(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
06 March 2025 (06.03.2025)

WIPO | PCT

(10) International Publication Number
WO 2025/046229 A1

(51) International Patent Classification:
C07K 14/315 (2006.01) *B82Y 40/00* (2011.01)
C07K 16/00 (2006.01) *C12N 15/10* (2006.01)
C07K 16/18 (2006.01) *C40B 40/10* (2006.01)
C07K 16/46 (2006.01)

(21) International Application Number:
PCT/GB2024/052240

(22) International Filing Date:
28 August 2024 (28.08.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2313175.8 30 August 2023 (30.08.2023) GB

(71) Applicant: **OXFORD UNIVERSITY INNOVATION LIMITED** [GB/GB]; Buxton Court, 3 West Way, Oxford, Oxfordshire OX2 0JB (GB).

(72) Inventors: **HOWARTH, Mark**; Department of Biochemistry, University of Oxford, South Parks Rd, Oxford, Oxfordshire OX1 3QU (GB). **DRISCOLL, Claudia**; Department of Biochemistry, University of Oxford, South Parks Rd, Oxford, Oxfordshire OX1 3QU (GB).

(74) Agent: **DEHNS**; St Bride's House, 10 Salisbury Square, London, Greater London EC4Y 8JD (GB).

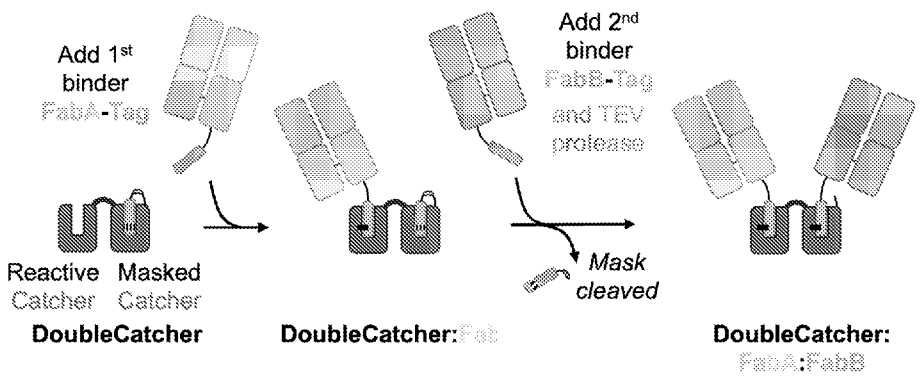
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: A BIVALENT POLYPEPTIDE SCAFFOLD AND USES THEREOF

Figure 1



(57) Abstract: The present invention relates to a bivalent polypeptide scaffold that facilitates sequential (e.g. stepwise) assembly of molecules to form a single structure. Nucleic acid molecules encoding the polypeptide scaffold, vectors comprising said nucleic acid molecules, and host cells comprising said vectors and nucleic acid molecules are also provided. Kits comprising said polypeptide scaffold and/or nucleic acid molecules/vectors are also provided. Further products comprising said polypeptide scaffold and uses of the polypeptide scaffold of the invention are also provided, e.g. methods of assembling bispecific binding structures.

WO 2025/046229 A1

A bivalent polypeptide scaffold and uses thereof

FIELD OF THE INVENTION

The present invention relates in one aspect to a bivalent polypeptide scaffold that facilitates the sequential (e.g. stepwise) assembly of molecules to form a single structure, e.g. to produce a bifunctional structure, such as a structure with bispecific binding properties. In particular, the polypeptide scaffold may be viewed as a bimodular polypeptide comprising two peptide-binding modules, wherein the peptide binding activity of one module is activated by a protease. Alternatively viewed, the peptide binding activity of one module is masked (i.e. blocked) such that only one module is available to bind its cognate peptide and the other module is unmasked via a proteolytic cleavage to allow the unmasked module to bind its cognate peptide. Each peptide-binding module may be viewed as a first part of a peptide tag and polypeptide binding partner cognate pair that can be conjugated via a covalent bond when contacted under conditions that allow the spontaneous formation of an isopeptide bond between the peptide-binding module and its cognate peptide tag. The cognate peptide tag for each peptide-binding module is separately linked to a molecule to be assembled in the structure (e.g. each peptide tag may be linked to a molecule with specific binding activity for a target). Nucleic acid molecules encoding the polypeptide scaffold, vectors comprising said nucleic acid molecules, and host cells comprising said vectors and nucleic acid molecules are also provided. Kits comprising said polypeptide scaffold and/or nucleic acid molecules/vectors are also provided. Further products comprising said polypeptide scaffold and uses of the polypeptide scaffold of the invention are also provided, e.g. methods of assembling bispecific binding structures.

BACKGROUND TO THE INVENTION

Bispecific antibodies are biological matchmakers, bringing together two different antigens or two different parts of the same antigen. Bispecifics have had clinical success in cancer therapy by directing cytotoxic T cells to cancer-specific antigens. There are also approved bispecific therapeutics that bridge proteins on the surface of the same cell. Cell signalling often naturally depends on co-localization of particular ligands at the cell-surface. Synthetic control of such signalling has been successful using bifunctional small molecules, DNA, and non-antibody protein-based binders. The response of the cell is often exquisitely

sensitive to the timing and spatial orientation of this bridging between surface targets. Therefore, despite the numerous successes from bispecifics, there remains a huge range of unexplored combinations of bispecifics that could illuminate the principles of cell signalling and provide valuable therapeutic leads. The provision of efficient systems to facilitate the screening of bispecific combinations represents an unmet challenge in protein engineering.

The most common routes to generate bispecifics depend upon rearrangement of IgG architecture, particularly through knob-into-hole generation of heterodimeric Fc domains, or disulfide bond shuffling in the Fab antibody fragments. Alternative bispecific pairing has been achieved by conjugating fragments with sortase, transglutaminase, click chemistry or split inteins. However, these routes require each binder to be fused to the respective conjugation domain, such that each binder must be re-cloned, re-expressed, and re-purified in two formats to explore all possible bispecific combinations. Such a requirement is a major obstacle to the objective of exploring thousands of binders across the proteome.

Typically, covalent protein interactions are mediated through disulfide bonds, but disulfides are reversible, inapplicable in reducing cellular compartments, and can interfere with protein folding. Peptide tags are convenient tools for protein analysis and modification because their small size minimises the perturbation to protein function. Peptide tags are simple to genetically encode and their small size reduces disruption from (i) interfering with other interactions, (ii) cost of biosynthesis, and (iii) introduction of immunogenicity. However, interactions between peptide tags and their peptide or polypeptide binding partners are rarely of high affinity, which limits their utility in the formation of stable complexes.

Proteins that are capable of spontaneous isopeptide bond formation (so-called “isopeptide proteins”) have been advantageously used to develop peptide tag/polypeptide binding partner pairs (i.e. two-part linkers) which covalently bind to each other and provide irreversible interactions (see e.g. WO2011/098772, WO 2016/193746, WO 2018/197854, WO2020/183198 and WO2022/214795 all herein incorporated by reference). In this respect, proteins that are capable of spontaneous isopeptide bond formation may be expressed as separate fragments, to give a peptide tag and a polypeptide binding partner for the peptide tag (a “catcher”), where the two fragments are capable of covalently reconstituting by

isopeptide bond formation, thereby linking any molecules that are joined to the peptide tag and its polypeptide binding partner.

Isopeptide bonds are amide bonds formed between carboxyl/carboxamide and amino groups, where at least one of the carboxyl or amino groups is outside of the protein main-chain (the backbone of the protein). Such bonds are chemically irreversible under typical biological conditions and they are resistant to most proteases. Since isopeptide bonds are covalent in nature, they result in the some of the strongest measured protein interactions. The isopeptide bond formed by a peptide tag and its polypeptide binding partner is stable under conditions where non-covalent interactions would rapidly dissociate, e.g. over long periods of time (e.g. weeks), at high temperature (to at least 95 ºC), at high force, or with harsh chemical treatment (e.g. pH 2-11, organic solvent, detergents or denaturants).

The present inventors previously developed various two-part peptide tag/polypeptide binding partner systems, most notably SpyTag/SpyCatcher which based on the CnaB2 domain of the *Streptococcus pyogenes* FbaB protein (Zakeri et al., 2012, Proc Natl Acad Sci U S A 109, E690-E697). The most recent iteration of this system, SpyTag003/SpyCatcher003, has been established to have the fastest reactivity for protein-protein reaction at termini (see WO 2020/183198 incorporated herein by reference).

The SpyTag/SpyCatcher system has been proposed for use in the assembly of multi-component polypeptides, such as fusion proteins (e.g. WO2016/193746) and bispecific binders (e.g. SpyLock, Rapid Generation of Bispecific Antibodies for High-Throughput Screening with SpyLock Technology, Bio-Rad, 2023). However, these systems require the use of multiple two-part linker systems with orthogonal activity or reducing agents, respectively, which reduces their application in high-throughput systems and limits their utility as universal systems. In particular, systems that require the use of a reducing agent to assemble multi-component polypeptide structures are limited to assembling components whose structures or activities will not be affected or damaged by contact with said agent.

Thus, there is a need for alternative and improved strategies for assembling bispecific molecules, particularly for systems that facilitate high-throughput and combinatorial assembly of bispecific molecules.

SUMMARY OF THE INVENTION

In work leading to the present invention, the inventors have adapted the SpyTag/SpyCatcher system to develop a 'one-pot' stepwise, controllable reaction pipeline suitable for assembling bispecific antibodies from any two SpyTag-linked antigen binders. The system provides a monospecific, bivalent 'homodimer' scaffold that enables stepwise assembly via the attachment of a protease-cleavable peptide mask to one of the modules in the homodimer.

As described in detail in the Examples, the invention is predicated on the determination that the isopeptide bond formation activity of SpyCatcher003 can be specifically, comprehensively and reversibly masked using a non-reactive SpyTag peptide (SpyTag003D10A, SEQ ID NO: 17) fused to the C-terminus of SpyCatcher003 via a flexible linker containing a Tobacco Etch Virus (TEV) protease cleavage site. It was found that, in the absence of TEV protease, the non-reactive SpyTag peptide binds tightly to SpyCatcher003 (K_D = 21 nM for the non-linked pair), thus masking the reactive Lys31 of SpyCatcher003, i.e. the fusion results in the provision of a "masked" SpyCatcher polypeptide. Whilst not wishing to be bound by theory, it is thought that, as the SpyCatcher003/SpyTag003D10A interaction is intramolecular in nature, this interaction is entropically favoured over a competing intermolecular interaction with SpyCatcher003 and SpyTag003 in solution. However, following cleavage of Masked SpyCatcher003 at the TEV cleavage site, the SpyTag003D10A peptide is free to dissociate from SpyCatcher003, unmasking the reactive Lys31 to enable reaction with a SpyTag peptide. Unexpectedly and advantageously, the inventors found that their design led to a masked SpyCatcher003 where reaction of SpyTag003-fusions was inhibited by over 95%, thereby providing a robust protease-activatable SpyCatcher module.

To design a platform that facilitates the conjugation of two SpyTagged proteins, Masked SpyCatcher003 was genetically fused to the C-terminus of a native SpyCatcher003 domain to form a 'DoubleCatcher'. This bimodular scaffold enables the assembly of bispecific molecules by first saturating the reactive SpyCatcher003 moiety of the DoubleCatcher with SpyTag-BinderA, then cleaving off the peptide that masks the second SpyCatcher003 moiety with TEV protease, which is then saturated by addition of SpyTag-BinderB. Unexpectedly and advantageously, it was found that the design also enabled a rapid displacement of the masking peptide to allow covalent coupling by SpyTag-BinderB so that bispecific molecules could be obtained rapidly and in high yield.

The inventors further have advantageously determined that the DoubleCatcher bimodular scaffold can be "locked" in various configurations by engineering the modules to contain cysteine residues that react to form disulfide bonds and then direct the binders to project in different directions from the scaffold. This panel of DoubleCatchers, in combination with a generalized methodology for one-pot assembly and purification of bispecifics, provides a simple and scalable tool to discover synergy in bispecific activity, through modulating receptor organization and geometry.

Thus, the bimodular polypeptide scaffold provided herein solves numerous problems associated with existing platforms for assembling bispecific molecules. In this respect, limitations of existing platforms include: the formation of undesired homodimer side products; a long assembly reaction time; the need to re-clone, re-express, and re-purify each binder arm into a specialised half-chain scaffold; the need for custom assembly and purification setups; and single format with little to no ability to accommodate structural changes.

In contrast, the bimodular polypeptide scaffold provided herein: provides a core dimerisation module that enables format tunability; minimises the formation of monospecific, bivalent biproducts ('homodimers'); facilitates bispecific antibody assembly reaction compatible with a range of temperatures, pH values, buffers, and salt concentrations; avoids the instability and the need for time-consuming optimisation that may arise from bispecific assembly relying on disulfide bond reduction and oxidation; and is simple and scalable. Moreover, the scaffold may be completely genetically encoded, with no use of unnatural amino acids, and can be expressed to high yields in *Escherichia coli*.

While the system has been exemplified using the SpyTag/SpyCatcher system, it will be understood that this is merely a representative protease-activatable bimodular system. It will be evident to the person skilled in the art that the SpyCatcher modules in the polypeptide scaffold exemplified herein may be substituted with any polypeptide binding partner of a two-part linker system (e.g. any of those disclosed in WO2011/098772, WO 2016/193746, WO 2018/197854, WO2020/183198 and WO2022/214795 all herein incorporated by reference), whilst retaining the advantageous properties of the system described in the Examples.

Accordingly, in one aspect the invention may be seen to provide a bivalent polypeptide scaffold comprising:

(a) a first peptide-binding domain linked to a second peptide-binding domain by a first linker domain; and

(b) a blocking peptide linked to the first or second peptide-binding domain by a second linker domain comprising a proteolytic cleavage site,

wherein:

(i) the first peptide-binding domain comprises a binding site for a first cognate peptide and is capable of spontaneously forming an isopeptide bond with the first cognate peptide;

(ii) the second peptide-binding domain comprises a binding site for a second cognate peptide and is capable of spontaneously forming an isopeptide bond with the second cognate peptide; and

(iii) the blocking peptide binds selectively and reversibly to the binding site of the cognate peptide in the peptide-binding domain to which the blocking peptide is linked via the second linker domain.

Alternatively, the bivalent polypeptide scaffold defined herein may be viewed as a protease-activatable bimodular system for the sequential (stepwise) assembly of molecules to form a single structure or an assembled complex (e.g. for the assembly of a bispecific).

As described in more detail below, the blocking peptide functions to mask or block the activity of the peptide-binding domain in the module linked to the blocking peptide. Thus, under conditions in which the scaffold is used (i.e. conditions suitable for the peptide-binding domains to spontaneously form an isopeptide bond with their cognate peptides), the blocking peptide binds selectively and reversibly to the binding site of the cognate peptide in the peptide-binding domain to which the blocking peptide is linked via the second linker domain, such that the blocking peptide inhibits the peptide-binding domain from spontaneously forming an isopeptide bond with its cognate peptide. Alternatively viewed, the blocking peptide inhibits the peptide-binding domain to which the blocking peptide is linked via the second linker domain from spontaneously forming an isopeptide bond with its cognate peptide, i.e. by binding selectively and reversibly to the binding site of the cognate peptide in the peptide-binding domain.

Moreover, under conditions in which the scaffold is used (i.e. conditions suitable for the peptide-binding domains to spontaneously form an isopeptide bond with their cognate peptides), cleavage of the second linker domain via the proteolytic cleavage site allows the peptide-binding domain that was linked to

peptide to spontaneously form an isopeptide bond with its cognate peptide. In other words, cleavage of the second linker domain via the proteolytic cleavage site allows the blocking peptide to dissociate from the peptide-binding domain, such that the peptide-binding domain is available to spontaneously form an isopeptide bond with its cognate peptide.

As noted above, the polypeptide scaffold can be completely genetically encoded, with no use of unnatural amino acids, and can be expressed to high yields in *E. coli*.

Accordingly, a further aspect provided herein is a nucleic acid molecule encoding a bivalent polypeptide scaffold as defined herein.

Thus, also provided is a vector comprising the nucleic acid molecule defined herein.

Furthermore, a cell is provided (e.g. host cell, such a prokaryotic host cell) comprising the nucleic acid molecule or vector as defined herein, e.g. a cell expressing a bivalent polypeptide scaffold as defined herein.

Another aspect provided herein is a process for producing or expressing the bivalent polypeptide scaffold as defined herein comprising the steps of:

a) providing a cell as defined herein (e.g. transforming or transfecting a cell with a vector as defined herein);

b) culturing the cell under conditions which allow the expression of the bivalent polypeptide scaffold; and optionally

c) isolating the bivalent polypeptide scaffold.

As discussed elsewhere herein, the polypeptide scaffold may comprise a purification tag or moiety, e.g. a peptide purification tag such as a His-tag, to facilitate its isolation.

Also provided herein is the use of a bivalent polypeptide scaffold as defined herein to conjugate a first molecule and a second molecule, wherein each molecule is linked to a cognate peptide capable of spontaneously forming an isopeptide bond with a peptide-binding domain in the polypeptide scaffold.

Alternatively viewed, provided herein is the use of a bivalent polypeptide scaffold as defined herein to sequentially assemble (i.e. to assemble stepwise) a first molecule and a second molecule into a single structure (assembled complex), wherein each molecule is linked to a cognate peptide capable of spontaneously forming an isopeptide bond with a peptide-binding domain in the polypeptide scaffold.

Still further provided herein is a method for conjugating two molecules comprising:

(i) contacting a bivalent polypeptide scaffold as defined herein with a first molecule linked to a cognate peptide under conditions that enable the spontaneous formation of an isopeptide bond between the cognate peptide and a peptide-binding domain, thereby conjugating the first molecule to a peptide-binding domain of the polypeptide scaffold to form a first reaction product;

(ii) contacting the first reaction product with a protease under conditions that enable the specific cleavage of the second linker domain to form a second reaction product;

(iii) contacting the second reaction product with a second molecule linked to a cognate peptide under conditions that enable the spontaneous formation of an isopeptide bond between the cognate peptide and the other peptide-binding domain, thereby conjugating the second molecule to the other peptide-binding domain of the polypeptide scaffold to form a complex (assembled complex or structure) comprising the molecules conjugated via the polypeptide scaffold.

Alternatively viewed, provided herein is a method for sequentially assembling two molecules (stepwise assembly of two molecules) into a single structure (assembled complex), the method comprising:

(i) contacting a bivalent polypeptide scaffold as defined herein with a first molecule linked to a cognate peptide under conditions that enable the spontaneous formation of an isopeptide bond between the cognate peptide and a peptide-binding domain, thereby conjugating the first molecule to a peptide-binding domain of the polypeptide scaffold to form a first reaction product;

(ii) contacting the first reaction product with a protease under conditions that enable the specific cleavage of the second linker domain to form a second reaction product;

(iii) contacting the second reaction product with a second molecule linked to a cognate peptide under conditions that enable the spontaneous formation of an isopeptide bond between the cognate peptide and the other peptide-binding domain, thereby conjugating the second molecule to the other peptide-binding domain of the polypeptide scaffold to assemble a single structure (assembled complex) comprising the molecules conjugated via the polypeptide scaffold.

The methods may further comprise a step of isolating or purifying the complex (assembled complex) or single structure.

Yet another aspect provided herein is a kit comprising:

(a) a plurality of bivalent polypeptide scaffolds as defined herein, wherein each polypeptide scaffold comprises a different first linker domain and/or the peptide-binding domains (modules) of each polypeptide scaffold are covalently linked by a disulfide bonds at different positions, and wherein each bivalent polypeptide scaffold is provided separately (e.g. separately contained or packaged); and optionally

(b) means for isolating the polypeptide scaffold, e.g. unreacted polypeptide scaffold.

Alternatively viewed, provided herein is a kit comprising:

(a) a plurality of bivalent polypeptide scaffolds as defined herein, wherein each polypeptide scaffold allows molecules conjugated to the peptide-binding domains to be displayed in a different spatial configuration to the other polypeptide scaffolds, and wherein each bivalent polypeptide scaffold is provided separately (e.g. separately contained or packaged); and optionally

(b) means for isolating the polypeptide scaffold, e.g. unreacted polypeptide scaffold.

Further provided herein is the use of bivalent polypeptide scaffold or the kit as defined herein to prepare a library of assembled complexes or structures (e.g. bifunctional molecules).

Also provided herein is a library or plurality of assembled complexes or structures (e.g. bifunctional molecules), containing an assembled complex or structure (e.g. bifunctional molecule) comprising a first peptide-binding domain linked to a second peptide-binding domain by a linker domain (i.e. the first linker domain as defined herein), wherein each peptide-binding domain is conjugated to its cognate peptide via an isopeptide bond and wherein each cognate peptide is linked to a different molecule, and wherein:

(a) the linker domain is a first linker domain as defined herein, preferably wherein the linker domain comprises a rigid peptide, optionally a peptide comprising or consisting of an alpha helix structure, such as an alpha helix structure from a ribosomal protein, optionally the *Bacillus stearothermophilus* ribosomal protein L9; and/or

(b) an amino acid in the first peptide-binding domain is linked to an amino acid in the second peptide-binding domain by a covalent bond, optionally a disulfide bond.

DETAILED DESCRIPTION

A “bivalent polypeptide scaffold” as used herein refers to a polypeptide-based structure or construct capable of independently interacting with two ligands (e.g. cognate peptides), such that the ligands may be assembled to form a single structure (assembled complex) whereby they are indirectly linked to each other via the polypeptide-based structure. The bivalent activity of the polypeptide scaffold is provided by two polypeptide modules or subunits (e.g. monomers) that are covalently linked or conjugated via a linker domain, e.g. a peptide linker, wherein each module is able to bind to a ligand. Thus, the polypeptide scaffold may be viewed a bimodular polypeptide system for the assembly of a molecules to form a single structure. The terms “bivalent polypeptide scaffold”, “bimodular polypeptide scaffold” and “polypeptide scaffold” are used interchangeably herein.

The peptide-binding domains of the polypeptide modules must be capable of binding to the same ligand (i.e. cognate peptide). Thus, the polypeptide scaffold may be viewed as being “monospecific” and/or comprising a “homodimer”. However, as discussed in more detail below, the peptide-binding domains may be capable of binding specifically to more than one ligand, e.g. variant peptides having a common or consensus sequence. Thus, while it is preferred that the peptide-binding domains are capable of binding to the same ligand(s) with the same activity, it is not essential that the same ligands are used when assembling the single structure (assembled complex, e.g. bispecific). In fact, an advantage of the system provided herein is that existing molecules containing the ligands may be used with the polypeptide scaffold to generate new libraries of assembled complexes, e.g. bifunctional structures. Similarly, it will be evident that the amino acid sequences of the peptide-binding domains may differ to impart the polypeptide scaffold with particular structural properties. Thus, the peptide-binding domains may not form a structural homodimer insofar as the sequences of the domains may not be identical, although structural homodimers are contemplated herein. Accordingly, the terms monospecific and homodimer as used herein refer particularly to the functional properties of the polypeptide scaffold, i.e. the peptide-binding domains have the same activity or essentially the same activity as each other with respect to their ability to interact with their ligands, e.g. form isopeptide bonds with their cognate peptides.

It will be evident from the description above that the terms “ligand” and “cognate peptide”, “peptide” or “peptide tag” may be used interchangeably herein.

As discussed in more detail below, the polypeptide scaffold may contain other domains (e.g. peptide or polypeptide domains) that are capable of interacting with a binding partner. For instance, the scaffold may contain a purification tag or moiety, such as a purification peptide tag (e.g. His-tag), that can interact specifically with a binding partner to facilitate the purification or isolation of the polypeptide scaffold or assembled complex. However, it will be understood that the interactions of these other domains are non-covalent and/or reversible interactions, such that the polypeptide scaffold contains only two polypeptide modules that are capable of forming irreversible covalent bonds (e.g. amide bonds) with their ligands. Thus, where the polypeptide scaffold contains an additional domain (e.g. peptide domain) that is directly adjacent to a peptide-binding domain, the polypeptide module may be viewed as comprising a peptide-binding domain. Alternatively, where the scaffold contains an additional domain (e.g. peptide domain) that is indirectly adjacent to a peptide binding domain (e.g. separated by a spacer or linker domain), the polypeptide module may be viewed as consisting of a peptide-binding domain.

The term “polypeptide-based” means that at least the peptide-binding domains of the polypeptide scaffold consist of polypeptides. In this respect, it is contemplated that other domains in the scaffold may not be peptides or polypeptides. For instance, a purification tag or moiety may be a small organic molecule, such as biotin. Furthermore, it is possible that the modules comprising the peptide-binding domains are linked by synthetic or chemical linkers (e.g. linker groups), as defined further below. However, an advantageous aspect of the polypeptide scaffold of the invention is that it may be fully genetically encoded. Thus, in a preferred aspect, all domains of the polypeptide scaffold consist of peptides or polypeptides that are linked by peptide bonds (i.e. the polypeptide scaffold consists of amino acids). As it is contemplated that the polypeptide scaffold may comprise post-translational modifications, e.g. *in vivo* modifications made during protein synthesis in a host cell or *in vitro* modifications, such as the addition of a purification tag or moiety, the polypeptide scaffold may be viewed as consisting essentially of peptides or polypeptides that are linked by peptide bonds (i.e. the polypeptide scaffold consists essentially of amino acids). Thus, polypeptide-based may be viewed as meaning at least 60%, 70%, 80%, 85%, 90% or 95% of the polypeptide scaffold is comprised of amino acids linked by peptide bonds (e.g. assessed via molecular weight). In preferred aspects, the polypeptide scaffold is a polypeptide.

The term “peptide-binding domain” as used herein refers to a polypeptide binding partner of a two-part linker as mentioned above. In brief, a two-part linker, i.e. a peptide tag and its polypeptide binding partner (a so-called peptide tag/binding partner pair) may be derived from a protein capable of spontaneously forming an isopeptide bond (an isopeptide protein), wherein the domains of the protein are expressed separately to produce a peptide tag that comprises one of the residues involved in the isopeptide bond (e.g. an aspartate or asparagine) and a polypeptide binding partner (or “catcher”) that comprises the other residue involved in the isopeptide bond (e.g. a lysine) and at least one other residue required to form the isopeptide bond (e.g. a glutamate). Mixing the peptide tag and polypeptide binding partner results in the spontaneous formation of an isopeptide bond between the tag and binding partner. Thus, by separately fusing the peptide tag and binding partner to different molecules, e.g. proteins, it is possible to covalently link said molecules together via an isopeptide bond formed between the peptide tag and polypeptide binding partner, i.e. to form a polypeptide-based linker between the molecules fused to the peptide tag and polypeptide binding partner. The peptide tag and polypeptide binding partner that react with each other (e.g. specifically and efficiently) to form an isopeptide bond may be defined as a pair of peptide linkers, particularly a cognate pair of peptide linkers. Alternatively, the polypeptide binding partner may be viewed as a polypeptide comprising a peptide-binding domain that reacts with its cognate peptide to form an isopeptide bond.

Thus, a peptide-binding domain refers to a polypeptide binding partner of a two-part linker that may be designed or derived directly from an isopeptide protein, e.g. the polypeptide may be a fragment of a native isopeptide protein or a variant thereof (i.e. a functional variant that retains the isopeptide bond forming activity). Thus, as mentioned above, a peptide-binding domain must comprise at least one amino acid residue, e.g. lysine, that is involved in the formation of an isopeptide bond via the complementary, reactive amino acid residue in the cognate peptide (e.g. asparagine or aspartic acid). Typically, the cognate peptide is designed or derived from the same isopeptide protein as the polypeptide binding partner to which it binds covalently, i.e. its corresponding polypeptide binding partner.

Generally, a polypeptide binding partner is larger than its corresponding peptide tag and comprises or consists of a larger fragment or portion of the isopeptide protein compared to the peptide tag. In particular, in addition to comprising a residue that is involved in the formation of the isopeptide bond (i.e. a

lysine or asparagine/aspartate) the peptide binding partner comprises a glutamic acid or aspartic acid residue that facilitates or induces the formation of the isopeptide bond between the peptide tag and polypeptide binding partner.

Thus, a polypeptide binding partner may comprise a fragment of an isopeptide protein which overlaps with a fragment designed to constitute a peptide tag or may comprise a discrete and separate fragment of the isopeptide protein compared to that of the peptide tag. Thus, the sequence of the polypeptide binding partner may overlap with that of the designed peptide tag or the peptide tag and polypeptide binding partner may comprise or consist of two discrete fragments of the isopeptide protein. It is contemplated that the peptide tag may not be based on the sequence of the isopeptide protein, e.g. the peptide tag (peptide linker) may be designed *de novo*.

Whilst there is no particular limit on the size of a polypeptide binding partner (i.e. a peptide-binding domain), practically it is preferable to minimise the size of the domains in the polypeptide scaffold.

Thus, the peptide-binding domain (i.e. polypeptide binding partner) may be between 50-300 amino acids in length, e.g. 60-250, 70-200, 80-150 amino acids in length, e.g. it may comprise or consist of 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190 or 200 amino acids.

A cognate peptide (i.e. a peptide tag) may be between 6-50 amino acids in length, e.g. 7-45, 8-40, 9-35, 10-30, 11-25 amino acids in length, e.g. it may comprise or consist of 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 amino acids.

Accordingly, peptide-binding domain is the polypeptide binding partner (i.e. “catcher”) part of a pair of peptide linkers comprising a peptide tag and a polypeptide binding partner (“catcher”), wherein said peptide linkers react spontaneously to form an isopeptide bond. Thus, a peptide-binding domain may be viewed as one part of a two-part linker, i.e. the peptide-binding domain may be referred to herein as a peptide linker. Similarly, a cognate peptide refers to the peptide tag part of a two-part linker and thus may also be referred to herein as a peptide linker.

Thus, a peptide linker (e.g. peptide tag and/or peptide binding partner) does not consist of the entire protein sequence of an isopeptide protein and is shorter in length. For instance, a peptide linker may comprise less than 5, 10, 20, 30, 40 or 50% of the number of amino acid residues present in the isopeptide protein.

Whilst the component parts of a two-part linker can be based upon a sequence of an isopeptide protein (particularly one or more fragments thereof), it will be readily understood by the skilled person that the sequence of each peptide linker may differ from the sequence of the portion of the isopeptide protein from which it is derived. Thus, each part of the two-part peptide linker may comprise mutations or alterations as compared to the sequence of the isopeptide protein from which it is derived. As discussed below, some mutations may be introduced to the peptide linker sequence to improve the stability and/or function of the peptide linker, e.g. to improve the reaction rate of the spontaneous isopeptide bond formation between the peptide linkers.

Thus, in some embodiments, a peptide linker may comprise or consist of a fragment of an isopeptide protein, wherein the fragment fulfils the size criteria set forth above and comprises at least 70, 75, 80, 85, 90, 95, 96, 97, 98, 99 or 100% sequence identity to a comparable region of the isopeptide protein from which is was derived.

Isopeptide proteins may be identified by searching for structural homologues of known isopeptide proteins, i.e. proteins with sequence similarity or identity to known isopeptide proteins. Such homologues may be viewed as functionally equivalent proteins and may find utility in the production of two-part linker systems that can be used in the polypeptide scaffold provided herein.

