile Ser Glu
 1              5                   10                  15

Glu Trp Lys Cys Asp Asn Asp Asn Asp Cys Gly Asp Gly Ser Asp Glu
            20                  25                  30

Met Glu Ser Val Cys
        35


<210> SEQ ID NO 142
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 142

Cys Ser Pro Thr Ala Phe Thr Cys Ala Asn Gly Arg Cys Val Gln Tyr
 1              5                   10                  15

Ser Tyr Arg Cys Asp Tyr Tyr Asn Asp Cys Gly Asp Gly Ser Asp Glu
            20                  25                  30

Ala Gly Cys
        35


<210> SEQ ID NO 143
<211> LENGTH: 38
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 143

Cys Asn Ala Thr Thr Glu Phe Met Cys Asn Asn Arg Arg Cys Ile Pro
 1              5                   10                  15

Arg Glu Phe Ile Cys Asn Gly Val Asp Asn Cys His Asp Asn Asn Thr
            20                  25                  30

Ser Asp Glu Lys Asn Cys
        35


<210> SEQ ID NO 144
<211> LENGTH: 38
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
```

-continued

```
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 144

Cys Gln Ser Gly Tyr Thr Lys Cys His Asn Ser Asn Ile Cys Ile Pro
 1              5                  10                  15

Arg Val Tyr Leu Cys Asp Gly Asp Asn Asp Cys Gly Asp Asn Ser Asp
            20                  25                  30

Glu Asn Pro Thr Tyr Cys
        35


<210> SEQ ID NO 145
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 145

Cys Ser Ser Ser Glu Phe Gln Cys Ala Ser Gly Arg Cys Ile Pro Gln
 1              5                  10                  15

His Trp Tyr Cys Asp Gln Glu Thr Asp Cys Phe Asp Ala Ser Asp Glu
            20                  25                  30

Pro Ala Ser Cys
        35


<210> SEQ ID NO 146
<211> LENGTH: 38
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 146

Cys Leu Ala Asp Glu Phe Lys Cys Asp Gly Gly Arg Cys Ile Pro Ser
 1              5                  10                  15

Glu Trp Ile Cys Asp Gly Asp Asn Asp Cys Gly Asp Met Ser Asp Glu
            20                  25                  30

Asp Lys Arg His Gln Cys
        35


<210> SEQ ID NO 147
<211> LENGTH: 41
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 147

Cys Ser Asp Ser Glu Phe Leu Cys Val Asn Asp Arg Pro Pro Asp Arg
 1              5                  10                  15

Arg Cys Ile Pro Gln Ser Trp Val Cys Asp Gly Asp Val Asp Cys Thr
            20                  25                  30

Asp Gly Tyr Asp Glu Asn Gln Asn Cys
        35                  40
```

-continued

```
<210> SEQ ID NO 148
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 148

Cys Ser Glu Asn Glu Phe Thr Cys Gly Tyr Gly Leu Cys Ile Pro Lys
 1               5                  10                  15

Ile Phe Arg Cys Asp Arg His Asn Asp Cys Gly Asp Tyr Ser Asp Glu
            20                  25                  30

Arg Gly Cys
        35


<210> SEQ ID NO 149
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 149

Cys Gln Gln Asn Gln Phe Thr Cys Gln Asn Gly Arg Cys Ile Ser Lys
 1               5                  10                  15

Thr Phe Val Cys Asp Glu Asp Asn Asp Cys Gly Asp Gly Ser Asp Glu
            20                  25                  30

Leu Met His Leu Cys
        35


<210> SEQ ID NO 150
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 150

Cys Pro Pro His Glu Phe Lys Cys Asp Asn Gly Arg Cys Ile Glu Met
 1               5                  10                  15

Met Lys Leu Cys Asn His Leu Asp Asp Cys Leu Asp Asn Ser Asp Glu
            20                  25                  30

Lys Gly Cys
        35


<210> SEQ ID NO 151
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 151

Cys Ser Ser Thr Gln Phe Leu Cys Ala Asn Asn Glu Lys Cys Ile Pro
 1               5                  10                  15

Ile Trp Trp Lys Cys Asp Gly Gln Lys Asp Cys Ser Asp Gly Ser Asp
            20                  25                  30
```

-continued

```
Glu Leu Ala Leu Cys
        35


<210> SEQ ID NO 152
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 152

Cys Arg Leu Gly Gln Phe Gln Cys Ser Asp Gly Asn Cys Thr Ser Pro
 1               5                  10                  15

Gln Thr Leu Cys Asn Ala His Gln Asn Cys Pro Asp Gly Ser Asp Glu
            20                  25                  30

Asp Arg Leu Leu Cys
        35


<210> SEQ ID NO 153
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 153

Cys Asp Ser Asn Glu Trp Gln Cys Ala Asn Lys Arg Cys Ile Pro Glu
 1               5                  10                  15

Ser Trp Gln Cys Asp Thr Phe Asn Asp Cys Glu Asp Asn Ser Asp Glu
            20                  25                  30

Asp Ser Ser His Cys
        35


<210> SEQ ID NO 154
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 154

Cys Arg Pro Gly Gln Phe Arg Cys Ala Asn Gly Arg Cys Ile Pro Gln
 1               5                  10                  15

Ala Trp Lys Cys Asp Val Asp Asn Asp Cys Gly Asp His Ser Asp Glu
            20                  25                  30

Pro Ile Glu Glu Cys
        35


<210> SEQ ID NO 155
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 155

Cys Asp Asn Phe Thr Glu Phe Ser Cys Lys Thr Asn Tyr Arg Cys Ile
 1               5                  10                  15
```

-continued

```
Pro Lys Trp Ala Val Cys Asn Gly Val Asp Asp Cys Arg Asp Asn Ser
            20                  25                  30

Asp Glu Gln Gly Cys
        35


<210> SEQ ID NO 156
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 156

Cys His Pro Val Gly Asp Phe Arg Cys Lys Asn His His Cys Ile Pro
 1               5                   10                  15

Leu Arg Trp Gln Cys Asp Gly Gln Asn Asp Cys Gly Asp Asn Ser Asp
            20                  25                  30

Glu Glu Asn Cys
        35


<210> SEQ ID NO 157
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 157

Cys Thr Glu Ser Glu Phe Arg Cys Val Asn Gln Gln Cys Ile Pro Ser
 1               5                   10                  15

Arg Trp Ile Cys Asp His Tyr Asn Asp Cys Gly Asp Asn Ser Asp Glu
            20                  25                  30

Arg Asp Cys
        35


<210> SEQ ID NO 158
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 158

Cys His Pro Glu Tyr Phe Gln Cys Thr Ser Gly His Cys Val His Ser
 1               5                   10                  15

Glu Leu Lys Cys Asp Gly Ser Ala Asp Cys Leu Asp Ala Ser Asp Glu
            20                  25                  30

Ala Asp Cys
        35


<210> SEQ ID NO 159
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain
```

-continued

```
<400> SEQUENCE: 159

Cys Gln Ala Thr Met Phe Glu Cys Lys Asn His Val Cys Ile Pro Pro
 1              5                   10                  15

Tyr Trp Lys Cys Asp Gly Asp Asp Asp Cys Gly Asp Gly Ser Asp Glu
            20                  25                  30

Glu Leu His Leu Cys
        35


<210> SEQ ID NO 160
<211> LENGTH: 38
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 160

Cys Asn Ser Pro Asn Arg Phe Arg Cys Asp Asn Asn Arg Cys Ile Tyr
 1              5                   10                  15

Ser His Glu Val Cys Asn Gly Val Asp Asp Cys Gly Asp Gly Thr Asp
            20                  25                  30

Glu Thr Glu Glu His Cys
        35


<210> SEQ ID NO 161
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 2 (LRP2) A domain

<400> SEQUENCE: 161

Cys Thr Glu Tyr Glu Tyr Lys Cys Gly Asn Gly His Cys Ile Pro His
 1              5                   10                  15

Asp Asn Val Cys Asp Asp Ala Asp Asp Cys Gly Asp Trp Ser Asp Glu
            20                  25                  30

Leu Gly Cys
        35


<210> SEQ ID NO 162
<211> LENGTH: 38
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 162

