Claims

1. An isolated protein comprising an immunoglobulin variable region comprising at least two cysteine residues positioned within framework region (FR) 1, wherein if at least two of the cysteine residues are not conjugated to a compound then a disulphide bond is capable of forming between the cysteine residues.

2. An isolated protein comprising an immunoglobulin heavy chain variable region (V_H) and an immunoglobulin light chain variable region (V_L), wherein at least one of the variable regions comprises at least two cysteine residues positioned within framework region (FR) 1, wherein if at least two of the cysteine residues are not conjugated to another compound a disulphide bond is capable of forming between the cysteine residues.

3. The protein according to claim 1 or 2, wherein the disulphide bond is present under non-reducing conditions.

4. The protein according to any one of claims 1 to 3, wherein the cysteine residues are positioned between residue 2 numbered according to the Kabat numbering system and complementary determining region (CDR) 1.

5. The protein according to any one of claims 2 to 4, wherein the cysteine residues are within the V_H and are positioned between residues 2 to 30 according to the Kabat numbering system.

6. The protein according to claim 5, wherein the cysteine residues are positioned between residues 7 to 20 and/or residues 24 to 30 numbered according to the Kabat numbering system.

7. The protein according to claim 6, wherein the cysteine residues are positioned between residues 7 to 20 numbered according to the Kabat numbering system.

8. The protein according to any one of claims 2 to 4, wherein the cysteine residues are within the V_L and are positioned between residues 2 to 22 numbered according to the Kabat numbering system.

9. The protein according to claim 8, wherein the cysteine residues are positioned between residues 7 to 20 numbered according to the Kabat numbering system.

10. The protein according to any one of claims 1 to 9, wherein the cysteine residues are additional to the conserved cysteine residue in the variable region.

11. The protein according to claim 10, wherein the conserved cysteine residue is at residue 23 in the V_L and/or residue 22 in the V_H numbered according to the Kabat numbering system.

12. The protein according to claim 10 or 11, wherein the cysteine residues are positioned N-terminal to the conserved cysteine residue.

13. The protein according to any one of claims 1 to 12, wherein the cysteine residues are positioned at one or more of the following:

(i) residue 8 and residue 11 of a κ V_L numbered according to the Kabat numbering system;

(ii) residue 14 and residue 17 of a κ V_L numbered according to the Kabat numbering system;

(iii) residue 7 and residue 11 of a λ V_L numbered according to the Kabat numbering system;

(iv) residue 14 and residue 17 of a λ V_L numbered according to the Kabat numbering system;

(v) residue 8 and residue 12 of a λ V_L numbered according to the Kabat numbering system;

(vi) residue 7 and residue 10 of a V_H numbered according to the Kabat numbering system.

(vii) residue 13 and residue 16 of a V_H numbered according to the Kabat numbering system.

14. The protein according to any one of claims 1 to 12, wherein the cysteine residues are positioned at one or more of the following:

(i) residue 13 and residue 19 of a κ V_L numbered according to the Kabat numbering system;

(ii) residue 13 and residue 19 of a λ V_L numbered according to the Kabat numbering system;

(vi) residue 6 and residue 9 of a V_H numbered according to the Kabat numbering system; and/or

(vii) residue 12 and residue 18 of a V_H numbered according to the Kabat numbering system.

15. An isolated protein comprising a Fv comprising at least one protein according to any one of claims 2 to 14 in which at least one V_L binds to at least one V_H to form an antigen binding site.

16. The protein according to claim 15, wherein the V_L and the V_H which form the antigen binding site are in a single polypeptide chain.

17. The protein according to claim 16, which is:

(i) a single chain Fv fragment (scFv);

(ii) a dimeric scFv (di-scFv); or

(iii) at least one of (i) and/or (ii) linked to a Fc or a heavy chain constant domain (C_H) 2 and/or C_H3.

18. The protein according to claim 15, wherein the V_L and V_H which form the antigen binding site are in different polypeptide chains.

19. The protein according to claim 18, which is:

(i) a diabody;

(ii) a triabody; or

(iii) a tetrabody.

20. The protein according to claim 18, which is an immunoglobulin.

21. The protein according to any one of claims 1 to 20, wherein the cysteine residues are linked by a disulphide bond.

22. The protein according to any one of claims 1 to 20 comprising a compound conjugated to at least one of the cysteine residues, wherein conjugation of the compound does not prevent binding of the protein to an antigen.