Thus, a peptide-binding domain and its cognate peptide (i.e. two-part linker) for use in the polypeptide scaffold provided herein may be derived from any suitable isopeptide protein. As mentioned above, various isopeptide proteins are known in the art and numerous two-part linkers (i.e. tag/catcher pairs) have been developed, see e.g. WO 2011/098772, WO 2016/193746, WO 2018/197854, WO 2020/183198 and WO 2022/214795 (all herein incorporated by reference), SdyTag/SdyCatcher (Tan et al., 2016, PloS ONE 11(10): e0165074, incorporated herein by reference) and GalacDock and GalacTag. Any such linkers may provide the peptide-binding domain of the polypeptide scaffold provided herein and the cognate peptide for use with the polypeptide scaffold provided herein.

In a preferred aspect, the peptide-binding domain and cognate peptide are derived from the isopeptide protein FbaB, which is from *Streptococcus pyogenes*. The FbaB protein comprises a domain, CnaB2, which spontaneously forms one isopeptide bond. A portion of the amino acid sequence of the FbaB protein containing the CnaB2 domain is set out in SEQ ID NO: 39 (the CnaB2 domain

consists of residues 17-118 of SEQ ID NO: 39). The isopeptide bond in the CnaB2 domain forms between a lysine at position 31 of SEQ ID NO: 39 and an aspartic acid residue at position 117 of SEQ ID NO: 39. The glutamic acid residue which induces the isopeptide bond is at position 77 of SEQ ID NO: 39. Thus, a peptide-binding domain and cognate peptide developed from the isopeptide protein domain set forth in SEQ ID NO: 39 may comprise a peptide-binding domain comprising a fragment of the FbaB protein comprising the reactive lysine at position 31 and the glutamic acid residue at position 77, and a cognate peptide comprising a fragment of the protein comprising the reactive aspartic acid at position 117. The peptide-binding domain and cognate peptide may contain sequences from outside the CnaB2 domain of the FbaB protein.

The RrgA protein is an adhesion protein from *Streptococcus pneumoniae*, which has an amino acid sequence as set out in SEQ ID NO: 40. An isopeptide bond forms spontaneously between lysine at position 742 in SEQ ID NO: 40 and asparagine at position 854 in SEQ ID NO: 40. The bond is induced by a glutamic acid residue at position 803 in SEQ ID NO: 40. Thus, a peptide-binding domain and cognate peptide developed from the isopeptide protein set forth in SEQ ID NO: 40 will comprise a fragment of the protein comprising the reactive asparagine at position 854 and a fragment of the protein comprising the reactive lysine at position 742, wherein one of the fragment also will contains the glutamic acid residue at position 803.

ACE19, a domain of an adhesin protein from *Enterococcus faecalis*, also spontaneously forms an isopeptide bond. ACE19 has an amino acid sequence as set forth in SEQ ID NO: 41. The isopeptide bond occurs between a lysine residue at position 181 of SEQ ID NO: 41 and an asparagine residue at position 294 of SEQ ID NO: 41. The bond is induced by an aspartic acid residue at position 213 in SEQ ID NO: 41. Thus, a pair of peptide linkers developed from isopeptide protein set forth in SEQ ID NO: 41 will comprise a peptide linker comprising a fragment of the protein comprising the reactive asparagine residue at position 294 and a peptide linker comprising a fragment of the protein comprising the reactive lysine residue at position 181, wherein one of the peptide linkers will also contain the aspartic acid residue at position 213.

The collagen binding domain from *Staphylococcus aureus*, which has an amino acid sequence set out in SEQ ID NO: 42, comprises one spontaneously formed isopeptide bond. The isopeptide bond occurs between lysine at position

176 of SEQ ID NO: 42 and asparagine at position 308 of SEQ ID NO: 42. The aspartic acid residue which induces the isopeptide bond is at position 209 of SEQ ID NO: 42. Thus, a pair of peptide linkers developed from the isopeptide protein set forth in SEQ ID NO: 42 will comprise a peptide linker comprising a fragment of the protein comprising the reactive lysine at position 176 and a peptide linker comprising a fragment of the protein comprising the reactive asparagine at position 308, wherein one of the peptide linkers will also contain the aspartic acid residue at position 209.

The PsCs protein is a fragment of the por secretion system C-terminal sorting domain protein from *Streptococcus intermedius*, which has an amino acid sequence as set out in SEQ ID NO: 43. An isopeptide bond is formed between lysine at position 405 in SEQ ID NO: 43 and aspartate at position 496 in SEQ ID NO: 43. Thus, a pair of peptide linkers developed from the isopeptide protein set forth in SEQ ID NO: 43 will comprise a peptide linker comprising a fragment of the protein comprising the reactive aspartate at position 496 and a peptide linker comprising a fragment of the protein comprising the reactive lysine at position 405.

The major pilin protein Spy0128 has an amino acid sequence as set out in SEQ ID NO: 44. Two isopeptide bonds are formed in the protein. One isopeptide bond is formed between lysine at position 179 in SEQ ID NO: 44 and asparagine at position 303 in SEQ ID NO: 44. The glutamic acid residue which induces the spontaneous isopeptide bond is found at position 258 in SEQ ID NO: 44. Thus, a pair of peptide linkers developed from an isopeptide protein set forth in SEQ ID NO: 44 may comprise a peptide linker comprising a fragment of the protein comprising the reactive asparagine at position 303 and a peptide linker comprising a fragment of the protein comprising the reactive lysine at position 179, wherein one of the peptide linkers will also contain the glutamic acid residue at position 258.

The other isopeptide bond in the major pilin protein Spy0128 occurs between the lysine residue at position 36 of SEQ ID NO: 44 and the asparagine residue at position 168 of SEQ ID NO: 44. The glutamic acid residue which induces isopeptide formation is found at position 117 in SEQ ID NO: 44. Thus, a pair of peptide linkers developed from an isopeptide protein set forth in SEQ ID NO: 44 may comprise a peptide linker comprising a fragment of the protein comprising the reactive lysine residue at position 36 and a peptide linker comprising a fragment of the protein comprising the reactive asparagine at position 168, wherein one of the peptide linkers will also contain the glutamic acid residue at position 117.

Thus, a peptide-binding domain for use in polypeptide scaffold and its cognate peptide may be derived from an isopeptide protein comprising an amino acid sequence as set forth in any one of SEQ ID NOs: 39-44 or a protein with at least 70% sequence identity to an amino acid sequence as set forth in any one of SEQ ID NOs: 39-44. As discussed above, it is preferred that the peptide-binding domain and cognate peptide are derived from the same isopeptide protein.

Alternatively viewed, a peptide-binding domain for use in polypeptide scaffold and its cognate peptide may each comprise a fragment or portion of an isopeptide protein comprising an amino acid sequence as set forth in any one of SEQ ID NOs: 39-44 or a functional variant thereof, wherein the functional variant comprises at least 70% sequence identity to the amino acid sequence of a corresponding fragment of a sequence as set forth in any one of SEQ ID NOs: 39-44. The term “functional variant” means that the variant fragment or portion must retain the ability to spontaneously form an isopeptide bond with its partner. For instance, a peptide-binding domain comprising a functional variant of a fragment of an isopeptide protein as defined above must be capable of spontaneously forming an isopeptide bond with a cognate peptide derived from the isopeptide protein (e.g. a cognate peptide comprising a fragment or portion of the isopeptide protein) under suitable conditions.

The isopeptide protein sequence, peptide-binding domain or cognate peptide may be at least 75, 80, 85, 90, 95, 96, 97, 98 or 99% identical to the sequence to which it is compared (e.g. SEQ ID NOs: 39-44 or a fragment or portion thereof). The sequence identity definition applies to all of the polypeptides and peptides disclosed herein.

It will be understood that the fragment or portion must satisfy the size criteria defined above for the peptide tag and polypeptide binding partner.

As a peptide tag (cognate peptide) derived from the isopeptide protein will be smaller than the polypeptide binding partner (peptide-binding domain), it may be preferable to define functional variants of the peptide tag by reference to the number of modifications rather than via sequence identity. Thus, a functional variant of a peptide tag may comprise 1-10, 1-9, 1-8, 1-7, 1-6 or 1-5 modifications, such as 1, 2, 3, 4 or 5 modifications (e.g. substitutions, additions and/or deletions, preferably substitutions and/or additions) relative to the sequence to which it is compared.

Sequence identity may be determined by any suitable means known in the art, e.g. using the SWISS-PROT protein sequence databank using FASTA pep-cmp

with a variable pamfactor, and gap creation penalty set at 12.0 and gap extension penalty set at 4.0, and a window of 2 amino acids. Other programs for determining amino acid sequence identity include the BestFit program of the Genetics Computer Group (GCG) Version 10 Software package from the University of Wisconsin. The program uses the local homology algorithm of Smith and Waterman with the default values: Gap creation penalty - 8, Gap extension penalty = 2, Average match = 2.912, Average mismatch = -2.003.

Preferably said comparison is made over the full length of the sequence (which may be a corresponding fragment or portion), but may be made over a smaller window of comparison, e.g. less than 200, 100 or 50 contiguous amino acids.

Preferably such sequence identity-related proteins, polypeptides and peptides are functionally equivalent to the protein, polypeptides and peptides which are set forth in the recited SEQ ID NOs. In other words, the sequence identity-related proteins, polypeptides and peptides are functional variants of the reference proteins, polypeptides and peptides. As referred to herein, "functional equivalence" refers to variants of the proteins, polypeptides and peptides discussed herein that may show some reduced efficacy in their function (e.g. spontaneously forming isopeptide bonds) relative to the parent molecule (i.e. the molecule with which it shows sequence identity), but preferably are as efficient or are more efficient.

As discussed above, numerous peptide tag/polypeptide binding partner pairs (i.e. tag/catcher pairs or two-part linkers) have been developed from the isopeptide proteins discussed above and any such pairs may find utility in the present invention. For instance, a pair has been developed from the CnaB2 domain of the FbaB protein described above, which is known as SpyTag/SpyCatcher (SEQ ID NO: 13/SEQ ID NO: 45), and functional variants of this pair with improved activity have been developed, which are known as SpyTag002/SpyCatcher002 (SEQ ID NO: 14/SEQ ID NO: 46) and SpyTag003/SpyCatcher003 (SEQ ID NO: 15/SEQ ID NO: 1), see e.g. WO2011/098772, WO 2018/197854 and WO2020/183198 (all herein incorporated by reference). Notably, all of the SpyTag variants are capable of reacting with all of the SpyCatcher variants. Accordingly, the first and second peptide-binding domains may each comprise or consist of a SpyCatcher polypeptide or a functional variant thereof and the first and second cognate peptides may each comprise or consist of a SpyTag peptide or a functional variant thereof. Any SpyTag peptide and SpyCatcher polypeptide may find utility in

the present invention. In particular, the SpyTag peptide may be SpyTag003 (SEQ ID NO: 15) or a functional variant thereof and the SpyCatcher polypeptide may be SpyCatcher003 (SEQ ID NO: 1) or a functional variant thereof. Thus, the first and second peptide-binding domains may each comprise or consist of a Catcher polypeptide derived from a FbaB protein or a functional variant thereof and the first and second cognate peptides may each comprise or consist of a cognate peptide of the Catcher polypeptide derived from a FbaB protein or a functional variant thereof.

Similarly, numerous peptide tag/polypeptide binding partner pairs (i.e. tag/catcher pairs or two-part linkers) have been developed from the RrgA protein described above. These are known as RrgATag/RrgACatcher (SEQ ID NO: 47/SEQ ID NO: 48), SnoopTag/SnoopCatcher (SEQ ID NO: 49/SEQ ID NO: 50) and DogTag/DogCatcher (SEQ ID NO: 51/SEQ ID NO: 52), and functional variants of these pairs have been developed, see e.g. WO 2016/193746 and WO 2022/214795 (both herein incorporated by reference). Thus, the first and second peptide-binding domains may each comprise or consist of a Catcher polypeptide derived from RrgA or a functional variant thereof and the first and second cognate peptides may each comprise or consist of a cognate peptide of the Catcher polypeptide derived from RrgA or a functional variant thereof.

In particular, the first and second peptide-binding domains may each comprise or consist of a SnoopCatcher polypeptide (SEQ ID NO: 50) or a functional variant thereof and the first and second cognate peptides may each comprise or consist of a SnoopTag peptide (SEQ ID NO: 49) or a functional variant thereof. Any SnoopTag peptide and SnoopCatcher polypeptide may find utility in the present invention.

Furthermore, the first and second peptide-binding domains may each comprise or consist of a DogCatcher polypeptide (SEQ ID NO: 52) or a functional variant thereof and the first and second cognate peptides may each comprise or consist of a DogTag peptide (SEQ ID NO: 51) or a functional variant thereof. Any DogTag peptide and DogCatcher polypeptide may find utility in the present invention.

As noted above, the term “functional variant” means that the variant polypeptide or peptide must retain the ability to spontaneously form an isopeptide bond with its partner. For instance, a peptide-binding domain comprising a functional variant of “Catcher” as defined above must be capable of spontaneously forming an isopeptide bond with its cognate peptide tag under suitable conditions.

A functional variant “Catcher” or “Tag” may have at least 70, 75, 80, 85, 90, 95, 96, 97, 98 or 99% sequence identity to the sequence to which it is compared (e.g. SEQ ID NOs: 1-15 or 45-52). As noted above, a functional variant “Tag” may be defined by reference to the number of modifications relative to the reference sequence as defined above.

Thus, the first and second peptide-binding domains each may comprise an amino acid sequence with at least 70% sequence identity to a sequence as set forth in any one of SEQ ID NOs: 1-12, 45, 46, 48, 50 or 52, wherein the peptide-binding domain is capable of spontaneously forming an isopeptide bond with its cognate peptide tag under suitable conditions.

The cognate peptide tag for SEQ ID NOs: 1-12 and sequence identity-related variants thereof may comprise or consist of an amino acid sequence as set forth in SEQ ID NO: 15. The cognate peptide tag for SEQ ID NO: 45 and sequence identity-related variants thereof may comprise or consist of an amino acid sequence as set forth in SEQ ID NO: 13. The cognate peptide tag for SEQ ID NO: 46 and sequence identity-related variants thereof may comprise or consist of an amino acid sequence as set forth in SEQ ID NO: 14. However, as noted above, all of the SpyTag peptides are capable of reacting with all of the SpyCatcher polypeptides. Thus, any of the SpyTag peptides (e.g. SEQ ID NOs: 13-15) may be viewed as the cognate peptide for any of the SpyCatcher polypeptides.

The cognate peptide tag for SEQ ID NO: 48 and sequence identity-related variants thereof comprises or consists of an amino acid sequence as set forth in SEQ ID NO: 47. The cognate peptide tag for SEQ ID NO: 50 and sequence identity-related variants thereof comprises or consists of an amino acid sequence as set forth in SEQ ID NO: 49. The cognate peptide tag for SEQ ID NO: 52 and sequence identity-related variants thereof comprises or consists of an amino acid sequence as set forth in SEQ ID NO: 51.

In a preferred aspect, the first and second peptide-binding domains each may comprise an amino acid sequence with at least 70% sequence identity (e.g. at least 80% sequence identity or more as defined elsewhere herein) to a sequence as set forth in any one of SEQ ID NOs: 1-12, 45 or 46 (preferably a sequence as set forth in any one of SEQ ID NOs: 1-12), wherein the amino acid sequence comprises a lysine at position 31 and a glutamic acid at position 77, wherein the specified amino acid residues are at positions equivalent to the positions in SEQ ID NO: 1. As noted above, the peptide-binding domain is capable of spontaneously

forming an isopeptide bond with its cognate peptide tag under suitable conditions, e.g. a peptide tag comprising or consisting of an amino acid sequence as set forth in SEQ ID NO: 15.

As discussed above, the polypeptide scaffold is monospecific, meaning that the peptide-binding domains are capable of spontaneously forming an isopeptide bond with the same cognate peptide, preferably with the same activity or essentially the same activity as each other under the same conditions. Thus, the peptide-binding domains may be defined by reference to the cognate peptide with which they form an isopeptide bond. Thus, while the amino acid sequences of the first and second peptide-binding domains may differ, both domains must be capable of spontaneously forming an isopeptide bond, preferably with the same cognate peptide, with the same activity or essentially the same activity as each other under the same conditions. Alternatively viewed, the first and second peptide-binding domains may be independently selected from the peptide-binding domains as defined herein, wherein both domains must be capable of spontaneously forming an isopeptide bond with the same cognate peptide, preferably with the same activity or essentially (substantially) the same activity as each other under the same conditions.

First and second peptide-binding domains with the same activity or essentially (substantially) the same activity may be viewed as domains which have activity that is “equivalent" to each other. Similarly, functional variants as defined herein may have the same activity or substantially the same activity as the reference molecule (i.e. peptide or polypeptide, e.g. a peptide-binding domain as defined herein) from which they are derived. Thus, a peptide-binding domain and/or functional variant with activity that is “equivalent" to the activity of a reference molecule may have activity that is similar (i.e. comparable) to the activity of the reference molecule (e.g. a peptide-binding domain comprising or consisting of an amino acid sequence as set forth in SEQ ID NO: 1), i.e. such that the practical applications of the peptide-binding domain and/or functional variant are not significantly affected, e.g. within a margin of experimental error. Thus, an equivalent peptide-binding domain and/or functional variant activity means that the peptide-binding domain or variant polypeptide is capable of spontaneously forming an isopeptide bond with a cognate peptide tag (e.g. comprising or consisting of an amino acid sequence as set forth in one of SEQ ID NOs: 13-15, preferably SEQ ID NO: 15) with a similar reaction rate (i.e. rate constant as discussed below) and/or

yield to a reference peptide-binding domain, such as a peptide-binding domain comprising or consisting of an amino acid sequence as set forth in SEQ ID NO: 1 under the same conditions.

The activities of different peptide-binding domains (e.g. SEQ ID NO: 1 versus a variant thereof) are measured under the same reaction conditions, e.g. temperature, substrates (i.e. peptide tag) and their concentration, buffer, salt etc., and can be readily compared to determine whether the activity for each peptide-binding domain is higher, lower or equivalent. A peptide-binding domain or functional variant with equivalent activity may show some reduced efficacy in the spontaneous formation of an isopeptide bond with its cognate peptide (e.g. lower reaction rate or activity in a limited range of reaction conditions (e.g. narrower temperature range, such as 10-30 °C etc.)) relative to the reference molecule (i.e. the molecule with which it shows sequence identity), but preferably is as efficient or is more efficient.

In particular, the peptide-binding domains and functional variants defined herein may have an equivalent rate constant to a Catcher polypeptide as discussed above, such as a SpyCatcher polypeptide, particularly SpyCatcher003 (SEQ ID NO: 1). The rate constant refers to the coefficient of proportionality relating the rate of the reaction (the formation of an isopeptide bond) at a given temperature to the product of the concentrations of reactants (i.e. the product of the concentration of the peptide tag and peptide binding partner).

Thus, the activity, e.g. rate constant, of the peptide-binding domain or functional variant as defined herein may be at least 60%, e.g. at least 70, 75, 80, 85 or 90% of the activity, e.g. rate constant, of a polypeptide comprising or consisting of an amino acid sequence as set forth in SEQ ID NO: 1-12, 45, 46, 48, 50 or 52, such as at least 91, 92, 93, 94, 95, 96, 97, 98 or 99% of the activity, e.g. rate constant, of a polypeptide comprising or consisting of an amino acid sequence as set forth in SEQ ID NO: 1-12, 45, 46, 48, 50 or 52. Alternatively viewed, the activity of the peptide-binding domain or functional variant may be no more than 40% lower than the activity, e.g. rate constant, of a polypeptide comprising or consisting of an amino acid sequence as set forth in SEQ ID NO: 1-12, 45, 46, 48, 50 or 52, e.g. no more than 35, 30, 25 or 20% lower than the activity, e.g. rate constant, of a polypeptide comprising or consisting of an amino acid sequence as set forth in SEQ ID NO: 1-12, 45, 46, 48, 50 or 52, such as no more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1% lower than the activity, e.g. rate constant, of a polypeptide comprising or

consisting of an amino acid sequence as set forth in SEQ ID NO: 1-12, 45, 46, 48, 50 or 52.

Notably, the rate constant of the reaction of a peptide-binding domain and its cognate peptide may be lower than the values described in the Examples or prior art when the cognate peptides are fused to large molecules (e.g. proteins), which diffuse slower than the isolated peptide tag and peptide binding partner. Moreover, the rate constant may be reduced if the molecules to which the cognate peptides are fused cause steric hindrance to the reaction. Similarly, the rate constant of the peptide-binding domain and its cognate peptide may be reduced when the peptide-binding domain is part of the polypeptide scaffold relative to an isolated peptide-binding domain. Accordingly, when measuring the rate constant of the reaction of the peptide-binding domains and functional variants disclosed herein, it is preferred that measurement is performed using isolated cognate peptides and peptide-binding domains, i.e. cognate peptides that are not fused or conjugated to other molecules (or fused to small molecules) and peptide-binding domains that are not part of the polypeptide scaffold. Thus, the values above preferably refer to the activity of isolated peptide-binding domains, i.e. domains that are not part of the polypeptide scaffold.

However, it will be understood that it is often convenient to measure the rate constant of the reaction of the peptide-binding domains and functional variants using a cognate peptide that is fused to a polypeptide. Thus, when measuring and comparing the rate constants of different peptide-binding domains and functional variants using a cognate peptide that is fused to a polypeptide, it is preferred that a polypeptide fused to the peptide tag is the same size, preferably the same sequence, in all reactions.

As noted above, the peptide-binding domain may be a SpyCatcher polypeptide or functional variant thereof, particularly at SpyCatcher003 polypeptide or functional variant thereof. In this respect, it has been demonstrated that various amino acids may be substituted in SpyCatcher (SEQ ID NO: 45) to improve its activity or other properties, e.g. rate of reaction. In particular, positions 2, 9, 13, 19, 34, 37, 50, 59, 62, 69, 86, 89, 91, 94, 97, 103, 105, 108 and 113 of SEQ ID NO: 45 may be substituted with any of the amino acids defined below to improve the function of the polypeptide:

1) threonine at position 2;

2) glycine at position 9;

3) proline at position 13;
4) threonine at position 19;
5) glutamic acid at position 34;
6) arginine at position 37;
7) serine at position 50;
8) threonine at position 59;
9) histidine at position 62;
10) tyrosine or serine at position 69;
11) valine at position 86;
12) proline at position 89;
13) glutamic acid at position 91;
14) isoleucine at position 94;
15) aspartic acid at position 97;
16) aspartic acid at position 103;
17) glutamic acid at position 105;
18) glutamic acid at position 108; and
19) threonine at position 113.

Notably, the presence of an aspartic acid residue at position 2 of SpyCatcher (SEQ ID NO: 45) may result in the formation of an unwanted side-reaction – a polypeptide dimer wherein the polypeptides are conjugated via an isopeptide bond. The presence of threonine or alanine at position 2 eliminates the unwanted side-reaction and may further improve the rate of reaction the polypeptide with its cognate peptide tag. Thus, it is preferred that the amino acid at position 2 of the SpyCatcher peptide-binding domain is not aspartic acid or glutamic acid. In particular, the amino acid at position 2 of the SpyCatcher peptide-binding domain may be threonine or alanine, preferably threonine.

Thus, it is contemplated that the peptide-binding domain may comprise a polypeptide having amino acid sequence with at least 70% sequence identity to an amino acid sequence as set forth in any one of SEQ ID NOs: 1-12, 45 or 46, wherein the amino acid sequence comprises a lysine at position 31 and a glutamic acid at position 77 and any combination of the specified amino acid residues defined above. However, some particularly preferred combinations include:

a) 1) threonine or alanine at position 2;
2) proline at position 13;
3) lysine at position 31;

4) arginine at position 37;
5) histidine at position 62;
6) glutamic acid at position 77;
7) glutamic acid at position 91;
8) aspartic acid at position 103;
9) glutamic acid at position 105;
10) glutamic acid at position 108; and
11) threonine at position 113;

b) 1) threonine or alanine at position 2;
2) proline at position 13;
3) lysine at position 31;
4) arginine at position 37;
5) histidine at position 62;
6) glutamic acid at position 77;
7) proline at position 89;
8) glutamic acid at position 91;
9) aspartic acid at position 103;
10) glutamic acid at position 105;
11) glutamic acid at position 108; and
12) threonine at position 113;

c) 1) threonine or alanine at position 2;
2) proline at position 13;
3) lysine at position 31;
4) arginine at position 37;
5) histidine at position 62;
6) glutamic acid at position 77;
7) proline at position 89;
8) glutamic acid at position 91;
9) aspartic acid at position 97;
10) aspartic acid at position 103;
11) glutamic acid at position 105;
12) glutamic acid at position 108; and
13) threonine at position 113; and

d) 1) threonine or alanine at position 2;
2) proline at position 13;

3) lysine at position 31;
4) arginine at position 37;
5) threonine at position 59
6) histidine at position 62;
7) glutamic acid at position 77;
8) proline at position 89;
9) glutamic acid at position 91;
10) isoleucine at position 94;
11) aspartic acid at position 97;
12) aspartic acid at position 103;
13) glutamic acid at position 105;
14) glutamic acid at position 108; and
15) threonine at position 113;

wherein the specified amino acid residues are at positions equivalent to the positions in SEQ ID NO: 1.

Thus, a peptide-binding domain may be a polypeptide comprising an amino acid sequence with at least 70% (e.g. at least 75, 80, 85, 90, 95, 96, 97, 98 or 99%) sequence identity to an amino acid sequence as set forth in SEQ ID NOs: 1-12, 45 or 46 (preferably SEQ ID NOs: 1-12 or 1-4), wherein the peptide-binding domain comprises a lysine at position 31, a glutamic acid at position 77 and all of the following:

1) threonine or alanine (preferably threonine) at position 2;
2) glycine at position 9;
3) proline at position 13;
4) threonine at position 19;
5) glutamic acid at position 34;
6) arginine at position 37;
7) serine at position 50;
8) threonine at position 59;
9) histidine at position 62;
10) tyrosine at position 69;
11) glycine at position 83;
12) valine at position 86;
13) proline at position 89;
14) glutamic acid at position 91;

15) isoleucine at position 94;
16) aspartic acid at position 97;
17) aspartic acid at position 103;
18) glutamic acid at position 105;
19) glutamic acid at position 108; and
20) threonine at position 113;

wherein the specified amino acid residues are at positions equivalent to the positions in SEQ ID NO: 1.

If the peptide-binding domain does not contain all of the residues specified above, it is preferred that in the specified positions the variants contain the amino acid residues at the equivalent positions in the SpyCatcher or SpyCatcher002 polypeptides (SEQ ID NOs: 45 or 46), preferably SpyCatcher002. The equivalent positions can readily be determined by comparing the amino acid sequence of the peptide-binding domain with SEQ ID NO: 45 or 46, e.g. using the BLASTP algorithm.

Thus, by way of example, if peptide-binding domain comprises an amino acid sequence with at least 70% sequence identity to a sequence as set forth in SEQ ID NO: 1 and the residue at position 91 (or the equivalent position) is not glutamic acid, it is preferred that the residue is threonine. Similarly, if the residue at position 103 (or the equivalent position) is not aspartic acid, it is preferred that the residue is asparagine. If the residue at position 108 (or the equivalent position) is not glutamic acid, it is preferred that the residue is lysine. If the residue at position 89 (or the equivalent position) is not proline, it is preferred that the residue is alanine. If the residue at position 97 (or the equivalent position) is not aspartic acid, it is preferred that the residue is glutamine. This applies to other residues specified below.

Thus, a peptide-binding domain functional variant may differ from SEQ ID NOs: 1-12, 45 or 46 (particularly SEQ ID NOs: 1-4) by, for example, 1 to 50, 1 to 45, 1 to 40, 1 to 35, 1 to 30, 1 to 25, 1 to 20, 1 to 15, 1 to 10, 1 to 8, 1 to 6, 1 to 5, 1 to 4, e.g. 1, 2 or 3 amino acid substitutions, insertions and/or deletions, preferably 1 to 30, 1 to 25, 1 to 23, 1 to 20, 1 to 15, 1 to 10, 1 to 8, 1 to 6, 1 to 5, 1 to 4, e.g. 1, 2 to 3 amino acid substitutions and/or 1 to 30, 1 to 25, 1 to 20, 1 to 15, 1 to 10, 1 to 8, 1 to 6, 1 to 5, 1 to 4, e.g. 1, 2 or 3 amino acid deletions. It is preferred that deletions are at the N- and/or C-terminus, i.e. truncations, preferably N-terminal truncations, thereby generating polypeptide portions of SEQ ID NOs: 1-12, 45 or 46.

Any substitutions that are present in the peptide-binding domain relative to the exemplified polypeptides (e.g. SEQ ID NOs: 1-12, 45 or 46) may be conservative amino acid substitutions, unless specified otherwise. A conservative amino acid substitution refers to the replacement of an amino acid by another which preserves the physicochemical character of the polypeptide (e.g. D may be replaced by E or vice versa, N by Q, or L or I by V or vice versa). Thus, generally the substituting amino acid has similar properties, e.g. hydrophobicity, hydrophilicity, electronegativity, bulky side chains etc. to the amino acid being replaced. Isomers of the native L-amino acid e.g. D-amino acids may be incorporated.

The peptide-binding domain functional variants defined above (e.g. variants having the residues 1-11, 1-12, 1-13, 1-15 or 1-20 as defined above) may also comprise one or more (i.e. two, three, four or five) of the following substitutions relative to SEQ ID NO: 1:

1) substitution of lysine with glutamine at position 52 (i.e. glutamine at position 52);

2) substitution of valine with aspartic acid at position 63 (i.e. aspartic acid at position 63);

3) substitution of tyrosine with serine at position 69 (i.e. serine at position 69);

4) substitution of threonine with leucine at position 88 (i.e. leucine at position 88); and

5) substitution of glutamic acid with proline at position 96 (i.e. proline at position 96);

wherein the specified amino acid residues are at positions equivalent to the positions in SEQ ID NO: 1.

The polypeptide variants of the invention may comprise amino acid deletions relative to SEQ ID NO: 1. While N-terminal deletions are preferred, in some embodiments the polypeptide variant may comprise a deletion at the position equivalent to 17 of SEQ ID NO: 1, i.e. deletion of a methionine at position 17 of SEQ ID NO: 1.

As mentioned above, it has been determined that the modules containing the peptide-binding domains (e.g. the peptide-binding domains) can be “locked” in various configurations (spatial orientations) by engineering each peptide-binding domain to contain amino acids (e.g. cysteine residues) that react with each other to

form a covalent bond (e.g. disulfide bond). This facilitates the control of the orientation of the molecules assembled on the polypeptide scaffold. Accordingly, it may be advantageous to substitute some of the preferred residues specified above with cysteine. In particular, the amino acids at positions 50, 52, 62 and 69 may be cysteine residues instead of the residues specified above or found in the SpyCatcher or SpyCatcher002 polypeptides. As shown in the Examples below, additional residues in the SpyCatcher-based peptide-binding domains may be substituted with cysteine, particularly residues 4, 12, 14 and 51.

While the residues specified above apply particularly to SpyCatcher-based peptide-binding domains (i.e. peptide-binding domains comprising or consisting of a SpyCatcher polypeptide or functional variant thereof), it will be evident that equivalent substitutions may be identified in other peptide-binding domains defined herein and thus the teaching applies generally to all peptide-binding domains.