Cys Asp Pro Gly Glu Phe Leu Cys His Asp His Val Thr Cys Val Ser
 1              5                   10                  15

Gln Ser Trp Leu Cys Asp Gly Asp Pro Asp Cys Pro Asp Asp Ser Asp
            20                  25                  30

Glu Ser Leu Asp Thr Cys
        35


<210> SEQ ID NO 163
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
```

-continued

```
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 163

Cys Pro Leu Asn His Ile Ala Cys Leu Gly Thr Asn Lys Cys Val His
 1               5                  10                  15

Leu Ser Gln Leu Cys Asn Gly Val Leu Asp Cys Pro Asp Gly Tyr Asp
            20                  25                  30

Glu Gly Val His Cys
        35


<210> SEQ ID NO 164
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 164

Cys Lys Ala Gly Glu Phe Arg Cys Lys Asn Arg His Cys Ile Gln Ala
 1               5                  10                  15

Arg Trp Lys Cys Asp Gly Asp Asp Asp Cys Leu Asp Gly Ser Asp Glu
            20                  25                  30

Asp Ser Val Asn Cys
        35


<210> SEQ ID NO 165
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 165

Cys Pro Asp Asp Gln Phe Lys Cys Gln Asn Asn Arg Cys Ile Pro Lys
 1               5                  10                  15

Arg Trp Leu Cys Asp Gly Ala Asn Asp Cys Gly Ser Asn Glu Asp Glu
            20                  25                  30

Ser Asn Gln Thr Cys
        35


<210> SEQ ID NO 166
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 166

Cys Gln Val Asp Gln Phe Ser Cys Gly Asn Gly Arg Cys Ile Pro Arg
 1               5                  10                  15

Ala Trp Leu Cys Asp Arg Glu Asp Asp Cys Gly Asp Gln Thr Asp Glu
            20                  25                  30

Met Ala Ser Cys
        35
```

-continued

```
<210> SEQ ID NO 167
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 167

Cys Glu Pro Leu Thr Gln Phe Val Cys Lys Ser Gly Arg Cys Ile Ser
 1               5                  10                  15

Ser Lys Trp His Cys Asp Ser Asp Asp Asp Cys Gly Asp Gly Ser Asp
            20                  25                  30

Glu Val Gly Cys
        35


<210> SEQ ID NO 168
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 168

Cys Phe Asp Asn Gln Phe Arg Cys Ser Ser Gly Arg Cys Ile Pro Gly
 1               5                  10                  15

His Trp Ala Cys Asp Gly Asp Asn Asp Cys Gly Asp Phe Ser Asp Glu
            20                  25                  30

Ala Gln Ile Asn Cys
        35


<210> SEQ ID NO 169
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 169

Cys Asn Gly Asn Glu Phe Gln Cys His Pro Asp Gly Asn Cys Val Pro
 1               5                  10                  15

Asp Leu Trp Arg Cys Asp Gly Glu Lys Asp Cys Glu Asp Gly Ser Asp
            20                  25                  30

Glu Lys Gly Cys
        35


<210> SEQ ID NO 170
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 170

Cys Asp His Lys Thr Lys Phe Ser Cys Trp Ser Thr Gly Arg Cys Ile
 1               5                  10                  15

Asn Lys Ala Trp Val Cys Asp Gly Asp Ile Asp Cys Glu Asp Gln Ser
            20                  25                  30
```

-continued

```
Asp Glu Asp Asp Cys
        35


<210> SEQ ID NO 171
<211> LENGTH: 38
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 171

Cys Gly Pro Pro Lys His Pro Cys Ala Asn Asp Thr Ser Val Cys Leu
 1               5                  10                  15

Gln Pro Glu Lys Leu Cys Asn Gly Lys Lys Asp Cys Pro Asp Gly Ser
            20                  25                  30

Asp Glu Gly Tyr Leu Cys
        35


<210> SEQ ID NO 172
<211> LENGTH: 38
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 172

Cys Asn Ala Tyr Ser Glu Phe Glu Cys Gly Asn Gly Glu Cys Ile Asp
 1               5                  10                  15

Tyr Gln Leu Thr Cys Asp Gly Ile Pro His Cys Lys Asp Lys Ser Asp
            20                  25                  30

Glu Lys Leu Leu Tyr Cys
        35


<210> SEQ ID NO 173
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 173

Cys Arg Arg Gly Phe Lys Pro Cys Tyr Asn Arg Arg Cys Ile Pro His
 1               5                  10                  15

Gly Lys Leu Cys Asp Gly Glu Asn Asp Cys Gly Asp Asn Ser Asp Glu
            20                  25                  30

Leu Asp Cys
        35


<210> SEQ ID NO 174
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 174

Cys Ala Thr Val Glu Phe Arg Cys Ala Asp Gly Thr Cys Ile Pro Arg
 1               5                  10                  15
```

```
Ser Ala Arg Cys Asn Gln Asn Ile Asp Cys Ala Asp Ala Ser Asp Glu
            20                  25                  30

Lys Asn Cys
        35


<210> SEQ ID NO 175
<211> LENGTH: 45
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 175

Cys Thr His Phe Tyr Lys Leu Gly Val Lys Thr Thr Gly Phe Ile Arg
 1               5                  10                  15

Cys Asn Ser Thr Ser Leu Cys Val Leu Pro Thr Trp Ile Cys Asp Gly
            20                  25                  30

Ser Asn Asp Cys Gly Asp Tyr Ser Asp Glu Leu Lys Cys
        35                  40                  45


<210> SEQ ID NO 176
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 176

Cys Glu Glu Asn Tyr Phe Ser Cys Pro Ser Gly Arg Cys Ile Leu Asn
 1               5                  10                  15

Thr Trp Ile Cys Asp Gly Gln Lys Asp Cys Glu Asp Gly Arg Asp Glu
            20                  25                  30

Phe His Cys
        35


<210> SEQ ID NO 177
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 177

Cys Ser Trp Asn Gln Phe Ala Cys Ser Ala Gln Lys Cys Ile Ser Lys
 1               5                  10                  15

His Trp Ile Cys Asp Gly Glu Asp Asp Cys Gly Asp Gly Leu Asp Glu
            20                  25                  30

Ser Asp Ser Ile Cys
        35


<210> SEQ ID NO 178
<211> LENGTH: 39
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain
```

-continued

```
<400> SEQUENCE: 178

Cys Ala Ala Asp Met Phe Ser Cys Gln Gly Ser Arg Ala Cys Val Pro
 1               5                  10                  15

Arg His Trp Leu Cys Asp Gly Glu Arg Asp Cys Pro Asp Gly Ser Asp
            20                  25                  30

Glu Leu Ser Thr Ala Gly Cys
        35


<210> SEQ ID NO 179
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 179

Cys Asp Glu Asn Ala Phe Met Cys His Asn Lys Val Cys Ile Pro Lys
 1               5                  10                  15

Gln Phe Val Cys Asp His Asp Asp Asp Cys Gly Asp Gly Ser Asp Glu
            20                  25                  30

Ser Pro Gln Cys
        35


<210> SEQ ID NO 180
<211> LENGTH: 40
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 180

Cys Gly Thr Glu Glu Phe Ser Cys Ala Asp Gly Arg Cys Leu Leu Asn
 1               5                  10                  15

Thr Gln Trp Gln Cys Asp Gly Asp Phe Asp Cys Pro Asp His Ser Asp
            20                  25                  30

Glu Ala Pro Leu Asn Pro Lys Cys
        35                  40


<210> SEQ ID NO 181
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 181

Cys Asn Ser Ser Phe Phe Met Cys Lys Asn Gly Arg Cys Ile Pro Ser
 1               5                  10                  15

Gly Gly Leu Cys Asp Asn Lys Asp Asp Cys Gly Asp Gly Ser Asp Glu
            20                  25                  30

Arg Asn Cys
        35


<210> SEQ ID NO 182
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
```

-continued

```
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 182

Cys Thr Ala Ser Gln Phe Arg Cys Lys Thr Asp Lys Cys Ile Pro Phe
 1               5                  10                  15

Trp Trp Lys Cys Asp Thr Val Asp Asp Cys Gly Asp Gly Ser Asp Glu
            20                  25                  30

