- *Classification:*
- BIOLOGICAL SCIENCES: Biochemistry

- *Title:*
- **Programmable polyproteams built using twin peptide superglues**
-
- Short title: Programmable polyproteams
-
- *Author affiliation:*
- Gianluca Veggiani[a], Tomohiko Nakamura[a,b], Michael D. Brenner[a], Raphaël V. Gayet[a,c], Jun Yan[d], Carol V. Robinson[d], and Mark Howarth[a,1]
-
- [a]Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, UK. [b]LOC Development Department, R&D Division, Medical Business Unit, Sony Corporation, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8510, Japan. [c]Ecole Normale Supérieure de Lyon, 46 Allée d'Italie, 69364 Lyon Cedex 07, France. [d]Department of Chemistry, Physical and Theoretical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford, OX1 3QZ, UK.
-
- *Corresponding author:*
- Mark Howarth
- Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, UK.
- +44 1865 613242
- mark.howarth@bioch.ox.ac.uk
-
- *Keywords:*
- bionanotechnology, synthetic biology, antibody, protein engineering

Abstract
Programmed connection of amino acids or nucleotides into chains introduced a revolution in control of biological function. Reacting proteins together is more complex because of the number of reactive groups and delicate stability. Here we achieved sequence-programmed irreversible connection of protein units, forming polyprotein teams by sequential amidation and transamidation. SpyTag peptide is engineered to spontaneously form an isopeptide bond with SpyCatcher protein. By engineering the adhesin RrgA from *Streptococcus pneumoniae*, we developed the peptide SnoopTag, which formed a spontaneous isopeptide bond to its protein partner SnoopCatcher with >99% yield and no cross-reaction to SpyTag/SpyCatcher. Solid-phase attachment followed by sequential SpyTag or SnoopTag reaction between building-blocks enabled iterative extension. Linear, branched and combinatorial polyproteins were synthesized, identifying optimal combinations of ligands against death receptors and growth factor receptors for cancer cell-death signal-activation. This simple and modular route to programmable "polyproteams" should enable exploration of a new area of biological space.

Significance Statement
Many biological events depend on proteins working together as a team. Here we establish how to program protein team-formation, covalently linking protein modules step by step. We split a domain from *Streptococcus pneumoniae* to form a peptide and protein pair, SnoopTag and SnoopCatcher, which form an isopeptide bond when mixed together. SnoopTag/SnoopCatcher reacted with each other but not with an alternative peptide/protein pair, SpyTag/SpyCatcher. We formed polyprotein chains by alternating SpyTag reaction with SnoopTag reaction. Cellular signaling often relies on integrated activation of different receptors, so we built diverse polyprotein teams to stimulate Death Receptor and Growth Factor receptors, finding an optimal combination for cell-death induction in cancer cells. Programmable "polyproteams" provide a simple route to investigate or harness biological team-work.

Text
\body
Biological events usually depend on the cooperative activity of multiple proteins. Clustering a single kind of protein often greatly enhances biological signals (1), for example in the repeating antigen structures on vaccines (2). However, clustering different kinds of proteins into programmed polyprotein teams ("polyproteams") is an unmet challenge (3, 4). Protein units can be joined genetically into one long open reading frame, but errors in protein synthesis and misfolding soon become limiting (5, 6). Expressing modules individually and then linking the modules together would overcome these challenges, as well as allowing independent post-translational modification of each module. Even the best non-covalent linkages (7-9) or reversible covalent linkages, including disulfide bonds (10, 11), would allow rearrangement of polyproteams, so irreversible covalent linkage is required. There are a limited number of mutually-unreactive (orthogonal) chemical reactions (12), therefore it is impractical to link more than a few building-blocks in a one-pot reaction. However, elongating one step at a time allows chain growth using a small number of orthogonal connections (13). If the growing chain is attached to a solid-phase, the reacting module can be added in large excess (driving reaction to completion), with unreacted building blocks simply washed away (so separation is unnecessary at each step). Establishing such solid-phase chemistry for connecting amino acids underpinned the breakthroughs in the biological understanding and therapeutic use of peptides (14, 15), while

the solid-phase synthesis of DNA primers underpinned the revolution in gene amplification and re-engineering (16, 17). Solid-phase reaction has also enabled the ligation of peptide fragments to make synthetic proteins (18, 19).

Other important features of a system for synthesizing polyproteams are molecularly-defined connections, independence from any template (20, 21), and simple expression of each module. Nearly quantitative yield for each reaction is required; otherwise after a few steps the incomplete chains generate hopelessly heterogeneous products. In addition, it is preferable for modules to be modified with peptide tags rather than protein fusion domains (e.g. HaloTag or SNAP-tag), for minimal disruption to module function (22, 23). For unbreakable linkage to a peptide, we previously developed the use of spontaneous isopeptide bond formation (24, 25). SpyTag is a 13 amino acid peptide which can be genetically fused to the protein of interest and, upon mixing with its protein partner SpyCatcher, an Asp of SpyTag forms a spontaneous isopeptide bond with a Lys of SpyCatcher (26). Reaction occurs with good specificity (27) and under a wide range of conditions, with a connection resistant to boiling in sodium dodecyl sulfate (SDS) and high force (26).