Accordingly, the polypeptide scaffold defined herein may comprise an amino acid in the first peptide-binding domain that is linked to an amino acid in the second peptide-binding domain by a covalent bond, such as a disulfide bond. Thus, the first and second peptide-binding domains each may contain a cysteine residue, wherein the cysteine residues react with each other to form a disulfide bond.

In particular, the first and second peptide-binding domains may each comprise an amino acid sequence with at least 70% sequence identity to a sequence as set forth in any one of SEQ ID NOs: 1-4, wherein the amino acid sequence comprises a lysine at position 31 and a glutamic acid at position 77 and a cysteine residue at position 4, 12, 14, 50, 51, 52, 62 or 69, wherein the specified amino acid residues are at positions equivalent to the positions in SEQ ID NO: 1.

As demonstrated in the Examples, the position of the cysteine residue in each peptide-binding domain may be selected to produce a polypeptide scaffold with a specific spatial configuration. In particular, the peptide-binding domains defined above may be selected to provide any one of the following combinations:

(a) the first peptide-binding domain has a cysteine residue at position 50 and the second peptide-binding domain has a cysteine residue at position 50;

(b) the first peptide-binding domain has a cysteine residue at position 4 and the second peptide-binding domain has a cysteine residue at position 14;

(c) the first peptide-binding domain has a cysteine residue at position 12 and the second peptide-binding domain has a cysteine residue at position 51;

(d) the first peptide-binding domain has a cysteine residue at position 69 and the second peptide-binding domain has a cysteine residue at position 69; or

(e) the first peptide-binding domain has a cysteine residue at position 52 and the second peptide-binding domain has a cysteine residue at position 62.

Alternatively viewed, the polypeptide scaffold may contain:

(a) a first peptide-binding domain comprising:

(i) an amino acid sequence selected from a sequence as set forth in any one of SEQ ID NOs: 1-12, 45 or 46, preferably any one of SEQ ID NOs: 1, 4-6, 8, 10 or 11; or

(ii) an amino acid sequence selected from a sequence as set forth in any one of SEQ ID NOs: 1-4, 45 or 46 wherein:

(1) the leucine residue at position 4 is substituted with cysteine;

(2) the glycine residue at position 12 is substituted with cysteine;

(3) the serine residue at position 14 is substituted with cysteine;

(4) the serine residue at position 50 is substituted with cysteine;

(5) the glycine residue at position 51 is substituted with cysteine;

(6) the lysine residue at position 52 is substituted with cysteine;

(7) the histidine or glutamine residue at position 62 is substituted with cysteine; or

(8) the tyrosine residue at position 69 is substituted with cysteine; and/or

(b) a second peptide-binding domain comprising:

(i) an amino acid sequence selected from a sequence as set forth in any one of SEQ ID NOs: 1-12, 45 or 46, preferably any one of SEQ ID NOs: 1, 4, 5, 7, 9, 10 or 12; or

(ii) an amino acid sequence selected from a sequence as set forth in any one of SEQ ID NOs: 1-4, 45 or 46, wherein:

(1) the leucine residue at position 4 is substituted with cysteine;

(2) the glycine residue at position 12 is substituted with cysteine;

(3) the serine residue at position 14 is substituted with cysteine;

(4) the serine residue at position 50 is substituted with cysteine;

(5) the glycine residue at position 51 is substituted with cysteine;

(6) the lysine residue at position 52 is substituted with cysteine;

(7) the histidine or glutamine residue at position 62 is substituted with cysteine; or

(8) the tyrosine residue at position 69 is substituted with cysteine.

The term “linker” as used herein refers to molecules that function to link, i.e. conjugate or join, two domains, molecules or components together, preferably by a covalent bond, e.g. an amide bond, particularly a peptide bond. As noted above, the peptide-binding domains of the polypeptide scaffold and its cognate peptide may be viewed as a two-part linker, wherein formation of the isopeptide bond between the first part, i.e. peptide-binding domain, and second part, i.e. cognate peptide, reconstitutes the linker, thereby joining a molecule conjugated to second part of the linker to the polypeptide scaffold, i.e. assembling the molecule conjugated to second part of the linker on the polypeptide scaffold. Alternatively stated, the peptide-binding domain and its cognate peptide may be viewed as a cognate pair that functions as a linker, i.e. a peptide tag and polypeptide cognate pair or a peptide tag and binding partner cognate pair. These terms are used interchangeably throughout the description.

The term “cognate” refers to components that function together. Thus, in the context of the present invention, a cognate pair refers to a peptide tag and a peptide-binding domain that react together spontaneously to form an isopeptide bond. Thus, a two-part linker comprising a peptide tag and polypeptide that react together efficiently to form an isopeptide bond under conditions that enable the spontaneous formation of said isopeptide bond can also be referred to as being a “complementary pair”, i.e. a peptide tag and polypeptide complementary pair.

Thus, when the peptide-binding domain is a SpyCatcher polypeptide or functional variant thereof, a cognate peptide refers to a SpyTag peptide or functional variant thereof (e.g. a peptide comprising an amino acid sequence set forth in one of SEQ ID NOs: 13-15) with which the peptide-binding domain reacts spontaneously to form an isopeptide bond. The cognate peptide may be a peptide comprising an amino acid sequence with at least 80% (e.g. at least 85, 90, 91, 92, 93, 94, 95, 96, 97, 98 or 99%) sequence identity to an amino acid sequence as set forth in one of SEQ ID NOs: 13-15 that is capable of spontaneously forming an isopeptide bond with a polypeptide comprising an amino acid sequence as set forth in SEQ ID NO: 1, e.g. between an aspartic acid in the cognate peptide tag (i.e. an aspartic acid at position equivalent to position 7 in SEQ ID NO: 13, position 8 in SEQ ID NO: 14 or position 10 in SEQ ID NO: 15) and the lysine residue at position 31 of SEQ ID NO: 1.

Functional variants of SpyTag are described in the art (see e.g. WO2020/183198 herein incorporated by reference) and elsewhere herein. Any such variant may find utility in the present invention.

Thus, the cognate peptide tag may comprise or consist of an amino acid sequence as set forth in SEQ ID NO: 13, 14 or 15.

The lysine residue at position 31 of a SpyCatcher peptide-binding domain defined herein (e.g. SEQ ID NO: 1-12, 45 or 46) spontaneously forms an isopeptide bond with the aspartic acid residue at position 8 in SpyTag002 (SEQ ID NO: 14), at position 7 of SpyTag (SEQ ID NO: 13) or position 10 of SpyTag003 (SEQ ID NO: 15) under various conditions including those explained below that are suitable for the formation of an isopeptide bond between said peptide tag and peptide-binding domain. It is known in the art that SpyCatcher polypeptides (and other "Catcher" polypeptides defined herein) are active under a range of conditions and capable of reacting with a variety of peptide tags (particularly SEQ ID NOs: 13-15).

For instance, the peptide-binding domain (and hence the polypeptide scaffold disclosed herein) may be active (i.e. capable of spontaneously forming an isopeptide bond with a peptide tag as described herein) in a variety of buffers including phosphate buffered saline (PBS), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), HEPES buffered saline (HBS), and Tris buffered saline (TBS), both with and without EDTA. The peptide-binding domain (and hence the polypeptide scaffold disclosed herein) may be active at a pH of about 3.0-8.0, e.g. 4.0-7.0, 5.0-7.0, such as about 5.5-6.5, over a wide range of temperatures, e.g. 0-40 °C, e.g. 1, 2, 3, 4, 5, 10, 12, 15, 18, 20, 22, 25, 28, 30, 35 or 37 °C, preferably about 25-35 °C, e.g. about 25 °C. The peptide-binding domain (and hence the polypeptide scaffold disclosed herein) also may be active in the presence of the commonly used detergents, such as Tween 20 and Triton X-100, e.g. up to a concentration of about 1% (v/v), and in the presence of urea, e.g. up to a concentration of about 3 M. The skilled person would readily be able to determine other suitable conditions.

Thus, conditions that are suitable for the formation of an isopeptide bond between a peptide-binding domain (and hence the polypeptide scaffold disclosed herein) and its cognate peptide (e.g. a peptide comprising or consisting of an amino acid sequence as set forth in SEQ ID NOs: 13-15) includes any conditions in which contacting the cognate peptide and peptide-binding domain (and hence the polypeptide scaffold disclosed herein) results in the spontaneous formation of an

isopeptide bond between said cognate peptide and peptide-binding domain (and hence the polypeptide scaffold disclosed herein), particularly between the aspartic acid residue at position 10 of SEQ ID NO: 15 and the lysine residue at position 31 of SpyCatcher or a functional variant thereof (e.g. SEQ ID NO: 1). For instance, contacting the cognate peptide and peptide-binding domain (and hence the polypeptide scaffold disclosed herein) in buffered conditions, e.g. in a buffered solution or on a solid phase (e.g. column) that has been equilibrated with a buffer, such as PBS. The step of contacting may be at any suitable pH, such as pH 3.0-8.0, e.g. 4.0-7.0, such as pH 4.2, 4.4, 4.6, 4.8, 5.0, 5.2, 5.4, 5.6, 5.8, 6.0, 6.2, 6.4, 6.6, 6.8 or 7.0. Additionally or alternatively, the step of contacting may be at any suitable temperature, such as about 0-40°C, e.g. about 1-39, 2-38, 3-37, 4-36, 5-35, 6-34, 7-33, 8-32, 9-31 or 10-30°C, e.g. about 10, 12, 15, 18, 20, 22, 25, 28, 30, 33, 35 or 37 °C, preferably about 25-35 °C, e.g. about 25 °C.

Contacting a cognate peptide and a peptide-binding domain (and hence the polypeptide scaffold disclosed herein) “under conditions that enable the spontaneous formation of an isopeptide bond” includes contacting the cognate peptide and peptide-binding domain (and hence the polypeptide scaffold disclosed herein) in the presence of a chemical chaperone, e.g. a molecule that enhances or improves the reactivity of the cognate peptide and peptide-binding domain). The chemical chaperone may be TMAO (trimethylamine N-oxide), which may be present in the reaction at a concentration of at least about 0.2 M, e.g. at least 0.3, 0.4, 0.5, 1.0, 1.5, 2.0 or 2.5 M, e.g. about 0.2-3.0 M, 0.5-2.0 M, 1.0-1.5 M.

As noted above, the formation of the isopeptide bond between the cognate peptide and a peptide-binding domain is spontaneous. For instance, when the peptide-binding domain is a SpyCatcher polypeptide or functional variant thereof, it comprises a glutamic acid at position 77 (or an equivalent position, based on the numbering of SEQ ID NO: 1) that facilitates, e.g. induces, promotes or catalyses, the formation of the isopeptide bond between the aspartate and lysine residues in the cognate peptide and peptide-binding domain, respectively.

The term "spontaneous" as used herein refers to an isopeptide bond, which can form in a protein or between peptides or proteins (e.g. between two peptides or a peptide and a protein, i.e. the peptide tag and polypeptide (peptide-binding domain)) without any other agent (e.g. an enzyme catalyst) being present and/or without chemical modification of the protein or peptide, e.g. without native chemical ligation or chemical coupling using 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide

(EDC). Thus, native chemical ligation to modify a peptide or protein having a C-terminal thioester is not carried out.

Thus, a spontaneous isopeptide bond can form between a cognate peptide and the polypeptide scaffold via a peptide-binding domain (an unmasked or uncaged peptide-binding domain) when in isolation and without chemical modification of the cognate peptide and/or peptide-binding domain. A spontaneous isopeptide bond may therefore form of its own accord in the absence of enzymes or other exogenous substances and without chemical modification of the cognate peptide and/or peptide-binding domain.

A spontaneous isopeptide bond may form almost immediately after contact of the cognate peptide and polypeptide scaffold, e.g. within 1, 2, 3, 4, 5, 10, 15, 20, 25 or 30 minutes, or within 1, 2, 4, 8, 12, 16, 20 or 24 hours.

The speed of isopeptide bond formation will be dependent on the concentration of the cognate peptide and polypeptide scaffold and the conditions of the reaction, e.g. temperature. Spontaneous isopeptide bond formation may complete for about 50% or more of the reactants in about 5 minutes or less, e.g. where the reactants are each present at a concentration of about 2-10 μM at a reaction temperature of about 25-35°C. Spontaneous isopeptide bond formation may complete for about 90% or more of the reactants in about 120 minutes or less, e.g. where the reactants are each present at a concentration of about 2-10 μM at a reaction temperature of about 25-35°C. It will be evident that speed of isopeptide bond formation may differ for each module of the polypeptide scaffold. However, isopeptide bond formation typically forms within the parameters defined above, e.g. within about 12-15 hours for at about 80% of the reactants at a reaction temperature of about 25-35°C.

The other reaction conditions, e.g. buffer, pH etc. used to determine the speed of reaction defined above may be any conditions defined herein. In some embodiments, the reaction conditions are those used in the Examples.

As discussed above, peptide-binding domain may be mutant form of the polypeptide binding partner (i.e. “Catcher” polypeptide), also referred to herein as homologues, variants or derivatives. Such mutant forms may be truncated versions of the polypeptide binding partners (i.e. “Catcher” polypeptides) disclosed herein. Such mutant forms are functionally equivalent to the parent molecule as defined above.

Thus, it is contemplated that the peptide-binding domain may be a truncated form of a polypeptide binding partner (i.e. “Catcher” polypeptide) as defined herein. In this respect, it is known in the art that the N-terminus and/or C-terminus of SpyCatcher polypeptides may be deleted without substantially reducing the activity of the polypeptide. In particular, the SEQ ID NOs: 1-12, 45 or 46 may be truncated by up to 21 amino acids at the N-terminus (e.g. 5, 10, 15 or 20 amino acids) and/or by up to 9 amino acids at the C-terminus (e.g. 1, 2, 3, 4, 5, 6, 7, 8 or 9 amino acids), preferably 5 amino acids or fewer. Thus, the term variant as used herein includes truncation variants of the disclosed polypeptides.

As referred to herein a "portion" comprises at least an amino acid sequence comprising residues 22-104 of SEQ ID NO: 1-12, 45 or 46, i.e. at least 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 100, 105, 110 or more amino acids of SEQ ID NOs: 1-12, 45 or 46 containing at least residues 22-104 of said sequences. Thus, said portion may be obtained from a central or N-terminal or C-terminal portion of the sequence. Preferably said portion is obtained from the central portion, i.e. it comprises an N-terminal and/or C-terminal truncation, preferably an N-terminal truncation, as defined above. Notably, "portions" as described herein are polypeptides and therefore satisfy the identity (relative to a comparable region) conditions and functional equivalence conditions mentioned herein. Thus, a portion may be viewed as a functional variant as defined herein.

A cognate peptide for use with the polypeptide scaffold may be a variant of the sequences described herein, e.g. may differ from SEQ ID NOs: 13-15 by, for example, 1 to 5, 1 to 4, e.g. 1, 2 to 3 amino acid substitutions, insertions and/or deletions, preferably substitutions, as defined above. However, the cognate peptide variants must retain their functional activity, i.e. their ability to spontaneously form an isopeptide bond with their cognate binding partner (i.e. the peptide-binding domain).

The terms “polypeptide binding partner”, “binding partner” or “catcher” as used herein generally refer to a polypeptide.

There is no standard definition regarding the size boundaries between what is meant by peptide. Typically, a peptide may be viewed as comprising between 2-39 amino acids. Accordingly, a polypeptide may be viewed as comprising at least 40 amino acids, preferably at least 50, 60, 70, 80, 90, 100 or 110 amino acids.

Thus, a cognate peptide as defined herein may be viewed as comprising at least 12 amino acids, e.g. 12-39 amino acids, such as e.g. 13-35, 14-34, 15-33, 16-

31, 17-30 amino acids in length, e.g. it may comprise or consist of 12, 13, 14, 15, 16, 17, 18, 19 or 20 amino acids.

A polypeptide as defined herein may be viewed as comprising at least 80 amino acids, e.g. 80-500 amino acids, such as e.g. 80-400, 80-350, 80-300 amino acids in length, e.g. it may comprise or consist of 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119 or 120 amino acids. The term "polypeptide" is used herein interchangeably with the term "protein". As noted above, the polypeptide scaffold contains more than one polypeptide domain and thus the size of the polypeptide scaffold may be at least double the sizes specified above. Thus, the term polypeptide or protein also may be viewed as any amino acid sequence comprising 150-1000, 50-900, 60-800, 70-700, 80-600, 90-500, 100-400 amino acids.

The polypeptide scaffold contains a blocking peptide linked to the first or second peptide-binding domain by a linker domain comprising a proteolytic cleavage site. The blocking peptide functions to mask or block the activity of the peptide-binding domain in the module linked to the blocking peptide. Alternatively viewed, the blocking peptide cages or inhibits the isopeptide formation activity of the peptide-binding domain to which it is linked. Thus, under conditions in which the scaffold is used (i.e. conditions suitable for the peptide-binding domains to spontaneously form an isopeptide bond with their cognate peptides), the blocking peptide binds selectively and reversibly to the binding site of the cognate peptide in the peptide-binding domain to which the blocking peptide is linked via the second linker domain such that it inhibits the peptide-binding domain from spontaneously forming an isopeptide bond with its cognate peptide. Alternatively viewed, the blocking peptide inhibits the peptide-binding domain to which the blocking peptide is linked via the second linker domain from spontaneously forming an isopeptide bond with its cognate peptide, i.e. by binding selectively and reversibly to the binding site of the cognate peptide in the peptide-binding domain.

Moreover, under conditions in which the scaffold is used (i.e. conditions suitable for the peptide-binding domains to spontaneously form an isopeptide bond with their cognate peptides), cleavage of the second linker domain via the proteolytic cleavage site allows the peptide-binding domain that was linked to peptide to spontaneously form an isopeptide bond with its cognate peptide. In other words, cleavage of the second linker domain via the proteolytic cleavage site allows

the blocking peptide to dissociate from the peptide-binding domain (i.e. to be displaced from the binding site), such that the peptide-binding domain is available to spontaneously form an isopeptide bond with its cognate peptide.

The term "binding site" refers to the part or region of the peptide-binding domain with which the cognate peptide interacts, e.g. the part or region containing the amino acid residue that reacts to form an isopeptide bond with the cognate peptide. For instance, the binding site in a SpyCatcher polypeptide contains the lysine residue at position 31.

As shown in the Examples below, a blocking peptide may be produced by substituting the amino acid residue in the cognate peptide with an amino acid that cannot react with the residue in the binding site of the peptide-binding domain to form an isopeptide bond. Thus, the blocking peptide may be a non-functional or non-reactive variant of the cognate peptide, i.e. a peptide that cannot react with the peptide-binding domain to spontaneously form an isopeptide bond (e.g. under conditions that are suitable for isopeptide bond formation for a cognate peptide). For example, where the residue in the cognate peptide that reacts to form an isopeptide bond is aspartic acid, the blocking peptide comprises non-functional or non-reactive variant of the cognate peptide in which the aspartic acid is substituted with a non-acidic amino acid, e.g. alanine, serine, threonine or valine. Alternatively, where the residue in the cognate peptide that reacts to form an isopeptide bond is lysine, the blocking peptide comprises non-functional or non-reactive variant of the cognate peptide in which the aspartic acid is substituted with a non-basic amino acid, e.g. alanine, serine, threonine or valine.

Thus, the blocking peptide may be a peptide that is capable of competing with the cognate peptide for the binding site of the peptide-binding domain to which it is linked but is not capable of forming an isopeptide bond with the peptide-binding domain. A peptide that is capable of competing with the cognate peptide for the binding site of the peptide-binding domain refers to a peptide that binds to the binding site of cognate peptide with the same, similar or greater affinity than the cognate peptide. Any suitable competition assay known in the art may be used to determine whether a blocking peptide competes with a cognate peptide.

A person of skill in the art readily could design and produce a blocking peptide of a peptide-binding domain as defined herein. When the peptide-binding domain comprises or consists of a SpyCatcher polypeptide or functional variant thereof the blocking peptide may comprise or consist of a non-reactive or non-

functional SpyTag peptide (i.e. a non-functional SpyTag variant). In particular, the non-reactive SpyTag variant may be a peptide that does not contain an aspartic acid residue that reacts to form the isopeptide bond with a lysine residue in a SpyCatcher polypeptide, preferably wherein the aspartic acid residue has been substituted with a non-acidic amino acid.

Thus, the blocking peptide may be a non-functional or non-reactive SpyTag variant peptide comprising or consisting of:

(a) an amino acid sequence as set forth in SEQ ID NO: 16, wherein X at position 10 is not an acidic amino acid, preferably wherein X is selected from alanine, serine, threonine and valine; or

(b) a functional variant of (a) comprising 1-5 amino acid substitutions and/or deletions.

In particular, the blocking peptide may be a non-functional or non-reactive SpyTag variant peptide comprising or consisting of an amino acid sequence as set forth in SEQ ID NO: 17, 18 or 19.

As set out above, a linker refers to molecules that function to link, i.e. conjugate or join, two domains, molecules or components together, preferably by a covalent bond, e.g. an amide bond, particularly a peptide bond. Thus, a “linker domain” refers to a part of the polypeptide scaffold that joins two domains together via covalent bonds, particularly peptide bonds (i.e. the linker domain is covalently bonded to each domain to be joined). Thus, the first linker domain is covalently bonded to each module comprising or consisting of a peptide-binding domain such that the modules (e.g. the peptide-binding domains) are linked. As discussed further below, in addition to linking the domains (i.e. parts) of the polypeptide scaffold, the linker domain also controls the spatial configuration of the domains of the polypeptide scaffold. Thus, a linker domain has a dual function of linking domains together and spacing them apart (e.g. to ensure each domain is able to perform its function). Thus, a linker domain alternatively may be viewed as a spacer or spacer domain and these terms may be used interchangeably herein, e.g. a first peptide-binding domain is linked to a second peptide-binding domain by a first spacer domain.

The linker domain that joins the peptide-binding domains (i.e. the modules of the polypeptide scaffold comprising or consisting of the peptide-binding domains) may be used to control the configuration of the peptide-binding domains in the polypeptide scaffold. This may be achieved by alternating the size (e.g. length) of

the linker domain, its shape and/or its rigidity, as shown in the Examples. Thus, any suitable linker domain that enables the peptide-binding domains of the polypeptide scaffold to react spontaneously with their cognate peptide may be used to link the peptide-binding domains. It is contemplated that the first linker domain may comprise or consist of a flexible linker or a rigid linker. The second linker domain may comprise a flexible linker.

Whilst it is preferred that the polypeptide scaffold can be produced recombinantly, it is contemplated that the domains of the polypeptide scaffold may be conjugated to each other by non-peptide bonds. In other words, the domains of the polypeptide scaffold may be produced separately by any suitable means, e.g. recombinantly, and subsequently conjugated (joined) to form the polypeptide scaffold. For instance, the peptide-binding domains may be produced recombinantly and conjugated to each other via a non-peptide linker or spacer, e.g. a chemical linker or spacer group.

Thus, domains of the polypeptide scaffold may be joined together either directly through a bond or indirectly through a linking group. Where linking groups are employed, such groups may be chosen to provide for covalent attachment of the domains through the linking group. Linking groups of interest may vary widely depending on the nature of the other entity, e.g. protein. The linking group, when present, may be biologically inert. Thus, the linking domain, particularly the first linking domain, may be a linking group.

Many linking groups are known to those of skill in the art and find use in the invention. In representative embodiments, the linking group is generally at least about 50 daltons, usually at least about 100 daltons and may be as large as 1000 daltons or larger, for example up to 1000000 daltons if the linking group contains a spacer, but generally will not exceed about 500 daltons and usually will not exceed about 300 daltons. Generally, such linking groups will comprise a spacer group terminated at either end with a reactive functionality capable of covalently bonding to the domains of the polypeptide scaffold.

Spacer groups of interest may include aliphatic and unsaturated hydrocarbon chains, spacers containing heteroatoms such as oxygen (ethers such as polyethylene glycol) or nitrogen (polyamines), peptides, polypeptides, carbohydrates, cyclic or acyclic systems that may possibly contain heteroatoms. Spacer groups may also be comprised of ligands that bind to metals such that the presence of a metal ion coordinates two or more ligands to form a complex.

Specific spacer elements include: 1,4-diaminohexane, xylylenediamine, terephthalic acid, 3,6-dioxaoctanedioic acid, ethylenediamine-N,N-diacetic acid, 1,1'-ethylenebis(5-oxo-3-pyrrolidinecarboxylic acid), 4,4'-ethylenedipiperidine, oligoethylene glycol and polyethylene glycol. Potential reactive functionalities include nucleophilic functional groups (amines, alcohols, thiols, hydrazides), electrophilic functional groups (aldehydes, esters, vinyl ketones, epoxides, isocyanates, maleimides), functional groups capable of cycloaddition reactions, forming disulfide bonds, or binding to metals. Specific examples include primary and secondary amines, hydroxamic acids, N-hydroxysuccinimidyl esters, N-hydroxysuccinimidyl carbonates, oxycarbonylimidazoles, nitrophenylesters, trifluoroethyl esters, glycidyl ethers, vinylsulfones, and maleimides. Specific linker groups include heterofunctional compounds, such as azidobenzoyl hydrazide, N-[4-(p-azidosalicylamino)butyl]-3'-[2'-pyridyldithio]propionamid), bis-sulfosuccinimidyl suberate, dimethyladipimidate, disuccinimidyltartrate, N-maleimidobutyryloxysuccinimide ester, N-hydroxy sulfosuccinimidyl-4-azidobenzoate, N-succinimidyl [4-azidophenyl]-1,3'-dithiopropionate, N-succinimidyl [4-iodoacetyl]aminobenzoate, glutaraldehyde, and succinimidyl-4-[N-maleimidomethyl]cyclohexane-1-carboxylate, 3-(2-pyridyldithio)propionic acid N-hydroxysuccinimide ester (SPDP), 4-(N-maleimidomethyl)-cyclohexane-1-carboxylic acid N-hydroxysuccinimide ester (SMCC), and the like. For instance, a spacer may be formed with an azide reacting with an alkyne or formed with a tetrazine reacting with a trans-cyclooctene or a norbornene.

Thus, it may be useful to modify one or more residues in the domains of the polypeptide scaffold to facilitate their conjugation. Thus, it is contemplated that polypeptide scaffold may comprise unnatural or non-standard amino acids, although this is a less preferred aspect of the invention.

The polypeptide scaffold or domains of the polypeptide scaffold may comprise one or more, e.g. at least 1, 2, 3, 4, 5 non-conventional amino acids, such as 10, 15, 20 or more non-conventional, i.e. amino acids which possess a side chain that is not coded for by the standard genetic code, termed herein "non-coded amino acids". Such amino acids are well known in the art and may be selected from amino acids which are formed through metabolic processes such as ornithine or taurine, and/or artificially modified amino acids such as 9H-fluoren-9-ylmethoxycarbonyl (Fmoc), (tert)-(B)utyl (o)xy (c)arbonyl (Boc), 2,2,5,7,8-

pentamethylchroman-6-sulphonyl (Pmc) protected amino acids, or amino acids having the benzyloxy-carbonyl (Z) group.

Examples of non-standard or structural analogue amino acids which may be used in the polypeptide scaffold are D amino acids, amide isosteres (such as N-methyl amide, retro-inverse amide, thioamide, thioester, phosphonate, ketomethylene, hydroxymethylene, fluorovinyl, (E)-vinyl, methyleneamino, methylenethio or alkane), L-N methylamino acids, D-α methylamino acids, D-N-methylamino acids. Further non-standard amino acids which may be used are disclosed in Willis and Chin, Nat Chem. 2018; 10(8):831-837, in Table 1 of WO2018/189517 and WO2018/197854, all of which are herein incorporated by reference.

As discussed above, it is preferred that the polypeptide scaffold can be produced recombinantly, i.e. it is fully genetically encoded. As such, it is preferred that the linker domains comprise or consist of a peptide or polypeptide. In particular, each linker domain consists of a peptide or polypeptide. Thus, the linker domains may be viewed as peptide or polypeptide linkers, or peptide or polypeptide spacers.

The sequence of the peptide linker is not limited to particular sequences. It is sufficient that the peptide linker enables each domain of the polypeptide scaffold to perform its function. Thus, the first linker domain peptide or polypeptide must enable each peptide-binding domain to react spontaneously with its cognate peptide under suitable conditions. The second linker domain peptide or polypeptide (preferably peptide) must allow the blocking peptide to bind to the peptide-binding domain to which it is linked and inhibit its function. It must also allow the proteolytic cleavage site to be cleaved by an enzyme. Thus, the length and/or sequence of each linker peptide or polypeptide may be varied. For instance, a linker domain peptide or polypeptide (i.e. peptide or polypeptide linker) may consist of about 5-300 amino acids, such as about 5-250, 5-200, 5-150 or 5-100 amino acids.

It will be understood that the length of the first linker domain peptide may vary more than the second linker domain peptide in order to provide greater options for controlling the configuration of the peptide-binding domains. Thus, the first linker domain peptide or polypeptide may consist of about 5-300 amino acids, optionally about 5-250, 5-200, 5-150 or 5-100 amino acids, such as about 6-300, 7-290, 8-280 or 9-270, 10-260, 11-250, 12-240 amino acids. However, the first linker domain peptide or polypeptide nay consist of about 6-90, 7-80, 8-70, 9-60 amino acids, such as about 8-50 or 9-40 amino acids.

In contrast, the second linker domain may consist of about 5-100 amino acids, optionally about 5-75 or 5-50 amino acids, such as about 6-49, 7-48, 8-47, 9-46, 10-45, 11-43, 12-40 amino acids, optionally about 13-35, 14-30 or 15-25 amino acids.

The nature of the residues is not critical and they may for example be any amino acid, e.g. a neutral amino acid, or an aliphatic amino acid, or alternatively they may be hydrophobic, or polar or charged or structure-forming e.g. proline. Exemplary peptide linker sequences include any combination of amino acid residues, particularly selected from S, G, L, V, P, R, H, M, A and E. It is preferred that linker does not contain C residues, particularly when C residues are present in the peptide-binding domains.

As noted above, the linker domains may be flexible or rigid. In particular, the first and/or second linker domains comprise or consist of a flexible peptide or polypeptide. Flexible peptides and polypeptides are well known in the art and any suitable flexible peptide or polypeptide may be used in the polypeptide scaffold. For instance, the flexible peptide or polypeptide may comprise or consist of a serine-rich and/or glycine-rich sequence.

Thus, the first linker domain may comprise or consist of an amino acid sequence selected from a sequence as set forth in SEQ ID NO: 20.

The second linker domain comprises a proteolytic cleavage site. Thus, the second linker domain may comprise or consist of a flexible peptide comprising a proteolytic cleavage site, optionally wherein the flexible peptide comprises a serine-rich and/or glycine-rich sequence.

As shown in the Examples, the proteolytic cleavage site may be contained in the flexible peptide, such that the amino acid sequences upstream (N-terminal) and downstream (C-terminal) of the proteolytic cleavage site comprise or consist of a flexible peptide or part thereof. Alternatively viewed, the flexible peptide may be split, such that the proteolytic cleavage site forms an internal sequence of the linker domain, i.e. an internal sequence of the flexible peptide. Thus, the second linker domain may consist of a peptide comprising a proteolytic cleavage site bordered by flexible peptides or parts thereof, e.g. serine-rich and/or glycine-rich sequences.