Pro Asp Asp Cys
        35


<210> SEQ ID NO 183
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 183

Cys Gln Pro Gly Arg Phe Gln Cys Gly Thr Gly Leu Cys Ala Leu Pro
 1               5                  10                  15

Ala Phe Ile Cys Asp Gly Glu Asn Asp Cys Gly Asp Asn Ser Asp Glu
            20                  25                  30

Leu Asn Cys
        35


<210> SEQ ID NO 184
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 184

Cys Leu Ser Gly Gln Phe Lys Cys Thr Lys Asn Gln Lys Cys Ile Pro
 1               5                  10                  15

Val Asn Leu Arg Cys Asn Gly Gln Asp Asp Cys Gly Asp Glu Glu Asp
            20                  25                  30

Glu Arg Asp Cys
        35


<210> SEQ ID NO 185
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 185

Cys Ser Pro Asp Tyr Phe Gln Cys Lys Thr Thr Lys His Cys Ile Ser
 1               5                  10                  15

Lys Leu Trp Val Cys Asp Glu Asp Pro Asp Cys Ala Asp Ala Ser Asp
            20                  25                  30

Glu Ala Asn Cys
        35
```

-continued

```
<210> SEQ ID NO 186
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 186

Cys Gly Pro His Glu Phe Gln Cys Lys Asn Asn Asn Cys Ile Pro Asp
 1               5                  10                  15

His Trp Arg Cys Asp Ser Gln Asn Asp Cys Ser Asp Asn Ser Asp Glu
            20                  25                  30

Glu Asn Cys
        35


<210> SEQ ID NO 187
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 187

Cys Thr Leu Lys Asp Phe Leu Cys Ala Asn Gly Asp Cys Val Ser Ser
 1               5                  10                  15

Arg Phe Trp Cys Asp Gly Asp Phe Asp Cys Ala Asp Gly Ser Asp Glu
            20                  25                  30

Arg Asn Cys
        35


<210> SEQ ID NO 188
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 188

Cys Ser Lys Asp Gln Phe Arg Cys Ser Asn Gly Gln Cys Ile Pro Ala
 1               5                  10                  15

Lys Trp Lys Cys Asp Gly His Glu Asp Cys Lys Tyr Gly Glu Asp Glu
            20                  25                  30

Lys Ser Cys
        35


<210> SEQ ID NO 189
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 189

Cys Ser Ser Arg Glu Tyr Ile Cys Ala Ser Asp Gly Cys Ile Ser Ala
 1               5                  10                  15

Ser Leu Lys Cys Asn Gly Glu Tyr Asp Cys Ala Asp Gly Ser Asp Glu
            20                  25                  30
```

-continued

```
Met Asp Cys
        35


<210> SEQ ID NO 190
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 190

Cys Lys Glu Asp Gln Phe Arg Cys Lys Asn Lys Ala His Cys Ile Pro
 1               5                  10                  15

Ile Arg Trp Leu Cys Asp Gly Ile His Asp Cys Val Asp Gly Ser Asp
            20                  25                  30

Glu Glu Asn Cys
        35


<210> SEQ ID NO 191
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 191

Cys Arg Ala Asp Glu Phe Leu Cys Asn Asn Ser Leu Cys Lys Leu His
 1               5                  10                  15

Phe Trp Val Cys Asp Gly Glu Asp Asp Cys Gly Asp Asn Ser Asp Glu
            20                  25                  30

Ala Pro Asp Met Cys
        35


<210> SEQ ID NO 192
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 192

Cys Pro Ser Thr Arg Pro His Arg Cys Arg Asn Asn Arg Ile Cys Leu
 1               5                  10                  15

Gln Ser Glu Gln Met Cys Asn Gly Ile Asp Glu Cys Gly Asp Asn Ser
            20                  25                  30

Asp Glu Asp His Cys
        35


<210> SEQ ID NO 193
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: human low-density lipoprotein receptor (LDLR)
      related protein 1B (LR1B) A domain

<400> SEQUENCE: 193

Cys Lys Lys Asp Glu Phe Ala Cys Ser Asn Lys Lys Cys Ile Pro Met
 1               5                  10                  15
```

```
Asp Leu Gln Cys Asp Arg Leu Asp Asp Cys Gly Asp Gly Ser Asp Glu
            20                  25                  30

Gln Gly Cys
        35


<210> SEQ ID NO 194
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: O75851 A domain

<400> SEQUENCE: 194

Cys Ala Glu Gly Glu Ala Leu Cys Gln Glu Asn Gly His Cys Val Pro
 1               5                  10                  15

His Gly Trp Leu Cys Asp Asn Gln Asp Asp Cys Gly Asp Gly Ser Asp
            20                  25                  30

Glu Glu Gly Glu Cys
        35


<210> SEQ ID NO 195
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: O75851 A domain

<400> SEQUENCE: 195

Cys Gly Glu Gly Gln Met Thr Cys Ser Ser Gly His Cys Leu Pro Leu
 1               5                  10                  15

Ala Leu Leu Cys Asp Arg Gln Asp Asp Cys Gly Asp Gly Thr Asp Glu
            20                  25                  30

Pro Ser Tyr Pro Cys
        35


<210> SEQ ID NO 196
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: O75851 A domain

<400> SEQUENCE: 196

Cys Pro Gln Gly Leu Leu Ala Cys Ala Asp Gly Arg Cys Leu Pro Pro
 1               5                  10                  15

Ala Leu Leu Cys Asp Gly His Pro Asp Cys Leu Asp Ala Ala Asp Glu
            20                  25                  30

Glu Ser Cys
        35


<210> SEQ ID NO 197
<211> LENGTH: 37
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: O75851 A domain

<400> SEQUENCE: 197

Cys Val Pro Gly Glu Val Ser Cys Val Asp Gly Thr Cys Leu Gly Ala
 1               5                  10                  15
```

-continued

```
Ile Gln Leu Cys Asp Gly Val Trp Asp Cys Pro Asp Gly Ala Asp Glu
            20                  25                  30

Gly Pro Gly His Cys
        35


<210> SEQ ID NO 198
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: O75851 A domain

<400> SEQUENCE: 198

Cys Pro Gly Leu Phe Pro Cys Gly Val Ala Pro Gly Leu Cys Leu Thr
 1               5                  10                  15

Pro Glu Gln Leu Cys Asp Gly Ile Pro Asp Cys Pro Gln Gly Glu Asp
            20                  25                  30

Glu Leu Asp Cys
        35


<210> SEQ ID NO 199
<211> LENGTH: 39
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: O75851 A domain

<400> SEQUENCE: 199

Cys Pro Glu Tyr Thr Cys Pro Asn Gly Thr Cys Ile Gly Phe Gln Leu
 1               5                  10                  15

Val Cys Asp Gly Gln Pro Asp Cys Gly Arg Pro Gly Gln Val Gly Pro
            20                  25                  30

Ser Pro Glu Glu Gln Gly Cys
        35


<210> SEQ ID NO 200
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: O75851 A domain

<400> SEQUENCE: 200

Cys Ser Pro Ser Gln Leu Ser Cys Gly Ser Gly Glu Cys Leu Ser Ala
 1               5                  10                  15

Glu Arg Arg Cys Asp Leu Arg Pro Asp Cys Gln Asp Gly Ser Asp Glu
            20                  25                  30

Asp Gly Cys
        35


<210> SEQ ID NO 201
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: O75851 A domain

<400> SEQUENCE: 201

Cys Glu Pro Gly Val Gly Leu Arg Cys Ala Ser Gly Glu Cys Val Leu
 1               5                  10                  15
```

-continued

```
Arg Gly Gly Pro Cys Asp Gly Val Leu Asp Cys Glu Asp Gly Ser Asp
            20                  25                  30

Glu Glu Gly Cys
        35


<210> SEQ ID NO 202
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: O75851 = ENSP00000262089 A domain

<400> SEQUENCE: 202

Cys Gly Pro Phe Glu Phe Arg Cys Gly Ser Gly Glu Cys Thr Pro Arg
 1               5                  10                  15

Gly Trp Arg Cys Asp Gln Glu Glu Asp Cys Ala Asp Gly Ser Asp Glu
            20                  25                  30

Arg Gly Cys
        35


<210> SEQ ID NO 203
<211> LENGTH: 38
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: ENSP00000262089 A domain