23. The protein according to claim 22, wherein the compound is selected from the group consisting of a radioisotope, a detectable label, a therapeutic compound, a colloid, a toxin, a nucleic acid, a peptide, a protein, a compound that increases the half life of the protein in a subject and mixtures thereof.

24. The protein of any one of claims 1 to 23 additionally comprising at least one N-terminal threonine or serine residue.

25. The protein according to claim 24 comprising a compound conjugated to the threonine or serine residue.

26. The protein according to claim 25, wherein the compound is selected from the group consisting of a radioisotope, a detectable label, a therapeutic compound, a colloid, a toxin, a nucleic acid, a peptide, a protein, a compound that increases the half life of the protein in a subject and mixtures thereof.

27. The protein according to claim 25 or 26 comprising a first compound conjugated to at least one of the cysteine residues and a second compound conjugated to the threonine or serine residue, wherein the second compound is different to the first compound.

28. The protein of claim 22, wherein the protein comprises at least one polypeptide comprising a sequence at least about 80% identical to a sequence set forth in SEQ ID NO: 57, and wherein the compound is a monodispersed PEG.

29. The protein of claim 28, wherein the monodispersed PEG has no more than 48 ethylene glycol units.

30. The protein of claim 28 or 29, wherein the monodispersed PEG has about 24 ethylene glycol units.

31. The protein according to any one of claims 1 to 30 that specifically binds to human epidermal growth factor (Her) 2, tumour associated glycoprotein TAG72, MUC1 or prostate specific membrane antigen (PSMA).

32. The protein according to claim 31, wherein the protein comprises a sequence at least about 80% identical to a sequence set forth in any one of SEQ ID NOs: 55, 59, 61 or 109, modified to include the two or more cysteine residues positioned within framework region (FR) 1.

33. The protein according to claim 32 comprising a sequence at least about 80% identical to a sequence set forth in any one or more of SEQ ID NOs: 57, 63, 65, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103 or 105, 119, 121, 123, 125, 127, 129, 131 or 133.

34. An isolated protein comprising an immunoglobulin heavy chain variable region (V_H) and an immunoglobulin light chain variable region (V_L), wherein at least one of the variable regions comprises a N-terminal threonine residue or serine residue.

35. An isolated protein comprising a Fv comprising at least one protein according to claim 31, wherein at least one V_L binds to at least one V_H to form an antigen binding site.

36. The protein according to claim 35, wherein the V_L and the V_H which form the antigen binding site are in a single polypeptide chain.

37. The protein according to claim 36, which is:

(i) a single chain Fv fragment (scFv);

(ii) a dimeric scFv (di-scFv); or

(iii) at least one of (i) and/or (ii) linked to a Fc or a heavy chain constant domain (C_H) 2 and/or C_H3.

38. The protein according to claim 35, wherein the V_L and the V_H which form the antigen binding site are in different polypeptide chains.

39. The protein according to claim 38, which is:

(i) a diabody;

(ii) a triabody; or

(iii) a tetrabody.

40. The protein according to claim 38, which is an immunoglobulin.

41. The protein according to any one of claims 34 to 40 additionally comprising a compound conjugated to the threonine or serine residue.

42. The protein according to claim 41, wherein the compound is selected from the group consisting of a radioisotope, a detectable label, a therapeutic compound, a colloid, a toxin, a nucleic acid, a peptide, a protein, a compound that increases the half life of the protein in a subject and mixtures thereof.

43. The protein according to any one of claims 34 to 43 comprising a sequence at least about 80% identical to a sequence set forth in any one of SEQ ID NOs: 55, 59, 61 or 109, modified to include the N-terminal threonine or serine residue.

44. A composition comprising the protein according to any one of claims 1 to 43 and a pharmaceutically acceptable carrier.

45. An isolated nucleic acid encoding the protein of any one of claims 1 to 21, 24, 31 to 40 or 43.

46. An expression construct comprising the nucleic acid according to claim 45 operably linked to a promoter.

47. An isolated cell comprising an exogenous nucleic acid according to claim 45 and/or the expression construct of claim 46 wherein the cell expresses the protein.

48. The cell according to claim 47, wherein the cell is a bacterial cell, a yeast cell, a mammalian cell or an insect cell.