To enable synthesis of polyproteams, here we first developed a covalent peptide/protein pair (SnoopTag/SnoopCatcher) orthogonal to SpyTag/SpyCatcher. Using iterative Snoop and Spy reactions, we then established the synthesis of linear and branched polyproteams. Polyproteams were analyzed by electrophoresis, chromatography and mass spectrometry. We then synthesized combinations of polyproteins for sensitive activation of cancer cell death.

Results

Design of an orthogonal covalent peptide-protein interaction. RrgA is an adhesin from *Streptococcus pneumoniae*, a Gram-positive bacterium which can cause septicemia, pneumonia and meningitis in humans. A spontaneous isopeptide bond forms in the D4 immunoglobulin-like domain of RrgA between residues Lys742 and Asn854 (**Fig. 1a**) (28). We split the D4 domain in the loop following the N-terminal β-strand and, after exploring various extensions and truncations of the two partners, settled on the peptide tag we termed SnoopTag (residues 734-745) and the protein partner which we named SnoopCatcher (residues 749-860, **Fig. 1b**).To optimize reaction, SnoopCatcher included the mutations G842T, designed to stabilize a β-strand, and D848G, designed to stabilize a hairpin-turn close to the reaction site (**Supplementary Fig. 1a**). SnoopTag fused to maltose-binding protein (MBP) and SnoopCatcher were expressed efficiently as soluble proteins in the cytosol of *Escherichia coli* and purified by Ni-NTA affinity chromatography. SnoopTag-MBP and SnoopCatcher, simply upon mixing, formed a complex stable to boiling in SDS (**Fig. 1c**). Mutations in the putative reactive Lys742 of SnoopTag (SnoopTag KA-MBP) and the putative reactive Asn854 of SnoopCatcher (SnoopCatcher NA) abolished reaction (**Fig. 1c**). Electrospray ionization mass spectrometry was consistent with the loss of NH_3 resulting from isopeptide bond formation between SnoopCatcher and SnoopTag; acetylated and gluconylated side-products common for *E. coli* overexpression were also observed (**Fig. 1d**). With 1:1 SnoopCatcher to SnoopTag-MBP, ~80% SnoopTag-MBP reacted. However, with a two-fold excess of SnoopCatcher, SnoopTag-MBP reacted quantitatively (**Fig. 1e** and **Supplementary Fig. 1b**). Similarly with an excess of SnoopTag-MBP, SnoopCatcher was ~100% consumed (**Fig. 1f** and **Supplementary Fig. 1c**). We established that reaction proceeded efficiently from pH 6-9 (**Fig. 2a**) and 4-37 °C (**Fig. 2b**). Cysteine is absent from SnoopTag and SnoopCatcher so, as expected, the reaction was insensitive to dithiothreitol (DTT) (**Supplementary Fig. 1d**). No specific buffer component was required, with reaction in PBS as

well as in the presence of the detergents Triton X-100 and Tween-20, or high salt (1 M NaCl) (**Supplementary Fig. 1d**). The chemical chaperone trimethylamine N-oxide (TMAO) (29) modestly enhanced reaction (**Supplementary Fig. 1e**).

Spontaneous hydrolysis of an amide bond normally takes years under neutral conditions (30) but we tested if hydrolysis was accelerated in this protein environment. We looked for reversal of the SnoopTag-MBP/SnoopCatcher interaction, by competing with excess of an alternative SnoopTag-linked protein or ammonia, but we did not observe reversibility (**Supplementary Fig. 2**).

SnoopTag has a reactive Lys (**Fig. 1b**), whereas SpyTag has a reactive Asp (26), so we hypothesized that SnoopTag/SnoopCatcher and SpyTag/SpyCatcher pairs would be fully orthogonal. Each partner reacted efficiently with its cognate pair but we found no trace of cross-reaction, even after overnight incubation (**Fig. 2c**). SpyTag and SpyCatcher also reacted to more than >95% conversion with an excess of their cognate partner (**Supplementary Fig. 3**). Therefore SnoopTag/SnoopCatcher provides a robust system for irreversible protein linkage and can be used in partnership with SpyTag/SpyCatcher.