<400> SEQUENCE: 203

Cys Ala Pro His His Ala Pro Cys Ala Arg Gly Pro His Cys Val Ser
 1               5                  10                  15

Pro Glu Gln Leu Cys Asp Gly Val Arg Gln Cys Pro Asp Gly Ser Asp
            20                  25                  30

Glu Gly Pro Asp Ala Cys
        35


<210> SEQ ID NO 204
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: ENSP00000262089 A domain

<400> SEQUENCE: 204

Cys Gly Pro Gly Gln Thr Pro Cys Glu Val Leu Gly Cys Val Glu Gln
 1               5                  10                  15

Ala Gln Val Cys Asp Gly Arg Glu Asp Cys Leu Asp Gly Ser Asp Glu
            20                  25                  30

Arg His Cys
        35


<210> SEQ ID NO 205
<211> LENGTH: 33
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: C18oRF1 A domain

<400> SEQUENCE: 205

Cys Lys Phe Thr Cys Thr Ser Gly Lys Cys Leu Tyr Leu Gly Ser Leu
 1               5                  10                  15
```

-continued

```
Val Cys Asn Gln Gln Asn Asp Cys Gly Asp Asn Ser Asp Glu Glu Asn
            20                  25                  30

Cys


<210> SEQ ID NO 206
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: AAH07083 A domain

<400> SEQUENCE: 206

Cys Pro Pro Thr Lys Phe Gln Cys Arg Thr Ser Gly Leu Cys Val Pro
 1               5                  10                  15

Leu Thr Trp Arg Cys Asp Arg Asp Leu Asp Cys Ser Asp Gly Ser Asp
            20                  25                  30

Glu Glu Glu Cys
        35


<210> SEQ ID NO 207
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: AAH07083 A domain

<400> SEQUENCE: 207

Cys Leu Ala Gly Glu Leu Arg Cys Thr Leu Ser Asp Asp Cys Ile Pro
 1               5                  10                  15

Leu Thr Trp Arg Cys Asp Gly His Pro Asp Cys Pro Asp Ser Ser Asp
            20                  25                  30

Glu Leu Gly Cys
        35


<210> SEQ ID NO 208
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Q9HBX9 A domain

<400> SEQUENCE: 208

Cys Ser Leu Gly Tyr Phe Pro Cys Gly Asn Ile Thr Lys Cys Leu Pro
 1               5                  10                  15

Gln Leu Leu His Cys Asn Gly Val Asp Asp Cys Gly Asn Gln Ala Asp
            20                  25                  30

Glu Asp Asn Cys
        35


<210> SEQ ID NO 209
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Q9BY79 A domain

<400> SEQUENCE: 209

Cys Ala His Asp Glu Phe Arg Cys Asp Gln Leu Ile Cys Leu Leu Pro
 1               5                  10                  15
```

-continued

Asp Ser Val Cys Asp Gly Phe Ala Asn Cys Ala Asp Gly Ser Asp Glu
            20   25   30

Thr Asn Cys
        35

<210> SEQ ID NO 210
<211> LENGTH: 34
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Q9BY79 A domain

<400> SEQUENCE: 210

Cys Gly Pro Ser Glu Leu Ser Cys Gln Ala Gly Gly Cys Lys Gly Val
1   5   10   15

Gln Trp Met Cys Asp Met Trp Arg Asp Cys Thr Asp Gly Ser Asp Asp
            20   25   30

Asn Cys

<210> SEQ ID NO 211
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: BAB55257 = ENSP00000239367 A domain

<400> SEQUENCE: 211

Cys Ser Arg Tyr His Phe Phe Cys Asp Asp Gly Cys Cys Ile Asp Ile
1   5   10   15

Thr Leu Ala Cys Asp Gly Val Gln Gln Cys Pro Asp Gly Ser Asp Glu
            20   25   30

Asp Phe Cys
        35

<210> SEQ ID NO 212
<211> LENGTH: 32
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: O95518 = ENSP00000255793 A domain

<400> SEQUENCE: 212

Cys Pro Gly Glu Phe Leu Cys Ser Val Asn Gly Leu Cys Val Pro Ala
1   5   10   15

Cys Asp Gly Val Lys Asp Cys Pro Asn Gly Leu Asp Glu Arg Asn Cys
            20   25   30

<210> SEQ ID NO 213
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: ENSP00000255793 A domain

<400> SEQUENCE: 213

Cys Arg Ala Thr Phe Gln Cys Lys Glu Asp Ser Thr Cys Ile Ser Leu
1   5   10   15

Pro Lys Val Cys Asp Gly Gln Pro Asp Cys Leu Asn Gly Ser Asp Glu
            20   25   30

Glu Gln Cys

-continued

```
        35


<210> SEQ ID NO 214
<211> LENGTH: 36
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: ENSP00000255793 A domain

<400> SEQUENCE: 214

Cys Gly Thr Phe Thr Phe Gln Cys Glu Asp Arg Ser Cys Val Lys Lys
 1               5                  10                  15

Pro Asn Pro Gln Cys Asp Gly Arg Pro Asp Cys Arg Asp Gly Ser Asp
            20                  25                  30

Glu Glu His Cys
        35


<210> SEQ ID NO 215
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: beta-Propeller domain repeat sequence

<400> SEQUENCE: 215

Tyr Trp Thr Asp
 1


<210> SEQ ID NO 216
<211> LENGTH: 27
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: consensus sequence of portion of A domain
      beginning at third Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (2)...(2)
<223> OTHER INFORMATION: Xaa = Val, Ile, Leu, Met or Ala
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (3)...(7)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (9)...(9)
<223> OTHER INFORMATION: Xaa = Asp, Asn or His
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (10)...(12)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (13)...(13)
<223> OTHER INFORMATION: Xaa = Asp, Glu, Asn, Gln, His or Thr
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (15)...(18)
<223> OTHER INFORMATION: Xaa = any amino acid, Xaa at position 18 may be
      present or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (19)...(19)
<223> OTHER INFORMATION: Xaa = Ser, Thr, Ala, Asp or Glu
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (20)...(20)
<223> OTHER INFORMATION: Xaa = Asp, Glu or His
<220> FEATURE:
```

-continued

```
<221> NAME/KEY: MOD_RES
<222> LOCATION: (21)...(21)
<223> OTHER INFORMATION: Xaa = Asp or Glu
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (22)...(26)
<223> OTHER INFORMATION: Xaa = any amino acid, Xaa at positions 23-26
      may be present or absent

<400> SEQUENCE: 216

Cys Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa
 1               5                  10                  15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys
            20                  25


<210> SEQ ID NO 217
<211> LENGTH: 64
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: consensus sequence of A domain spanning all six
      Cys residues
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (2)...(16)
<223> OTHER INFORMATION: Xaa = any amino acid, Xaa at positions 5-16 may
      be present or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (18)...(32)
<223> OTHER INFORMATION: Xaa = any amino acid, Xaa at positions 22-32
      may be present or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (34)...(40)
<223> OTHER INFORMATION: Xaa = any amino acid, Xaa at position 40 may be
      present or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (43)...(45)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (46)...(46)
<223> OTHER INFORMATION: Xaa = Asp, Glu, Asn, Gln, His, Ser or Thr
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (48)...(53)
<223> OTHER INFORMATION: Xaa = any amino acid, Xaa at positions 52-53
      may be present or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (56)...(63)
<223> OTHER INFORMATION: Xaa = any amino acid, Xaa at positions 58-63
      may be present or absent

<400> SEQUENCE: 217

Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1               5                  10                  15

Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
            20                  25                  30

Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Asx Xaa Xaa Xaa Xaa Cys Xaa
        35                  40                  45

Xaa Xaa Xaa Xaa Xaa Asp Glu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys
    50                  55                  60