49. A method for producing the protein according to any one of claims 1 to 21, 24, 31 to 40 or 43 the method comprising maintaining the expression construct according to claim 46 under conditions sufficient for the encoded protein to be produced.

50. The method according to claim 49 comprising culturing the cell according to claim 47 or 48 under conditions sufficient the encoded for the protein to be produced.

51. The method according to claim 49 or 50 comprising isolating the protein.

52. A method for producing the protein according to any one of claims 22 to 30, the method comprising:

(i) obtaining the protein according to any one of claims 1 to 21, 24 or 31 to 33; and

(ii) conjugating a compound to at least one of the cysteine residues in FR1 of the protein to thereby produce the protein.

53. The method according to claim 52, wherein the cysteine residues in the protein obtained at (i) are linked by a disulphide bond and the method additionally comprises reducing or otherwise breaking the disulphide bond prior to conjugating the compound to the cysteine residue(s).

54. The method according to claim 53, wherein reducing or otherwise breaking the disulphide bond generates a free thiol group in the protein and the compound has a thiol reactive group permitting conjugation of the compound to the protein.

55. The method according to any one of claims 52 to 54, wherein the protein comprises at least one N-terminal serine or threonine residue and the method additionally comprises conjugating a compound to the serine or threonine residue.

56. A method for producing the protein of any one of claims 26, 27 or 41 to 43, the method comprising:

(i) obtaining the protein according to any one of claims 24 or 34 to 40; and

(ii) conjugating a compound to at least one serine or threonine residue at the N-terminus of said protein to thereby produce the protein.

57. Use of the protein according to any one of claims 1 to 43 or the composition according to claim 44 in medicine.

58. A method of treating or preventing a condition in a subject, the method comprising administering the protein according to any one of claims 1 to 43 or the composition according to claim 44 to a subject in need thereof.

59. A method for delivering a compound to a cell, the method comprising contacting the cell with the protein according to any one of claims 22 to 30, 41 or 42 or the composition according to claim 44.

60. The method according to claim 59, wherein the cell is contacted by administering the protein or composition to a subject.

61. A method for localizing or detecting an antigen in a subject, said method comprising:

(i) administering to a subject the protein according to any one of claims 22 to 30, 41 or 42 or the composition according to claim 44 for a time and under conditions sufficient for the protein to bind to an antigen, wherein the protein is conjugated to a detectable label; and

(ii) detecting or localizing the detectable label *in vivo*.

62. A method for diagnosing or prognosing a condition in a subject, the method comprising contacting a sample from the subject with the protein according to any one of claims 1 to 43 or the composition according to claim 44 for a time and under conditions sufficient for the protein to bind to an antigen and form a complex and detecting the complex, wherein detection of the complex is diagnostic or prognostic of the condition in the subject.

63. The method according to claim 62, comprising determining the level of the complex, wherein an enhanced or reduced level of said complex is diagnostic or prognostic of the condition in the subject.

64. An isolated protein comprising a plurality of polypeptides each comprising an immunoglobulin heavy chain variable region (V_H) and an immunoglobulin light chain variable region (V_L) linked by a region comprising an insufficient number of amino acids to permit the V_H and the V_L to associate with one another, wherein:

(i) at least one of the polypeptides comprises a V_H comprising a sequence set forth in SEQ ID NO: 111 or a sequence at least about 60% identical thereto; and

(ii) at least another of the polypeptides comprises a V_L comprising a sequence set forth in SEQ ID NO: 113 or a sequence at least about 60% identical thereto,

wherein the V_H of the polypeptide at (i) and the V_L of the polypeptide at (ii) associate to form a Fv capable of specifically binding to tumour antigen TAG-72.

65. The protein of claim 64, wherein the region linking the V_H and the V_L comprises 6 or fewer amino acids.

66. The protein of claim 64 or 65, wherein the V_L comprises a threonine at position 5 numbered according to the Kabat numbering system and/or a threonine at position 53 according to the Kabat numbering system and/or a glutamic acid at position 79 numbered according to the Kabat numbering system.

67. The protein of any one of claims 64 to 66, wherein the V_H comprises a leucine at position 80 numbered according to the Kabat numbering system.