A proteolytic cleavage site refers to an amino acid sequence that is recognised and cleaved by an enzyme, i.e. a protease. Thus, a proteolytic cleavage site may be viewed as the recognition and/or cleavage site for a protease. It will be understood that the proteolytic cleavage site and cleavage enzyme (e.g. protease)

used in the present invention are specific, i.e. the cleavage enzyme recognises and cleaves a particular amino acid sequence (which may be a consensus amino acid sequence). This is to ensure that the cleavage enzyme does not cleavage other domains in the polypeptide scaffold and/or the molecules to be assembled on the polypeptide scaffold. Thus, any cleavage enzyme (e.g. protease) that recognises and cleaves a specific (i.e. defined) proteolytic cleavage site may be used in the invention. In particular, it is contemplated that the cleavage enzyme is a Tobacco Etch Virus (TEV) protease or thrombin protease or a functional variant thereof (i.e. a variant that retains the specific proteolytic activity). Thus, the proteolytic cleavage site may be Tobacco Etch Virus (TEV) protease recognition site or a thrombin recognition site. Accordingly, the proteolytic cleavage site may comprise or consist of an amino acid sequence as set forth in SEQ ID NO: 38 or 55. Alternatively viewed, the second linker domain may comprise an amino acid sequence as set forth in SEQ ID NO: 38 or 55. More particularly, the second linker domain may comprise or consist of an amino acid sequence as set forth in SEQ ID NO: 53, 54 or 56-58.

In view of the comments above, it will be evident that the polypeptide scaffold may be viewed as a protease-activatable bimodular system for the sequential (stepwise) assembly of molecules to form a single structure (e.g. for the assembly of a bispecific).

As noted above, it is contemplated that the first linker domain may comprise or consist of a rigid linker, particularly a rigid peptide or polypeptide linker. Rigid peptide and polypeptide structures are known in the art and any suitable rigid peptide or polypeptide may be used in the polypeptide scaffold. For instance, the rigid peptide or polypeptide may comprise or consist of an alpha helix structure or a proline-rich sequence. In particular, the alpha helix structure is from a ribosomal protein, optionally the *Bacillus stearothermophilus* ribosomal protein L9. Thus, the first linker domain may comprise or consist of an amino acid sequence selected from a sequence as set forth in SEQ ID NO: 21.

The polypeptide scaffold may also comprise purification moieties or tags to facilitate its purification. Purification may occur prior to use in the methods and uses of the invention discussed below, i.e. prior to assembly of molecules on the polypeptide scaffold. Alternatively or additionally, purification via a purification moiety or tag may occur after assembly has occurred, i.e. to purify or isolate assembled structure from unreacted components, e.g. unreacted polypeptide

scaffold. Any suitable purification moiety or tag may be incorporated into the polypeptide and such moieties are well known in the art. For instance, the polypeptide scaffold may comprise a peptide purification tag or moiety, e.g. a His-tag, FLAG-tag, Calmodulin-tag, E-tag, HA-tag, Myc-tag, S-tag, SBP-tag, Softag 1, Softag 3, Strep-tag, Strep-tag II, TC tag, V5 tag, VSV-tag or Xpress tag sequence. Such purification moieties or tags may be incorporated at any position within the polypeptide. In some preferred embodiments, the purification moiety is located at or towards (i.e. within 5, 10, 15, 20 amino acids of) the N- or C-terminus of the polypeptide scaffold. Where the peptide purification tag is used to purify or isolate the assembled structure, it is preferred that the peptide tag is not directly linked to the domain comprising the blocking peptide.

It may be useful to include one or more spacers, e.g. a peptide spacer, between the peptide purification tag and the adjacent domain of polypeptide scaffold. Suitable spacers are discussed above. For example, a spacer may comprise or consist of 1-40, more particularly 2-20, 1-15, 1-12, 1-10, 1-8, or 1-6 amino acid residues, e.g. 6, 7, 8, 9, 10 or more residues. By way of a representative example, the spacer sequence, if present, may have 1-15, 1-12, 1-10, 1-8 or 1-6 residues etc.

In particular, it may be advantageous to use a spacer comprising a proteolytic cleavage site as defined herein to facilitate cleavage of the peptide purification tag from the polypeptide scaffold. Removal of the peptide purification tag (e.g. via cleavage) may occur before or after assembly of the structure. It may be particularly important to remove the purification tag when the assembled structure is for *in vivo* use, e.g. as a medicament.

It will be evident from the Examples and the sequences of the exemplified polypeptide scaffolds that the polypeptide scaffold may contain additional sequences, particularly N-terminal or C-terminal sequences. As noted above, these sequences may contain purification tags or immobilisation domains (see below). Further terminal sequences may be added to facilitate the recombinant expression of the polypeptide scaffold. Any suitable additional sequence may be used.

It may be useful to fuse or conjugate the polypeptide scaffold to a solid substrate (i.e. a solid phase or solid support) and it will be evident that this may be achieved in any convenient way. Thus, the manner or means of immobilisation and the solid support may be selected, according to choice, from any number of immobilisation means and solid supports as are widely known in the art and

described in the literature. Thus, the polypeptide scaffold may be directly bound to the support, for example via a domain or moiety of the polypeptide scaffold (e.g. chemically cross-linked). The polypeptide scaffold may be bound indirectly by means of a linker group, or by an intermediary binding group(s) (e.g. by means of a biotin-streptavidin interaction). Thus, the polypeptide scaffold may be covalently or non-covalently linked to the solid support. The linkage may be a reversible (e.g. cleavable) or irreversible linkage. Thus, the linkage may be cleaved enzymatically, chemically, or with light, e.g. the linkage may be a light-sensitive linkage.

Thus, the polypeptide scaffold may be provided with means for immobilisation (e.g. an affinity binding partner, e.g. biotin or a hapten, capable of binding to its binding partner, i.e. a cognate binding partner, e.g. streptavidin or an antibody) provided on the support. The interaction between the polypeptide scaffold and the solid support must be robust enough to allow for washing steps, i.e. the interaction between the polypeptide scaffold and solid support is not disrupted (significantly disrupted) by the washing steps. For instance, it is preferred that with each washing step, less than 5%, preferably less than 4, 3, 2, 1, 0.5 or 0.1% of the polypeptide scaffold is removed or eluted from the solid phase.

The solid support (phase or substrate) may be any of the well-known supports or matrices which are currently widely used or proposed for immobilisation, separation etc. These may take the form of particles (e.g. beads which may be magnetic, para-magnetic or non-magnetic), sheets, resins, gels, filters, membranes, fibres, capillaries, slides, arrays or microtitre strips, tubes, plates or wells etc.

The support may be made of glass, silica, latex or a polymeric material, e.g. a polysaccharide polymer material, such as agarose (e.g. sepharose). Suitable are materials presenting a high surface area for binding of the polypeptide of the invention. Such supports may have an irregular surface and may be for example porous or particulate, e.g. particles, fibres, webs, sinters or sieves. Particulate materials, e.g. beads are useful due to their greater binding capacity, particularly polymeric beads. The solid support may be a resin, e.g. an amylose resin.

Conveniently, a particulate solid support used according to the invention will comprise spherical beads. The size of the beads is not critical, but they may for example be of the order of diameter of at least about 1 μm and preferably at least about 2 μm, 5 μm, 10 μm or 20 μm and have a maximum diameter of preferably not more than about 500 μm, and e.g. not more than about 100 μm.

Monodisperse particles, that is those which are substantially uniform in size (e.g. size having a diameter standard deviation of less than 5%) have the advantage that they provide very uniform reproducibility of reaction.

However, to aid manipulation and separation, magnetic beads are advantageous. The term "magnetic" as used herein means that the support is capable of having a magnetic moment imparted to it when placed in a magnetic field, and thus is displaceable under the action of that field. In other words, a support comprising magnetic particles may readily be removed by magnetic aggregation, which provides a quick, simple and efficient way of separating the particles following the isopeptide bond formation steps.

Similarly, it will be understood that the polypeptide scaffold may be fused or conjugated to a label, e.g. a radiolabel, a fluorescent label, luminescent label, a chromophore label as well as to substances and enzymes which generate a detectable substrate, e.g. horseradish peroxidase, luciferase or alkaline phosphatase. This detection may be applied in numerous assays where antibodies are conventionally used, including Western blotting/immunoblotting, histochemistry, enzyme-linked immunosorbent assay (ELISA), or flow cytometry (FACS) formats. Labels for magnetic resonance imaging, positron emission tomography probes and boron 10 for neutron capture therapy may also be conjugated to the polypeptide scaffold.

It may be useful to introduce a cysteine residue into the polypeptide scaffold to couple the label or a solid substrate. For instance, the introduction of a cysteine residue would allow the polypeptide scaffold to be coupled to another molecule or component, such as a label, e.g. a fluorescent label, containing a maleimide functional group. Suitable residues in the polypeptide scaffold that may be substituted with a cysteine residue are disclosed above. A cysteine residue may be included in the polypeptide scaffold in a linker or spacer domain or an additional domain (e.g. N-terminal domain) as defined above.

It will be understood that the order of the domains in the polypeptide scaffold is not critical. In this respect, the references to first and second peptide-binding domains and first and second linker domains are used for the purpose of clarity and are not intended to indicate the order of the domains in the polypeptide scaffold or the order of their function during the assembly of the structure. It will be evident that the unblocked (i.e. uncaged, unmasked, uninhibited) peptide-binding domain will react with its cognate peptide first during the assembly of the structure and this may

be the first or second peptide-binding domain, i.e. the domain that is not linked to the blocking peptide.

Thus, the polypeptide scaffold may comprise from N-terminal to C-terminal:

(i) the first peptide-binding domain;

(ii) the first linker domain;

(iii) the second peptide-binding domain;

(iv) the second linker domain comprising a proteolytic cleavage site; and

(v) the blocking peptide.

Alternatively, the polypeptide scaffold may comprise from N-terminal to C-terminal:

(i) the blocking peptide;

(ii) the second linker domain comprising a proteolytic cleavage site;

(iii) the first peptide-binding domain;

(iv) the first linker domain; and

(v) second peptide-binding domain.

The above configurations are not intended to define all of the domains in the polypeptide scaffold, although the polypeptide scaffold may consist of the above domains. As noted above, the polypeptide scaffold may contain other domains and sequences, such as purification tags or moieties, e.g. purification peptide tags. The skilled person readily could determine suitable locations for other domains in the polypeptide scaffold.

In particular, the polypeptide scaffold may comprise an amino acid sequence with at least 80% sequence identity to a sequence as set forth in any one of SEQ ID NOs: 22-28, 37 or 62-69, wherein the amino acid sequence comprises:

(i) a lysine at position 31 and 122;

(ii) a glutamic acid at position 77 and 199;

(iii) a non-acidic amino acid at position 259, preferably selected from alanine, serine, threonine or valine; and optionally

(iv) a cysteine residue at a position selected from position 4, 12, 14, 50, 51, 52, 62 and 69, preferably a position selected from position 4, 12, 50, 52 and 69; and

(v) a cysteine residue at a position selected from position 126, 134, 136, 172, 173, 174, 184 and 191, preferably a position selected from position 136, 172, 173, 184 and 191;

wherein the specified amino acid residues are at positions equivalent to the positions in SEQ ID NO: 22.

More particularly, the polypeptide scaffold may comprise or consist of an amino acid sequence as set forth in any one of SEQ ID NOs: 22-37 or 62-69.

As noted above, an advantage of the present invention arises from the fact that the polypeptide scaffold may be completely genetically encoded. Thus, further provided herein is a nucleic acid molecule encoding a polypeptide scaffold as defined above.

The nucleic acid molecule may be codon-optimised for expression in a host cell. In particular, the nucleic acid molecule may be codon-optimised for expression in a bacterial cell, such as *E. coli*. Alternatively, the nucleic acid molecule may be codon-optimised for expression in a mammalian cell, such as a human cell, e.g. a HEK cell.

The nucleic acid molecules may be made up of ribonucleotides and/or deoxyribonucleotides as well as synthetic residues, e.g. synthetic nucleotides, that are capable of participating in Watson-Crick type or analogous base pair interactions. Preferably, the nucleic acid molecule is DNA or RNA.

The nucleic acid molecules described above may be operatively linked to an expression control sequence, or a recombinant DNA cloning vehicle or vector containing such a recombinant DNA molecule. This allows cellular expression of the polypeptide scaffold as a gene product, the expression of which is directed by the gene(s) introduced into cells of interest. Gene expression is directed from a promoter active in the cells of interest and may be inserted in any form of linear or circular nucleic acid (e.g. DNA) vector for incorporation in the genome or for independent replication or transient transfection/expression. Suitable transformation or transfection techniques are well described in the literature. Alternatively, the naked nucleic acid (e.g. DNA or RNA, which may include one or more synthetic residues, e.g. base analogues) molecule may be introduced directly into the cell for the production of the polypeptide scaffold. Alternatively, the nucleic acid may be converted to mRNA by *in vitro* transcription and the relevant proteins may be generated by *in vitro* translation.

Appropriate expression vectors include appropriate control sequences such as, for example, translational (e.g. start and stop codons, ribosomal binding sites) and transcriptional control elements (e.g. promoter-operator regions, termination stop sequences) linked in matching reading frame with the nucleic acid molecules of the invention. Appropriate vectors may include plasmids and viruses (including both bacteriophage and eukaryotic viruses). Suitable viral vectors include

baculovirus and also adenovirus, adeno-associated virus, lentivirus, herpes and vaccinia/pox viruses. Many other viral vectors are described in the art. Examples of suitable vectors include bacterial and mammalian expression vectors pGEX-KG, pEF-neo and pEF-HA.

As noted above, the polypeptide scaffold may comprise additional sequences (e.g. peptide/polypeptides tags to facilitate purification of the polypeptide scaffold) and thus the nucleic acid molecule may conveniently be fused with DNA encoding an additional peptide or polypeptide as defined above to produce a fusion protein on expression.

Thus, further provided herein is a vector, preferably an expression vector, comprising a nucleic acid molecule as defined above.

Methods for preparing recombinant nucleic acid molecules are provided, said methods comprising inserting the nucleic acid molecule defined herein into vector nucleic acid.

Nucleic acid molecules provided herein, preferably contained in a vector, may be introduced into a cell by any appropriate means. Suitable transformation or transfection techniques are well described in the literature. Numerous techniques are known and may be used to introduce such vectors into prokaryotic or eukaryotic cells for expression. Preferred host cells for this purpose include insect cell lines, yeast, mammalian cell lines or *E. coli*, such as strain C41. Thus, transformed or transfected prokaryotic or eukaryotic host cells containing a nucleic acid molecule, particularly a vector as defined above, are also provided herein.

Alternatively viewed, there is provided a recombinant host cell containing a nucleic acid molecule and/or vector as described above.

By "recombinant" is meant that the nucleic acid molecule and/or vector has been introduced into the host cell. The host cell may or may not naturally contain an endogenous copy of the nucleic acid molecule, but it is recombinant in that an exogenous or further endogenous copy of the nucleic acid molecule and/or vector has been introduced.

Further provided herein is a method of preparing a polypeptide scaffold as hereinbefore defined, which comprises culturing a host cell containing a nucleic acid molecule as defined above, under conditions whereby said nucleic acid molecule encoding said polypeptide scaffold is expressed and recovering the polypeptide scaffold that is produced.

The polypeptide scaffold may be generated synthetically, e.g. by ligation of amino acids or smaller synthetically generated peptides, or more conveniently by recombinant expression of a nucleic acid molecule encoding said polypeptide scaffold as described hereinbefore.

Nucleic acid molecules provided herein may be generated synthetically by any suitable means known in the art.

Thus, the polypeptide scaffold may be an isolated, purified, recombinant or synthesised polypeptide.

Similarly, the nucleic acid molecules provided herein may be an isolated, purified, recombinant or synthesised nucleic acid molecule.

Thus, alternatively viewed, the polypeptides and nucleic acid molecules provided herein are non-native, i.e. non-naturally occurring, molecules.

Standard amino acid nomenclature is used herein. Thus, the full name of an amino acid residue may be used interchangeably with one letter code or three letter abbreviations. For instance, lysine may be substituted with K or Lys, isoleucine may be substituted with I or Ile, and so on. Moreover, the terms aspartate and aspartic acid, or glutamate and glutamic acid are used interchangeably herein and may be replaced with Asp or D, or Glu or E, respectively.

As shown in the Examples, the polypeptide scaffold has been exemplified using antibodies to assemble bispecific antibody structures. However, it will be understood by the skilled person that the utility of the polypeptide scaffold is not limited to the assembly of antibody structures or even the assembly of binding molecules. It has been demonstrated in the art that cognate peptides as defined herein may be covalently linked to a wide variety of molecules and retain their ability to spontaneously form isopeptide bonds with their polypeptide binding partner (i.e. peptide-binding domain of the polypeptide scaffold). Thus, the polypeptide scaffold may find utility in assembling structures containing any two molecules that can be linked to a cognate peptide as defined herein, i.e. the assembly of any two molecules for which it is desirable to be contained in a single structure. The molecules assembled on the scaffold and/or their functions may be the same or different. Thus, the polypeptide scaffold may find utility in the assembly of bifunctional structures, i.e. structures with two distinct and separable functions. Nevertheless, the polypeptide scaffold is particularly useful in the assembly of bispecific structures, i.e. structures that are able to bind specifically to two or more

targets, e.g. molecules or epitopes, particularly different target molecules or different epitopes on the same target molecule.

It will be understood that each binding molecule assembled on the polypeptide scaffold may bind specifically to more than one target molecule or epitope. For instance, a module (peptide-binding domain) of the polypeptide scaffold may be conjugated via its cognate peptide to a molecule, e.g. an antibody or growth factor, that binds two different types (i.e. variants) of a receptor. Thus, an assembled complex may be capable of binding to more than two target molecules or epitopes, particularly when both modules are conjugated to molecules with more than one target molecule or epitope. Thus, a single assembled complex may be capable of binding to different combinations of targets, e.g. combinations of two targets. Furthermore, if the molecule conjugated to a module of the polypeptide scaffold is itself capable of binding more than one target simultaneously, e.g. the molecule conjugated to a module of the polypeptide scaffold is bispecific, the assembled complex may be capable of binding more than two targets simultaneously. Thus, the assembled complex may be bispecific, trispecific or tetraspecific. Alternatively viewed, the assembled complex may be at least bispecific.

In this respect, the clinical success of bispecific antibodies – which bind two epitopes simultaneously – continues to grow rapidly. The mechanism of action of most bispecifics is to bridge two cell types together, but there are potent examples of bispecific therapeutics that act as molecular bridges between proteins on the surface of the same cell. As the cell-surface interactome continues to become annotated and new ligands for disease-associated receptors are discovered, there is a great potential for using bispecific ligands to probe signalling pathway pathways downstream of such interactions or to search for synergy in novel ones. However, protein-protein interactions, such as receptor dimerization for trans-phosphorylation of cytosolic domains, require precise spatial control and rigidity to effect a downstream response. Similarly, the cellular response following effector cell-disease cell interactions depends on the proximity of the two cell membranes. Bispecific binding geometries must be considered during the assembly of bispecific molecules to induce desired interactions. For bispecific molecules to achieve such broad functionality, diversity between targets must be addressed: each pair of targets will vary in relative antigen size and cell-surface density, their relative locations, and their proximity and geometric requirements for function. Therefore,

there is no 'one-size-fits-all' format for bispecific antibodies and the polypeptide scaffold provided herein therefore provides a manipulable or controllable format for assembling a plurality, e.g. library, of bispecific antibodies that vary in one or more molecular properties (e.g. affinity, epitope specificity and/or orientation of binders) that can be screened to identify and select molecules with the required function.

Thus, for example, in the context of bispecific structures alone, the polypeptide scaffold may find utility in:

1. Medium-to-high throughput assembly of bispecific binder matrices in which monospecific binder arms are combined in all possible orientations, and in a variety of different architectures;
2. Identifying novel cell-surface pairs to be targeted for therapeutic development;
3. Enhancing an existing proximity-inducing therapeutic by assembling and screening format variants which vary in size, distance between epitopes, relative epitope orientation, and rigidity;
4. Investigating bridging of novel cell-surface targets associated with disease or with unknown signalling mechanisms, and identifying tools for controlling or synergising their activity;
5. Understanding the dependencies of signalling pathways on ligand affinity, receptor organisation, and receptor oligomerisation;
6. Discovering signalling pathway activation downstream of non-heterogeneous receptor pairs, bridged by bispecific antibodies; and
7. Controlling the identity and geometry of molecules at the cell surface to fine-tune signalling, for example to control signalling potency or to induce pathway bias.

Further examples of the uses of the polypeptide scaffold include: assembling enzyme networks for catalysis; recruiting a drug (such as an antibody) to a target; recruiting stem cells to damaged tissue for tissue repair; molecular imaging and diagnostics using labelled polypeptide scaffold.

Thus, provided herein is the use of a bivalent polypeptide scaffold as defined herein to conjugate a first molecule and a second molecule, wherein each molecule is linked to a cognate peptide capable of spontaneously forming an isopeptide bond with a peptide-binding domain in the polypeptide scaffold.

Alternatively viewed, provided herein is the use of a polypeptide scaffold as defined herein to sequentially assemble (i.e. to assemble stepwise) a first molecule

and a second molecule into a single structure, wherein each molecule is linked to a cognate peptide capable of spontaneously forming an isopeptide bond with a peptide-binding domain in the polypeptide scaffold.

Still further provided herein is a method for conjugating two molecules comprising:

(i) contacting a polypeptide scaffold as defined herein with a first molecule linked to a cognate peptide under conditions that enable the spontaneous formation of an isopeptide bond between the cognate peptide and a peptide-binding domain, thereby conjugating the first molecule to a peptide-binding domain of the polypeptide scaffold to form a first reaction product;

(ii) contacting the first reaction product with a protease under conditions that enable the specific cleavage of the second linker domain to form a second reaction product;

(iii) contacting the second reaction product with a second molecule linked to a cognate peptide under conditions that enable the spontaneous formation of an isopeptide bond between the cognate peptide and the other peptide-binding domain, thereby conjugating the second molecule to the other peptide-binding domain of the polypeptide scaffold to form a complex comprising the molecules conjugated via the polypeptide scaffold.

Alternatively viewed, provided herein is a method for sequentially assembling two molecules (stepwise assembly of two molecules) into a single structure, the method comprising:

(i) contacting a polypeptide scaffold as defined herein with a first molecule linked to a cognate peptide under conditions that enable the spontaneous formation of an isopeptide bond between the cognate peptide and a peptide-binding domain, thereby conjugating the first molecule to a peptide-binding domain of the polypeptide scaffold to form a first reaction product;

(ii) contacting the first reaction product with a protease under conditions that enable the specific cleavage of the second linker domain to form a second reaction product;

(iii) contacting the second reaction product with a second molecule linked to a cognate peptide under conditions that enable the spontaneous formation of an isopeptide bond between the cognate peptide and the other peptide-binding domain, thereby conjugating the second molecule to the other peptide-binding

domain of the polypeptide scaffold to assemble a single structure comprising the molecules conjugated via the polypeptide scaffold.

It will be understood that the first molecule and a second molecule are indirectly conjugated via the polypeptide scaffold.

As mentioned above, the cognate peptide may be linked (e.g. fused or conjugated) to any molecule, component or entity (wherein the conjugate retains the isopeptide bond forming activity of the cognate peptide). Such molecules or components (i.e. entities) may be a nucleic acid molecule, protein (e.g. antibody or antigen-binding fragment thereof), peptide, small-molecule organic compound, fluorophore, metal-ligand complex, polysaccharide, nanoparticle, 2D monolayer (e.g. graphene), nanotube, polymer, cell, virus, virus-like particle or any combination of these. The component or entity to which the cognate peptide linked may be a solid support, i.e. solid substrate or solid phase, as above.

The cell may be a prokaryotic or eukaryotic cell. In some embodiments, the cell is a prokaryotic cell, e.g. a bacterial cell. In some embodiments, the cell is a eukaryotic cell, such as an animal cell, e.g. a human cell, or a yeast cell, e.g. *Pichia pastoris* (also known as *Komagataella phaffii*).

The cognate peptide may be linked to a compound or molecule which has a therapeutic or prophylactic effect, e.g. an antibiotic, antiviral, vaccine, antitumour agent, e.g. a radioactive compound or isotope, cytokines, toxins, oligonucleotides and nucleic acids encoding genes or nucleic acid vaccines.

The cognate peptide may be linked to a label as defined above.

The cognate peptide may be linked to a molecule that binds specifically and/or selectively to one or more targets (e.g. an antigen, such as a protein (e.g. receptor) or epitope thereof). Any suitable molecule with specific binding activity may be used, such as a polypeptide (e.g. antibody or antigen-binding fragment thereof, a receptor protein or ligand-binding fragment thereof) or an aptamer. Particularly, the cognate peptide may be fused or produced with a polypeptide (e.g. recombinantly produced as discussed above). For example, the cognate peptide may be fused or produced with a polypeptide that binds specifically and/or selectively to one or more targets (e.g. an antigen, such as a protein (e.g. receptor) or epitope thereof). Thus, the polypeptide that binds specifically and/or selectively to a target (e.g. an antigen, such as a protein (e.g. receptor) or epitope thereof) may be an antibody or target-binding fragment thereof.

Each binding molecule linked to a cognate peptide may specifically and/or selectively bind to a different target or different epitopes on the same target. Thus, the method and use may be viewed as producing a molecule or structure with bispecific binding activity. In particular, the first and second molecules are antibodies or binding fragments thereof, or antibody mimetics, that bind to different targets or epitopes of the same target.

As will be understood by those in the art, the immunological binding reagents encompassed by the term "antibody" extend to all antibodies and antigen (e.g. target) binding fragments thereof, including whole antibodies, dimeric, trimeric and multimeric antibodies; bispecific antibodies; chimeric antibodies; recombinant and engineered antibodies, and fragments thereof.

The term "antibody" is thus used to refer to any antibody-like molecule that has an antigen binding region, and this term includes antibody fragments that comprise an antigen binding domain such as Fab', Fab, F(ab')2, single domain antibodies (dAbs, also known as nanobodies), TandAbs dimer, Fv, scFv (single-chain Fv), dsFv, ds-scFv, Fd, linear antibodies, minibodies, diabodies, bispecific antibody fragments, bibody, tribody (scFv-Fab fusions, bispecific or trispecific, respectively); sc-diabody; kappa(lambda) bodies (scFv-CL fusions); Bispecific T-cell Engager (BiTE) (scFv-scFv tandems to attract T cells); dual variable domain (DVD)-Ig (bispecific format); small immunoprotein (SIP) (kind of minibody); SMIP ("small modular immunopharmaceutical" scFv-Fc dimer; DART (disulfide-stabilized diabody "Dual Affinity ReTargeting"); small antibody mimetics comprising one or more complementarity determining regions (CDRs) and the like.

The techniques for preparing and using various antibody-based constructs and fragments are well known in the art.

Antibodies can be fragmented using conventional techniques which are well-known and described in the art. For example, F(ab')2 fragments can be generated by treating the antibody with pepsin. The resulting F(ab')2 fragment can be treated to reduce disulfide bridges to produce Fab' fragments. Papain digestion can lead to the formation of Fab fragments. Fab, Fab' and F(ab')2, scFv, Fv, dsFv, Fd, sdAbs (nanobodies), TandAbs, ds-scFv, dimers, minibodies, diabodies, bispecific antibody fragments and other fragments can also be synthesized by recombinant techniques or can be chemically synthesized.

The antibodies or antibody fragments can be produced naturally or can be wholly or partially synthetically produced. Thus, the antibody may be from any

appropriate source, for example recombinant sources and/or produced in transgenic animals or transgenic plants, or in eggs using the IgY technology. Thus, the antibody molecules can be produced *in vitro* or *in vivo*.

Preferably, the antibody or antibody fragment comprises an antibody light chain variable region (VL) that comprises three CDR domains and an antibody heavy chain variable region (VH) that comprises three CDR domains. Said VL and VH generally form the antigen binding site.

An "Fv" fragment is the minimum antibody fragment that contains a complete antigen-recognition and binding site. This region has a dimer of one heavy chain and one light chain variable domain in tight, non-covalent association. It is in this configuration that the three hypervariable regions (CDRs) of each variable domain interact to define an antigen-binding site on the surface of the VH-VL dimer. Collectively, the six hypervariable regions (CDRs) confer antigen-binding specificity to the antibody.

However, it is well-documented in the art that the presence of three CDRs from the light chain variable domain and three CDRs from the heavy chain variable domain of an antibody is not necessary for antigen binding. Thus, constructs smaller than the above classical antibody fragment are known to be effective at binding to their target.

For example, camelid antibodies have an extensive antigen binding repertoire but are devoid of light chains. Also, results with single domain antibodies comprising VH domains alone or VL domains alone show that these domains can bind to antigen with acceptably high affinities. Thus, three CDRs can effectively bind antigen. Furthermore, it is also known that a single CDR, or two CDRs, can effectively bind antigen.

Thus, although preferred antibodies of the invention might comprise six CDR regions (three from a light chain and three from a heavy chain), antibodies with fewer than six CDR regions and as few as one or two CDR regions are encompassed by the invention. In addition, antibodies with CDRs from only the heavy chain or light chain are also contemplated, e.g. nanobodies.

The antibody or fragment thereof may be derived from any animal, such as a mouse, rat, rabbit, goat, primate, or human.

The term “antibody mimetic” refers to molecules that can bind antigens specifically but that are not structurally related to antibodies. Examples of antibody mimetics include affibodies, affilins, affimers, affitins, alphabodies, anticalins,

avimers, DARPins, Fynomers, Gastrobodies, monobodies, nanoCLAMPs, Optimers, Repebodies, Pronectins and Centryins.

The term “target” refers to any molecule or epitope thereof to which a binding molecule may bind selectively and/or specifically. In a representative example, a target may be a cell surface molecule, such as a protein, e.g. a receptor and/or glycoprotein.

Any suitable means may be used to link the cognate peptide to the molecule to be assembled on the polypeptide scaffold. As noted above, the term link or conjugate refers to a covalent linkage comprising one or more covalent bonds. Linker groups useful for linking a cognate peptide to the molecule to be assembled on the polypeptide scaffold as described above and a suitable linker group may be selected by the skilled person based on the molecule. The linker group may be a peptide, particularly when the molecule to be assembled on the polypeptide scaffold is a peptide or polypeptide. In this context, the cognate peptide may be fused to the peptide or polypeptide, i.e. the molecule to be assembled on the polypeptide scaffold may be a fusion protein (e.g. a recombinant protein) comprising the cognate peptide.

The cognate peptide used in the methods and uses provided herein may be any suitable peptide that spontaneously reacts with the peptide-binding domains of the polypeptide scaffold to form an isopeptide bond. As noted above, the polypeptide scaffold may be monospecific meaning that both peptide-binding domains are capable of reacting with the same cognate peptide. However, it will be understood that this does not mean that the molecules to be assembled on the polypeptide scaffold must contain the same cognate peptide (i.e. cognate peptides with the same amino acid sequence). Peptide-binding domains may react with a variety of related peptides. Thus, each cognate peptide may be independently selected from peptides capable of reacting with the peptide-binding domains of the polypeptide scaffold. Accordingly, is contemplated that when the peptide-binding domain comprises a SpyCatcher polypeptide or variant thereof, each cognate peptide is independently selected from a SpyTag peptide or a functional variant thereof. For instance, the cognate peptide may comprise or consist of an amino acid sequence as set forth in any one of SEQ ID NOs: 13-15, preferably SEQ ID NO: 15.