Establishment of solid-phase polyproteam synthesis. Having dual orthogonal covalent reactions, we set out to achieve solid-phase iterative synthesis of polyproteins (**Fig. 3a**). The interaction of *E. coli* MBP with amylose resin is widely used in affinity purification: MBP-fusions typically express well, with low non-specific resin binding and selective mild elution using maltose. The affinity of wild-type MBP for maltose is 1.2 μM (31), which is satisfactory for protein purification but insufficient for multiple rounds of washing and chain extension in polyproteam synthesis. Therefore, we combined mutations enhancing MBP's maltose-binding stability (A312V, I317V and deletion at 172, 173, 175 and 176) (**Supplementary Fig. 4a**) (31, 32). We then tandemly linked this MBP mutant, to generate MBPx-SpyCatcher (His_6-MBPmt-linker-MBPmt-SpyCatcher) for anchoring to amylose. For initial chain building, we incorporated affibodies, non-immunoglobulin scaffolds expressed efficiently in *E. coli*. The affibody to the receptor tyrosine kinase and proto-oncogene HER2 (33) was linked at its N-terminus with SnoopTag and at its C-terminus with SpyTag (SnoopTag-AffiHER2-SpyTag). Affibody units were bridged using SpyCatcher connected through a helical spacer to SnoopCatcher (SpyCatcher-SnoopCatcher also expressed efficiently as a soluble protein in *E. coli*) (**Fig. 3a**). Since each linkage is covalent, chain synthesis could be easily followed by adding maltose to elute from the resin and then SDS-PAGE (**Fig. 3b**). MBPx-SpyCatcher reacted quantitatively with SnoopTag-AffiHER2-SpyTag (**Fig. 3b**, lane 5). This construct then reacted quantitatively with SpyCatcher-SnoopCatcher (**Fig. 3b**, lane 6). Sequential addition of SnoopTag-AffiHER2-SpyTag and SpyCatcher-SnoopCatcher enabled chain growth with high purity, extending to a product 10 units long (a decamer, **Fig. 3b**, lane 13). To demonstrate solid-phase extension with a different solid-phase, we used biotin-SpyCatcher linked to monomeric avidin beads: polyproteams were assembled to a decamer and eluted with free biotin (**Supplementary Fig. 4b**).

Analysis of the decamer polyproteam. Since chain units are connected via internal isopeptide bonds, the SDS-PAGE mobility of the chain diverges from the predicted molecular weight. To validate the identity of the assembled decamer, we initially performed electrospray ionization mass spectrometry, showing good correspondence between observed mass (285 kDa with a standard deviation of 3 kDa) and expected mass (282.5 kDa) (**Fig. 4a**). Size-exclusion

chromatography of the decamer showed one major peak centred at ~287 kDa, based on calibration with protein standards, consistent with the expected monomeric mass and indicating that there was minimal self-association under these conditions (**Fig. 4b**). To assess thermostability, the decamer was heated at a range of temperatures and centrifuged to remove aggregates: the decamer remained largely soluble even at 70 °C (**Supplementary Fig. 5a**). We also tested decamer integrity to storage: after 1 or 4 days at 25 °C we saw little degradation and little loss of solubility (**Supplementary Fig. 5b**).

Expanding from the initial incorporation of AffiHER2 modules into chains, we generated fluorescent protein polyproteams (**Supplementary Fig. 6a**). We also generated bottle-brush polyproteams, by joining a tandemly-linked affibody against HER2 where both the tags were at the N-terminus (SnoopTag-SpyTag-AffiHER2×3), so that the binding units branched from a central core (**Supplementary Fig. 6b**) (34).

Application of polyproteins to combinatorial cell signaling. We established combinatorial synthesis of polyproteams, using a simple 96-well plate format. These polyproteams were used to probe the spatial integration between different signaling pathways. Cancer cells are frequently sensitized to cell death induced by Death Receptor signaling, but clinical results have been disappointing using bivalent IgG, which is not potent at cell death induction (35, 36). We set out to explore higher multivalency and how Death Receptor signaling interacts with signaling by growth factor receptors frequently over-expressed in cancer. Therefore we created combinations of polyproteams with repeats of a nanobody agonist for Death Receptor 5 (DR5) and an affibody to either epidermal growth factor receptor (EGFR), HER2, or type I insulin-like growth factor receptor (IGF1R) (**Fig. 5**). These pro-apoptotic polyproteams were synthesized with good purity (**Fig. 5a**). Each of the combinatorial polyproteams bearing 4 nanobodies and one affibody unit at each of 5 positions was tested for effects on viability of the breast cancer cell-line MDA-MB-231. Affibody position affected cytotoxicity, with optimal killing by the chain with 4 nanobodies (N) followed by an affibody to EGFR (E) (NNNNE chain, **Fig. 5b**). Negative-control polyproteams containing only affibodies to HER2 (HHHHH), EGFR (EEEEE), IGF1R (IIIII) or Taq polymerase (i.e. a ligand not present on the cell surface) (TTTTT) showed no toxicity (**Supplementary Fig. 7**). We further assessed the dose-dependence of NNNNE cytotoxicity (**Fig. 5c**). Most of the cytotoxicity of NNNNE was apparent at 24 hours (**Fig. 5d**). NNNNE induced potent activation of caspase signaling, similarly to tumor necrosis factor-related apoptosis-inducing ligand (TRAIL, the endogenous DR5 agonist), and NNNNE's effect was fully blocked by the caspase inhibitor Z-VAD-FMK. HHHHH did not activate caspases (**Fig. 5e**).

To test the role of valency in polyproteam activation of cell death, we synthesized decamers having varying numbers of anti-DR5 nanobodies alongside either affibodies to EGFR or negative-control affibodies to Taq polymerase (**Supplementary Fig. 7**). Polyproteams with one or two nanobodies did not induce cell death, but killing gradually increased as the nanobody valency increased from three to five. With three or four nanobodies in the chain, compared to the anti-Taq control, targeting to EGFR enhanced the polyproteam potency (**Supplementary Fig. 7b**).