68. The protein of any one of claims 64 to 67 comprising at least two polypeptides, each comprising:

(i) a V_H comprising a sequence set forth in SEQ ID NO: 111 or a sequence at least about 60% identical thereto; and

(ii) a V_L comprising a sequence set forth in SEQ ID NO: 113 or a sequence at least about 60% identical thereto,

wherein the V_H of one polypeptide and the V_L of another polypeptide associate to form a Fv capable of specifically binding to TAG72.

69. The protein of any one of claims 64 to 68, which is a diabody, a triabody or a tetrabody.

70. The protein of any one of claims 64 to 69, wherein:

(i) at least one of the polypeptides comprises a V_H comprising complementarity determining regions (CDRs) of a V_H comprising a sequence set forth in SEQ ID NO: 111; and

(ii) at least another of the polypeptides comprises a V_L comprising CDRs of a V_L comprising a sequence set forth in SEQ ID NO: 113.

71. The protein of any one of claims 64 to 70 comprising at least two polypeptides both of which comprise:

(i) a V_H comprising a sequence set forth in SEQ ID NO: 111 or a humanized or de-immunized form thereof; and

(ii) a V_L comprising a sequence set forth in SEQ ID NO: 113 or a humanized or de-immunized form thereof.

72. The protein of any one of claims 64 to 71 comprising at least two polypeptides each comprising a sequence at least 80% identical to a sequence set forth in SEQ ID NO: 55.

73. The protein of any one of claims 64 to 73, wherein at least one of the polypeptides comprises at least two cysteine residues positioned within framework region (FR) 1, wherein if at least two of the cysteine residues is not conjugated to a compound then a disulphide bond is capable of forming between the cysteine residues.

74. The protein of any one of claims 64 to 73, wherein at least one of the polypeptides comprises a N-terminal serine and/or threonine residue.

75. The protein of claim 73 or 74 comprising a sequence set forth in any one or more of SEQ ID NOs: 57, 63, 75, 77, 99, 101, 103, 119, 121, 123, 125, 127, 129, 131 or 133.

76. The protein of any one of claims 64 to 75 comprising a compound conjugated thereto.

77. The protein of claim 76, wherein the compound is conjugated to a cysteine residue or a serine residue or a lysine residue in at least one of the polypeptides.

78. The protein of claim 76 or 77, wherein at least one of the polypeptides comprises at least two cysteine residues positioned within framework region (FR) 1 positioned such that if at least two of the cysteine residues is not conjugated to a compound then a disulphide bond is capable of forming between the cysteine residues and/or, wherein at least one of the polypeptides comprises a N-terminal serine and/or threonine residue, and wherein the compound is conjugated to at least one of the cysteine residues and/or to the serine residue.

79. The protein of claim 78, wherein the polypeptide comprises the cysteine residues and the N-terminal serine and/or threonine residue, and wherein the compound is conjugated to at least one of the cysteine residues and optionally, a different compound is conjugated to the serine residue.

80. The protein of any one of claims 76 to 79, wherein the compound is selected from the group consisting of a radioisotope, a detectable label, a therapeutic compound, a colloid, a toxin, a nucleic acid, a peptide, a protein, a compound that increases the half life of the protein in a subject and mixtures thereof.

81. The protein of claim any one of claims 76 to 80, wherein the compound is a physiologically acceptable polymer.

82. The protein of claim any one of claims 76 to 81, wherein the compound is polyethylene glycol (PEG).

83. The protein of claim 82, wherein the PEG is monodispersed PEG.

84. The protein of claim 82 or 83, wherein the PEG has a molecular weight no greater than 4000 Da.

85. The protein of claim 82 to 84, wherein the PEG has a molecular weight no greater than 2000 Da.

86. The protein of claim 82 to 85, wherein the PEG has a molecular weight no greater than 1,500 Da.

87. A composition comprising the protein of any one of claims 64 to 86 and a pharmaceutically acceptable carrier.

88. An isolated nucleic acid encoding one or more of the polypeptides in the protein of any one of claims 64 to 75.

89. An expression construct comprising the nucleic acid of claim 88 operably linked to a promoter.

90. An isolated cell expressing the protein of any one of claims 64 to 75.

91. The cell of claim 90 comprising the nucleic acid of claim 88 and/or the expression construct of claim 89.

92. A method for producing the protein of any one of claims 64 to 75 the method comprising maintaining the expression construct of claim 89 such that the encoded polypeptide and the protein are produced.