As noted above, the protease may be any protease that specifically cleaves the second linker domain of the polypeptide scaffold. In particular, the protease is a TEV protease, thrombin or a functional variant or derivative thereof.

By a “functional variant or derivative” is meant that a mutant or modified protease (i.e. containing one or more amino acid substitutions, deletions or additions relative to the protease from which is was derived), may show some reduced protease activity (e.g. cleaving the second linker domain less efficiently) relative to the activity of the protease from which it is derived in conditions that are optimum for the enzyme, e.g. in the buffer, salt and temperature conditions recommended by the manufacturer. Thus, a variant or derivative protease may be considered to be functional if it has at least 50%, e.g. at least 60, 70, 80, 85, 90, 95, 96, 97, 98, 99 or 100%, activity relative to the activity of the protease from which it was derived in conditions that are optimum for the enzyme.

It will be understood that a protease may be modified without affecting the sequence of the polypeptide. Modifications that do not affect the sequence of the polypeptide include, e.g. chemical modification, including by deglycosylation or glycosylation. Such polypeptides may be prepared by post-synthesis/isolation modification of the polypeptide without affecting functionality, e.g. certain glycosylation, methylation etc. of particular residues.

As referred to herein, to achieve "functional equivalence" the protease may show some increased or reduced efficacy in protease activity relative to the parent molecule (i.e. the molecule from which it was derived, e.g. by amino acid substitution), but preferably is as efficient or is more efficient. Thus, functional equivalence relates to a protease that has protease activity capable of cleaving the second linker domain in a reaction mixture. This may be tested by comparison of the protease activity of the derivative polypeptide relative to the polypeptide from which it is derived in a quantitative manner. The derivative is preferably at least 30, 50, 70 or 90% as effective as the parent polypeptide in the methods defined herein.

Functionally-equivalent proteins which are related to or derived from the naturally-occurring protein, may be obtained by modifying the native amino acid sequence by single or multiple amino acid substitution, addition and/or deletion (providing they satisfy the above-mentioned sequence identity requirements), but without destroying the molecule's function. Preferably the native sequence has less than 20 substitutions, additions or deletions, e.g. less than 10, 5, 4, 3, 2, or 1 such modifications. Such proteins are encoded by "functionally-equivalent nucleic acid

molecules" which are generated by appropriate substitution, addition and/or deletion of one or more bases.

Any suitable conditions that enable the specific cleavage of the second linker domain to form a second reaction product may be used in the method defined herein. Typical conditions will depend on the protease used and may be determined using routine methods known in the art. Thus, alternatively viewed, suitable conditions may be conditions (e.g. buffer, salt, temperature conditions) under which the protease is functional, e.g. displays protease activity, particularly protease activity against the second linker domain.

By “functional” is meant that the protease may show some reduced activity in degrading the second linker domain relative to the activity of the protease on the second linker domain in conditions that are optimum for the enzyme, e.g. in the buffer, salt and temperature conditions recommended by the manufacturer. Thus, the protease may be considered to be functional if it has at least 50%, e.g. at least 60, 70, 80, 85, 90, 95, 96, 97, 98, 99 or 100%, activity relative to the activity of the protease in conditions that are optimum for the enzyme.

The reaction mixture (solution) comprising the protease may contain other components, e.g. buffer, salt, etc. sufficient to ensure that the protease is functional. In a representative example, the reaction mixture comprises a TEV protease or thrombin enzyme in a buffered solution (e.g. Tris-HCl, PBS, MES, acetate buffer) having a pH of about 4.0-9.0, e.g. about 6.0-8.0 or about 7.0-8.0. Additionally or alternatively, the reaction mixture may be used at any suitable temperature, such as about 10-45°C, e.g. about 10-44, 11-43, 12-42, 13-41, 14-40, 15-39, 16-38, 17-37°C, e.g. about 10, 12, 15, 18, 20, 22, 25, 28, 30, 33, 35 or 37 °C, preferably about 30-40 °C, e.g. about 34-37 °C.

The methods may further comprise a step of isolating or purifying the complex or single structure. Any suitable means may be used to isolate or purify the complex of single structure, e.g. using a purification tag or moiety that forms part of the polypeptide scaffold. Alternatively, one or both of the molecules assembled in the complex (structure) may contain features that facilitate their purification.

In a representative example, the step of purifying the complex or single structure may comprise removing unreacted molecules comprising a cognate peptide. This may be achieved, for example, by contacting the reaction mix containing the assembled complex with an immobilised polypeptide comprising a peptide-binding domain of the polypeptide scaffold or a functional variant thereof

(e.g. bead covalently linked to the polypeptide). When the immobilised polypeptide is provided in excess, unreacted molecules comprising a cognate peptide will react with the immobilised polypeptide. Separating the immobilised polypeptide from the reaction mix (e.g. removing beads linked to the polypeptide), will therefore result in a mixture which is enriched for the assembled complex. Thus, when the peptide-binding domains of the polypeptide scaffold are SpyCatcher polypeptides, the step of purifying the complex or single structure may comprise contacting the reaction mixture (e.g. solution) containing the assembled complex with an immobilised SpyCatcher polypeptide or functional variant thereof (e.g. SpyCatcher003 or functional variant thereof). The contacting step is under conditions suitable to allow the cognate peptides linked to the molecules to spontaneously form isopeptide bonds with the immobilised SpyCatcher polypeptide or functional variant thereof (e.g. SpyCatcher003 or functional variant thereof). The method may comprise a step of separating the immobilised SpyCatcher polypeptide or functional variant thereof from the mixture containing the assembled complex. Specifically, the SpyCatcher polypeptide or functional variant thereof (e.g. SpyCatcher003 or functional variant thereof) may be immobilised on a bead or resin as defined herein.

As shown in the Examples, the methods defined herein may be multiplexed to produce a plurality of assembled complexes (e.g. bifunctional or bispecific structures). The reactions may be performed together, e.g. the method may involve contacting a plurality of polypeptide scaffolds with a first molecule linked to a cognate peptide. Alternatively, the reactions may be performed separately, e.g. in separate wells of a multiwell plate or equivalent multiplex reaction vessel. Where reactions are performed separately, the assembled complexes may be combined to form a library of assembled complexes.

Thus, another aspect provided herein is a kit comprising:

(a) a plurality of bivalent polypeptide scaffolds as defined herein, wherein each polypeptide scaffold comprises a different first linker and/or the peptide-binding domains (modules) of each polypeptide scaffold are covalently linked by a disulfide bonds at different positions, and wherein each bivalent polypeptide scaffold is provided separately (e.g. separately contained or packaged); and optionally

(b) means for isolating the polypeptide scaffold, e.g. unreacted polypeptide scaffold.

Alternatively viewed, provided herein is a kit comprising:

(a) a plurality of bivalent polypeptide scaffolds as defined herein, wherein each polypeptide scaffold allows molecules conjugated to the peptide-binding domains to be displayed in a different spatial configuration to the other polypeptide scaffolds, and wherein each bivalent polypeptide scaffold is provided separately (e.g. separately contained or packaged); and optionally

(b) means for isolating the polypeptide scaffold, e.g. unreacted polypeptide scaffold.

Any suitable means for isolating the polypeptide scaffold may be provided in the kit. Such means are described herein, e.g. immobilised polypeptides comprising the peptide-binding domain of the polypeptide scaffold, means for binding to a purification tag or moiety on the polypeptide scaffold.

Further provided herein is the use of bivalent polypeptide scaffold or the kit as defined herein to prepare a library of assembled complexes (e.g. bifunctional molecules).

Also provided herein is a library or plurality of assembled complexes (e.g. bifunctional molecules), containing an assembled complex (e.g. bifunctional molecule) comprising first peptide-binding domain linked to a second peptide-binding domain by a linker domain (i.e. the first linker domain as defined herein), wherein each peptide-binding domain is conjugated to its cognate peptide via an isopeptide bond and wherein each cognate peptide is linked to a different molecule, and wherein:

(a) the linker domain is a first linker domain as defined herein, preferably wherein the linker domain comprises a rigid peptide, optionally a peptide comprising or consisting of an alpha helix structure, such as an alpha helix structure from a ribosomal protein, optionally the *Bacillus stearothermophilus* ribosomal protein L9; and/or

(b) an amino acid in the first peptide-binding domain is linked to an amino acid in the second peptide-binding domain by a covalent bond, optionally a disulfide bond.

As used herein, the term “plurality” means two or more, e.g. at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 20 or 30 or more, such as 50, 100, 150, 200, 250, 500, 1000, 10000 or more depending on the context of the invention. For instance, a plurality of assembled complexes or polypeptides scaffolds used in the methods and kits described above, may include 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10} or more assembled complexes or polypeptides, e.g. 10^{11}, 10^{12}, 10^{13} or more.

The invention will now be described in more detail in the following non-limiting Examples with reference to the following drawings:

Figure 1 shows a cartoon depicting the DoubleCatcher technology. The DoubleCatcher bimodular scaffold shown contains an accessible N-terminal Catcher ("Reactive Catcher") and a C-terminal Catcher masked by the unreactive, blocking peptide ("Masked Catcher"). The first cognate peptide-linked binder (FabA-Tag) reacts with the N-terminal Catcher but is blocked from reacting at the Masked Catcher by the unreactive blocking peptide. The blocking peptide contains a mutation, indicated by a striped line, that prevents isopeptide bond formation with the Catcher. Addition of TEV protease to the reaction mixture results in the cleavage of the unreactive blocking peptide, thus allowing the now unmasked Catcher to bind the second cognate peptide-linked binder (FabB-Tag). Black rectangles indicate isopeptide bonds formed between Tag and Catcher. Colons indicate proteins linked by an isopeptide bond.

Figure 2 shows (A) a graph depicting the percentage of 5 μM Masked SpyCatcher003 or Unmasked SpyCatcher003 reacted with 5 μM SpyTag003-MBP for up to 60 minutes. All reactions were analysed by SDS-PAGE and visualised with Coomassie staining. The half-time of reaction is indicated. Each reaction was carried out in triplicate, with a line connecting the mean. (B) An SDS-PAGE gel visualised with Coomassie staining showing the reactivity of 5 μM Masked SpyCatcher, unmasked by the presence of TEV protease, and incubated with 5 μM SpyTag003-MBP at 37 °C for up to 60 minutes. (C) An SDS-PAGE gel visualised with Coomassie staining showing the minimal reactivity of 5 μM Masked SpyCatcher incubated with 5 μM SpyTag003-MBP at 37 °C for up to 60 minutes.

Figure 3 shows a bar chart depicting the reactivity of DoubleCatcher polypeptide scaffolds containing different SpyTag003DX variant blocking peptides (where X = A, V, E, S or T). 10 μM of each DoubleCatcher masked with each SpyTag003DX variant was reacted with 5-fold excess SpyTag003-MBP (50 μM) for 4 h at 37 °C in PBS pH 7.4. All reactions were analysed by SDS-PAGE and visualised with Coomassie staining, to determine the percentage of DoubleCatcher with only one of the Catcher sites reacted. The data shown are the mean of reactions carried out in triplicate ± 1 s.d.

Figure 4 shows an SDS-PAGE gel visualised with Coomassie staining showing the masking ability of SpyTag003DS as a DoubleCatcher blocking peptide relative to SpyTag003DA. "A" and "B" refer to SpyTag-AffiHER2$_3$ and (MBP$_X$)$_2$-

SpyTag003, respectively. DC refers to DoubleCatcher. Where indicated, the sample was incubated with SpyCatcher003 resin, to deplete proteins containing an unreacted SpyTag or SpyTag003.

Figure 5 shows a bar chart depicting the reactivity of 10 μM DoubleCatcher variants reacted with 50 μM SUMO-SpyTag003 for 4 hours at 37 °C in PBS pH 7.4. The DoubleCatcher variants were linker variants comprising a TEV cleavage site: WT (SEQ ID NO: 53), PGP (SEQ ID NO: 58), BeforeTEV (SEQ ID NO: 56) and AfterTEV (SEQ ID NO: 57), or variants with modifications to the blocking peptide, Delta RGV (SEQ ID NO: 61), and the masked SpyCatcher003 module of the DoubleCatcher protein (S59T V94I, SEQ ID NO: 4). The percentage of DoubleCatcher variant coupled to a single SUMO-SpyTag003 (DC:SUMO-SpyTag003) was quantified by SDS-PAGE with Coomassie staining. The data shown represent the mean of reactions carried out in triplicate ± 1 s.d.

Figure 6 shows heterodimerization using a thrombin protease-activatable DoubleCatcher variant. First, the reactive SpyCatcher003 site in the DoubleCatcher-Thrombin (DC-Th) was saturated with SpyTag003-MBP, abbreviated as (MBP), to form DoubleCatcher-Thrombin:SpyTag003-MBP (DC-Th:MBP). DC-Th:MBP was then incubated with 0.2 μM bovine thrombin for 2 hours at 34 °C, to cleave the SpyTag003D10A peptide mask from the scaffold. 2 μM of the cleaved product was then reacted with 10 μM SpyTag003-SUMO, abbreviated as (SUMO), at 25 °C for the indicated time. The reaction was visualised by SDS-PAGE with Coomassie staining. The formation of the band corresponding to the DoubleCatcher-Thrombin:SpyTag003-MBP:SpyTag003-SUMO heterodimer, DC-Th:(SUMO)(MBP), confirms the unmasking and subsequent reaction of the second SpyCatcher003 moiety within the DoubleCatcher scaffold. All reactions were carried out in PBS at pH 7.4.

Figure 7 shows (A) a graph depicting the reactivity of 2.5 μM DoubleCatcher (DC) with 5 μM SpyTag003-mClover3 for various durations in minutes up to overnight ("O/N") incubation. All reactions were analysed by SDS-PAGE and visualised with Coomassie staining (gel shown in B). The loss of unreacted DoubleCatcher was quantified by densitometry. (C) depicts the reactivity of the Unmasked SpyCatcher003 module in DoubleCatcher:SpyTag003-MBP following incubation with TEV protease for 2 hours at 34 °C. 2 μM of the resulting DoubleCatcher:SpyTag003-MBP was reacted with 10 μM of SUMO-SpyTag003 for various durations in minutes up to overnight ("O/N") incubation. All reactions were

analysed by SDS-PAGE and visualised with Coomassie staining (D). Loss of DoubleCatcher:SpyTag003-MBP was quantified by densitometry. For (A) and (C), all triplicate data points are shown, with a line connecting the mean. The half-time for each reaction is written on the graphs.

Figure 8 presents (A) an SDS-PAGE gel visualised with Coomassie staining showing the reactions occurring within the one-pot bispecific assembly method. DoubleCatcher ("DC") was first incubated with 1.25-fold excess of SpyTag-AffiHER2$_3$ ("A") or (MBP$_X$)$_2$-SpyTag003 ("B") for 2 hours at 37 °C. The second binder (either 5-fold excess of SpyTag-AffiHER2$_3$ or (MBP$_X$)$_2$-SpyTag003) was added in the presence of superTEV ("TEV") protease to cleave the unreactive SpyTag003DA peptide to produce Unmasked SpyCatcher003 for 4 h at 34 °C. In the Purification lane, DC+A+TEV+B O/N was purified using bead-immobilized SpyCatcher003 to deplete unreacted binders linked to SpyTag or SpyTag003 from the solution (16 h at 4 °C); (B) a graph depicting the mass photometry of the reaction shown in (A) (sample corresponding to the Purification lane). Peaks are shown for the product of coupling, as in (A), confirming the identity of the target bispecific assembly. Peaks were fit by Gaussians with the mean indicated above (TEV: 70 kDa; DC:(A)(B): 144 kDa) and expected mass below (TEV: 71.2 kDa; DC:(A)(B): 144.6 kDa); and (C) a graph depicting the mass photometry of the reaction shown in (A), providing the data for TEV alone.

Figure 9 shows schematics depicting the AlphaFold2-predicted structure of (A) DoubleCatcher (left) or DoubleCatcher with a stable α-helical linker (right, DoubleCatcher H-Lock); (B) DoubleCatcher α-Lock; and (C) DoubleCatcher β-Lock, DoubleCatcher γ-Lock, DoubleCatcher δ-Lock and DoubleCatcher ε-Lock. The Tag is marked as a vector from the Tag's N-terminus to its C-terminus. Cysteines forming a disulfide bond are marked as spheres.

Figure 10 shows (A) a graph depicting the results of size-exclusion chromatography of DoubleCatcher variants normalized by A_{280}, wherein the chromatograms are staggered from adjacent curves in the y-axis. The elution peak from molecular weight (MW) markers is indicated above the graph with arrows. (B) an SDS-PAGE gel visualised with Coomassie staining showing the analysis of size-exclusion chromatography-purified DoubleCatcher variants. The disulfide-containing variants were incubated with 0.4 mM $CuSO_4$ after elution to drive oxidation; (C) a graph depicting the results of dynamic light scattering of DoubleCatcher variants in PBS, wherein the shown traces are staggered from

adjacent curves in the y-axis; and (D) a graph showing the quantification of the dynamic light scattering experiment of (C), showing mean hydrodynamic radius (R_h) ± 1 s.d. (n = 10) for each DoubleCatcher variant.

Figure 11 depicts the effect of the bispecific DoubleCatcher anti-HER2 panel on metabolic activity in (A) SKBR3 cells and (B) BT474 cells. The rows indicate the binder at the N-terminal site of DoubleCatcher. The columns show the binder at the C-terminal site of DoubleCatcher. The effect on metabolic activity of SKBR3 and BT474 cells with 100 nM bispecific after 4 days from resazurin assay is shown as the percentage of metabolic activity (mean ± 1 s.d., n = 3) compared to the 100 nM DoubleCatcher alone control at 100%.

EXAMPLES

A schematic of monospecific, bivalent 'homodimer' scaffold (Double Catcher) disclosed herein is shown in Figure 1, which depicts a representative example of the stepwise assembly of a bispecific complex comprising two Fabs. As described in detail below, the polypeptide scaffold disclosed herein was exemplified using the SpyCatcher003 construct, which facilitates the irreversible dimerization of any two molecules (e.g. antigen-binding units) fused to SpyTag or its variants (e.g. SpyTag002 or SpyTag003). However, as described above, the skilled person would understand from the disclosures herein that the concept of using a protease-cleavable peptide mask to one of the modules in the homodimer is not limited to SpyTag/SpyCatcher system and could be applied to any Catcher/Tag system.

Example 1 – Development of a "masked" SpyCatcher003 polypeptide

Initial experiments were conducted using a single SpyCatcher003 moiety to determine whether the activity of the polypeptide could be masked using a linked non-reactive SpyTag peptide.

SpyTag003D10A (SEQ ID NO: 17, also known as SpyTag003DA), a non-reactive mutant of SpyTag003 (SEQ ID NO: 15), was fused to the C-terminus of SpyCatcher003 using a flexible linker containing a tobacco etch virus (TEV) protease cleavage site (ENLYFQ/G, SEQ ID NO: 55) to function as a blocking peptide. It was determined that the blocking peptide binds tightly to SpyCatcher003 (K_d = 21 nM) in the SpyCatcher003/SpyTag003DA construct (SEQ ID NO: 59, termed SpyMask), thereby masking the reactive Lys of SpyCatcher003. The inventors hypothesised that the intramolecular SpyCatcher003:SpyTag003DA

interaction may be entropically favoured over a competing intermolecular interaction with SpyTag003 in solution, thus preventing premature displacement.

To test the hypothesis and confirm the potential utility of a protease-cleavable peptide mask, the SpyMask construct was cleaved with superTEV protease and incubated with a model fusion protein of SpyTag003 linked to maltose-binding protein (MBP). The reaction between unmasked SpyCatcher003 and SpyTag003-MBP proceeded efficiently to >99% completion, with a reaction time half-life ($t_{1/2}$) of 2.7 minutes (**Figure 2A-B**).

As the disassociation of SpyTag003DA from the binding domain of TEV-cleaved SpyCatcher003 limits the subsequent rate of reaction with Tagged protein of interest (POI) e.g. SpyTag003-MBP, the inventors sought to accelerate the rate of disassociation of SpyTag003DA. It was found that elevating the temperature of the reaction to 37 °C and increasing the molar excess of SpyTag003-MBP relative to Masked SpyCatcher003 by 5-fold reduced the limitation that SpyTag003DA disassociation had on the rate of reaction.

To confirm the muted reactivity of Masked SpyCatcher003 (SpyMask) prior to TEV cleavage, equimolar amounts of SpyMask were incubated with SpyTag003-MBP and the disappearance of SpyMask was monitored over time by SDS-PAGE and densitometry (**Figure 2B-C**). Disappearance of the SpyMask band on an SDS-PAGE gel is indicative of the formation of a construct with a larger mass (e.g. SpyCatcher003:SpyTag003-MBP). The results of these experiments demonstrated that only 3.7% of the Masked SpyCatcher003 reacted with SpyTag003-MBP after 60 minutes. These results clearly demonstrate that SpyTag003DA-mediated silencing of SpyCatcher003 via a protease-cleavable linker is an effective method of preventing SpyCatcher003 reactivity.

<u>Example 2 – Development of a bivalent polypeptide scaffold that enables the stepwise assembly of a bispecific/functional complex</u>

Following the successful generation of a Masked SpyCatcher003 moiety (SpyMask), the inventors developed a platform to facilitate the conjugation of two distinct proteins linked to a Tag (e.g. SpyTag003, i.e. a cognate peptide) able to react with SpyCatcher003. SpyMask was genetically fused to the C-terminus of a native, unmasked SpyCatcher003 protein via a flexible Gly/Ser linker, to give DoubleCatcher (SEQ ID NO: 29) in which the N-terminal SpyCatcher003 moiety

was unmasked and able to bind SpyTag003-linked proteins of interest, while the C-Terminal SpyCatcher003 moiety was masked by SpyTag003DA (see Figure 1).

The reactivity of each SpyCatcher003 moiety within the DoubleCatcher construct was determined by monitoring the reaction of DoubleCatcher with SpyTag003-linked model proteins before and after cleavage by superTEV protease. First, DoubleCatcher was incubated with 2-fold excess of SpyTag003 linked to the green fluorescent protein mClover3 in the absence of superTEV protease. 91% reactivity of the N-terminal SpyCatcher003 was observed after 1 h, with a $t_{1/2}$ of 0.58 min (**Figure 7A-B**).

Following confirmation that the N-terminal SpyCatcher003 module was able to reliable react with Tagged proteins, further experiments were conducted to determine if two distinct proteins were able to be bound by a single DoubleCatcher construct. DoubleCatcher was incubated with SpyTag003-MBP for 2 h at 37 °C to saturate the N-terminal SpyCatcher003 moiety. superTEV protease was then added to the crude mixture for 4 h at 34 °C to induce cleavage of the SpyTag003DA mask (i.e. blocking peptide) from the C-terminal SpyCatcher003 module. 2 μM of the cleaved construct was then reacted with a 5-fold molar excess of SUMO-SpyTag003 and the reaction was monitored over time by SDS-PAGE and densitometry (**Figure 7C-D**). SUMO-SpyTag003 rapidly reacted with the unmasked C-terminal SpyCatcher003 module with a $t_{1/2}$ of 3.6. This data clearly demonstrates that SpyTag003DA can be efficiently cleaved and replaced with a reactive Tag linked to a molecule, e.g. POI, thus establishing DoubleCatcher as robust tool for the controlled combination of distinct molecules, e.g. proteins.

Example 3 – Testing the efficacy of DoubleCatcher variants

While the use of SpyTag003DA was found to be highly effective for masking the reactive domain of SpyCatcher003, it was determined that alternative mutations to the aspartic acid residue at position 10 of SpyTag003 were also able to successfully mask the reactive lysine residue of SpyCatcher003.

SpyTag003 variants with mutations to the aspartic acid residue at position 10 to valine (SpyTag003DV), glutamic acid (SpyTag003DE), serine (SpyTag003DS, SEQ ID NO: 18) and threonine (SpyTag003DT SEQ ID NO: 19) were utilised to produce DoubleCatcher variants. 10 μM of each DoubleCatcher variant was incubated with 50 μM SpyTag003-MBP for 4h at 37 °C in PBS at pH 7.4. All SpyTag003 variants (SpyTag003DX) were able to successfully mask the reactive

lysine residue in SpyCatcher003. Interestingly, it was determined that SpyTag003DS (SEQ ID NO: 18) substantially improved the masking ability of the non-covalent SpyTag003 peptide, relative to the SpyTag003DA mask (83.3% compared to 75.9%; **Figure 3B**). SpyTag003DT (SEQ ID NO: 19) showed similar activity to SpyTag003DA (SEQ ID NO: 17).

A direct comparison of the masking ability of SpyTag003DS compared to SpyTag003DA was then performed to monitor the reaction efficacy of each site individually. 2.5 µM of each DoubleCatcher variant was reacted with 5 µM SpyTag003-MBP for 2 h at 37 °C before 0.21 µM MBP-SuperTEV protease and 12.5 µM $(MBP_X)_2$-SpyTag003 was added to the reaction mixture for 4 h at 34 °C. 20 µL of SpyCatcher003 immobilised on resin was added overnight at 4 °C to remove any remaining unreacted excess SpyTag003-fusion in the reaction mixture. All reactions were carried out in PBS at pH 7.4. It was determined that relative to the original DoubleCatcher containing SpyTag003DA (SEQ ID NO: 29), the DoubleCatcher variant containing the SpyTag003DS blocking peptide reacted to a lesser extent following TEV cleavage (**Figure 4**). Interestingly, analysis of the control lane, which shows the results of monosaturated DoubleCatcher incubated with TEV alone, suggests TEV cleavage was occurring (**Figure 4**). Therefore, in view of the fact that TEV cleavage is occurring, it is likely that the SpyTag003DS mask is dissociating with a very slow off-rate leading to the reactive residue of the Catcher remaining masked for a longer period of time. In view of this slow disassociation, SpyTag003DA was chosen going forward as the preferred blocking peptide.

As premature disassociation of the blocking peptide (e.g. SpyTag003DA) may result in undesired reactivity of the C-terminal SpyCatcher003 module, further experiments were conducted to determine if alternative linker configurations for the linker that connects the blocking peptide to the polypeptide scaffold (i.e. DoubleCatcher) would reduce the risk of premature disassociation. Four linker variants comprising a TEV cleavage site were investigated: WT (SEQ ID NO: 53, as used previously in Example 1), PGP (SEQ ID NO: 58), BeforeTEV (SEQ ID NO: 56) and AfterTEV (SEQ ID NO: 57). Moreover, the WT linker (SEQ ID NO: 53) was also tested in combination with modifications to the blocking peptide and the masked SpyCatcher003 domain of the DoubleCatcher protein. The modified blocking peptide, termed Delta RGV (SEQ ID NO: 61), contained an N-terminal deletion of the RGV residues. The modified masked SpyCatcher003 module contained

substitutions at positions 59 and 94 (S59T and V94I, i.e. the sequence of SEQ ID NO: 4).

10 μM of each DoubleCatcher variant containing the aforementioned linker configurations or modified blocking peptide or modified masked SpyCatcher003 module, was incubated with 50 μM SpyTagged Small Ubiquitin-like Modifier (SUMO-SpyTag003) for 4 h at 37 °C in PBS (pH 7.4). The percentage of DoubleCatcher saturated to a single SUMO-SpyTag003 (DC:SUMO-SpyTag003) was quantified as a percentage of all reaction products. These experiments demonstrated that the S59T V94I (78.2%) and BeforeTEV (77.3%) variants performed marginally better than native DoubleCatcher ('WT'; 76.4%) (**Figure 5**).

As detailed above, whilst the use of a cleavable TEV site allows the Masked SpyCatcher003 module within DoubleCatcher to be efficiently and reliably unmasked, it may be beneficial to employ alternative protease cleavage mechanisms, if required. Accordingly, a DoubleCatcher polypeptide was produced in which the SpyTag003DA blocking peptide was fused to the C-terminus of SpyCatcher003 using a flexible linker containing a Thrombin cleavage site to form a DoubleCatcher-Thrombin variant (SEQ ID NO: 36). Following reaction of the DoubleCatcher-Thrombin variant with SpyTag003-MBP as described above, 2 μM DoubleCatcher-Thrombin:SpyTag003-MBP was reacted with 0.2 μM Bovine thrombin for 2h at 34 °C. 10 μM SpyTag003-SUMO was then added to the reaction mixture for up to 60 minutes at 25 °C. All reactions were carried out in PBS at pH 7.4. It was found that thrombin was able to successfully cleave the thrombin cleavage site to unmask the reactive domain of the C-terminus SpyCatcher003, thus demonstrating that DoubleCatcher reactivity can be controlled using alternative proteases (**Figure 6**).

Example 4 - Establishment of scalable heterodimerization of proteins using DoubleCatcher

A universal bispecific assembly platform for high-throughput screening requires a simple and scalable methodology. In particular, procedures involving multiple rounds of column chromatography will provide a bottleneck to the combinatorial complexity of bispecifics that can be interrogated. To address this need, the inventors next sought to optimize a one-pot reaction in which any two Tagged binders can be conjugated to any DoubleCatcher construct.

For optimisation experiments, a SpyTag construct linked to three tandemly-repeated affibodies against HER2 (SpyTag-AfiiHER2$_3$) and SpyTag003 linked to high stability MBP ((MBP$_X$)$_2$-SpyTag003) was used as a model (Veggiani et al., 2016, Proc. Natl. Acad. Sci. *113*, 1202–1207). The chosen ligands lack free Cys which avoids impurity from inter-binder disulfides, and have a difference in mass that permits unambiguous detection of reactant and product bands by SDS-PAGE. The intact masses of these purified reactants in PBS pH 7.4 were additionally validated by electrospray-ionization mass spectrometry.

In the first step of the one-pot assembly method, DoubleCatcher was reacted with a small molar excess (1.25-fold) of SpyTag-AffiHER2$_3$ for 2 h at 37 °C, to ensure negligible concentration of unreacted DoubleCatcher remained in the solution. SDS-PAGE revealed 97% of the DoubleCatcher constructs within the reaction were the desired DoubleCatcher:SpyTag-AffiHER2$_3$ species (**Figure 8A**). As observed in the Masked SpyCatcher003 reactivity tests (**Figure 2A**), 3% of the constructs present were DoubleCatcher conjugated to two SpyTag-AffiHER2$_3$ molecules, which was visible in SDS-PAGE ("DC:A2" lane in **Figure 8A**). To confirm where the homodimeric DoubleCatcher:(SpyTag-AffiHER2$_3$)2 would result in a band on the SDS-PAGE gel, superTEV protease (1/12th molar equivalent) was added to the crude mixture. It was found that the level of the DoubleCatcher:(SpyTag-AffiHER2$_3$)2 homodimer in the reaction mixture increased to 30% ("DC + A + TEV" lane in **Figure 8A**). The identity of this by-product was corroborated by the increase in the density of the band corresponding to AffiHER2$_3$ homodimer in a control reaction containing twice the concentration of SpyTag003-AffiHER2$_3$ ('DC + A + TEV + A' lane in **Figure 8A**).