Discussion

We have generated a modular approach to synthesis of programmed polyproteams, through spontaneous isopeptide bond formation between peptide tags. The polyproteams are linked

through irreversible amide bonds, so are stable over time (if protected from proteases) and allow easy analysis by SDS-PAGE. The initiation, extension and release steps use mild conditions, independent of redox state, so should be applicable to a wide range of proteins. With only a single way for the chain to grow, products are molecularly defined, favoring reproducibility and precise tuning of function. Also, subunits do not need to be connected in an N- to C- orientation, as we show with bottle-brush polymer architectures. No chemical modification of the module is required, avoiding time-consuming and hard-to-control bioconjugation steps, so our method is accessible to any laboratory able to express recombinant proteins. Combinatorial synthesis of polyproteams allows rapid testing of patterns able to give potent cellular signal activation.

Spontaneous isopeptide bond formation has the advantage of a simple reaction pathway between two functional groups having low intrinsic reactivity (an amine with a carboxylic acid or a carboxamide) so there is little side-reaction (26, 37) and yields for each step were almost quantitative (38). This enabled a high purity of polyproteams after 9 consecutive reactions. Polyproteam assembly is not traceless, leaving a SpyCatcher-SnoopCatcher unit between each module. In future work, it will be interesting to explore SpyCatcher-SnoopCatcher units with alternative sizes and orientations, as well as using other linkage chemistries for solid-phase synthesis. Sortase, subtiligase, transglutaminase and split inteins enable covalent protein-protein interaction that is either traceless or leaves a short peptide tag (39-43) and these approaches will be valuable to test in the future with MBPx attachment and maltose elution. However, those reactions pass through a (thio)ester intermediate, so hydrolysis may compete with ligation (39, 40). Artificial amino acids for bio-orthogonal reaction would enable minimal modification, but with some issues: the increased complexity of module expression and competing reactions (such as azide reduction (44), alkyne reaction with thiols (45) and spontaneous tetrazine degradation (46)), as well as competition from suppression of stop codons by normal amino acids (47). Recent work demonstrated the synthesis of ubiquitin homo-polymers by azide-alkyne (48), thiol-ene (49), or native chemical ligation reactions (50). A potential disadvantage of our approach is that Tag/Catchers are not human, so are likely to be immunogenic, although non-human connectors are advantageous for vaccination. Also each module must have a single SpyTag and SnoopTag, which is straight-forward for monomers or hetero-multimers but needs careful attention for homo-multimers. As one makes larger polyproteams, their solubility and colloidal characteristics will change, just as multi-domain extracellular matrix proteins such as fibronectin cannot be treated in the same way as small globular proteins. However, the affibody polyproteam showed good stability over time and to heating. Polyproteams should be a powerful platform to dissect the spatial requirements for cellular signaling, such as in immunity and differentiation (51-53). Other applications of this simple route to new biological architectures may include vaccination (2), biomaterials (34, 54-56), multi-enzyme organization (9), and enhancing capture of circulating tumor cells (57).

Materials and Methods

Isopeptide bond reconstitution reactions. SI Materials and Methods provides a full description of cloning, protein expression and purification, MS, size-exclusion chromatography, cell culture, stability testing of chains, combinatorial assembly of chains, and caspase activation assay. To assess the formation of a covalent bond between SnoopTag and SnoopCatcher, proteins were mixed each at 10 μM final concentration in TBS pH 8.0 containing 1.5 M trimethylamine N-oxide (TMAO, Sigma-Aldrich). Reactions were stopped 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). Samples were

subsequently heated using a Bio-Rad C1000 thermal cycler at 95 °C for 5 min, before SDS-PAGE on 16% polyacrylamide gels using an XCell SureLock gel container (Life Technologies) at 200 V. SDS-PAGE running buffer was Tris-glycine, except Tris-acetate buffer was used to improve resolution of high molecular weight products, as previously (7). Gels were stained with Instant Blue Coomassie stain (Triple Red Ltd.) and bands were densitometrically analyzed using a Gel Doc XR imager and Image Lab 3.0 software (Bio-Rad).

Solid-phase synthesis of chains. 40 µL slurry amylose resin (New England BioLabs) was applied to a 1 mL poly-prep column (Bio-Rad), rinsed with 1 mL MilliQ water and equilibrated with 1 mL TBS pH 8.0. 320 pmol MBPx-SpyCatcher in TBS pH 8.0 in a final volume of 80 µL was added to the resin and incubated at 25 °C for 1 h with 700 rpm shaking on a ThermoMixer comfort (Eppendorf). Unreacted protein was removed by gravity flow and resin was washed with 1 mL Wash Buffer (50 mM Tris HCl pH 8.0 with 500 mM NaCl). 3 nmol monomer containing a SnoopTag and a SpyTag in TBS pH 8.0 in a final volume of 80 µL was added to the resin and incubated at 25 °C for 1 h with 700 rpm shaking. Unreacted SnoopTag/SpyTag monomer was removed from the column by gravity flow and resin washed with 1 mL Wash Buffer. 4 nmol SpyCatcher-SnoopCatcher in TBS pH 8.0 with 1.5 M TMAO was added to the resin and incubated at 25 °C for 2 h with 700 rpm shaking. Unreacted SpyCatcher-SnoopCatcher was removed by gravity flow and resin was washed with 1 mL Wash Buffer. Chains were produced by sequential addition of SnoopTag/SpyTag monomer and SpyCatcher-SnoopCatcher, according to the conditions described above. Chains were eluted, after resin washing, by applying onto the column 40 µL TBS pH 8.0 containing 50 mM D-maltose (Sigma-Aldrich) and incubating at 25 °C for 10 min with 700 rpm shaking. Chains were collected by centrifuging the column in a 1.5 mL microcentrifuge tube for 10 s at 17,000 g. For SDS-PAGE testing after each step, samples were eluted as previously described, mixed with 6×SDS-loading buffer and heated at 95 °C for 5 min. SDS-PAGE was performed on 10% and 4% Tris-acetate gels at 150 V.