<210> SEQ ID NO 218
<211> LENGTH: 123
```

-continued

<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: exemplary C2 domain

<400> SEQUENCE: 218

Tyr Ser His Lys Phe Thr Val Val Val Leu Arg Ala Thr Lys Val Thr
1 5 10 15

Lys Gly Ala Phe Gly Asp Met Leu Asp Thr Pro Asp Pro Tyr Val Glu
20 25 30

Leu Phe Ile Ser Thr Thr Pro Asp Ser Arg Lys Arg Thr Arg His Phe
35 40 45

Asn Asn Asp Ile Asn Pro Val Trp Asn Glu Thr Phe Glu Phe Ile Leu
50 55 60

Asp Pro Asn Gln Glu Asn Val Leu Glu Ile Thr Leu Met Asp Ala Asn
65 70 75 80

Tyr Val Met Asp Glu Thr Leu Gly Thr Ala Thr Phe Thr Val Ser Ser
85 90 95

Met Lys Val Gly Glu Lys Lys Glu Val Pro Phe Ile Phe Asn Gln Val
100 105 110

Thr Glu Met Val Leu Glu Met Ser Leu Glu Val
115 120

<210> SEQ ID NO 219
<211> LENGTH: 5
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: peptide linker repeat

<400> SEQUENCE: 219

Gly Gly Gly Gly Ser
1 5

<210> SEQ ID NO 220
<211> LENGTH: 15
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: 15mer peptide linker

<400> SEQUENCE: 220

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
1 5 10 15

<210> SEQ ID NO 221
<211> LENGTH: 5
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: simple peptide linker
<220> FEATURE:
<221> NAME/KEY: REPEAT
<222> LOCATION: (1)...(5)
<223> OTHER INFORMATION: amino acids Gly-4Ser may be repeated an
unspecified number of times

<400> SEQUENCE: 221

Gly Gly Gly Gly Ser
1 5

-continued

```
<210> SEQ ID NO 222
<211> LENGTH: 25
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: flexible peptide linker
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (2)...(25)
<223> OTHER INFORMATION: Gly at positions 2-25 may be present or absent

<400> SEQUENCE: 222

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
 1               5                  10                  15

Gly Gly Gly Gly Gly Gly Gly Gly Gly
            20                  25


<210> SEQ ID NO 223
<211> LENGTH: 20
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: flexible peptide linker
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (6)...(20)
<223> OTHER INFORMATION: Gly at positions 6-20 may be present or absent

<400> SEQUENCE: 223

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
 1               5                  10                  15

Gly Gly Gly Gly
            20


<210> SEQ ID NO 224
<211> LENGTH: 15
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: flexible peptide linker
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (6)...(15)
<223> OTHER INFORMATION: Gly at positions 6-15 may be present or absent

<400> SEQUENCE: 224

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
 1               5                  10                  15


<210> SEQ ID NO 225
<211> LENGTH: 12
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: flexible peptide linker
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (9)...(12)
<223> OTHER INFORMATION: Gly at positions 9-12 may be present or absent

<400> SEQUENCE: 225

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
 1               5                  10


<210> SEQ ID NO 226
<211> LENGTH: 17
```

-continued

```
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: peptide linker
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (2)...(5)
<223> OTHER INFORMATION: Gly at positions 2-5 may be present or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (6)...(6)
<223> OTHER INFORMATION: Xaa = Ala, Val, Leu, Ile, Met, Phe, Trp, Pro,
      Gly, Ser, Thr, Cys, Tyr, Asn, Gln, Lys, Arg, His, Asp or Glu
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (8)...(11)
<223> OTHER INFORMATION: Gly at positions 8-11 may be present or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (12)...(12)
<223> OTHER INFORMATION: Xaa = Ala, Val, Leu, Ile, Met, Phe, Trp, Pro,
      Gly, Ser, Thr, Cys, Tyr, Asn, Gln, Lys, Arg, His, Asp or Glu
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (14)...(17)
<223> OTHER INFORMATION: Gly at positions 14-17 may be present or absent

<400> SEQUENCE: 226

Gly Gly Gly Gly Gly Xaa Gly Gly Gly Gly Gly Xaa Gly Gly Gly Gly
 1               5                  10                  15

Gly


<210> SEQ ID NO 227
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: peptide linker
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (2)...(5)
<223> OTHER INFORMATION: Gly at positions 2-5 may be present or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (6)...(6)
<223> OTHER INFORMATION: Xaa = Ser, Ala or Thr
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (8)...(11)
<223> OTHER INFORMATION: Gly at positions 8-11 may be present or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (12)...(12)
<223> OTHER INFORMATION: Xaa = Ser, Ala or Thr
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (14)...(17)
<223> OTHER INFORMATION: Gly at positions 14-17 may be present or absent

<400> SEQUENCE: 227

Gly Gly Gly Gly Gly Xaa Gly Gly Gly Gly Gly Xaa Gly Gly Gly Gly
 1               5                  10                  15

Gly


<210> SEQ ID NO 228
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: peptide linker
```

-continued

```
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (2)...(5)
<223> OTHER INFORMATION: Gly at positions 2-5 may be present or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (8)...(11)
<223> OTHER INFORMATION: Gly at positions 8-11 may be present or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (14)...(17)
<223> OTHER INFORMATION: Gly at positions 14-17 may be present or absent

<400> SEQUENCE: 228

Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly
 1              5                   10                  15

Gly


<210> SEQ ID NO 229
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: peptide linker
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (4)...(4)
<223> OTHER INFORMATION: Xaa = Ala, Val, Leu, Ile, Met, Phe, Trp, Pro,
      Gly, Ser, Thr, Cys, Tyr, Asn, Gln, Lys, Arg, His, Asp or Glu
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (8)...(8)
<223> OTHER INFORMATION: Xaa = Ala, Val, Leu, Ile, Met, Phe, Trp, Pro,
      Gly, Ser, Thr, Cys, Tyr, Asn, Gln, Lys, Arg, His, Asp or Glu

<400> SEQUENCE: 229

Gly Gly Gly Xaa Gly Gly Gly Xaa Gly Gly Gly
 1              5                   10


<210> SEQ ID NO 230
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: peptide linker
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (4)...(4)
<223> OTHER INFORMATION: Xaa = Ser, Ala or Thr
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (8)...(8)
<223> OTHER INFORMATION: Xaa = Ser, Ala or Thr

<400> SEQUENCE: 230

Gly Gly Gly Xaa Gly Gly Gly Xaa Gly Gly Gly
 1              5                   10


<210> SEQ ID NO 231
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: peptide linker

<400> SEQUENCE: 231

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly
 1              5                   10
```

-continued

```
<210> SEQ ID NO 232
<211> LENGTH: 25
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: peptide linker
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (2)...(12)
<223> OTHER INFORMATION: Gly at positions 2-12 may be present or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (15)...(25)
<223> OTHER INFORMATION: Gly at positions 15-25 may be present or absent

<400> SEQUENCE: 232

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Cys Gly Gly Gly
 1               5                  10                  15

Gly Gly Gly Gly Gly Gly Gly Gly Gly
            20                  25


<210> SEQ ID NO 233
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: peptide linker

<400> SEQUENCE: 233

Gly Gly Gly Gly Gly Cys Gly Gly Gly Gly Gly
 1               5                  10


<210> SEQ ID NO 234
<211> LENGTH: 25
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: rigid proline-containing linker
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (2)...(12)
<223> OTHER INFORMATION: Pro at positions 2-12 may be present or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (15)...(25)
<223> OTHER INFORMATION: Pro at positions 15-25 may be present or absent

<400> SEQUENCE: 234

Pro Pro Pro Pro Pro Pro Pro Pro Pro Pro Pro Pro Cys Pro Pro Pro
 1               5                  10                  15

Pro Pro Pro Pro Pro Pro Pro Pro Pro
            20                  25


<210> SEQ ID NO 235
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: rigid proline-containing linker

<400> SEQUENCE: 235

Pro Pro Pro Pro Pro Cys Pro Pro Pro Pro Pro
 1               5                  10
```

-continued

```
<210> SEQ ID NO 236
<211> LENGTH: 19
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: peptide linker comprising N-glycosylation site
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (2)...(8)
<223> OTHER INFORMATION: Gly at positions 2-8 may be present or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (10)...(10)
<223> OTHER INFORMATION: Xaa = any amino acid except Pro
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (13)...(19)
<223> OTHER INFORMATION: Gly at positions 13-19 may be present or absent

<400> SEQUENCE: 236

Gly Gly Gly Gly Gly Gly Gly Gly Asn Xaa Thr Gly Gly Gly Gly Gly
 1               5                  10                  15

Gly Gly Gly


<210> SEQ ID NO 237
<211> LENGTH: 50
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric C2 domain assembly PCR
      oligonucleotide