For subsequent unmasking and saturation of the C-terminal SpyCatcher003 module, superTEV protease (1/12th molar equivalent) and (MBP$_X$)$_2$-SpyTag003 (5-fold molar excess; 'B') were added simultaneously to the singly saturated DoubleCatcher and allowed to react for 4 h at 34 °C. Thus, immediately upon cleavage of SpyTag003DA, the concentration of (MBP$_X$)$_2$-SpyTag003 in solution should be 20 times that of unreacted SpyTag-AffiHER2$_3$ to minimize homodimer formation. Indeed, AffiHER2$_3$ homodimer still accounted for only 3% of the total species in the reaction mixture ("DC + A + TEV + B" lane in **Figure 8A**). Trace (3%) formation of DoubleCatcher conjugated to two (MBP$_X$)$_2$-SpyTag003 molecules ("MBP homodimer") was observed by SDS-PAGE, validated by the formation of a

band of equivalent mobility in a reaction with DoubleCatcher, superTEV protease and $(MBP_X)_2$-SpyTag003 alone ("DC + B + TEV + B" lane in **Figure 8A**).

For depletion of excess SpyTag003-linked molecules from the desired product, the crude reaction mixture was incubated with bead-immobilized SpyCatcher003 by adding beads in a 1:1 suspension with PBS directly to the reaction mixture for 16 h at 4 °C. The purity of DoubleCatcher:SpyTag-AffiHER2$_3$:$(MBP_X)_2$-SpyTag003 following purification was assessed by SDS-PAGE densitometry as 95% (excluding superTEV protease) (Purification lane in **Figure 8A**). Based on SDS-PAGE, superTEV protease (70 kDa, **Figure 8B-C**) was found to be the major contaminant within the resultant mixture. Since TEV protease has exceptional specificity for its target sequence (TEV), which is not found in the human proteome, and has been expressed before without effect on cells, it was not considered necessary to remove it from the final product. However, in cases where removal is desirable, it should also be possible to add amylose-agarose to the well to deplete the MBP-linked superTEV protease.

The bispecific assembly of DoubleCatcher was also validated by mass photometry which allows for the single molecule-based assessment of each species in solution. The peak corresponding to DoubleCatcher:SpyTag-AffiHER2$_3$:$(MBP_X)_2$-SpyTag003 was observed with a mass of 144 kDa, which matched well to the predicted mass of 144.6 kDa (**Figure 8B**). The second peak of 70 kDa indicated the presence of MBP-superTEV protease (**Figure 8B**), as validated by running MBP-superTEV protease on its own (**Figure 8C**).

Example 5 – Customization of the DoubleCatcher Architecture

As with most bispecific antibodies, the format of a bispecific binder built using DoubleCatcher can be influenced by the identity of the SpyTag003-linked molecule and by the order in which the binders are coupled. However, the overall shape, size, and flexibility of the bispecific molecule can also be determined by the DoubleCatcher module itself. To amplify the format space to be screened, matrices of bispecific antibodies could therefore be developed in three dimensions by varying two binder arms against the targets of interest and, in parallel, varying the DoubleCatcher module. The highest-scoring AlphaFold2-predicted structure for DoubleCatcher is displayed in **Figure 9A** with the position and orientation (N to C) of the two SpyTag binding sites represented by arrows. Modifications to the DoubleCatcher architecture that change the relative spacing and orientation of the

SpyTag binding sites will reflect changes in the structural relationship between the two target-binding arms in the final bispecific.

It was determined that modifications to the architecture of DoubleCatcher could be achieved by varying the linker connecting the two SpyCatcher003 moieties. The amenability of DoubleCatcher to linker modification was tested by substituting the Gly/Ser linker with the central helix from the *Bacillus stearothermophilus* ribosomal protein L9. The DoubleCatcher variant with the helical linker is termed 'DoubleCatcher H-lock' (SEQ ID NO: 30). Size-exclusion chromatography analysis of Ni-NTA-purified DoubleCatcher H-lock revealed that its apparent molecular mass was 17 kDa larger than DoubleCatcher (**Figure 10A-B**), while the predicted increase in mass is only 3.5 kDa.

In order to increase the rigidity of the DoubleCatcher construct which overcomes the natural flexibility of the SpyCatcher003 termini, the inventors engineered the formation of a covalent linkage between the two Catcher domains of the DoubleCatcher construct. It was determined that a pair of residues (one residue from each SpyCatcher003) within the loop regions of the Catcher domains and in close proximity (1.2 nm) within each of the five highest-confidence predicted structures of DoubleCatcher could be mutated into cysteine residues to engineer the formation of a SpyCatcher003-SpyCatcher003 intramolecular disulfide bond. This engineered DoubleCatcher structure was termed "DoubleCatcher α-lock" (SEQ ID NO: 31). The substantially less variable average distance between Nε of the reactive Lys residues of each SpyCatcher003 in the top 5 AlphaFold2-predicted structures of DoubleCatcher α-lock (4.7 ± 0.04 nm) suggests the two SpyCatcher003 moieties may have low relative mobility. The DoubleCatcher double-cysteine variants designed from the four remaining predicted structures of DoubleCatcher are presented in **Figure 9C**. Therefore, the combination of SpyTagged antigen binders using these DoubleCatcher variants should provide conformationally distinct bispecific binders that bind targets in different orientations.

Size-exclusion chromatography profiles of Ni-NTA-purified double-cysteine DoubleCatcher variants reveal that each construct exists as two species: one main peak corresponding to monomer and one minor peak that corresponds to covalent dimer (**Figure 10A**). SDS-PAGE of size-exclusion chromatography fractions for each double-cysteine mutant revealed that the DoubleCatcher monomer in the main peak existed as a mixture of oxidized and reduced species. The DoubleCatcher double-cysteine monomer was incubated with 0.5 mM $CuSO_4$ for 2

h at 25 °C in HEPES (pH 7) to drive disulfide bond formation prior to analysis (**Figure 10B**). The identity and disulfide bond formation of each DoubleCatcher was validated by electrospray-ionization mass spectrometry. The DoubleCatcher variants were analyzed by dynamic light scattering (DLS) (**Figure 10C**), with quantification shown in **Figure 10D**. As expected, both the mean R_h and the mean s.d. of the double-cysteine mutants (3.5 ± 1.0 nm, mean ± 1 s.d., n = 10) were less than those for the cysteine-free variants (4.5 ± 1.7 nm, mean ± 1 s.d., n = 10), consistent with a higher degree of sample homogeneity, likely because of reduced flexibility within the architecture.

Example 6 - Anti-HER2 bispecifics assembled by DoubleCatcher induce various anti-proliferative or proliferative effects in HER2-addicted cancer cells

Having demonstrated that DoubleCatcher is a robust tool for the controlled and efficient combination of distinct proteins to produce bispecific ligands, the inventors sought to develop this platform for use in a screening approach of equal throughput of binders to the important cell-surface target of HER2 (also known as Receptor tyrosine-protein kinase erbB-2). The relative activity of each signalling pathway of HER2 and the resultant downstream physiological outcome has been attributed to different ligands stabilizing receptor dimers in distinct conformations, thus inducing distinct intracellular phosphorylation of the 17 unique sites in HER2's intracellular domain. Although the functional selectivity of native Receptor Tyrosine Kinase ligands has been extensively studied, this bias is yet to be exploited for personalized cancer therapies. The inventors therefore sought to assemble and screen a large matrix of anti-HER2 bispecifics for their ability to alter HER2 activity.

Binders specific to each subdomain of the HER2 extracellular region were identified, including Fabs based on the clinically used Pertuzumab and Trastuzumab: Fab0.11 Fab, Pertuzumab Fab, 39S Fab, MF3958 Fab, H2-18 Fab, HFS2 Fab, nano-HER2 nanobody and AffiHER2 affibody. Fab0.11 is a Trastuzumab variant with a point mutation to enhance binding stability. A variant of Fab0.11 with no linker to SpyTag003 was also expressed to determine the importance of linker flexibility on binding activity. Each binder was expressed as a SpyTag003-fusion in Expi293 cells, except for the nanobody (nanoHER2 and the affibody (AffiHER2) which were expressed in *E. coli*. The identity and disulphide bond formation of each binder was confirmed using electrospray-ionization mass spectrometry. Each binder was then dimerized with itself or another binder using

DoubleCatcher methodology as described herein. The anti-proliferative activity of the resulting matrix of anti-HER2 bispecifics and homodimers was quantified in HER2-addicted BT474 cells (a human ductal breast carcinoma cell-line) or SKBR3 cells (a human adenocarcinoma cell-line) using a resazurin assay.

Following incubation of BT474 or SKBR3 cells with 100 nM of the anti-HER2 assemblies for 96 h, the change in metabolic activity was calculated relative to untreated cells (**Figure 11A-B**). Interestingly, it was found that different anti-HER2 assemblies exhibited a range of antagonistic and agonistic activities. It was found that bispecific constructs displaying one or more arms binding to domain II (39S, HFS2, or Pertuzumab) were consistently anti-proliferative. It is hypothesised that as HER2 dimerization is facilitated predominantly by interactions between domain II, these bispecifics are likely to block HER2 dimerization and activation. Conversely, bispecifics harboring the small, non-domain II binders nanoHER2 (16 kDa) or AffiHER2 (15 kDa) display potent proliferative activity. This was further confirmed by data demonstrating that bispecifics with significant agonistic activity in both SKBR3 and BT474 all comprised at least one nanoHER2 or AffiHER2.

Where agonistic activity of biparatopic anti-HER2 molecules has been observed, it has been suggested to result from favouring HER2 homodimerization over heterodimerization, or from a shorter distance between antigen-binding sites. Considering their small size, DoubleCatchers coupled to either nanoHER2 or AffiHER2 may bridge HER2 molecules close enough for trans-phosphorylation by HER2 kinase domains.

It was also found that anti-HER2 bispecifics exhibited different potencies in SKBR3 cells relative to BT474 cells. It was found that a higher number of bispecifics had significant proliferative activity in BT474 (20 constructs) than in SKBR3 (9 constructs) and also that constructs with significant anti-proliferative activity were more potent in BT474 than in SKBR3. The inventors hypothesise that this result may be explained by the differential total expression of HER2 in BT474 compared to SKBR3 (~2.8×10^6 vs. ~1.2×10^6 copies per cell), as well as the expression of Epidermal Growth Factor Receptor (7×10^4 vs 2.2×10^5 copies per cell). Both cell types express HER3 to a similar level (1-1.2×10^6 copies each). As it is known that HER2-HER3 dimerization mediates activation of PI3K/Akt and the proliferation axis, while HER2-EGFR dimers favour MAPK/ERK and pro-survival signalling, it hypothesised that bispecifics that prevent HER2 homodimerization but do not occlude domain II of HER2 (for example, domain I-binding H2-18 or

MF3958) may allow heterodimerization to occur more favourably. This could result in a higher frequency of proliferative HER2-HER3 dimers in BT474 than SKBR3.

In both cell-lines, it was found that Fab0.11:NoLink-Fab0.11 was antagonistic (78 ± 4% for SKBR3; 74 ± 6% for BT474) but showed no activity in the opposite orientation Fab0.11:Fab0.11-NoLink (100 ± 9% for SKBR3; 103 ± 1% for BT474). The same orientation-dependence was observed in SKBR3 for bispecifics with Fab0.11-NoLink with H2-18, or nanoHER2 with 39S, H2-18, or AffiHER2. Examining the effect of removing the linker between binder and SpyTag003, it was found that the average difference between bispecifics with at least one Fab0.11 versus at least one Fab0.11-NoLink was only -3%, supporting the usage of linker-free SpyTag fusions in future screens. Together, these results highlight the importance of screening epitope-diverse bispecific binders in multiple formats and orientations.

Methods

Plasmids and Cloning

Constructs were cloned by PCR methods using Q5 High-Fidelity Polymerase (New England Biolabs) and fragments were assembled by Gibson assembly. All constructs were validated by Sanger Sequencing (Eurofins Genomics).

pET28a-SpyTag003-MBP (GenBank Accession no. MN433888, Addgene plasmid ID 133450), pET28-SpyTag003-mClover3 (Addgene plasmid ID 133453), pDEST14-SpyCatcher003 S49C (Addgene plasmid ID 133448), pET28-MBP-super TEV protease (Addgene plasmid ID 171782), pET28a-SnoopTag-AffiHER2-SpyTag (N-terminal His_6–SnoopTag–anti-HER2 Affibody–SpyTag) (GenBank accession no. KU296975), and pET28a-SnoopTag-SpyTag-$(AffiHER2)_3$ (GenBank accession no. KU296976) have been previously described. pDEST14-SpyCatcher003-TEVs-SpyTag003DA ("Masked SpyCatcher003" (SEQ ID NO: 59); N-terminal His_6 tag-SpyCatcher003-TEV protease cleavage site-SpyTag003D10A) was derived from pDEST14-SpyCatcher003 (GenBank Accession no. MN433887, Addgene plasmid ID 133447) by incorporating the TEV protease cleavage site (SEQ ID NO: 55) and SpyTag003 D10A (SEQ ID NO: 17) by Gibson assembly. pDEST14-SpyCatcher003-$(GSG)_3$-SpyCatcher003-TEVs-SpyTag003DA (DoubleCatcher (SEQ ID NO: 29); N-terminal His_6 tag-TEV protease cleavage site-two SpyCatcher003 moieties connected by a $(GSG)_3$ spacer-TEV protease cleavage site-SpyTag003DA) was synthesized as a gene fragment (Integrated DNA Technologies) and inserted into

pDEST14. The Gly/Ser spacer in this construct was then replaced with an α-helical spacer (SEQ ID NO: 21), and this construct was termed DoubleCatcher H-Lock (SEQ ID NO: 30). Double-cysteine substitution variants of DoubleCatcher were derived from pDEST14-SpyCatcher003-$(GSG)_3$-SpyCatcher003-TEVs-SpyTag003DA by Gibson assembly: DoubleCatcher α-Lock (SEQ ID NO: 31), DoubleCatcher β-Lock (SEQ ID NO: 32), DoubleCatcher γ-Lock (SEQ ID NO: 33), DoubleCatcher δ-Lock (SEQ ID NO: 34), and DoubleCatcher ε-Lock (SEQ ID NO: 35).

Sequences for the heavy and light chains of anti-HER2 Fabs (39S (Oganesyan *et al.*, 2018, J. Biol. Chem. *293*, 8439–8448.), MF3958 (Geuijen *et al.*, 2018, Cancer Cell *33,* 922-936.e10.), H2-18 (Hu *et al.*, 2015, Oncotarget *6*, 1695–1706), HFS2 (Salimi *et al.*, 2018, Mol. Biol. Rep. *45*, 2247–2256), Pertuzumab (WO2014172371A2) were synthesized (Twist Bioscience) and cloned into pcDNA3.1. Each Fab heavy chain had the final assembly: N-terminal signal peptide-Heavy variable domain-Human IgG1 C_H1 domain-EPKSC IgG1 upper hinge region residues-SpyTag003-His_6 tag, with the exception of HSF2 whose C_H1 domain was of mouse origin, and Pertuzumab for which the hinge region was omitted. Each Fab light chain had the final assembly: N-terminal signal peptide-Light variable domain-Human kappa light chain, with the exception of HSF2 whose kappa light chain was of mouse origin. pET28a-nanoHER2-SpyTag003, based on the 11A4 nanobody against HER2 (WO2013135655A1), was cloned with a C-terminal His_6 tag.

Bacterial Protein Expression

pDEST14 constructs were transformed into chemically competent *E. coli* C41 (DE3), except for the pDEST14-SpyCatcher003-SpyCatcher003-TEV site-SpyTag003DA disulfide-containing variants, which were transformed into chemically competent *E. coli* T7SHuffle (NEB). pET21 and pET28a constructs were transformed into chemically competent *E. coli* BL21 (DE3) (Agilent Technologies), with the exception of pET28a-SnoopTag-AffiHER2-SpyTag, which was transformed into chemically competent *E. coli* T7 Express (NEB). Colonies were picked into 10 mL LB containing either 100 μg/mL ampicillin (pDEST14) or 25-50 μg/mL kanamycin (pET28a/pET21), and grown at 37 °C at 200 rpm for 4-16 h. Starter cultures were diluted 100-fold into either 1 L LB supplemented with appropriate antibiotic and 0.8% glucose, 1 L LB supplemented with antibiotic alone (T7 Express only), or 2 × TY supplemented with antibiotic and 0.5% (v/v) glycerol (T7SHuffle only), and grown at 37 °C with shaking at 200 rpm. At OD_{600} 0.5-0.6, expression was induced with 0.42

mM isopropyl β-D-1-thiogalactopyranoside (IPTG; Fluorochem). pDEST14-transformed cultures were then incubated for 4 h at 30 °C and 200 rpm, while all other cultures were incubated for 16 h at 18 °C and 200 rpm post-induction. Cells were harvested by centrifugation at 4,000 g for 15 min at 4 °C and pellets were stored at -80 °C until purification.

Mammalian Protein Expression

Anti-HER2 Fabs were expressed in Expi293 cells (Thermo Fisher), which were maintained in Expi293 Expression Medium (Thermo Fisher) supplemented with 50 U/mL penicillin and 50 µg/mL streptomycin and grown in a humidified Multitron Cell incubator (Infors HT) at 37 °C with 8% (v/v) CO_2 and shaking at either 125 rpm (125 mL flasks; Corning) or 240 rpm (50 mL mini bioreactor tubes; Corning). After ≥ 3 passages, cells were diluted to 3 × 10^6 cells/mL in antibiotic-free Expi293 Expression Medium and transfected using an ExpiFectamine 293 Transfection Kit (Thermo Scientific). ExpiFectamine 293 Transfection Enhancers 1 and 2 were added 20 h post transfection, and cultures were incubated for 5 days post-transfection, when cell viability fell below 50%. Cell supernatants were supplemented with cOmplete, Mini, EDTA-free Protease Inhibitor Cocktail (Roche) and clarified by centrifugation at 4,000 *g* at 4 °C, followed by filtration through a 0.22 µm syringe filter (Thermo Fisher). One tenth supernatant volume of Ni-NTA wash buffer (50 mM Tris-HCl pH 7.8, 300 mM NaCl, 10 mM imidazole) was added to filtered supernatant, which was then pH adjusted to pH 7.6 using 1 M Tris-HCl pH 8.0. Supernatants were stored at -80 °C until purification. The typical yield for SpyTag003-fused Fabs ranged from 6 to 156 mg per L of culture.

Ni-NTA purification of His_6-tagged proteins

After thawing, bacterial cell pellets were resuspended in 1× Ni-NTA buffer (50 mM Tris-HCl pH 7.8, 300 mM NaCl) supplemented with cOmplete, Mini, EDTA-free Protease Inhibitor Cocktail (Roche) and 1 mM phenylmethylsulfonyl fluoride (PMSF), and lysed by sonication on ice. SpyCatcher003 S49C variants and DoubleCatcher double-cysteine variants were further supplemented with 10 mM 2-mercaptoethanol to break disulfide bonds to other proteins. Clarified cell lysates were centrifuged at 30,000 *g* for 30 min at 4 °C and adjusted to pH 7.6 using 1 M Tris-HCl pH 8.0, before purification by Ni-NTA affinity chromatography (Qiagen). Elution of all proteins was performed in the absence of 2-mercaptoethanol. For anti-HER2 binding proteins to

be heterodimerized by DoubleCatcher for treatment of cells, the following washes were performed directly prior to elution: 50 column volumes (CV) wash buffer containing 0.1% (v/v) Triton X-114 (50 mM Tris-HCl pH 7.8, 300 mM, 10 mM imidazole), followed by 10 CV wash buffer (50 mM Tris-HCl pH 7.8, 300 mM). Clarified mammalian supernatants were purified identically. Eluted proteins were dialyzed three times against PBS pH 7.4 (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4) in 3.5 kDa molecular weight cut-off Spectra/Por tubing (Spectrum Labs) at 4 °C. SpyCatcher003 S49C variants to be coupled to resin were instead dialyzed into coupling buffer (50 mM Tris-HCl pH 8.5, 5 mM EDTA). Protein concentrations were determined from A_{280} using their predicted extinction coefficient from ExPASy ProtParam. Typical yield for SpyCatcher003-containing variants was 10-20 mg per L of culture. A typical yield for affibody expression was ~3 mg per L of culture and 30 mg per L for nanobody expressions. Expression of SpyTag003-fused, non-antigen binding proteins yielded 5-15 mg protein per litre of culture.

Prior to assembly of bispecific molecules with DoubleCatcher for mammalian cell treatment, endotoxin was removed from DoubleCatcher scaffolds by phase separation with Triton X-114. Purified protein samples were transferred to 1.5 mL endotoxin-free microcentrifuge tubes (StarLab, cat. E1415−1510), to which 1% (v/v) Triton X-114 was added. Samples were incubated on ice for 5 min, until all Triton X-114 was dissolved. Samples were then incubated at 37 °C for 5 min and centrifuged for 1 min at 16,000 *g* at 37 °C. The supernatant was pipetted off, and the entire procedure was repeated twice. Endotoxin concentration of any bispecific constituent to be added to cells was determined using the Limulus amebocyte lysate (LAL) Chromogenic Endotoxin Quantitation Kit (Thermo Fisher) according to the manufacturer's instructions, which confirmed the endotoxin level to be below 1 endotoxin unit/mL.

SpySwitch purification of SpyTag003-tagged proteins

MBP_X-MBP_X-SpyTag003 cell pellets were lysed and clarified as outlined above, except thawed pellets were resuspended in 1 × SpySwitch buffer (50 mM Tris pH 7.5, 300 mM NaCl). SpySwitch purification was performed at 4 °C. Briefly, SpySwitch resin (0.75 mL per 1 L bacterial culture) was equilibrated with 2×10 CV SpySwitch buffer, prior to incubation with clarified cell lysate on an end-over-end rotator for 1 h at 4 °C, after which resin was washed with 4 × 10 CV SpySwitch buffer. pH-dependent elution was performed by incubating the resin with 6 × 1.5 CV

SpySwitch elution buffer (50 mM acetic acid, 150 mM NaCl, pH 5.0) for 5 min each, which was immediately neutralized by mixing the flow-through with 6 × 0.3 CV of 1 M Tris pH 8.0. Proteins were subsequently dialyzed into PBS pH 7.4.

Polyacrylamide gel electrophoresis

SDS-PAGE was performed using 8-16% polyacrylamide gels in an XCell SureLock system (Thermo Scientific). SDS-PAGE was run at 200 V in 25 mM Tris-HCl pH 8.5, 192 mM glycine, 0.1% (w/v) SDS. Gels were stained with Brilliant Blue G-250 and destained with Milli-Q H_2O prior to imaging on a ChemiDoc XRS+ imager. ImageLab 6.1.0 software (Bio-Rad) was used for densitometric quantification of bands. For imaging on an iBright FL1500 imaging system (Invitrogen), analysis was performed using iBright Analysis Software Versions 5.0.1 and 5.2.0 (Thermo Fisher).

Size-Exclusion Chromatography

Ni-NTA- or SpySwitch-purified protein samples (SpyCatcher003-containing variants, SpyTag003-linked non-antigen binding proteins) were injected onto a pre-equilibrated HiLoad 16/600 Superdex 200 pg column (GE Healthcare), or a HiLoad 16/600 Superdex 75 pg column (SUMO-SpyTag003 and affibodies; GE Healthcare). Samples were run on an ÄKTA Pure 25 (GE Healthcare) fast protein liquid chromatography machine at 1 mL/min in PBS pH 7.4 at 4 °C. The absorbance profile of column elutions was recorded at 230 nm, 260 nm, and 280 nm. Each column was calibrated using 660 kDa to 1.35 kDa molecular weight protein standards (thyroglobulin, IgG, ovalbumin, myoglobin, and vitamin B12) (Bio-Rad). Following peak analysis by SDS-PAGE, fractions containing purified protein at the expected molecular weight were concentrated using Vivaspin 20 centrifugal filters (Cytiva) at 4 °C, and stored at -80 °C.

Resin coupling

SpyCatcher003 S49C was attached to SulfoLink Coupling Resin (Thermo Fisher) according to the manufacturer's instructions, with the following changes: SpyCatcher003 S49C in coupling buffer was concentrated to 20 mg/mL using a Vivaspin 20 5 kDa MWCO centrifugal filter at 4 °C, reduced with 1 mM tris(2-carboxyethyl)phosphine (TCEP) for 30 min at 25 °C, and coupled to SulfoLink Coupling Resin at 20 mg reduced protein/mL packed resin. Following washes with 1 M NaCl, the coupled resin was washed with TP buffer (25 mM orthophosphoric acid

adjusted to pH 7.0 with Tris base) to remove SpyCatcher003 S49C that had bound non-covalently to the resin. The coupled resin was stored in 20% (v/v) ethanol in PBS pH 7.4 at 4 °C for ≤ 3 months.

Isopeptide bond reconstitution reactions

All reactions were carried out in triplicate in PBS pH 7.4. 2.5 μM SpyCatcher003 (at 25 °C), SpyMask, or DoubleCatcher variants (each at 37 °C) was reacted with 5 μM SpyTag003-linked protein. At each time-point or end-point indicated, reactions were quenched by adding 6 × SDS loading buffer [0.23 M Tris-HCl pH 6.8, 24% (v/v) glycerol, 120 μM bromophenol blue, 0.23 M SDS], with subsequent heating at 95 °C for 5 min in a Bio-Rad C1000 thermal cycler. Reactions were analysed by SDS-PAGE and the depletion of SpyCatcher003-derived reactant over time was quantified by densitometry. The extent of reaction between binding partners at each time point is defined as 100 × [1 – (SpyCatcher003-derived reactant)/(average density of SpyCatcher003-derived reactant in control lanes)]. The control lanes contain SpyCatcher003-containing partner alone.

Assembly of dimeric molecules using DoubleCatcher

With the exception of reconstitution assays, the assembly of bispecific molecules from two SpyTag003-linked proteins of interest, 'Protein 1' and 'Protein 2', was performed in 100 μL total volume per well in 96-well cell culture plates (Thermo Scientific, cat. 167008): 1) DoubleCatcher variant with concentration n was first mixed with 1.25 × n (1-75μM) Protein 1-SpyTag003 for 2 h at 25 °C with shaking at 550 rpm; 2) 5 × n Protein 2-SpyTag003 and 1/12 × n MBP-superTEV protease was added directly to the reaction mixture and incubated for 4 h at 34 °C with shaking at 550 rpm; 3) reaction mixture was incubated with SulfoLink resin-immobilized SpyCatcher003 S49C for 16 h at 4 °C with shaking at 550 rpm, to remove excess unreacted SpyTag003-linked binders from solution. All coupling reactions were carried out in PBS pH 7.4. For the assembly of anti-HER2 Fab 'homodimers', Protein 1-SpyTag003 and Protein 2-SpyTag003 were identical.

Dynamic Light Scattering

Size-exclusion chromatography-purified protein samples were centrifuged for 30 min at 16,900 × g at 4 °C to remove any aggregates. Samples were filtered through 0.33 mm diameter sterile syringe filters with a 0.22 μm hydrophilic

polyvinylidene difluoride (PVDF) membrane (Sigma-Aldrich) and diluted to a final concentration of 25 μM for protein samples > 35 kDa, or 75 μM for samples < 35 kDa into sterile filtered PBS pH 7.4. 20 μL diluted sample was loaded into a reusable quartz cuvette and measurements (10 scans of 10 s each) were recorded at 20 °C using an Omnisizer (Viscotek). Data were analyzed in duplicate using OmniSIZE 3.0. For each sample, the radius of hydration (R_h) was plotted with error bars ± 1 s.d.

Cell Culture

SKBR3 cells were from the American Type Culture Collection. BT474 cells were from Cancer Research UK, Lincoln's Inn Fields. SKBR3 cells were grown in complete DMEM: Dulbecco's Modified Eagle Medium–high glucose (DMEM) supplemented with 10% (v/v) fetal bovine serum (FBS), 100 U/mL penicillin and 100 μg/mL streptomycin (1 × pen/strep; Sigma-Aldrich), 1% (v/v) GlutaMAX (Thermo Fisher Scientific) at 37 °C and 5% (v/v) CO_2. BT474 cells were grown in complete RPMI: RPMI-1640 supplemented with 10% fetal bovine serum (FBS), 100 U/mL penicillin and 100 μg/mL streptomycin (1 × pen/strep; Sigma-Aldrich), and insulin (5 μg/mL; Sigma) at 37 °C and 5% (v/v) CO_2. Cells were passaged at 70-80% confluency and were sub-cultured for fewer than 3 months. Cell lines were tested for mycoplasma contamination by polymerase chain reaction (PCR) every 6 months, and were found to be mycoplasma-negative each time.

Cell Viability Assay

BT474 or SKBR3 cells were seeded into 96-well plates at 1.25×10^4 cells per well and grown in DMEM supplemented with 0.5% (v/v) FBS, 1% (v/v) GlutaMAX and 50 U/mL penicillin and 50 μg/mL streptomycin (0.5 × pen/strep) at 37 °C for 24 h. 100 nM and subsequent serial dilutions of anti-HER2 binders coupled to DoubleCatcher were prepared in 125 μL DMEM with 0.1% (v/v) FBS, 1% (v/v) GlutaMAX and 0.5 × pen/strep. Spent media was removed and protein solutions were added to cells. DoubleCatcher or MBP-superTEV protease alone were applied to cells at the same concentrations as the anti-HER2 binders as a negative control for cell killing. Cells were treated for 96 h, after which 40 μL 0.15 mg/mL Resazurin (Alamar Blue; Sigma-Aldrich) prepared in PBS pH 7.4 and dissolved in PBS and sterile-filtered through a 0.22 μM pore-sized membrane filter was added directly to each well, and plates were incubated for 4 h at 37 °C with 5% (v/v) CO_2. Fluorescence (λ_{ex} 544 nm, λ_{em} 590 nm) was measured using a FLUOstar Omega plate reader (BMG Labtech). The

fluorescence intensity of media, without cells, containing Resazurin ('background') was subtracted from each measurement. The percent metabolic activity relative to untreated cells was then defined as 100 × (background-subtracted fluorescence of cells)/(background-subtracted fluorescence of untreated cells). Untreated cells are defined as cells grown in DMEM + 1% (v/v) FBS, 0.5 × Pen/Strep alone.

Mass Spectrometry

Intact protein mass spectrometry in positive ion mode was performed using a RapidFire 365 jet-stream electrospray ion source (Agilent) coupled to a 6550 Accurate-Mass Quadrupole Time-of-Flight (Q-TOF) (Agilent) mass spectrometer. 50 µL protein samples at 10 µM in PBS pH 7.4 were prepared on a 384-well polypropylene plate (Greiner). Samples were acidified to 1% (v/v) formic acid, before aspiration under vacuum for 0.4 s and loading onto a C4 solid-phase extraction cartridge. Following washes with 0.1% (v/v) formic acid in water (1.5 mL/min flow rate for 5.5 s), samples were eluted to the Q-TOF detector with deionized water containing 85% (v/v) acetonitrile and 0.1% (v/v) formic acid (1.25 mL/min flow rate for 5.5 s). Data were analyzed using Mass Hunter Qualitative Analysis software B.07.00 (Agilent). Expected molecular weights for full-length proteins were calculated using the ExPASy ProtParam tool, with the N-terminal fMet (bacterial expression) or signal peptide sequence (mammalian expression) removed. -2 Da was calculated for each predicted disulfide bond. The small amount of +178 Da peak relates to gluconoylation, a common spontaneous post-translational modification for proteins overexpressed in E. coli BL21. Signal peptide cleavage was predicted by SignalP 6.0.