For biotin-SpyCatcher-based assembly, 40 µL of slurry monomeric avidin resin (Thermo Scientific) was applied to a 1 mL poly-prep column, rinsed with 1 mL MilliQ water and equilibrated with 1 mL TBS pH 8.0. Biotin-SpyCatcher in TBS pH 8.0 in a final volume of 80 µL at 4 µM was added to the resin and incubated at 25 °C for 1 h with 700 rpm shaking. Unreacted biotin-SpyCatcher was removed by gravity flow, resin was washed with 1 mL Wash Buffer and sequential addition of SnoopTag-AffiHER2-SpyTag and SpyCatcher-SnoopCatcher was performed as described above. After resin washing, chains were eluted by applying 40 µL 1 mM D-biotin in TBS pH 8.0 and incubating the solution at 25 °C for 4 h with 700 rpm shaking. Chains were collected as previously indicated and analyzed by SDS-PAGE on 16% and 8% Tris-glycine gels with Coomassie staining.

ACKNOWLEDGEMENTS. Funding was provided by the Medical Research Council (G.V. and J.Y.), Merton College Oxford (G.V.), Sony (T.N.), the Ecole Normale Supérieure de Lyon (R.G.), the Royal Society (C.V.R.) and the European Research Council (ERC-2013-CoG 615945-PeptidePadlock, M.D.B. and M.H.). We thank Chris Schoene (University of Oxford) for biotin-SpyCatcher, University of Oxford Department of Biochemistry Biophysical Facility for assistance, and Ario de Marco (University of Nova Gorica) for Erv1p and DsbC plasmids.

Author contributions: G.V., T.N., M.D.B., R.V.G. and M.H. designed the study. G.V., T.N., M.D.B., R.V.G., and J.Y. performed experiments and analyzed data. C.V.R analyzed mass spectrometry data. M.H. wrote the paper.

References

1. Kiessling LL, Gestwicki JE, & Strong LE (2006) Synthetic multivalent ligands as probes of signal transduction. *Angew Chem Int Ed Engl* 45(15):2348-2368.
2. Bachmann MF & Jennings GT (2010) Vaccine delivery: a matter of size, geometry, kinetics and molecular patterns. *Nat Rev Immunol* 10(11):787-796.
3. Gradisar H*, et al.* (2013) Design of a single-chain polypeptide tetrahedron assembled from coiled-coil segments. *Nat Chem Biol* 9(6):362-366.
4. Fletcher JM*, et al.* (2013) Self-Assembling Cages from Coiled-Coil Peptide Modules. *Science* 340(6132):595-599.
5. Zaher HS & Green R (2009) Fidelity at the molecular level: lessons from protein synthesis. *Cell* 136(4):746-762.
6. Baneyx F & Mujacic M (2004) Recombinant protein folding and misfolding in Escherichia coli. *Nat Biotechnol* 22(11):1399-1408.
7. Fairhead M*, et al.* (2014) SpyAvidin Hubs Enable Precise and Ultrastable Orthogonal Nanoassembly. *J Am Chem Soc* 136(35):12355-12363.
8. Pengo P*, et al.* (2010) Solid-phase preparation of protein complexes. *J Mol Recognit* 23(6):551-558.
9. Arfi Y, Shamshoum M, Rogachev I, Peleg Y, & Bayer EA (2014) Integration of bacterial lytic polysaccharide monooxygenases into designer cellulosomes promotes enhanced cellulose degradation. *Proc Natl Acad Sci USA* 111(25):9109-9114.
10. Russell J, Colpitts T, Holets-McCormack S, Spring T, & Stroupe S (2004) Defined protein conjugates as signaling agents in immunoassays. *Clin Chem* 50(10):1921-1929.
11. Rossi EA, Goldenberg DM, & Chang CH (2012) Complex and defined biostructures with the dock-and-lock method. *Trends Pharmacol Sci* 33(9):474-481.
12. Patterson DM, Nazarova LA, & Prescher JA (2014) Finding the right (bioorthogonal) chemistry. *ACS Chem Biol* 9(3):592-605.
13. Merrifield RB (1963) Solid Phase Peptide Synthesis. 1. The Synthesis of a Tetrapeptide. *J Am Chem Soc* 85(14):2149-2154.
14. Marshall GR (2003) Solid-phase synthesis: a paradigm shift. *J Pept Sci* 9(9):534-544.
15. Cheloha RW, Maeda A, Dean T, Gardella TJ, & Gellman SH (2014) Backbone modification of a polypeptide drug alters duration of action in vivo. *Nat Biotechnol* 32(7):653-655.
16. Letsinger RL & Mahadevan V (1965) Oligonucleotide Synthesis on a Polymer Support. *J Am Chem Soc* 87:3526-3527.
17. Kosuri S & Church GM (2014) Large-scale de novo DNA synthesis: technologies and applications. *Nat Methods* 11(5):499-507.
18. Raibaut L*, et al.* (2013) Highly efficient solid phase synthesis of large polypeptides by iterative ligations of bis(2-sulfanylethyl)amido (SEA) peptide segments. *Chem Sci* 4(10):4061-4066.
19. Jbara M, Seenaiah M, & Brik A (2014) Solid phase chemical ligation employing a rink amide linker for the synthesis of histone H2B protein. *Chem Commun* 50(83):12534-12537.
20. Delebecque CJ, Lindner AB, Silver PA, & Aldaye FA (2011) Organization of Intracellular Reactions with Rationally Designed RNA Assemblies. *Science* 333(6041):470-474.
21. Chen R*, et al.* (2014) Biomolecular scaffolds for enhanced signaling and catalytic efficiency. *Curr Opin Biotech* 28:59-68.