93. The method of claim 92 comprising culturing the cell of claim 90 or 91 under conditions sufficient the encoded for the protein to be produced.

94. A method for producing the protein of any one of claims 76 to 86 comprising a compound conjugated to at least one of the cysteine residues, the method comprising:

(i) obtaining the protein of any one of claims 73 or 75; and

(ii) conjugating a compound to at least one of the cysteine residues in FR1 of the polypeptide(s) to thereby produce the protein.

95. A method for producing the protein of claim 78 or 79 comprising a compound conjugated to a N-terminal serine or threonine residue, the method comprising:

(i) obtaining the protein of claim 74 or 75; and

(ii) conjugating a compound to at least one serine or threonine residue at the N-terminus of the polypeptide to thereby produce the protein.

96. The method of claim 94 or 95, wherein the compound is polyethylene glycol (PEG).

97. A method for localizing and/or detecting and/or diagnosing and/or prognosing a cancer in a subject, said method comprising:

(i) administering to a subject the protein of any one of claims 64 to 86 or the composition of claim 87 such that the protein binds to tumour antigen TAG-72, if present; and

(ii) detecting the protein bound to the TAG72 *in vivo*, wherein detection of the bound protein localizes and/or detects and/or diagnoses and/or prognoses the cancer.

98. A method of diagnosing or prognosing a cancer in a subject, the method comprising:

(i) contacting a sample from the subject with the protein of any one of claims 64 to 86 or the composition of claim 87 such that the protein binds to tumour antigen TAG-72, if present; and

(ii) detecting the protein bound to the TAG72, wherein detection of the bound protein is diagnostic or prognostic of the cancer.

99. The method of claim 97 or 98, wherein the protein is conjugated to a detectable label and the method comprises detecting the label to detect the protein bound to TAG72.

100. A method of treating a cancer, the method comprising administering the protein of any one of claims 64 to 86 or the composition of claim 87 such that it binds to tumour antigen TAG-72 on cancer cells and treats the cancer.

101. The method of claim 100, wherein the protein or a compound conjugated thereto induces death of the cancer cells.

102. The method of claim 100 or 101, wherein the protein is conjugated to a compound induces death of the cancer cells.

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. An isolated protein comprising an immunoglobulin variable region comprising at least two cysteine residues positioned within framework region (FR) 1, wherein if at least two of the cysteine residues are not conjugated to a compound then a disulphide bond is capable of forming within FR1 between the cysteine residues.

2. An isolated protein comprising an immunoglobulin heavy chain variable region (V_H) and an immunoglobulin light chain variable region (V_L), wherein at least one of the variable regions comprises at least two cysteine residues positioned within framework region (FR) 1, wherein if at least two of the cysteine residues are not conjugated to another compound a disulphide bond is capable of forming between the cysteine residues.

3. An isolated protein comprising a plurality of polypeptides each comprising an immunoglobulin heavy chain variable region (V_H) and an immunoglobulin light chain variable region (V_L) linked by a region comprising an insufficient number of amino acids to permit the V_H and the V_L to associate with one another, wherein:

(i) at least one of the polypeptides comprises a V_H comprising a sequence set forth in SEQ ID NO: 111 or a sequence at least about 60% identical thereto; and

(ii) at least another of the polypeptides comprises a V_L comprising a sequence set forth in SEQ ID NO: 113 or a sequence at least about 60% identical thereto,

wherein the V_H of the polypeptide at (i) and the V_L of the polypeptide at (ii) associate to form a Fv capable of specifically binding to tumour antigen TAG-72, and wherein at least one of the polypeptides comprises at least two cysteine residues positioned within framework region (FR) 1, wherein if at least two of the cysteine residues is not conjugated to a compound then a disulphide bond is capable of forming within FR1 between the cysteine residues.

4. The protein according to any one of claims 1 to 3, wherein the disulphide bond is present under non-reducing conditions.

5. The protein according to any one of claims 1 to 4, wherein the cysteine residues are positioned between residue 2 numbered according to the Kabat numbering system and complementary determining region (CDR) 1.

6. The protein according to any one of claims 2 to 5, wherein the cysteine residues are within the V_H and are positioned in one or more of the following regions:

(i) between residues 2 to 30 according to the Kabat numbering system;

(ii) between residues 7 to 20 and/or residues 24 to 30 numbered according to the Kabat numbering system; and/or

(iii) between residues 7 to 20 numbered according to the Kabat numbering system.