Mass Photometry

Microscope coverslips (24 × 50 mm, Menzel Gläser) were cleaned by sequential sonication in 50% (v/v) isopropanol and Milli-Q H_2O (5 min each), dried under a clean nitrogen stream and assembled into flow chambers with isopropanol/Milli-Q H_2O-rinsed and nitrogen-dried silicone gaskets (6 mm × 1 mm, GBL103280, Grace Bio-Labs). Data were acquired on a Refeyn TwoMP mass photometer. Immediately prior to mass photometry, protein samples were diluted into freshly prepared and degassed MP sample buffer [PBS pH 7.4, 300 mM NaCl, 0.5% (v/v) glycerol, 1 mM NaN_3, sterile-filtered through a 0.22 µm pore-size syringe filter (Thermo Fisher)]. A dynamin protein mass standard (MS1000 20-40×), diluted into

MP sample buffer, was used to calibrate the mass photometer each time a new cover slip was used. The focal position of the microscope was found in flow chambers containing 20 µL sample, after which 0.5-10 µL of MP sample buffer was replaced with an equivalent volume protein sample, to a final working concentration of 2-50 nM, to account for differences in the dissociation characteristics of different proteins. After ≤ 10 s, images were acquired for 60 s at 331 Hz. Each sample was measured at least twice independently. All data were processed using DiscoverMP v1.2.3 software (Refeyn Ltd). Masses were plotted as mass histogram (bin width = 2.5 kDa) and fitted to Gaussian non-linear regression curves to identify the mean masses of each protein subpopulation within the heterogeneous sample in GraphPad Prism 9 (GraphPad Software). Expected molecular weights for protein constituents were calculated using the ExPASy ProtParam tool, with the N-terminal fMet (bacterial expression) or signal peptide sequence (mammalian expression) removed.

CLAIMS

1. A bivalent polypeptide scaffold comprising:

(a) a first peptide-binding domain linked to a second peptide-binding domain by a first linker domain; and

(b) a blocking peptide linked to the first or second peptide-binding domain by a second linker domain comprising a proteolytic cleavage site,

wherein:

(i) the first peptide-binding domain comprises a binding site for a first cognate peptide and is capable of spontaneously forming an isopeptide bond with the first cognate peptide;

(ii) the second peptide-binding domain comprises a binding site for a second cognate peptide and is capable of spontaneously forming an isopeptide bond with the second cognate peptide; and

(iii) the blocking peptide binds selectively and reversibly to the binding site of the cognate peptide in the peptide-binding domain to which the blocking peptide is linked via the second linker domain.

2. The polypeptide scaffold of claim 1, wherein the first and second peptide-binding domains each comprise or consist of a SpyCatcher polypeptide or a functional variant thereof and the first and second cognate peptides each comprise or consist of a SpyTag peptide or a functional variant thereof.

3. The polypeptide scaffold of claim 2, wherein the SpyCatcher polypeptide is SpyCatcher003 or a functional variant thereof.

4. The polypeptide scaffold of any one of claims 1 to 3, wherein the first and second peptide-binding domains each comprise an amino acid sequence with at least 70% sequence identity to a sequence as set forth in any one of SEQ ID NOs: 1-12, wherein the amino acid sequence comprises a lysine at position 31 and a glutamic acid at position 77, wherein the specified amino acid residues are at positions equivalent to the positions in SEQ ID NO: 1.

5. The polypeptide scaffold of any one of claims 1 to 4, wherein an amino acid in the first peptide-binding domain is linked to an amino acid in the second peptide-binding domain by a covalent bond, optionally a disulfide bond.

6. The polypeptide scaffold of any one of claims 1 to 5, wherein the first peptide-binding domain contains a cysteine residue and the second peptide-binding domain contains a cysteine residue, wherein the cysteine residues react with each other to form a disulfide bond.

7. The polypeptide scaffold of any one of claims 1 to 6, wherein the first and second peptide-binding domains each comprise an amino acid sequence with at least 70% sequence identity to a sequence as set forth in any one of SEQ ID NOs: 1-4, wherein the amino acid sequence comprises a lysine at position 31, a glutamic acid at position 77 and a cysteine residue at position 4, 12, 14, 50, 51, 52, 62 or 69, wherein the specified amino acid residues are at positions equivalent to the positions in SEQ ID NO: 1.

8. The polypeptide scaffold of claim 7, wherein:

(a) the first peptide-binding domain has a cysteine residue at position 50 and the second peptide-binding domain has a cysteine residue at position 50;

(b) the first peptide-binding domain has a cysteine residue at position 4 and the second peptide-binding domain has a cysteine residue at position 14;

(c) the first peptide-binding domain has a cysteine residue at position 12 and the second peptide-binding domain has a cysteine residue at position 51;

(d) the first peptide-binding domain has a cysteine residue at position 69 and the second peptide-binding domain has a cysteine residue at position 69; or

(e) the first peptide-binding domain has a cysteine residue at position 52 and the second peptide-binding domain has a cysteine residue at position 62.

9. The polypeptide scaffold of any one of claims 1 to 8, wherein:

(a) the first peptide-binding domain comprises:

(i) an amino acid sequence selected from a sequence as set forth in any one of SEQ ID NOs: 1-12, 45, or 46, preferably any one of SEQ ID NOs: 1, 4-6, 8, 10 or 11; or

(ii) an amino acid sequence selected from a sequence as set forth in any one of SEQ ID NOs: 1-4, 45 or 46, wherein:

(1) the leucine residue at position 4 is substituted with cysteine;

(2) the glycine residue at position 12 is substituted with cysteine;

(3) the serine residue at position 14 is substituted with cysteine;

(4) the serine residue at position 50 is substituted with cysteine;

(5) the glycine residue at position 51 is substituted with cysteine;

(6) the lysine residue at position 52 is substituted with cysteine;

(7) the histidine residue at position 62 is substituted with cysteine; or

(8) the tyrosine residue at position 69 is substituted with cysteine;

and/or

(b) the second peptide-binding domain comprises:

(i) an amino acid sequence selected from a sequence as set forth in any one of SEQ ID NOs: 1-12, 45 or 46, preferably any one of SEQ ID NOs: 1, 4, 5, 7, 9, 10 or 12; or

(ii) an amino acid sequence selected from a sequence as set forth in any one of SEQ ID NOs: 1-4, 45 or 46, wherein:

(1) the leucine residue at position 4 is substituted with cysteine;

(2) the glycine residue at position 12 is substituted with cysteine;

(3) the serine residue at position 14 is substituted with cysteine;

(4) the serine residue at position 50 is substituted with cysteine;

(5) the glycine residue at position 51 is substituted with cysteine;

(6) the lysine residue at position 52 is substituted with cysteine;

(7) the histidine or glutamine residue at position 62 is substituted with cysteine; or

(8) the tyrosine residue at position 69 is substituted with cysteine.

10. The polypeptide scaffold of any one of claims 2 to 9, wherein the SpyTag peptide is SpyTag (SEQ ID NO: 13), SpyTag002 (SEQ ID NO: 14) or SpyTag003 (SEQ ID NO: 15).

11. The polypeptide scaffold of any one of claims 1 to 10, wherein the blocking peptide is capable of competing with the cognate peptide for the binding site of the peptide-binding domain to which it is linked.

12. The polypeptide scaffold of any one of claims 1 to 11, wherein the blocking peptide comprises a non-reactive SpyTag peptide, which does not contain an aspartic acid residue that reacts to form the isopeptide bond with a lysine residue in

a SpyCatcher polypeptide, preferably wherein the aspartic acid residue has been substituted with a non-acidic amino acid.

13. The polypeptide scaffold of claim 12, wherein the non-reactive SpyTag peptide comprises or consists of:

(a) an amino acid sequence as set forth in SEQ ID NO: 16, wherein X at position 10 is not an acidic amino acid, preferably wherein X is selected from alanine, serine, threonine and valine; or

(b) a functional variant of (a) comprising 1-5 amino acid substitutions and/or deletions.

14. The polypeptide scaffold of claim 12 or 13, wherein the non-reactive SpyTag peptide comprises or consists of an amino acid sequence as set forth in SEQ ID NO: 17, 18 or 19.

15. The polypeptide scaffold of any one of claims 1 to 14, wherein the first linker domain comprises a flexible peptide, optionally a peptide comprising or consisting of a serine and/or glycine-rich sequence.

16. The polypeptide scaffold of any one of claims 1 to 14, wherein the first linker domain comprises a rigid peptide, optionally a peptide comprising or consisting of an alpha helix structure.

17. The polypeptide scaffold of claim 16, wherein the alpha helix structure is from a ribosomal protein, optionally the *Bacillus stearothermophilus* ribosomal protein L9.

18. The polypeptide scaffold of any one of claims 1 to 17, wherein the first linker domain consists of about 5-300 amino acids, optionally about 6-300, 7-290, 8-280, 9-270, 10-260, 11-250 or 12-240 amino acids.

19. The polypeptide scaffold of any one of claims 1 to 18, wherein the first linker domain comprises or consists of an amino acid sequence selected from a sequence as set forth in SEQ ID NOs: 20 and 21.

20. The polypeptide scaffold of any one of claims 1 to 19, wherein the second linker domain consists of a flexible peptide comprising a proteolytic cleavage site, optionally wherein the flexible peptide comprises a serine and/or glycine-rich sequence.

21. The polypeptide scaffold of any one of claims 1 to 20, wherein the second linker domain consists of a peptide comprising a proteolytic cleavage site bordered by serine and/or glycine-rich sequences.

22. The polypeptide scaffold of any one of claims 1 to 21, wherein the second linker domain consists of about 12-40 amino acids, optionally about 13-35, 14-30 or 15-25 amino acids.

23. The polypeptide scaffold of any one of claims 1 to 22, wherein the proteolytic cleavage site is a Tobacco Etch Virus (TEV) protease recognition site or a thrombin recognition site.

24. The polypeptide scaffold of any one of claims 1 to 23, wherein the polypeptide scaffold comprises a purification tag, optionally a His-tag, FLAG-tag, Calmodulin-tag, E-tag, HA-tag, Myc-tag, S-tag, SBP-tag, Softag 1, Softag 3, Strep-tag, Strep-tag II, TC tag, V5 tag, VSV-tag or Xpress tag.

25. The polypeptide scaffold of any one of claims 1 to 24, wherein the polypeptide scaffold comprises from N-terminal to C-terminal:

(A)

(i) the first peptide-binding domain;

(ii) the first linker domain;

(iii) the second peptide-binding domain;

(iv) the second linker domain comprising a proteolytic cleavage site; and

(v) the blocking peptide; or

(B)

(i) the blocking peptide;

(ii) the second linker domain comprising a proteolytic cleavage site;

(iii) the first peptide-binding domain;

(iv) the first linker domain; and

(v) second peptide-binding domain.

26. The polypeptide scaffold of any one of claims 1 to 25, wherein the polypeptide scaffold comprises an amino acid sequence with at least 80% sequence identity to a sequence as set forth in any one of SEQ ID NOs: 22-28, wherein the amino acid sequence comprises:

(i) a lysine at position 31 and 122;

(ii) a glutamic acid at position 77 and 199;

(iii) a non-acidic amino acid at position 259, preferably selected from alanine, serine, threonine or valine; and optionally

(iv) a cysteine residue at a position selected from position 4, 12, 14, 50, 51, 52, 62 and 69, preferably a position selected from position 4, 12, 50, 52 and 69; and

(v) a cysteine residue at a position selected from position 126, 134, 136, 172, 173, 174, 184 and 191, preferably a position selected from position 136, 172, 173, 184 and 191;

wherein the specified amino acid residues are at positions equivalent to the positions in SEQ ID NO: 22.

27. The polypeptide scaffold of any one of claims 1 to 26, wherein the polypeptide scaffold comprises or consists of an amino acid sequence as set forth in any one of SEQ ID NOs: 22 to 37 or 62 to 69.

28. A nucleic acid molecule encoding a bivalent polypeptide scaffold as defined in any one of claims 1 to 27.

29. A vector comprising the nucleic acid molecule of claim 28.

30. Use of a bivalent polypeptide scaffold as defined in any one of claims 1 to 27 to conjugate a first molecule and a second molecule, wherein each molecule is linked to a cognate peptide capable of spontaneously forming an isopeptide bond with a peptide-binding domain in the polypeptide scaffold.

31. The use of claim 30 to produce a molecule with bispecific binding activity.

32. A method for conjugating two molecules comprising:

(i) contacting a bivalent polypeptide scaffold as defined in any one of claims 1 to 27 with a first molecule linked to a cognate peptide under conditions that enable the spontaneous formation of an isopeptide bond between the cognate peptide and a peptide-binding domain, thereby conjugating the first molecule to a peptide-binding domain of the polypeptide scaffold to form a first reaction product;

(ii) contacting the first reaction product with a protease under conditions that enable the specific cleavage of the second linker domain to form a second reaction product;

(iii) contacting the second reaction product with a second molecule linked to a cognate peptide under conditions that enable the spontaneous formation of an isopeptide bond between the cognate peptide and the other peptide-binding domain, thereby conjugating the second molecule to the other peptide-binding domain of the polypeptide scaffold to form a complex comprising the molecules conjugated via the polypeptide scaffold.

33. The method of claim 32 comprising a step of isolating or purifying the complex.

34. The method of claim 32 or 33, the method being a method for producing a molecule with bispecific binding activity, wherein each molecule linked to a cognate peptide selectively binds to a different target or epitope.

35. The use of claim 30 or 31, or the method of any one of claims 32 to 34, wherein the cognate peptide is a SpyTag peptide or a functional variant thereof, optionally wherein the cognate peptide comprises or consists of an amino acid sequence as set forth in any one of SEQ ID NOs: 13-15, preferably SEQ ID NO: 15.

36. The use of any one of claims 30, 31 or 35 or the method of any one of claims 32 to 35, wherein the first and second molecule are independently selected from protein (e.g. an antibody or binding fragment thereof), peptide, small-molecule organic compound (e.g. drug), nucleic acid molecule, fluorophore, metal-ligand complex, polysaccharide, nanoparticle, 2D monolayer (e.g. graphene), nanotube, polymer, cell, virus or virus-like particle.

37. The use or method of claim 36, wherein the first and second molecules are antibodies or binding fragments thereof that bind to different targets or epitopes.

38. A kit comprising:

(1)

(a) a plurality of bivalent polypeptide scaffolds as defined in any one of claims 1 to 27, wherein each polypeptide scaffold comprises a different first linker domain and/or the peptide-binding domains of each polypeptide scaffold are covalently linked by a disulfide bonds at different positions, and wherein each bivalent polypeptide scaffold is provided separately; and optionally

(b) means for isolating the polypeptide scaffold, e.g. unreacted polypeptide scaffold; or

(2)

(a) a plurality of bivalent polypeptide scaffolds as defined in any one of claims 1 to 27, wherein each polypeptide scaffold allows molecules conjugated to the peptide-binding domains to be displayed in a different spatial configuration to the other polypeptide scaffolds, and wherein each bivalent polypeptide scaffold is provided separately; and optionally

(b) means for isolating the polypeptide scaffold, e.g. unreacted polypeptide scaffold.

39. Use of bivalent polypeptide scaffold as defined in any one of claims 1 to 27 or the kit as defined in claim 38 to prepare a library of assembled structures.

40. A plurality of assembled structures containing an assembled structure comprising a first peptide-binding domain linked to a second peptide-binding domain by a linker domain, wherein each peptide-binding domain is conjugated to its cognate peptide via an isopeptide bond and wherein each cognate peptide is linked to a different molecule, and wherein:

(a) the linker domain is a first linker domain as defined in any preceding claim, preferably wherein the linker domain comprises a rigid peptide; and/or

(b) an amino acid in the first peptide-binding domain is linked to an amino acid in the second peptide-binding domain by a covalent bond, optionally a disulfide bond.

Figure 1

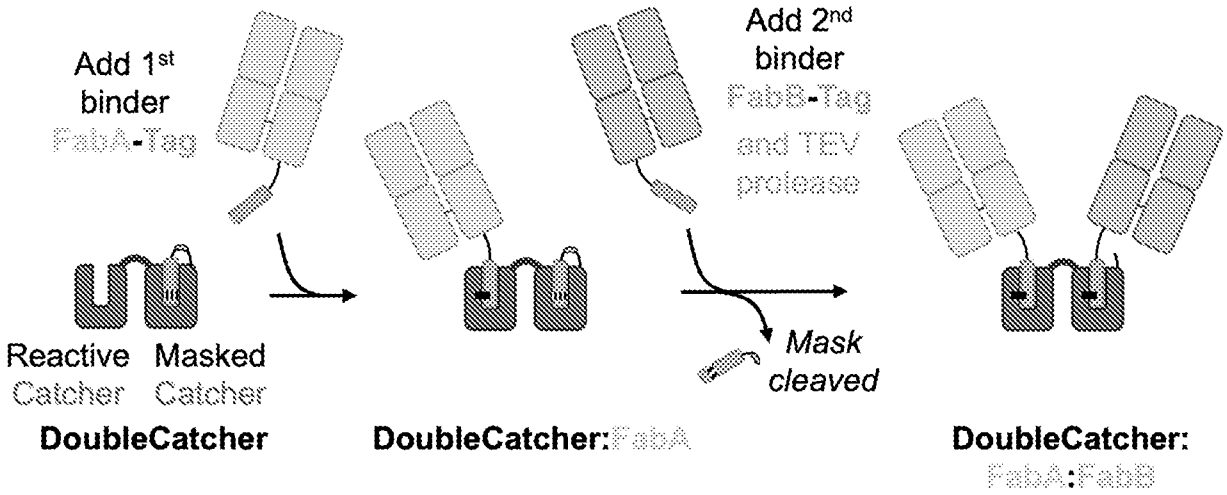

Add 1st binder
FabA-Tag
Reactive Catcher
Masked Catcher
DoubleCatcher
DoubleCatcher:FabA
Add 2nd binder
FabB-Tag
and TEV protease
Mask cleaved
DoubleCatcher:
FabA:FabB

Figure 2

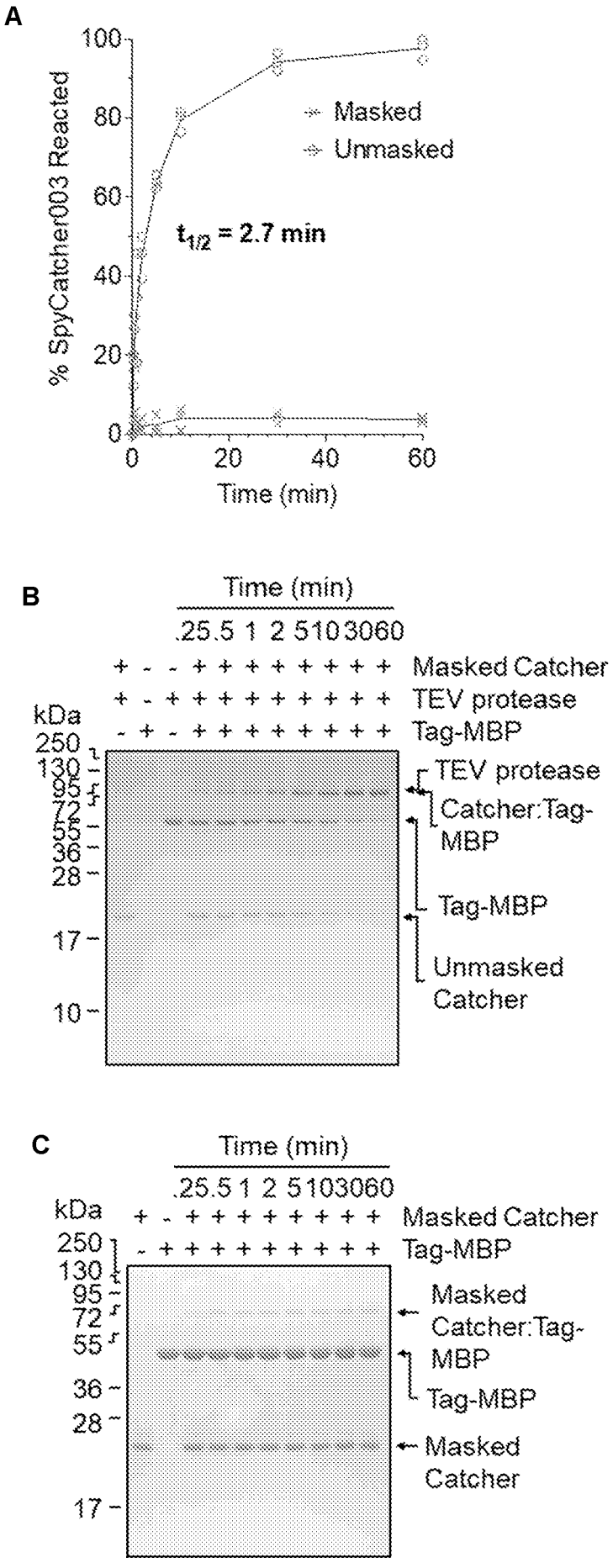

A
100
80
60
40
20
0
% SpyCatcher003 Reacted
Masked
Unmasked
$t_{1/2}$ = 2.7 min
0
20
40
60
Time (min)
B
Time (min)
.25.5 1 2 5 10 30 60
+ - - + + + + + + + + + Masked Catcher
+ - + + + + + + + + + + TEV protease
- + - + + + + + + + + + Tag-MBP
kDa
250
130
95
72
55
36
28
17
10
TEV protease
Catcher:Tag-MBP
Tag-MBP
Unmasked Catcher
C
Time (min)
.25.5 1 2 5 10 30 60
kDa
+ - + + + + + + + + + Masked Catcher
- + + + + + + + + + + Tag-MBP
250
130
95
72
55
36
28
17
Masked Catcher:Tag-MBP
Tag-MBP
Masked Catcher

Figure 3

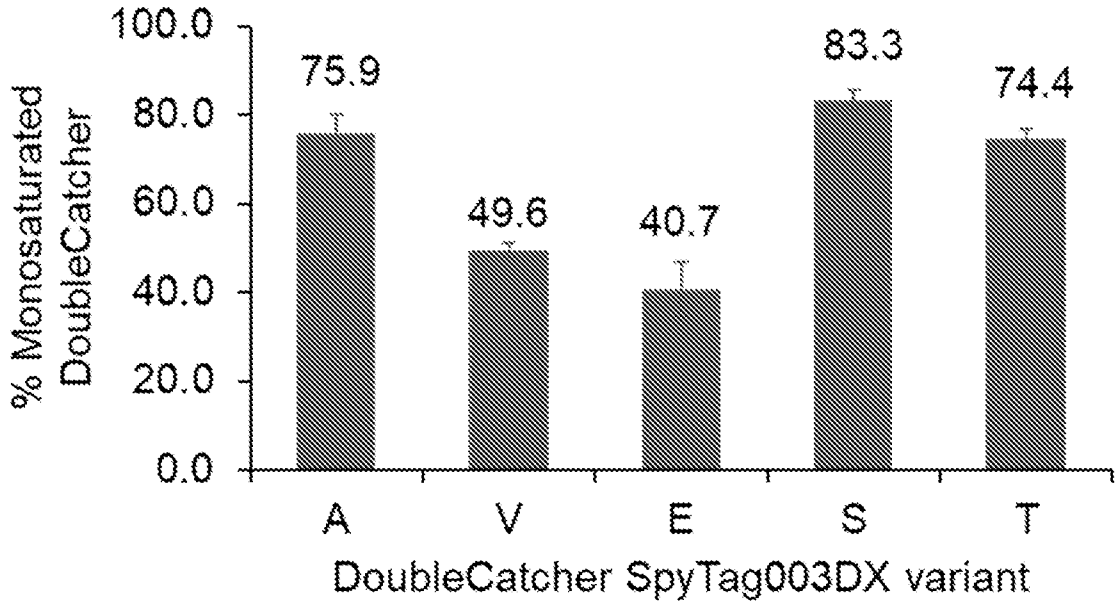

100.0
80.0
60.0
40.0
20.0
0.0
% Monosaturated DoubleCatcher
75.9
49.6
40.7
83.3
74.4
A
V
E
S
T
DoubleCatcher SpyTag003DX variant

Figure 4

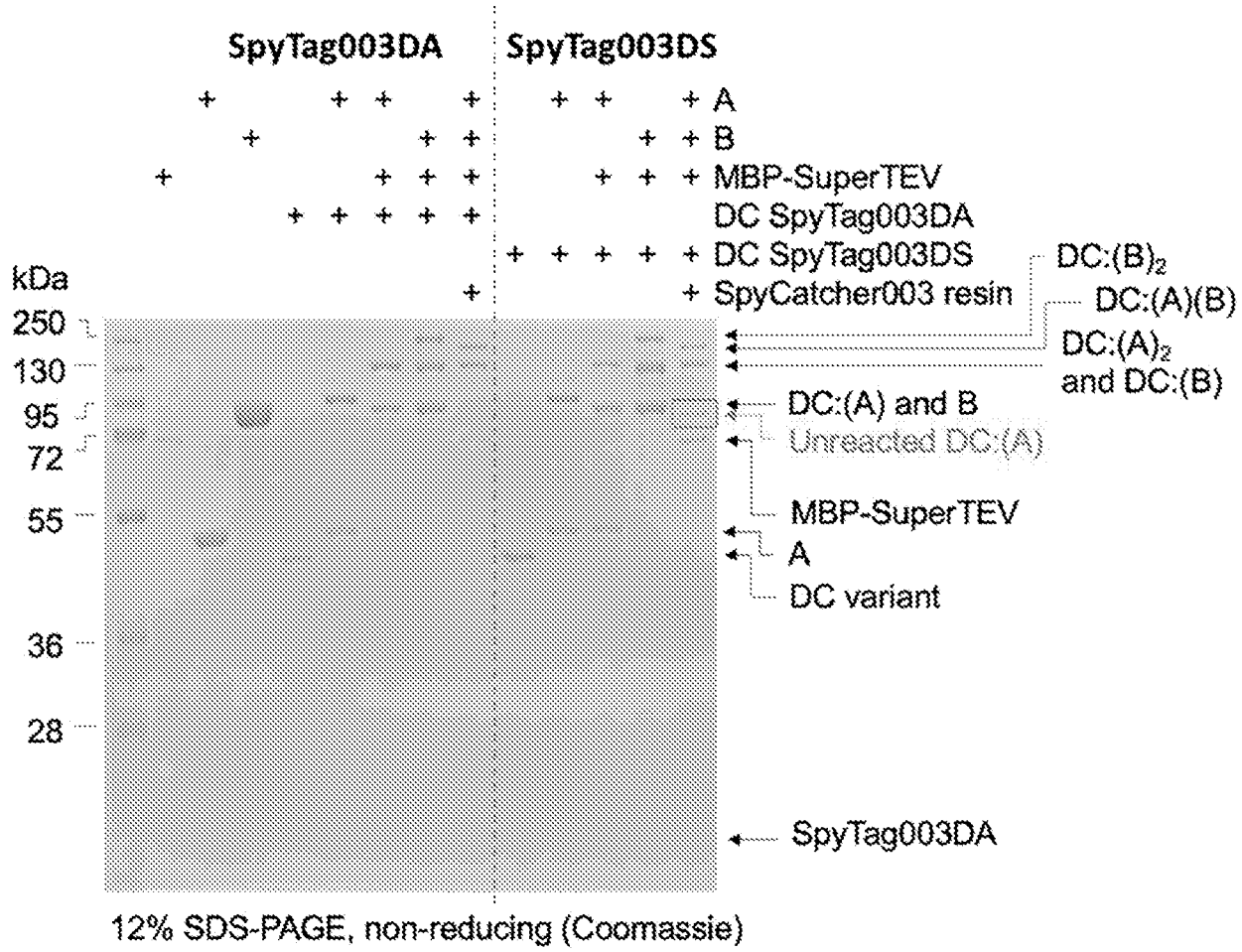

SpyTag003DA
SpyTag003DS
A
B
MBP-SuperTEV
DC SpyTag003DA
DC SpyTag003DS
SpyCatcher003 resin
kDa
250
130
95
72
55
36
28
DC:(B)2
DC:(A)(B)
DC:(A)2
and DC:(B)
DC:(A) and B
Unreacted DC:(A)
MBP-SuperTEV
A
DC variant
SpyTag003DA
12% SDS-PAGE, non-reducing (Coomassie)

Figure 5

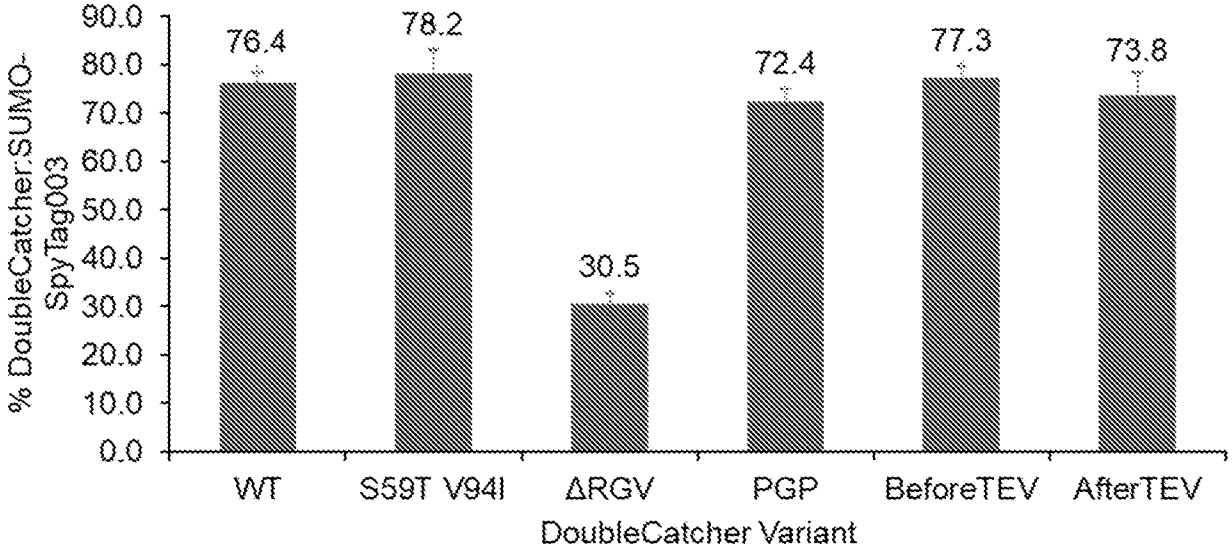
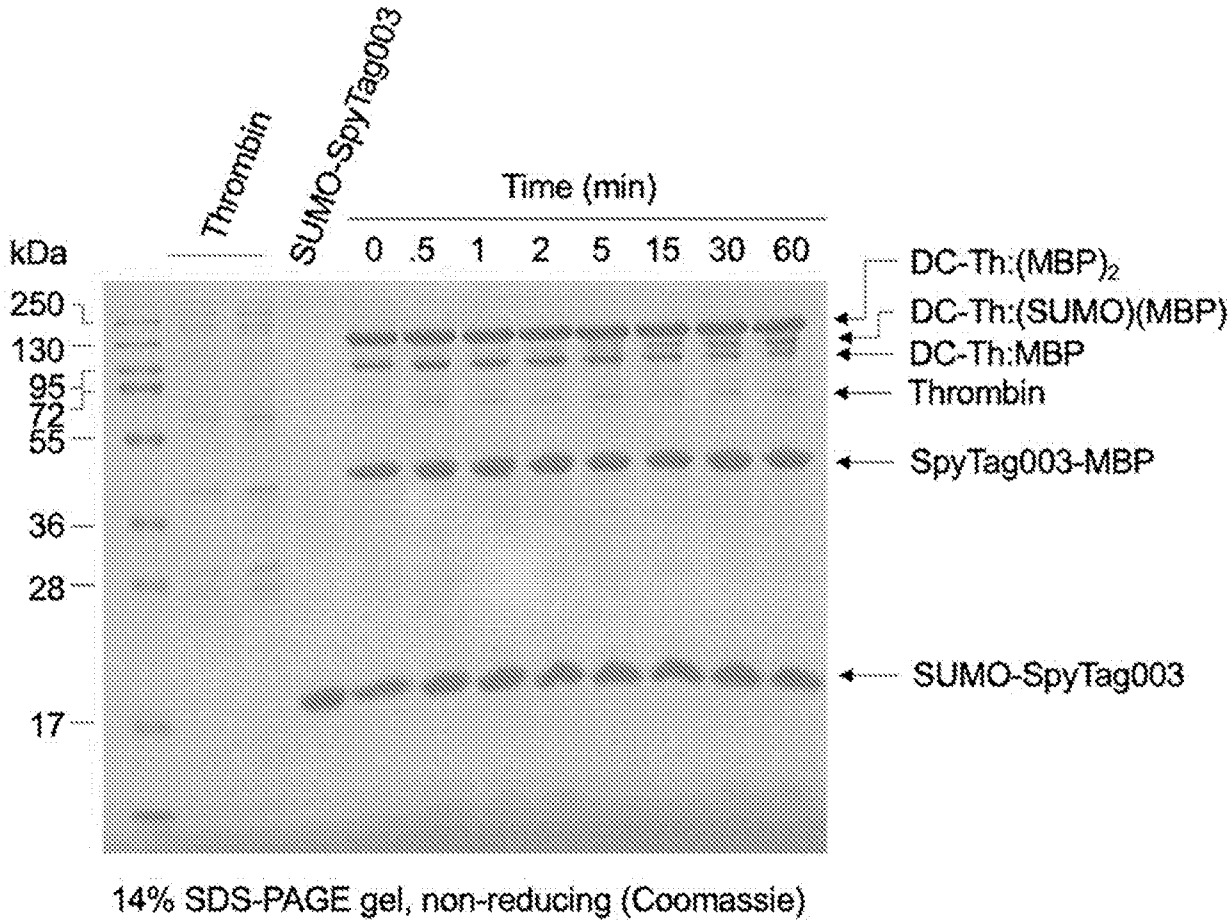