22. Modica JA, Skarpathiotis S, & Mrksich M (2012) Modular Assembly of Protein Building Blocks To Create Precisely Defined Megamolecules. *Chembiochem* 13(16):2331-2334.
23. Huh WK*, et al.* (2003) Global analysis of protein localization in budding yeast. *Nature* 425(6959):686-691.
24. Zakeri B & Howarth M (2010) Spontaneous intermolecular amide bond formation between side chains for irreversible peptide targeting. *J Am Chem Soc* 132(13):4526-4527.
25. Veggiani G, Zakeri B, & Howarth M (2014) Superglue from bacteria: unbreakable bridges for protein nanotechnology. *Trends Biotechnol* 32(10):506-512.
26. Zakeri B*, et al.* (2012) Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc Natl Acad Sci USA* 109(12):E690-697.
27. Bedbrook CN*, et al.* (2015) Genetically Encoded Spy Peptide Fusion System to Detect Plasma Membrane-Localized Proteins In Vivo. *Chem Biol* 22(8):1108-1121.
28. Izore T*, et al.* (2010) Structural basis of host cell recognition by the pilus adhesin from Streptococcus pneumoniae. *Structure* 18(1):106-115.
29. Yancey PH (2004) Compatible and counteracting solutes: protecting cells from the Dead Sea to the deep sea. *Sci Prog* 87(Pt 1):1-24.
30. Radzicka A & Wolfenden R (1996) Rates of uncatalyzed peptide bond hydrolysis in neutral solution and the transition state affinities of proteases. *J Am Chem Soc* 118(26):6105-6109.
31. Telmer PG & Shilton BH (2003) Insights into the conformational equilibria of maltose-binding protein by analysis of high affinity mutants. *J Biol Chem* 278(36):34555-34567.
32. Walker IH, Hsieh PC, & Riggs PD (2010) Mutations in maltose-binding protein that alter affinity and solubility properties. *Appl Microbiol Biot* 88(1):187-197.
33. Wikman M*, et al.* (2004) Selection and characterization of HER2/neu-binding affibody ligands. *Protein Eng Des Sel* 17(5):455-462.
34. Zhang WB, Sun F, Tirrell DA, & Arnold FH (2013) Controlling macromolecular topology with genetically encoded SpyTag-SpyCatcher chemistry. *J Am Chem Soc* 135(37):13988-13997.
35. Holland PM (2014) Death receptor agonist therapies for cancer, which is the right TRAIL? *Cytokine Growth Factor Rev* 25(2):185-193.
36. Huet HA*, et al.* (2014) Multivalent nanobodies targeting death receptor 5 elicit superior tumor cell killing through efficient caspase induction. *mAbs* 6(6):1560-1570.
37. Kang HJ & Baker EN (2011) Intramolecular isopeptide bonds: protein crosslinks built for stress? *Trends Biochem Sci* 36(4):229-237.
38. Shi X*, et al.* (2012) Quantitative fluorescence labeling of aldehyde-tagged proteins for single-molecule imaging. *Nat Methods* 9(5):499-503.
39. Shah NH & Muir TW (2014) Inteins: Nature's Gift to Protein Chemists. *Chem Sci* 5(1):446-461.
40. Rashidian M, Dozier JK, & Distefano MD (2013) Enzymatic Labeling of Proteins: Techniques and Approaches. *Bioconj Chem* 24(8):1277-1294.
41. Popp MWL & Ploegh HL (2011) Making and Breaking Peptide Bonds: Protein Engineering Using Sortase. *Angew Chem Int Ed Engl* 50(22):5024-5032.
42. Carvajal-Vallejos P, Pallisse R, Mootz HD, & Schmidt SR (2012) Unprecedented rates and efficiencies revealed for new natural split inteins from metagenomic sources. *J Biol Chem* 287(34):28686-28696.
43. Chang TK, Jackson DY, Burnier JP, & Wells JA (1994) Subtiligase: a tool for semisynthesis of proteins. *Proc Natl Acad Sci USA* 91(26):12544-12548.
44. Sasmal PK*, et al.* (2012) Catalytic Azide Reduction in Biological Environments. *Chembiochem* 13(8):1116-1120.
45. van den Bosch SM*, et al.* (2013) Evaluation of strained alkynes for Cu-free click reaction in live mice. *Nucl Med Biol* 40(3):415-423.