7. The protein according to any one of claims 2 to 5, wherein the cysteine residues are within the V_L and are positioned in one or more of the following regions:

(i) between residues 2 to 22 numbered according to the Kabat numbering system; and/or

(ii) between residues 7 to 20 numbered according to the Kabat numbering system.

8. The protein according to any one of claims 1 to 7, wherein the cysteine residues are additional to the conserved cysteine residue in the variable region at residue 23 in the V_L and/or residue 22 in the V_H numbered according to the Kabat numbering system.

9. The protein according to claim 8, wherein the cysteine residues are positioned N-terminal to the conserved cysteine residue.

10. The protein according to any one of claims 1 to 9, wherein the cysteine residues are positioned at one or more of the following:

(i) residue 8 and residue 11 of a κ V_L numbered according to the Kabat numbering system;

(ii) residue 14 and residue 17 of a κ V_L numbered according to the Kabat numbering system;

(iii) residue 7 and residue 11 of a λ V_L numbered according to the Kabat numbering system;

(iv) residue 14 and residue 17 of a λ V_L numbered according to the Kabat numbering system;

(v) residue 8 and residue 12 of a λ V_L numbered according to the Kabat numbering system;

(vi) residue 13 and residue 19 of a κ V_L numbered according to the Kabat numbering system;

(vii) residue 13 and residue 19 of a λ V_L numbered according to the Kabat numbering system;

(viii) residue 7 and residue 10 of a V_H numbered according to the Kabat numbering system;

(ix) residue 13 and residue 16 of a V_H numbered according to the Kabat numbering system;

(x) residue 6 and residue 9 of a V_H numbered according to the Kabat numbering system; and/or

(xi) residue 12 and residue 18 of a V_H numbered according to the Kabat numbering system.

11. The protein of claim 3, wherein the region linking the V_H and the V_L comprises 6 or fewer amino acids.

12. The protein of claim 3 or 11, wherein the V_L comprises a threonine at position 5 numbered according to the Kabat numbering system and/or a threonine at position 53 according to the Kabat numbering system and/or a glutamic acid at position 79 numbered according to the Kabat numbering system.

13. The protein of any one of claims 3, 11 or 12, wherein the V_H comprises a leucine at position 80 numbered according to the Kabat numbering system.

14. The protein of any one of claims 3 or 11 to 13 comprising at least two polypeptides, each comprising:

(i) a V_H comprising a sequence set forth in SEQ ID NO: 111 or a sequence at least about 60% identical thereto; and

(ii) a V_L comprising a sequence set forth in SEQ ID NO: 113 or a sequence at least about 60% identical thereto,

wherein the V_H of one polypeptide and the V_L of another polypeptide associate to form a Fv capable of specifically binding to TAG72.

15. The protein of any one of claims 3 or 11 to 14, wherein:

(i) at least one of the polypeptides comprises a V_H comprising complementarity determining regions (CDRs) of a V_H comprising a sequence set forth in SEQ ID NO: 111; and

(ii) at least another of the polypeptides comprises a V_L comprising CDRs of a V_L comprising a sequence set forth in SEQ ID NO: 113.

16. The protein of any one of claims 3 or 11 to 15 comprising at least two polypeptides both of which comprise:

(i) a V_H comprising a sequence set forth in SEQ ID NO: 111 or a humanized or de-immunized form thereof; and

(ii) a V_L comprising a sequence set forth in SEQ ID NO: 113 or a humanized or de-immunized form thereof.

17. An isolated protein comprising a Fv comprising at least one protein according to any one of claims 2 to 16 in which at least one V_L binds to at least one V_H to form an antigen binding site.

18. The protein according to claim 17, wherein the V_L and the V_H which form the antigen binding site are in a single polypeptide chain or wherein the V_L and V_H which form the antigen binding site are in different polypeptide chains.

19. The protein according to claim 18, wherein if the V_L and V_H are in a single polypeptide chain, the protein is:

(i) a single chain Fv fragment (scFv);

(ii) a dimeric scFv (di-scFv); or

(iii) at least one of (i) and/or (ii) linked to a Fc or a heavy chain constant domain (C_H) 2 and/or C_H3, or

if the V_L and V_H are in a separate polypeptide chains, the protein is:

(iv) a diabody;

(v) a triabody;

(vi) a tetrabody; or

(vii) an immunoglobulin.