% DoubleCatcher:SUMO-SpyTag003
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
76.4
78.2
30.5
72.4
77.3
73.8
WT
S59T V94I
ΔRGV
PGP
BeforeTEV
AfterTEV
DoubleCatcher Variant

Figure 6



14% SDS-PAGE gel, non-reducing (Coomassie)

Figure 7

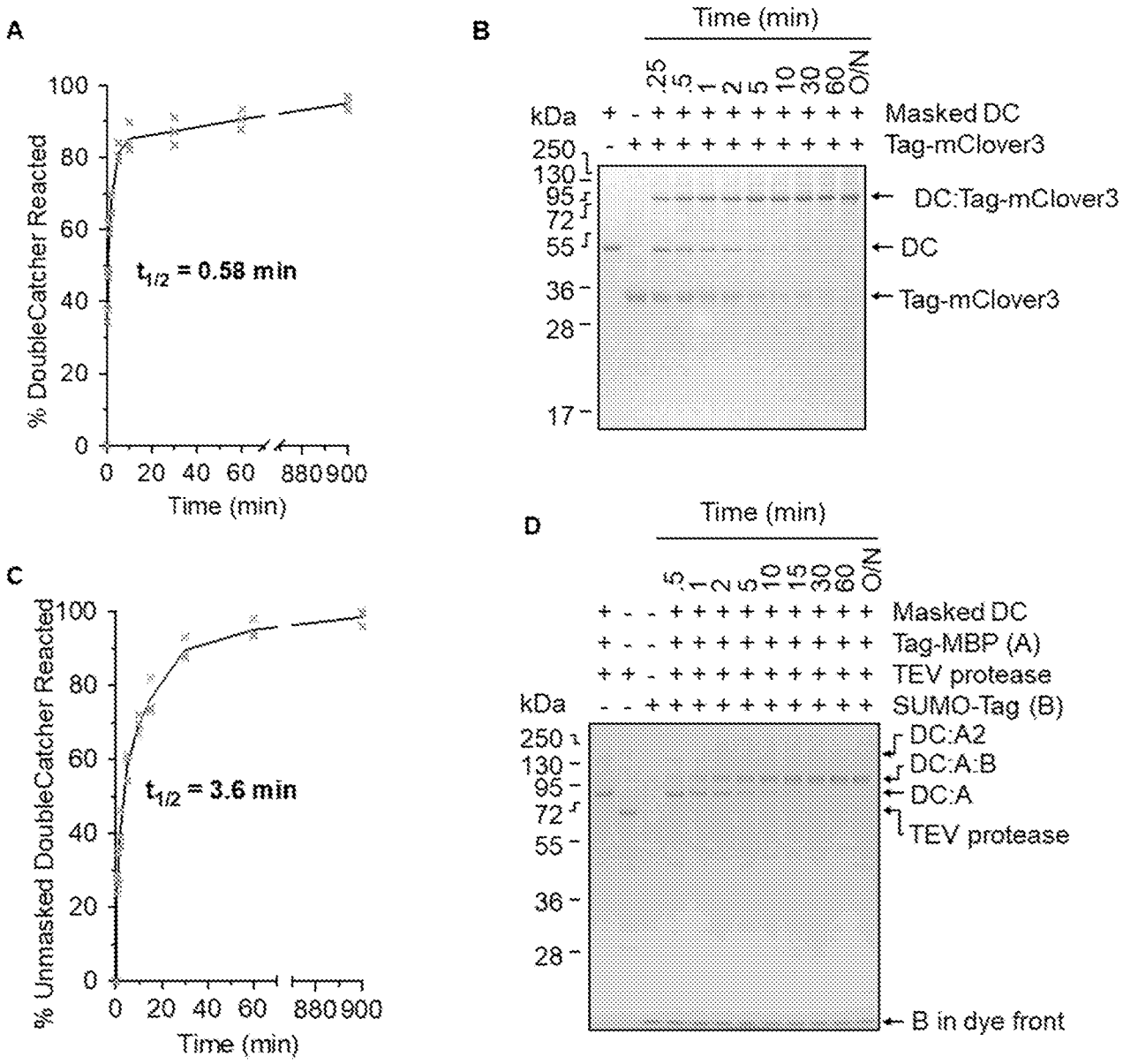

A
% DoubleCatcher Reacted
$t_{1/2}$ = 0.58 min
Time (min)
B
Time (min)
kDa
Masked DC
Tag-mClover3
DC:Tag-mClover3
DC
Tag-mClover3
C
% Unmasked DoubleCatcher Reacted
$t_{1/2}$ = 3.6 min
Time (min)
D
Time (min)
Masked DC
Tag-MBP (A)
TEV protease
SUMO-Tag (B)
DC:A2
DC:A:B
DC:A
TEV protease
B in dye front

Figure 8

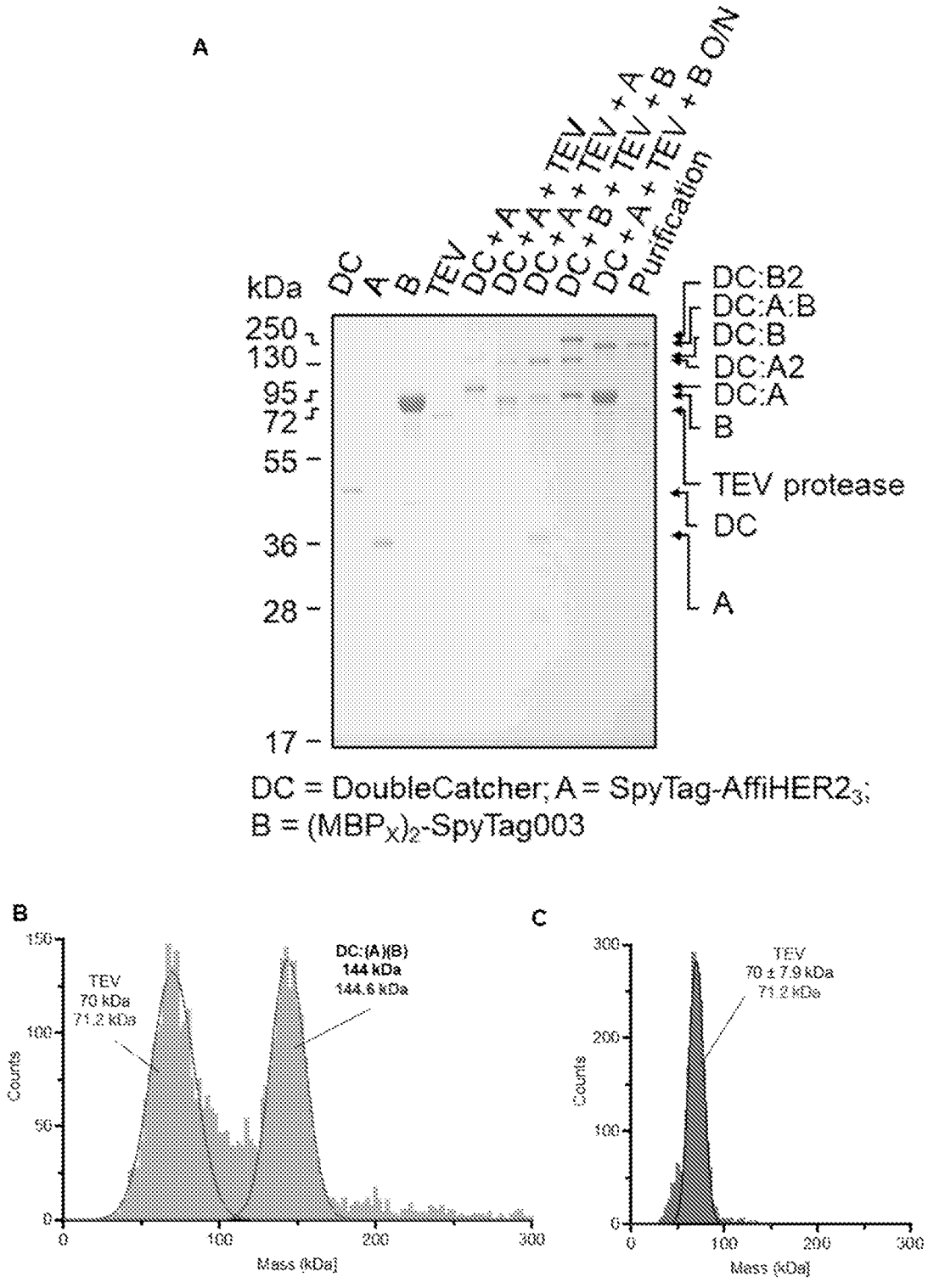

A
kDa
DC
A
B
TEV
DC + A
DC + A + TEV
DC + A + B
DC + A + B + TEV
Purification
250
130
95
72
55
36
28
17
DC:B2
DC:A:B
DC:B
DC:A2
DC:A
B
TEV protease
DC
A
DC = DoubleCatcher; A = SpyTag-AffiHER2$_3$;
B = (MBP$_X$)$_2$-SpyTag003
B
TEV
70 kDa
71.2 kDa
DC:(A)(B)
144 kDa
144.6 kDa
Counts
Mass (kDa)
C
TEV
70 ± 7.9 kDa
71.2 kDa
Counts
Mass (kDa)

Figure 9

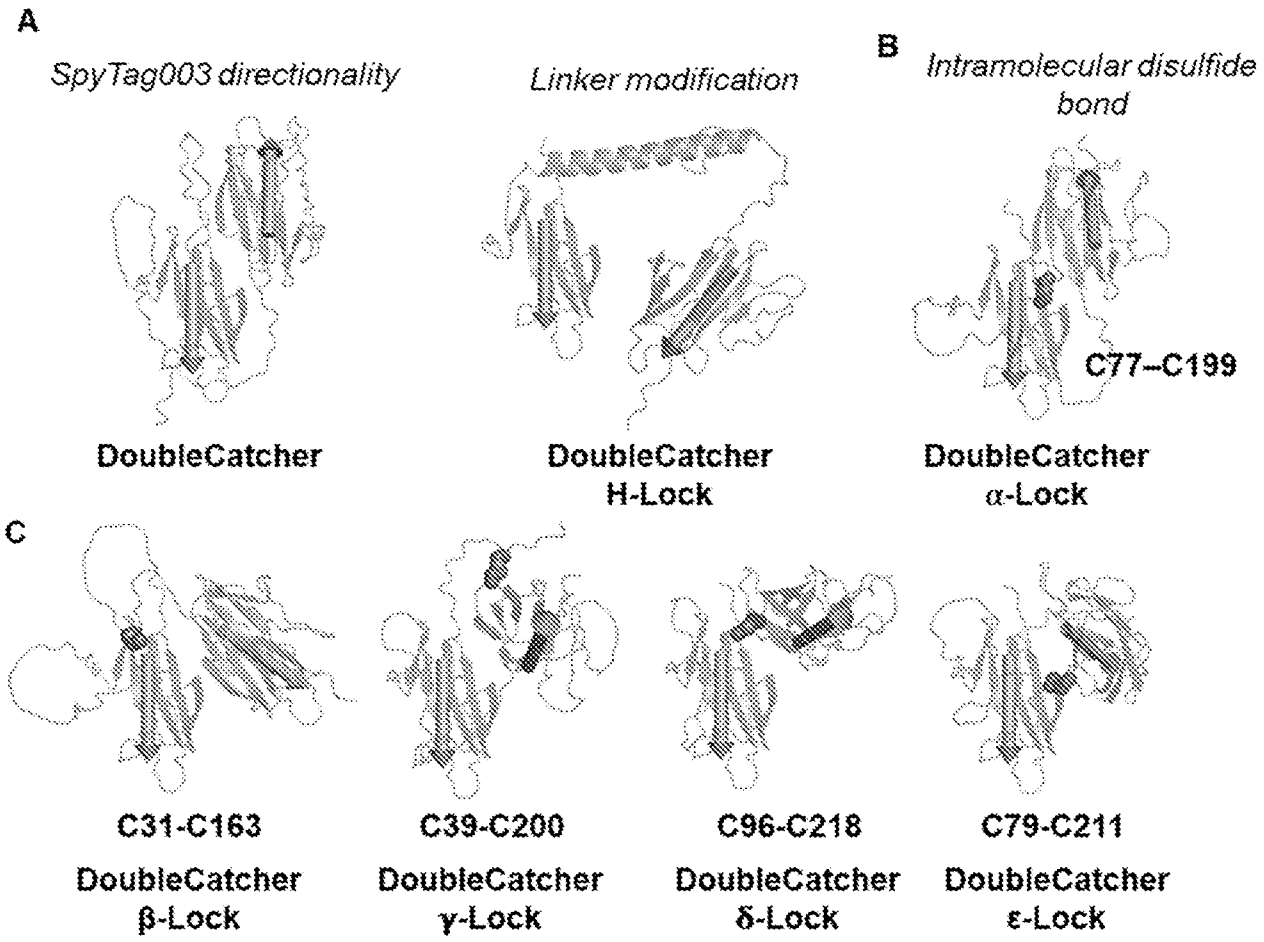

A
SpyTag003 directionality
Linker modification
B
Intramolecular disulfide bond
C77–C199
DoubleCatcher
DoubleCatcher H-Lock
DoubleCatcher α-Lock
C
C31-C163
C39-C200
C96-C218
C79-C211
DoubleCatcher β-Lock
DoubleCatcher γ-Lock
DoubleCatcher δ-Lock
DoubleCatcher ε-Lock

Figure 10

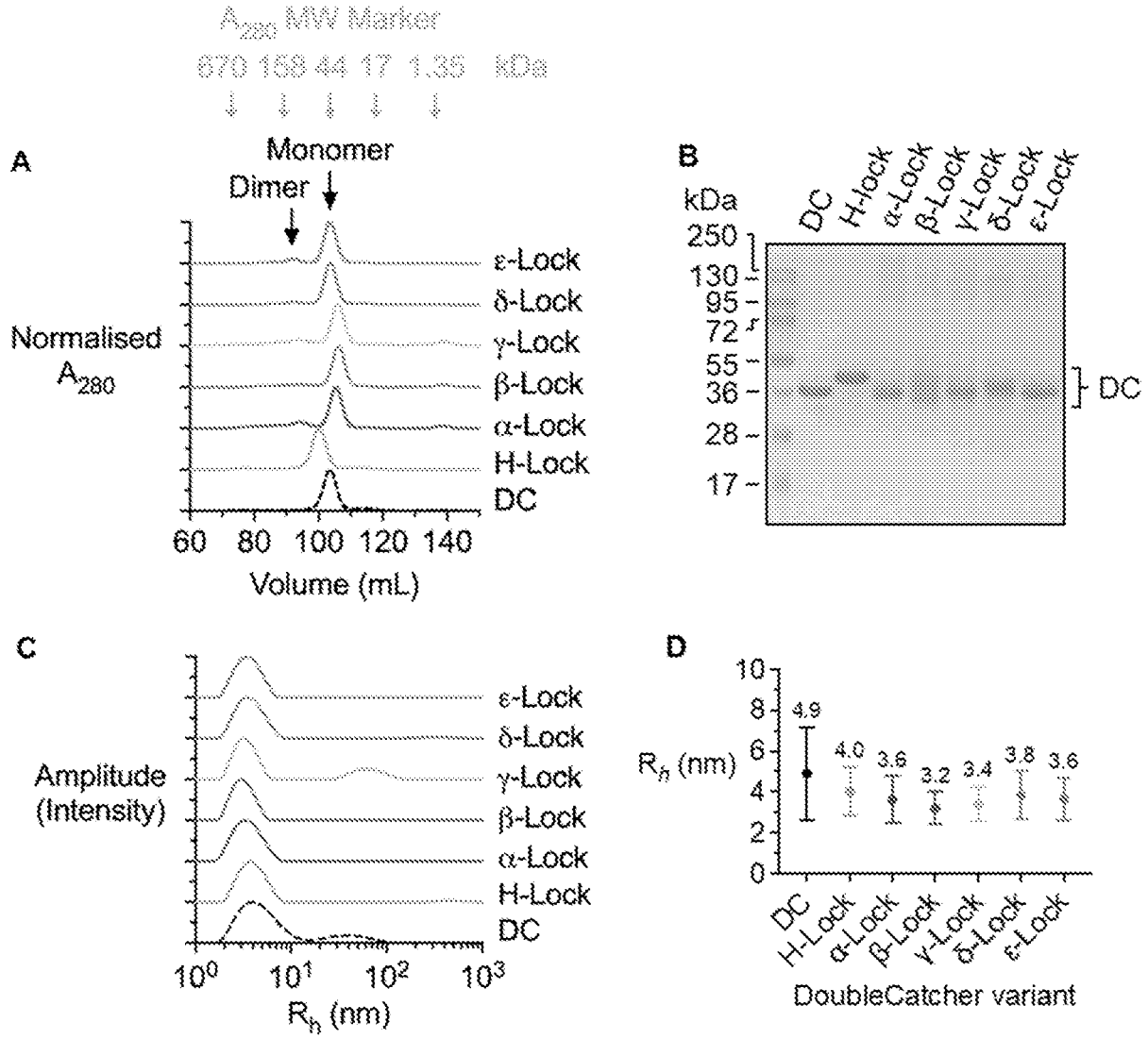

A280 MW Marker
670 158 44 17 1.35 kDa
A
Monomer
Dimer
ε-Lock
δ-Lock
γ-Lock
β-Lock
α-Lock
H-Lock
DC
Normalised A280
60 80 100 120 140
Volume (mL)
B
kDa
DC
H-Lock
α-Lock
β-Lock
γ-Lock
δ-Lock
ε-Lock
250
130
95
72
55
36
28
17
DC
C
ε-Lock
δ-Lock
γ-Lock
β-Lock
α-Lock
H-Lock
DC
Amplitude (Intensity)
10^0 10^1 10^2 10^3
Rh (nm)
D
Rh (nm)
10
8
6
4
2
0
4.9
4.0
3.6
3.2
3.4
3.8
3.6
DC
H-Lock
α-Lock
β-Lock
γ-Lock
δ-Lock
ε-Lock
DoubleCatcher variant

Figure 11

A

Metabolic Activity (% Untreated Control) in SKBR3 cells

N-terminally coupled binder \ C-terminally coupled binder	Fab0.11	39S	MF3958	H2-18	HFS2	AffiHER2	Fab0.11 NL	nanoHER2	Pertuzumab
Fab0.11	78.0 ± 6.6	61.2 ± 3.5	70.3 ± 4.3	82.7 ± 1.8	81.7 ± 2.6	154.5 ± 15.3	100.4 ± 8.7	155.7 ± 10.4	68.0 ± 1.5
39S	63.7 ± 6.6	94.1 ± 0.5	94.1 ± 4.0	94.0 ± 3.6	96.6 ± 4.4	98.3 ± 4.2	68.3 ± 2.7	119.8 ± 5.0	77.6 ± 4.7
MF3958	63.5 ± 6.4	86.6 ± 1.5	88.3 ± 1.3	91.1 ± 3.1	93.5 ± 4.1	83.8 ± 4.0	73.0 ± 6.5	95.9 ± 5.7	78.8 ± 5.9
H2-18	68.1 ± 20.8	86.6 ± 1.9	89.3 ± 3.1	93.8 ± 1.7	96.3 ± 6.5	92.6 ± 3.3	79.2 ± 2.2	109.8 ± 5.0	82.7 ± 1.8
HFS2	77.4 ± 7.9	87.4 ± 1.5	87.4 ± 5.6	89.9 ± 7.4	93.7 ± 7.0	94.8 ± 7.1	79.8 ± 4.1	119.7 ± 4.9	87.1 ± 3.1
AffiHER2	130.2 ± 7.2	80.5 ± 4.9	85.3 ± 5.3	91.8 3.2	94.7 2.6	92.7 2.5	120.2 ± 4.8	115.1 ± 2.6	72.9 ± 3.4
Fab0.11 NL	77.9 ± 3.8	59.6 ± 0.8	69.2 ± 3.1	79.9 ± 0.3	77.1 ± 1.0	141.9 ± 4.5	73.2 ± 3.1	158.4 ± 16.5	63.8 ± 4.5
nanoHER2	151.3 ± 16.2	68.5 ± 12.9	84.4 ± 0.4	84.7 ± 12.7	85.3 ± 18.4	84.1 ± 9.3	115.1 ± 43.6	97.3 ± 7.6	61.7 ± 2.8
Pertuzumab	66.3 ± 2.8	81.2 ± 6.1	73.1 ± 6.4	84.2 ± 2.6	86.2 ± 0.0	69.8 ± 4.0	58.8 ± 4.7	54.5 ± 6.0	104.1 ± 10.4

B

Metabolic Activity (% Untreated Control) in BT474 cells

N-terminally coupled binder \ C-terminally coupled binder	Fab0.11	39S	MF3958	H2-18	HFS2	AffiHER2	Fab0.11 NL	nanoHER2	Pertuzumab
Fab0.11	73.7 ± 4.1	28.5 ± 1.0	36.7 ± 3.7	68.2 ± 0.9	72.0 ± 0.6	136.5 ± 1.8	102.8 ± 1.0	146.4 ± 6.0	32.2 ± 0.9
39S	26.9 ± 1.5	98.2 ± 0.6	72.8 ± 1.8	93.0 ± 3.9	93.0 ± 4.7	72.6 ± 3.1	26.4 ± 0.9	114.9 ± 3.1	52.2 ± 1.0
MF3958	31.3 ± 1.3	86.3 ± 1.1	104.9 ± 2.3	103.9 ± 4.2	100.2 ± 4.8	70.1 ± 4.1	30.2 ± 1.0	101.9 ± 1.4	69.1 ± 2.7
H2-18	67.2 ± 1.9	95.6 ± 4.1	103.5 ± 2.6	102.6 ± 5.7	100.1 ± 4.2	96.6 ± 1.3	65.3 ± 0.7	114.5 ± 2.6	88.2 ± 3.7
HFS2	72.7 ± 2.0	102.6 ± 4.5	111.5 ± 2.3	106.4 ± 2.5	105.7 ± 2.5	112.6 ± 3.6	78.9 ± 1.9	131.9 ± 1.7	93.6 ± 2.5
AffiHER2	130.1 ± 5.4	75.9 ± 4.1	86.5 ± 4.8	104.1 ± 3.6	106.6 ± 6.2	112.6 ± 2.5	126.2 ± 2.0	131.9 ± 2.2	75.6 ± 1.6
Fab0.11 NL	74.1 ± 6.1	31.8 ± 2.0	34.0 ± 3.3	62.3 ± 5.7	68.1 ± 11.9	124.6 ± 11.1	62.2 ± 5.4	141.3 ± 2.5	30.2 ± 2.8
nanoHER2	142.5 ± 1.5	108.3 ± 1.2	98.4 ± 2.4	112.9 ± 1.5	123.2 ± 1.5	123.8 ± 0.9	146.6 ± 1.1	124.6 ± 1.8	102.7 ± 1.7
Pertuzumab	33.6 ± 0.5	65.4 ± 2.5	71.2 ± 2.1	96.9 ± 1.6	93.6 ± 2.0	73.9 ± 1.3	36.6 ± 2.2	106.3 ± 2.9	99.2 ± 3.7

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2024/052240

A. CLASSIFICATION OF SUBJECT MATTER

INV. C07K14/315 C07K16/00 C07K16/18 C07K16/46 B82Y40/00
C12N15/10 C40B40/10

ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS, COMPENDEX, Sequence Search, EMBASE, FSTA, INSPEC, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	WO 2016/193746 A1 (UNIV OXFORD INNOVATION LTD [GB]) 8 December 2016 (2016-12-08) cited in the application the whole document claim 1, Fig. 1, page 24, line 27 to page 26, line 13; claims 13, 14; page 25, lines 15-17 -----	1-11, 15-25, 28-40 1-40
A	US 2020/299358 A1 (KNAPPIK HANS JOACHIM [DE] ET AL) 24 September 2020 (2020-09-24) the whole document [0006],[0007], [0036], [0052], [0067]-]0068], [0075]; figure 4 ----- -/--	1-40

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance;; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance;; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
11 November 2024	20/11/2024

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Madruga, Jaime

1

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2024/052240

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KIMURA HAYATO ET AL: "Rapid and homogeneous electrochemical detection by fabricating a high affinity bispecific antibody-enzyme complex using two Catcher/Tag systems", BIOSENSORS AND BIOELECTRONICS, ELSEVIER SCIENCE LTD, UK, AMSTERDAM , NL, vol. 175, 15 December 2020 (2020-12-15), XP086435784, ISSN: 0956-5663, DOI: 10.1016/J.BIOS.2020.112885 [retrieved on 2020-12-15] the whole document figures 1, 2	1-40

A	MEI LINGHAN ET AL: "Rapid Production of Bispecific Antibodies from Off-the-Shelf IgGs with High Yield and Purity", BIOCONJUGATE CHEMISTRY, vol. 33, no. 1, 12 December 2021 (2021-12-12), pages 134-141, XP093220939, US ISSN: 1043-1802, DOI: 10.1021/acs.bioconjchem.1c00476 the whole document figure 1	1-40

A	KYOHEI YUMURA ET AL: "Use of SpyTag/SpyCatcher to construct bispecific antibodies that target two epitopes of a single antigen", JOURNAL OF BIOCHEMISTRY, vol. 162, no. 3, 16 June 2017 (2017-06-16) , pages 203-210, XP055719652, GB ISSN: 0021-924X, DOI: 10.1093/jb/mvx023 the whole document figure 1	1-40

A	HAYATO KIMURA ET AL: "Convenient and Universal Fabrication Method for Antibody-Enzyme Complexes as Sensing Elements Using the SpyCatcher/SpyTag System", ANALYTICAL CHEMISTRY, vol. 90, no. 24, 14 November 2018 (2018-11-14), pages 14500-14506, XP055669298, US ISSN: 0003-2700, DOI: 10.1021/acs.analchem.8b04344 the whole document figures 2, 3	1-40
	-/--	

1

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2024/052240

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	-& HAYATO KIMURA ET AL: "Convenient and Universal Fabrication Method for Antibody-Enzyme Complexes as Sensing Elements Using the SpyCatcher/SpyTag System", ANALYTICAL CHEMISTRY, vol. 90, no. 24, 14 November 2018 (2018-11-14), pages 14500-14506, XP055669301, US ISSN: 0003-2700, DOI: 10.1021/acs.analchem.8b04344 the whole document figure S4 -----	1-40
Y,P	HENTRICH CHRISTIAN ET AL: "Engineered reversible inhibition of SpyCatcher reactivity enables rapid generation of bispecific antibodies", NATURE COMMUNICATIONS, vol. 15, no. 1, 15 July 2024 (2024-07-15), XP093220727, UK ISSN: 2041-1723, DOI: 10.1038/s41467-024-50296-y Retrieved from the Internet: URL:https://www.nature.com/articles/s41467-024-50296-y.pdf> the whole document abstract figures 1, 5A -----	1-40
Y	Christian Hentrich ET AL: "SpyLock, Rapid Generation of Bispecific Antibodies for High-Throughput Screening with SpyLock Technology, Bio-Rad, 2023", BIO-RAD, 1 January 2023 (2023-01-01), page 1, XP093220723, US Retrieved from the Internet: URL:https://info.bio-rad.com/Rapid-generation-HT-Spylock-technology.html?WT.mc_id=231116039963&WT.srch=1&utm_source=google&utm_medium=cpc&utm_campaign=pioneer&gad_source=1 [retrieved on 2024-11-04] the whole document -----	1-40

1

INTERNATIONAL SEARCH REPORT

International application No.

PCT/GB2024/052240

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. [X] forming part of the international application as filed.

   b. [ ] furnished subsequent to the international filing date for the purposes of international search (Rule 13*ter*.1(a)).

      [ ] accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.

2. [ ] With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/GB2024/052240

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2016193746 A1	08-12-2016	AU 2016272543 A1	04-01-2018
		BR 112017026042 A2	14-08-2018
		CA 2987821 A1	08-12-2016
		CN 108026148 A	11-05-2018
		CN 116199733 A	02-06-2023
		DK 3303374 T3	05-07-2021
		EP 3303374 A1	11-04-2018
		ES 2880336 T3	24-11-2021
		JP 6883529 B2	09-06-2021
		JP 2018521640 A	09-08-2018
		KR 20180050640 A	15-05-2018
		US 2018244730 A1	30-08-2018
		WO 2016193746 A1	08-12-2016
US 2020299358 A1	24-09-2020	AU 2020243430 A1	24-09-2020
		CN 113811548 A	17-12-2021
		EP 3941485 A1	26-01-2022
		JP 2022525777 A	19-05-2022
		US 2020299358 A1	24-09-2020
		WO 2020188346 A1	24-09-2020