46. Versteegen RM, Rossin R, ten Hoeve W, Janssen HM, & Robillard MS (2013) Click to release: instantaneous doxorubicin elimination upon tetrazine ligation. *Angew Chem Int Ed Engl* 52(52):14112-14116.
47. Aerni HR, Shifman MA, Rogulina S, O'Donoghue P, & Rinehart J (2015) Revealing the amino acid composition of proteins within an expanded genetic code. *Nucleic Acids Res* 43(2):e8.
48. Schneider T*, et al.* (2014) Dissecting ubiquitin signaling with linkage-defined and protease resistant ubiquitin chains. *Angew Chem Int Ed Engl* 53(47):12925-12929.
49. Trang VH*, et al.* (2012) Nonenzymatic polymerization of ubiquitin: single-step synthesis and isolation of discrete ubiquitin oligomers. *Angew Chem Int Ed Engl* 51(52):13085-13088.
50. Moyal T, Bavikar SN, Karthikeyan SV, Hemantha HP, & Brik A (2012) Polymerization behavior of a bifunctional ubiquitin monomer as a function of the nucleophile site and folding conditions. *J Am Chem Soc* 134(38):16085-16092.
51. Krueger AT, Kroll C, Sanchez E, Griffith LG, & Imperiali B (2014) Tailoring chimeric ligands for studying and biasing ErbB receptor family interactions. *Angew Chem Int Ed Engl* 53(10):2662-2666.
52. Dushek O, Goyette J, & van der Merwe PA (2012) Non-catalytic tyrosine-phosphorylated receptors. *Immunol Rev* 250(1):258-276.
53. Shaw A*, et al.* (2014) Spatial control of membrane receptor function using ligand nanocalipers. *Nat Methods* 11(8):841-846.
54. Chen AY*, et al.* (2014) Synthesis and patterning of tunable multiscale materials with engineered cells. *Nat Mater* 13(5):515-523.
55. Sun F, Zhang WB, Mahdavi A, Arnold FH, & Tirrell DA (2014) Synthesis of bioactive protein hydrogels by genetically encoded SpyTag-SpyCatcher chemistry. *Proc Natl Acad Sci USA* 111(31):11269-11274.
56. Botyanszki Z, Tay PK, Nguyen PQ, Nussbaumer MG, & Joshi NS (2015) Engineered catalytic biofilms: Site-specific enzyme immobilization onto E. coli curli nanofibers. *Biotechnol Bioeng* 112(10):2016-2024.
57. Fierer JO, Veggiani G, & Howarth M (2014) SpyLigase peptide-peptide ligation polymerizes affibodies to enhance magnetic cancer cell capture. *Proc Natl Acad Sci USA* 111(13):E1176-1181.
58. Kuhlman B, Yang HY, Boice JA, Fairman R, & Raleigh DP (1997) An exceptionally stable helix from the ribosomal protein L9: implications for protein folding and stability. *J Mol Biol* 270(5):640-647.
59. Li L, Fierer JO, Rapoport TA, & Howarth M (2014) Structural analysis and optimization of the covalent association between SpyCatcher and a peptide Tag. *J Mol Biol* 426(2):309-317.
60. Eigenbrot C, Ultsch M, Dubnovitsky A, Abrahmsen L, & Hard T (2010) Structural basis for high-affinity HER2 receptor binding by an engineered protein. *Proc Natl Acad Sci USA* 107(34):15039-15044.
61. Friedman M*, et al.* (2007) Phage display selection of Affibody molecules with specific binding to the extracellular domain of the epidermal growth factor receptor. *Protein Eng Des Sel* 20(4):189-199.
62. Li J*, et al.* (2010) Selection of affibody molecules to the ligand-binding site of the insulin-like growth factor-1 receptor. *Biotechnol Appl Bioc* 55(2):99-109.
63. Lendel C, Dogan J, & Hard T (2006) Structural basis for molecular recognition in an affibody:affibody complex. *J Mol Biol* 359(5):1293-1304.
64. Veggiani G & de Marco A (2011) Improved quantitative and qualitative production of single-domain intrabodies mediated by the co-expression of Erv1p sulfhydryl oxidase. *Protein Expres Purif* 79(1):111-114.