20. The protein according to any one of claims 1 to 19, wherein the cysteine residues are linked by a disulphide bond.

21. The protein according to any one of claims 1 to 19 comprising a compound conjugated to at least one of the cysteine residues, wherein conjugation of the compound does not prevent binding of the protein to an antigen.

22. The protein of any one of claims 1 to 21 additionally comprising at least one N-terminal threonine or serine residue.

23. The protein according to claim 22 comprising a compound conjugated to the threonine or serine residue.

24. The protein according to claim 21 or 23, wherein the compound is selected from the group consisting of a radioisotope, a detectable label, a therapeutic compound, a colloid, a toxin, a nucleic acid, a peptide, a protein, a compound that increases the half life of the protein in a subject and mixtures thereof.

25. The protein of claim 21 or 23, wherein the compound is a physiologically acceptable polymer.

26. The protein of claim any one of claims 21 or 23, wherein the compound is polyethylene glycol (PEG).

27. The protein of claim 26, wherein the PEG is monodispersed PEG.

28. The protein according to claim 22 or 23 comprising a first compound conjugated to at least one of the cysteine residues and a second compound conjugated to the threonine or serine residue, wherein the second compound is different to the first compound.

29. The protein according to any one of claims 1 to 28 that specifically binds to human epidermal growth factor (Her) 2, tumour associated glycoprotein TAG72, MUC1 or prostate specific membrane antigen (PSMA).

30. The protein according to claim 29, wherein the protein comprises:

(i) a sequence at least about 80% identical to a sequence set forth in any one of SEQ ID NOs: 55, 59, 61 or 109, modified to include the two or more cysteine residues positioned within FR1; or

(ii) a sequence at least about 80% identical to a sequence set forth in any one or more of SEQ ID NOs: 57, 63, 65, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103 or 105, 119, 121, 123, 125, 127, 129, 131 or 133, wherein the sequence comprises the two or more cysteine residues positioned within FR1.

31. A composition comprising the protein according to any one of claims 1 to 30 and a pharmaceutically acceptable carrier.

32. An isolated nucleic acid encoding the protein of any one of claims 1 to 20, 22, 29 or 30.

33. An expression construct comprising the nucleic acid according to claim 32 operably linked to a promoter.

34. An isolated cell comprising an exogenous nucleic acid according to claim 32 and/or the expression construct of claim 33 wherein the cell expresses the protein.

35. A method for producing the protein according to any one of claims 1 to 20, 22, 29 or 30 the method comprising maintaining the expression construct according to claim 33 under conditions sufficient for the encoded protein to be produced and, optionally isolating the protein.

36. The method according to claim 35 comprising culturing the cell according to claim 34 under conditions sufficient the encoded for the protein to be produced.

37. A method for producing the protein according to any one of claims 21 to 28, the method comprising:

(i) obtaining the protein according to any one of claims 1 to 20, 22, 29 or 30; and

(ii) conjugating a compound to at least one of the cysteine residues in FR1 of the protein to thereby produce the protein.

38. The method according to claim 37, wherein the protein comprises at least one N-terminal serine or threonine residue and the method additionally comprises conjugating a compound to the serine or threonine residue.

39. A method for producing the protein of claim 23, the method comprising:

(i) obtaining the protein according to claim 22; and

(ii) conjugating a compound to at least one serine or threonine residue at the N-terminus of said protein to thereby produce the protein.

40. Use of the protein according to any one of claims 1 to 30 or the composition according to claim 31 in medicine.

41. A method for delivering a compound to a cell, the method comprising contacting the cell with the protein according to any one of claims 21 or 23 to 28 or the composition according to claim 31.

42. The method according to claim 41, wherein the cell is contacted by administering the protein or composition to a subject.

43. A method for localizing or detecting an antigen in a subject, said method comprising:

(i) administering to a subject the protein according to any one of claims 21 or 23 to 28 or the composition according to claim 31 for a time and under conditions sufficient for the protein to bind to an antigen, wherein the protein is conjugated to a detectable label; and

(ii) detecting or localizing the detectable label *in vivo*.