<400> SEQUENCE: 237

acactgcaat cgcgccttac ggctcccggg cggatcctcc cataagttca                50


<210> SEQ ID NO 238
<211> LENGTH: 72
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric C2 domain assembly PCR
      oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (1)...(72)
<223> OTHER INFORMATION: n = g, a, c or t

<400> SEQUENCE: 238

agctaccaaa gtgacannkn nknnknnknn knnknnknnk nnknnknnkn nkccatacgt     60

cgaattgttc at                                                         72


<210> SEQ ID NO 239
<211> LENGTH: 72
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric C2 domain assembly PCR
      oligonucleotide

<400> SEQUENCE: 239

agctaccaaa gtgacaaaag gtgcttttgg tgatatgttg gatactccag atccatacgt     60

cgaattgttc at                                                         72


<210> SEQ ID NO 240
<211> LENGTH: 62
```

-continued

```
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric C2 domain assembly PCR
      oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (1)...(62)
<223> OTHER INFORMATION: n = g, a, c or t

<400> SEQUENCE: 240

taggaagaga acacgtcatt ttnnknnknn kattaaccct gtttggaacg agacctttga     60

gt                                                                    62


<210> SEQ ID NO 241
<211> LENGTH: 62
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric C2 domain assembly PCR
      oligonucleotide

<400> SEQUENCE: 241

taggaagaga acacgtcatt ttaataatga tattaaccct gtttggaacg agacctttga     60

gt                                                                    62


<210> SEQ ID NO 242
<211> LENGTH: 58
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric C2 domain assembly PCR
      oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (1)...(58)
<223> OTHER INFORMATION: n = g, a, c or t

<400> SEQUENCE: 242

ttggaaatca ccctaatgnn knnknnknnk nnknnknnkn nkactctagg tacagcaa       58


<210> SEQ ID NO 243
<211> LENGTH: 58
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric C2 domain assembly PCR
      oligonucleotide

<400> SEQUENCE: 243

ttggaaatca ccctaatgga tgcaaattat gttatggacg aaactctagg tacagcaa       58


<210> SEQ ID NO 244
<211> LENGTH: 60
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric C2 domain assembly PCR
      oligonucleotide

<400> SEQUENCE: 244

aagaaggaag tcccatttat tttcaatcaa gttactgaaa tggtcttaga gatgtccctt     60


<210> SEQ ID NO 245
```

-continued

```
<211> LENGTH: 48
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric C2 domain assembly PCR
      oligonucleotide

<400> SEQUENCE: 245

tgtcactttg gtagctctta acacaactac agtgaactta tgggagga                  48


<210> SEQ ID NO 246
<211> LENGTH: 51
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric C2 domain assembly PCR
      oligonucleotide

<400> SEQUENCE: 246

acgtgttctc ttcctagaat ctggagttgt actgatgaac aattcgacgt a              51


<210> SEQ ID NO 247
<211> LENGTH: 62
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric C2 domain assembly PCR
      oligonucleotide

<400> SEQUENCE: 247

attagggtga tttccaaaac attttcttga ttaggatcta atataaactc aaaggtctcg     60

tt                                                                    62


<210> SEQ ID NO 248
<211> LENGTH: 64
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric C2 domain assembly PCR
      oligonucleotide

<400> SEQUENCE: 248

atgggacttc cttcttttct cccactttca ttgaagatac agtaaacgtt gctgtaccta     60

gagt                                                                  64


<210> SEQ ID NO 249
<211> LENGTH: 67
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric C2 domain assembly PCR
      oligonucleotide

<400> SEQUENCE: 249

gaccgatagc ttgccgattg cagtgtggcc acagaggcct cgagaacttc aagggacatc     60

tctaaga                                                               67


<210> SEQ ID NO 250
<211> LENGTH: 56
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: C2 domain amplification PCR oligonucleotide
```

-continued

```
<400> SEQUENCE: 250

acactgcaat cgcgccttac ggctcaggtg ctggtggttc ccataagttc actgta        56


<210> SEQ ID NO 251
<211> LENGTH: 81
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: C2 domain amplification PCR oligonucleotide

<400> SEQUENCE: 251

gaccgatagc ttgccgattg cagtcagcac ctgaaccacc accaccagaa ccaccaccac     60

caacttcaag ggacatctct a                                              81


<210> SEQ ID NO 252
<211> LENGTH: 227
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: C2 domain DNA ligation multimerization stop
      fragment Stop1

<400> SEQUENCE: 252

gaattcaacg ctactaccat tagtagaatt gatgccacct tttcagctcg cgccccaaat     60

gaaaaaatgg tcaaactaaa tctactcgtt cgcagaattg ggaatcaact gttacatgga    120

atgaaacttc cagacaccgt actttatgaa tatttatgac gattccgagg cgcgcccgga    180

ctacccgtat gatgttccgg attatgcccc gggatcctca ggtgctg                  227


<210> SEQ ID NO 253
<211> LENGTH: 173
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: C2 domain DNA ligation multimerization stop
      fragment Stop2

<400> SEQUENCE: 253

caggtgctgc actcgaggcc actgcggccg catattaacg tagatttttc ctcccaacgt     60

cctgactggt ataatgagcc agttcttaaa atcgcataac cagtacatgg tgattaaagt    120

tgaaattaaa ccgtctcaag agctttgtta cgttgatttg ggtaatgaag ctt           173


<210> SEQ ID NO 254
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Stop1-C2-C2-Stop2 fragment PCR amplification
      primer

<400> SEQUENCE: 254

aattcaacgc tactaccat                                                 19


<210> SEQ ID NO 255
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Stop1-C2-C2-Stop2 fragment PCR amplification
      primer
```

-continued

<400> SEQUENCE: 255

agcttcatta cccaaatcaa c 21

<210> SEQ ID NO 256
<211> LENGTH: 81
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric A domain assembly PCR oligonucleotide

<400> SEQUENCE: 256

cactatgcat ggactcagtg tgtccgataa gggcacacgg tgcctacccg tatgatgttc 60

cggattatgc cccgggcagt a 81

<210> SEQ ID NO 257
<211> LENGTH: 84
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric A domain assembly PCR oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (1)...(84)
<223> OTHER INFORMATION: n = g, a, c or t

<400> SEQUENCE: 257

cgccgtcgca tmscmagykc nsagraatac awyggccgyt wyygcacbka aattsgyyag 60

vcnsacaggt actgcccggg gcat 84

<210> SEQ ID NO 258
<211> LENGTH: 84
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric A domain assembly PCR oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (1)...(84)
<223> OTHER INFORMATION: n = g, a, c or t

<400> SEQUENCE: 258

cgccgtcgca tmscmatkcc nsagraatac awyggccgyt wyygcacbka aattsgyyag 60

vcnsacaggt actgcccggg gcat 84

<210> SEQ ID NO 259
<211> LENGTH: 79
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric A domain assembly PCR oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (1)...(79)
<223> OTHER INFORMATION: n = g, a, c or t

<400> SEQUENCE: 259

atgcgacggc gwwratgatt gtsvagatgg tagcgatgaa vwgrrttgtv mavnmvnmvg 60

ccvtacgggc tcggcctct 79

<210> SEQ ID NO 260
<211> LENGTH: 79

```
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric A domain assembly PCR oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (1)...(79)
<223> OTHER INFORMATION: n = g, a, c or t

<400> SEQUENCE: 260

atgcgacggc gwwccggatt gtsvagatgg tagcgatgaa vwgrrttgtv mavnmvnmvg     60

ccvtacgggc tcggcctct                                                   79


<210> SEQ ID NO 261
<211> LENGTH: 79
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric A domain assembly PCR oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (1)...(79)
<223> OTHER INFORMATION: n = g, a, c or t

<400> SEQUENCE: 261

atgcgacggc gwwratgatt gtsvagataa cagcgatgaa vwgrrttgtv mavnmvnmvg     60

ccvtacgggc tcggcctct                                                   79


<210> SEQ ID NO 262
<211> LENGTH: 79
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric A domain assembly PCR oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (1)...(79)
<223> OTHER INFORMATION: n = g, a, c or t

