Programmable polyproteams built using twin peptide superglues

Gianluca Veggiani, Tomohiko Nakamura, Michael D. Brenner, Raphaël V. Gayet, Jun Yan, Carol V. Robinson, and Mark Howarth

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 “polyproteins” should enable exploration of a new area of biological space.

BIOCHEMISTRY

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 (“polyproteins”) 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 allow independent posttranslational modification of each module. Even the best noncovalent linkages (7–9) or reversible covalent linkages, including disulfide bonds (10, 11), would allow rearrangement of polyproteins, 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), whereas the solid-phase synthesis of DNA primers underpinned the revolution in gene amplification and reengineering (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 polyproteins 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 that 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 SDS and high force (26).

To enable synthesis of polyproteins, 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 polyproteins. Polyproteins were analyzed by electrophoresis, chromatography, and mass spectrometry. We then synthesized combinations of polyproteins for sensitive activation of cancer cell death.

Significance

Many biological events depend on proteins working together as a team. Here we establish how to program 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 polyprotein teams to stimulate Death Receptor and Growth Factor receptors, finding an optimal combination for cell-death induction in cancer cells. Programmable “polyproteins” provide a simple route to investigate or harness biological teamwork.

Author contributions: G.V., T.N., R.V.G., J.Y., and M.H. designed research; G.V., T.N., M.D.B., J.Y., C.V.R., and M.H. analyzed data; and G.V., T.N., M.D.B., and J.Y. performed research; R.V.G. contributed new reagents/analytic tools; G.V., T.N., M.D.B., J.Y., C.V.R., and M.H. analyzed data; and G.V. and M.H. wrote the paper.

Conflict of interest statement: M.H. is an inventor on patent EP2534484, which applies to spontaneous isopeptide bond formation to a peptide tag, and United Kingdom Patent Application No. 1509782.7, which applies to iterative synthesis through isopeptide bonds. This article is a PNAS Direct Submission.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. KU296974, KU296974, KU296975, KU296976, KU356670, KU361182, KU361183, KUS006643, KUS006644, KUS006645, and KUS006646).

1To whom correspondence should be addressed. Email: mark.howarth@bioch.ox.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1519214113/-/DCSupplemental.
Results

Design of an Orthogonal Covalent Peptide–Protein Interaction. RrgA is an adhesin from *Streptococcus pneumoniae*, a Gram-positive bacterium that can cause septicemia, pneumonia, and meningitis in humans. A spontaneous isopeptide bond forms in the D4 Ig-like domain of RrgA between residues Lys742 and Asn854 (Fig. L4) (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 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 (Fig. S1A). SnoopTag fused to maltose-binding protein (MBP) and SnoopCatcher were expressed efficiently as soluble proteins in the cytosol of *Escherichia coli* and purified by nickel-nitrilotriacetic acid (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 formation between SnoopCatcher and SnoopTag; acetylated and glucosylated side products common for *Escherichia coli* overexpression were also observed (Fig. 1D). With a 1:1 SnoopCatcher to SnoopTag-MBP, ∼80% of SnoopTag-MBP reacted. However, with a twofold excess of SnoopCatcher, SnoopTag-MBP reacted quantitatively (Fig. 1E and Fig. S1B). Similarly with an excess of SnoopTag-MBP, SnoopCatcher was ∼100% consumed (Fig. 1F and Fig. S1C). 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 DTT (Fig. S1D). 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) (Fig. S1D). The chemical chaperone trimethylamine N-oxide (TMAO) (29) modestly enhanced reaction (Fig. S1E).

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 (Fig. S2).

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 (Fig. S3). 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 nonspecific 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) (Fig. S4A) (31, 32). We then tandemly linked this MBP mutant to generate MBPx-SpyCatcher (His9–MBPmt–linker–MBPmt–SpyCatcher) for anchoring to amylose. For initial chain building, we incorporated affibodies, nonimmunoglobulin scaffolds expressed efficiently in E. coli. The affibody to the receptor tyrosine kinase and proto-oncogene HER2 (human epidermal growth factor receptor 2) (33) was linked at its N terminus with SpyTag and at its C terminus with SpyTag (SpyCatcher–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). Because 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: Polyproteins were assembled to a decamer and eluted with free biotin (Fig. S4B).

Analysis of the Decamer Polyproteam. Because 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 SD of 3 kDa) and expected mass (282.5 kDa) (Fig. 4A). Size-exclusion

![Fig. 3. Solid-phase polyproteam synthesis. (A) Principle of polyproteam synthesis. Amylose resin is bound by modified MBP 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–SpyTag and SpyCatcher–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% (wt/vol) gels] with Coomassie staining.}
chromatography of the decamer showed one major peak centered 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 (Fig. S5A). We also tested decamer integrity to storage: after 1 or 4 d at 25 °C, we saw little degradation and little loss of solubility (Fig. S5B).