65. Fairhead M & Howarth M (2015) Site-specific biotinylation of purified proteins using BirA. *Method Mol Biol* 1266:171-184.
66. Castano-Cerezo S, Bernal V, Rohrig T, Termeer S, & Canovas M (2015) Regulation of acetate metabolism in Escherichia coli BL21 by protein N(epsilon)-lysine acetylation. *Appl Microbiol Biot* 99(8):3533-3545.
67. Geoghegan KF*, et al.* (1999) Spontaneous alpha-N-6-phosphogluconoylation of a "His tag" in Escherichia coli: the cause of extra mass of 258 or 178 Da in fusion proteins. *Anal Biochem* 267(1):169-184.
68. Pringle SD*, et al.* (2007) An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *Int J Mass Spectrom* 261(1):1-12.
69. Hernandez H & Robinson CV (2007) Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat Protoc* 2(3):715-726.

FIGURE LEGENDS

Figure 1 Establishing the covalently-reactive peptide/protein pair SnoopTag and SnoopCatcher. (**a**) Spontaneous isopeptide bond formation between Lys and Asn, releasing ammonia. (**b**) Cartoon of splitting RrgA D4 domain (based on PDB 2WW8) to make SnoopTag and SnoopCatcher. Reactive residues in cyan. (**c**) SnoopTag-MBP reaction with SnoopCatcher, each at 10 μM, after 2 h at 25 °C analyzed by SDS-PAGE with Coomassie staining, alongside controls with Ala mutation of SnoopTag's reactive Lys (KA) or SnoopCatcher's reactive Asn (NA). (**d**) Isopeptide bond formation between SnoopTag peptide and SnoopCatcher shown by mass spectrometry. (**e**) Time-course of SnoopTag reaction with 1:1 or 2:1 ratio of SnoopCatcher to SnoopTag-MBP, tested as in (c). (**f**) Time-course of SnoopCatcher reaction with 1:1, 2:1 or 4:1 ratio of SnoopTag-MBP to SnoopCatcher, tested as in (c). Error bars are mean ± 1 s.d., n=3. Some error bars are too small to be visible.

Figure 2 SnoopTag reaction is robust to conditions and orthogonal to SpyTag. (**a**) pH-dependence: 10 μM SnoopTag-MBP was incubated with 10 μM SnoopCatcher at the indicated pH for 15 min at 25 °C and analyzed by SDS-PAGE with Coomassie staining. (**b**) Temperature-dependence of SnoopTag/SnoopCatcher reaction tested as in (a). (**c**) SnoopTag/SnoopCatcher and SpyTag/SpyCatcher orthogonal reactivity, after incubation for 18 h at 25 °C, with each species at 10 μM, determined by SDS-PAGE with Coomassie staining. Error bars are all mean ± 1 s.d., n=3. Some error bars are too small to be visible.

Figure 3 Solid-phase polyproteam synthesis. (**a**) Principle of polyproteam synthesis. Amylose-resin is bound by modified maltose-binding protein linked to SpyCatcher (MBPx-SpyCatcher). A protein of interest bearing SpyTag and SnoopTag is added to the resin and reacts with SpyCatcher on the growing chain (red line represents an isopeptide bond). SpyCatcher-SnoopCatcher is added and reacts with SnoopTag of the growing chain. Extension is continued by sequential additions of SnoopTag-X-SpyTag and SpyCatcher-SnoopCatcher. After extension is completed, the polyproteam is eluted from the resin with maltose. (**b**) Analysis of solid-phase polyproteam synthesis. Lanes 1-3 show MBPx-SpyCatcher, SpyCatcher-SnoopCatcher and SnoopTag-AffiHER2-SpyTag in isolation. MBPx-SpyCatcher was bound to the amylose resin and stepwise reaction with SnoopTag-AffiHER2-SpyTag and SpyCatcher-SnoopCatcher was carried out. After each stage, one aliquot of sample was eluted from the resin with maltose (lanes 4-13). Samples, without further purification, were analyzed by SDS-PAGE (both 4% and 10% gels) with Coomassie staining.

Figure 4 Biophysical analysis of polyproteams. (**a**) Electrospray ionization mass spectrometry to test identity of decamer polyproteam, MBPx-SpyCatcher:(SnoopTag-AffiHER2-SpyTag:SpyCatcher-SnoopCatcher)$_4$:SnoopTag-AffiHER2-SpyTag. Red circles correspond to the decamer, with the charge state of the highest peak marked. (**b**) Size-exclusion chromatography of the same polyproteam. The inset shows the molecular weight standards.

Figure 5 Combinatorial synthesis of polyproteams. (**a**) Chain synthesis with four anti-DR5 nanobodies and one affibody-based unit at various positions, analyzed by SDS-PAGE with Coomassie staining. (**b**) Polyprotein effect on viability of MDA-MB-231 cells after 40 h incubation with each chain from (a) at 200 ng/mL, shown as a heat-map. (**c**) Viability of MDA-

MB-231 cells at 40 h with titration of NNNNE and HHHHH control (N = anti-DR5 nanobody, E = AffiEGFR, H = AffiHER2×3). (**d**) Time-course of cytotoxicity with 200 ng/mL NNNNE on MDA-MB-231 cells. (**e**) Caspase reporter activation in response to NNNNE, KillerTRAIL (a DR5 agonist) or HHHHH with or without Z-VAD-FMK pan-caspase inhibitor. (Error bars in each case are mean ± 1 s.d., n = 3. Some error bars are too small to be visible.)