<400> SEQUENCE: 262

atgcgacggc gwwccggatt gtsvagataa cagcgatgaa vwgrrttgtv mavnmvnmvg     60

ccvtacgggc tcggcctct                                                   79


<210> SEQ ID NO 263
<211> LENGTH: 81
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: monomeric A domain assembly PCR oligonucleotide

<400> SEQUENCE: 263

tcctggtagt acttatctac tactatttgt ctgtgtctgc tctgggttcc taacggttcg     60

gccacagagg ccgagcccgt a                                                81


<210> SEQ ID NO 264
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR oligonucleotide

<400> SEQUENCE: 264

aagcctcagc gaccgaa                                                     17
```

-continued

```
<210> SEQ ID NO 265
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR oligonucleotide

<400> SEQUENCE: 265

agcccaatag gaacccat                                                   18


<210> SEQ ID NO 266
<211> LENGTH: 228
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: A domain DNA ligation multimerization stop
      fragment Stop1

<400> SEQUENCE: 266

gaattcaacg ctactaccat tagtagaatt gatgccacct tttcagctcg cgccccaaat     60

gaaaaaatgg tcaaactaaa tctactcgtt cgcagaattg ggaatcaact gttacatgga    120

atgaaacttc cagacaccgt actttatgaa tatttatgac gattccgagg cgcgcccgga    180

ctacccgtat gatgttccgg attatgcccc gggcggatcc agtacctg                 228


<210> SEQ ID NO 267
<211> LENGTH: 176
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: A domain DNA ligation multimerization stop
      fragment Stop2

<400> SEQUENCE: 267

gccctacggg cctcgaggca cctggtgcgg ccgcatatta acgtagattt ttcctcccaa     60

cgtcctgact ggtataatga gccagttctt aaaatcgcat aaccagtaca tggtgattaa    120

agttgaaatt aaaccgtctc aagagctttg ttacgttgat ttgggtaatg aagctt        176


<210> SEQ ID NO 268
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Stop1-A-A-A-Stop2 fragment PCR amplification
      primer

<400> SEQUENCE: 268

agcttcatta cccaaatcaa c                                               21


<210> SEQ ID NO 269
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Stop1-A-A-A-Stop2 fragment PCR amplification
      primer