Expanding from the initial incorporation of AffiHER2 modules into chains, we generated fluorescent protein polyproteins (Fig. S6A). We also generated bottle-brush polyproteins by joining a tandemly linked affibody against HER2, where both the tags were at the N terminus (SpyTag–SpyTag–SpyCatcher), so that the binding units branched from a central core (Fig. S6B) (34).

Application of Polyproteins to Combinatorial Cell Signaling. We established combinatorial synthesis of polyproteins using a simple 96-well plate format. These polyproteins 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 overexpressed in cancer. Therefore, we created combinations of polyproteins 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 proapoptotic polyproteins were synthesized with good purity (Fig. 5A). Each of the combinatorial polyproteins bearing four nanobodies and one affibody unit at each of five 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 four nanobodies (N) followed by an affibody to EGFR (E) (NNNNE chain; Fig. 5B). Negative-control polyproteins containing only affibodies to HER2 (HHHHH), EGFR (EEEEE), IGF1R (IIIIII), or Taq polymerase (i.e., a ligand not present on the cell surface) (TTTTT) showed no toxicity (Fig. S7). We further assessed the dose dependence of NNNNE cytotoxicity (Fig. S5C). Most of the cytotoxicity of NNNNE was apparent at 24 h (Fig. S5D). 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 Benzyloxycarbonyl–Val-Ala-Asp–fluoromethylketone (Z-VADFMK). HHHHH did not activate caspases (Fig. 5E).

To test the role of valency in polyprotein 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 (Fig. S7). Polyproteins 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 with the anti-Taq control, targeting to EGFR enhanced the polyprotein potency (Fig. S7B).

Discussion We have generated a modular approach to synthesis of programmed polyproteins, through spontaneous isopeptide bond formation between peptide tags. The polyproteins are linked through irreversible amide bonds and 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, and therefore 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 polyproteins 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 carbonamide), so there is little side reaction (26, 37) and yields for each step were almost quantitative (38). This enabled a high purity of polyproteins after nine consecutive reactions. Polyprotein assembly is not traceless, leaving a SpyCatcher–SpyCatcher unit between each module. In future work, it will be interesting to explore SpyCatcher–SpyCatcher units with alternative sizes and orientations, as well as using other linkage chemistries for solid-phase synthesis. Sortase, subtiligase, transglutaminase, and split intein 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 bioorthogonal 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 and therefore likely to be immunogenic, although nonhuman connectors are advantageous for vaccination. Also, each module must have a single SpyTag and SnoopTag, which is straightforward for monomers or
heteromultimers but needs careful attention for homo-multimers. As one makes larger polyproteams, their solubility and colloidal characteristics will change, just as multdomain 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), multienzyme 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 Tris-buffered saline (TBS) pH 8.0 containing 1.5 M TMAO (Sigma-Aldrich). Reactions were stopped by adding 6x SDS loading buffer [0.23 M Tris HCl, pH 6.8, 24% (vol/vol) glycerol, 120 μM bromophenol blue, 0.2 M SDS]. Samples were subsequently heated using a Bio-Rad C1000 thermal cycler at 95 °C for 5 min, before SDS/PAGE on 16% (wt/vol) polyacrylamide gels with Coomassie staining. For SDS/PAGE testing after each step, samples were eluted as previously described, mixed with 6x SDS loading buffer, and heated at 95 °C for 5 min. SDS/PAGE was performed on 10% (wt/vol) and 4% (wt/vol) 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 polyprep column (Bio-Rad), rinsed with 1 mL MilliQ water, and equilibrated with 1 mL TBS pH 8.0. We added 320 pmol MB-231 cells. (D) Viability of MDA-MB-231 cells after 40 h of incubation with each chain from A at 200 ng/mL, shown as a heat-map. (E) Viability of MDA-MB-231 cells at 40 h with titration of NNNNE and HHHHHH control (N, anti-DRS 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 DRS agonist) or HHHHHH with or without Z-VAD-FMK pan-caspase inhibitor. (Error bars in each case are mean ± 1 SD; n = 3. Some error bars are too small to be visible.)

Fig. 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 of 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-DRS 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 DRS agonist) or HHHHHH with or without Z-VAD-FMK pan-caspase inhibitor. (Error bars in each case are mean ± 1 SD; n = 3. Some error bars are too small to be visible.)
ACKNOWLEDGMENTS. We thank Chris Schoene (University of Oxford) for biomimetic cellulose of the Superieur of Bireurde of Lyon (R.V.G.), the Royal Society (C.V.R.), and European Research Council Grant ERC-2013-CoG 615945-PeptidePadlock (to M.D.B. and M.H.).