<400> SEQUENCE: 269

aattcaacgc tactaccat                                                  19
```

1-92. (canceled)

93. A product comprising at least two monomer domains, wherein at least one monomer domain is a non-naturally occurring monomer domain and the monomer domains bind calcium.

94. The product of claim 93, wherein more than one of the monomer domains is a non-naturally-occurring monomer domain.

95. The product of claim 93, wherein each of the monomer domains is a non-naturally-occurring monomer domain.

96. The product of claim 93, wherein each of the monomer domains binds calcium.

97. The product of claim 93, wherein at least one of the monomer domains is derived from an LDL-receptor class A domain.

98. The product of claim 93, wherein at least one of the monomer domains is derived from an EGF-like domain.

99. The product of claim 93, wherein at least one domain has a binding specificity for a blood factor.

100. The product of claim 99, wherein the blood factor is serum albumin.

101. The product of claim 93, wherein the monomer domains are separated by a linker.

102. The product of claim 101, wherein the linker is a peptide linker.

103. The product of claim 102, wherein the linker is between 4 to 12 amino acids long.

104. The product of claim 93, wherein the product comprises a first monomer domain that binds a first molecule and a second monomer domain that binds a second molecule.

105. The product of claim 104, wherein the first and second molecules are different.

106. The product of claim 104, wherein the first and second molecules are different copies of the same molecule.

107. The product of claim 93, wherein the product comprises two monomer domains, each monomer domain having a binding specificity for a binding site on a first molecule.

108. The product of claim 107, wherein each of the two monomer domains have a binding specificity for a different binding site on the first molecule.

109. The product of claim 93, wherein the monomer domains are between 25 and 500 amino acids.

110. The product of claim 93, wherein the monomer domains are between 25 and 50 amino acids.

111. The product of claim 93, wherein the monomer domains are between 30 and 100 amino acids.

112. The product of claim 93, wherein the product comprises at least three monomer domains.

113. The product of claim 93, wherein the product comprises four monomer domains.

114. The product of claim 93, wherein the monomer domains are derived from human monomer domains.

115. The product of claim 93, wherein a polypeptide comprises the at least two monomer domains.

116. A product comprising 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.

117. The product of claim 116, wherein each monomer domain is between 30-100 amino acids and each of the monomer domains comprise at least one disulfide linkage.

118. The product of claim 116, wherein each monomer domain is between 30-100 amino acids and is derived from an extracellular protein.

119. The product of claim 116, wherein each monomer domain is between 30-100 amino acids and binds to a protein target.

120. The product of claim 116, wherein the monomer domains are derived from human monomer domains.

121. The product of claim 116, wherein more than one of the monomer domains is a non-naturally-occurring monomer domain.

122. The product of claim 116, wherein each of the monomer domains is a non-naturally-occurring monomer domain.

123. The product of claim 116, wherein each of the monomer domains binds a metal ion.

124. The product of claim 116, wherein each of the monomer domains binds a calcium ion.

125. The product of claim 116, wherein at least one of the monomer domains is derived from an LDL-receptor class A domain.

126. The product of claim 116, wherein at least one of the monomer domains is derived from an EGF-like domain.

127. The product of claim 116, wherein at least one domain has a binding specificity for a blood factor.

128. The product of claim 127, wherein the blood factor is serum albumin.

129. The product of claim 116, wherein the monomer domains are separated by a linker.

130. The product of claim 129, wherein the monomer domains are separated by a peptide linker.

131. The product of claim 130, wherein the linker is between 4 to 12 amino acids long.

132. The product of claim 116, wherein the product comprises a first monomer domain that binds a first molecule and a second monomer domain that binds a second molecule.

133. The product of claim 132, wherein the first and second molecules are different.

134. The product of claim 133, wherein the first and second molecules are different copies of the same molecule.

135. The product of claim 116, wherein the product comprises two monomer domains, each monomer domain having a binding specificity for a site on a first molecule.

136. The product of claim 135, wherein each of the two monomer domains have a binding specificity for a different binding site on the first molecule.

137. The product of claim 116, wherein a polypeptide comprises the at least four monomer domains.

138. A product comprising 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.

139. The product of claim 138, wherein each monomer domain is between 35-100 amino acids.

140. The product of claim 138, wherein each domain comprises at least one disulfide bond and is derived from a human protein and/or an extracellular protein.

141. The product of claim 138, wherein the monomer domains are derived from human monomer domains.

142. The product of claim 138, wherein more than one of the monomer domains is a non-naturally-occurring monomer domain.

143. The product of claim 138, wherein each of the monomer domains is a non-naturally-occurring monomer domain.

144. The product of claim 138, wherein each of the monomer domains binds a metal ion.

145. The product of claim 138, wherein each of the monomer domains binds calcium.

146. The product of claim 138, wherein at least one of the monomer domains is derived from an LDL-receptor class A domain.

147. The product of claim 138, wherein at least one of the monomer domains is derived from an EGF-like domain.

148. The product of claim 138, wherein at least one domain has a binding specificity for a blood factor.

149. The product of claim 148, wherein the blood factor is serum albumin.

150. The product of claim 138, wherein the monomer domains are separated by a linker.

151. The product of claim 150, wherein the monomer domains are separated by a peptide linker.

152. The product of claim 151, wherein the linker is between 4 to 12 amino acids long.

153. The product of claim 138, wherein the product comprises a first monomer domain that binds a first molecule and a second monomer domain that binds a second molecule.

154. The product of claim 153, wherein the first and second molecules are different.

155. The product of claim 154, wherein the first and second molecules are different copies of the same molecule.

156. The product of claim 138, wherein the product comprises two monomer domains, each monomer domain having a binding specificity for a different site on a first molecule.

157. The product of claim 157, wherein each of the two monomer domains have a binding specificity for a different binding site on the first molecule.

158. The product of claim 138, wherein the product comprises at least three monomer domains.

159. The product of claim 138, wherein the product comprises four monomer domains.

160. The product of claim 138, wherein a polypeptide comprises the at least four monomer domains.

161. A product comprising 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.

162. The product of claim 161, wherein each domain is 25-50 amino acids long and comprises at least one disulfide bond.

163. The product of claim 161, wherein each domain is 25-50 amino acids and binds to a protein target.

164. The product of claim 161, wherein each domain is 25-50 amino acids long and is derived from an extracellular protein.

165. The product of claim 161, wherein each domain is 35-50 amino acids long.

166. The product of claim 165, wherein each domain is 35-45 amino acids long.

167. The product of claim 161, wherein more than one of the monomer domains is a non-naturally-occurring monomer domain.

168. The product of claim 161, wherein each of the monomer domains is a non-naturally-occurring monomer domain.

169. The product of claim 161, wherein each of the monomer domains binds calcium.

170. The product of claim 161, wherein at least one of the monomer domains is derived from an LDL-receptor class A domain.

171. The product of claim 161, wherein at least one of the monomer domains is derived from an EGF-like domain.

172. The product of claim 161, wherein at least one domain has a binding specificity for a blood factor.

173. The product of claim 172, wherein the blood factor is serum albumin.

174. The product of claim 161, wherein the monomer domains are separated by a linker.

175. The product of claim 174, wherein the monomer domains are separated by a peptide linker.

176. The product of claim 175, wherein the linker is between 4 to 12 amino acids long.

177. The product of claim 161, wherein the product comprises a first monomer domain that binds a first molecule and a second monomer domain that binds a second molecule.

178. The product of claim 177, wherein the first and second molecules are different.

179. The product of claim 178, wherein the first and second molecules are different copies of the same molecule.

180. The product of claim 161, wherein the product comprises two monomer domains, each monomer domain having a binding specificity for a different site on a first molecule.

181. The product of claim 156, wherein each of the two monomer domains have a binding specificity for a different binding site on the first molecule.

182. The product of claim 161, wherein the product comprises at least three monomer domains.

183. The product of claim 161, wherein the product comprises four monomer domains.

184. The product of claim 161, wherein a polypeptide comprises the at least two monomer domains.

185. A product comprising at least two monomer domains, wherein at least one monomer domain is non-naturally-occurring and each monomer domain comprises at least two disulfide bonds.

186. The product of claim 185, wherein each monomer domain comprises at least three disulfide bonds.

187. The product of claim 185, wherein at least one monomer domain is derived from an extracellular protein.

188. The product of claim 185, wherein at least one monomer domain binds to a target protein.

189. The product of claim 185, wherein more than one of the monomer domains is a non-naturally-occurring monomer domain.

190. The product of claim 185, wherein each of the monomer domains is a non-naturally-occurring monomer domain.

191. The product of claim 185, wherein each of the monomer domains binds calcium.

192. The product of claim 185, wherein at least one of the monomer domains is derived from an LDL-receptor class A domain.

193. The product of claim 185, wherein at least one of the monomer domains is derived from an EGF-like domain.

194. The product of claim 185, wherein at least one domain has a binding specificity for a blood factor.

195. The product of claim 194, wherein the blood factor is serum albumin.

196. The product of claim 185, wherein the monomer domains are separated by a linker.

197. The product of claim 196, wherein the monomer domains are separated by a peptide linker.

198. The product of claim 197, wherein the linker is between 4 to 12 amino acids long.

199. The product of claim 185, wherein the product comprises a first monomer domain that binds a first molecule and a second monomer domain that binds a second molecule.

200. The product of claim 199, wherein the first and second molecules are different.

201. The product of claim 200, wherein the first and second molecules are different copies of the same molecule.

202. The product of claim 185, wherein the product comprises two monomer domains, each monomer domain having a binding specificity for a different site on a first molecule.

203. The product of claim 202, wherein each of the two monomer domains have a binding specificity for a different binding site on the first molecule.

204. The product of claim 185, wherein the product comprises at least three monomer domains.

205. The product of claim 185, wherein the product comprises four monomer domains.

206. The product of claim 185, wherein the product comprises the at least two monomer domains.

207. A method for identifying a monomer domain with affinity for a molecule, the method comprising,

providing a library of monomer domains, wherein each monomer domain:

is between 30-100 amino acids;

comprises at least one disulfide bond; and

binds an ion; and

screening the library of monomer domains for affinity to a first molecule; and

identifying at least one monomer domain that binds to at least one molecule.

208. The method of claim 207, wherein the ion is calcium.

209. The method of claim 207, further comprising,

screening the library of monomer domains for affinity to a second molecule;

identifying a monomer domain that binds to a second molecule;

linking at least one monomer domain with affinity for the first molecule with at least one monomer domain with affinity for the second molecule, thereby forming a multimer with affinity for the first and the second molecules.

210. The method of claim 207, further comprising,

linking the identified monomer domain to a library of monomer domains 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 the first molecule or a second molecule; and

identifying a multimer that binds to the first molecule or second molecule.

211. The method of claim 210, wherein the library of multimers is screened for the ability to bind to the first molecule.

212. The method of claim 210, wherein the library of multimers is screened for the ability to bind to the second molecule.

213. A method for identifying a multimer that binds to at least one molecule, the method comprising:

providing a library of multimers, wherein each multimer comprises at least two monomer domains, and wherein at least one monomer:

is between 30-100 amino acids;

comprises a t least one disulfide bond; and

binds an ion; and

screening the library of multimers for molecule-binding multimers.

214. The method of claim 213, wherein the ion is calcium.

215. A library of multimers, wherein the multimers comprise at least two monomer domains connected by a linker; and the monomer domains

are between 30-100 amino acids;

comprise a t least one disulfide bond; and

bind an ion.

216. The library of claim 215, wherein the ion is calcium.

* * * * *

专利名称(译)	单体结构域的组合文库		
公开(公告)号	US20050221384A1	公开(公告)日	2005-10-06
申请号	US10/966064	申请日	2004-10-15
[标]申请(专利权)人(译)	avidia研究所 麦克西根股份有限公司		
申请(专利权)人(译)	AVIDIA研究所 MAXYGEN INC.		
当前申请(专利权)人(译)	AMGEN MOUNTAIN VIEW，INC.		
[标]发明人	KOLKMAN JOOST A STEMMER WILLEM P C FRESKGARD PER OLA		
发明人	KOLKMAN, JOOST A. STEMMER, WILLEM P.C. FRESKGARD, PER-OLA		
IPC分类号	G01N33/53 C07B61/00 C07K1/04 C07K14/47 C12N15/09 C12Q1/02 C12Q1/68 C40B30/04 C40B40/10 G01N33/68 G01N37/00 G01N33/543		
CPC分类号	B01J2219/00659 B01J2219/00702 B01J2219/00725 C07K1/047 C40B30/04 C40B40/10 G01N33/6845		
优先权	60/337209 2001-11-19 US 60/333359 2001-11-26 US 60/286823 2001-04-26 US 60/374107 2002-04-18 US		
外部链接	Espacenet USPTO		

摘要(译)

提供了鉴定具有所需特性的离散单体结构域和免疫结构域的方法。还提供了由两种或更多种选择的离散单体结构域产生多聚体的方法，以及鉴定具有所需性质的多聚体的方法。还提供了呈递系统，其呈递离散的单体和/或免疫结构域，选择的单体和/或免疫结构域，多聚体和/或选择的多聚体以允许它们的选择。表达一个或多个文库成员的组合物，文库和细胞，以及试剂盒和整合系统也包括在本发明中。

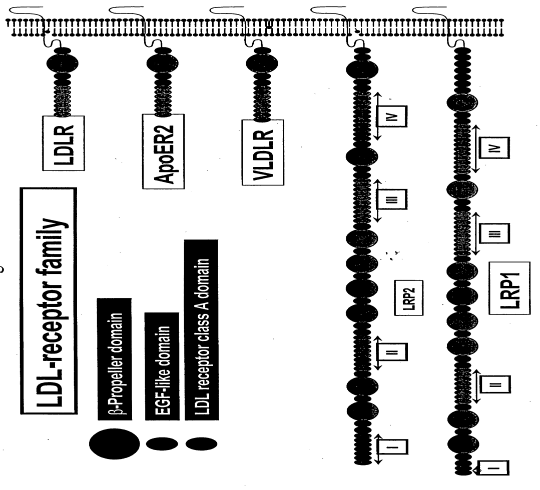


Figure 1