

Click biology highlights the opportunities from reliable biological reactions

Received: 7 May 2024

Accepted: 15 May 2025

Published online: 19 June 2025

 Check for updates

Mark R. Howarth ^{1,2} 

Click chemistry is a powerful concept that refers to a set of covalent bond-forming reactions with highly favorable properties. In this Perspective, I outline the analogous concept of click biology as a set of reactions derived from the regular building blocks of living cells, rapidly forming covalent bonds to specific partners under cell-friendly conditions. Click biology using protein components employs canonical amino acids and may react close to the diffusion limit, with selectivity in living cells amid thousands of components generated from the same building blocks. I discuss how the criteria for click chemistry can be applied or modified to fit the extra constraints of click biology and achieve favorable characteristics for biological research. Existing reactions that may be described as click biology include split intein reconstitution, spontaneous isopeptide bond formation by SpyTag and SpyCatcher and suicide enzyme reaction with small-molecule ligands (HaloTag and SNAP-tag). I also describe how click biology has created new possibilities in fields including molecular imaging, mechanobiology, vaccines and engineering cellular intelligence.

Click chemistry has been an important and popular vision, changing people's thinking about chemical reactivity. Click chemistry has empowered diverse areas of fundamental research and numerous applications¹. The functional groups used in click chemistry are nearly always absent from living organisms^{1,2}, so it is helpful to consider how genetically encoded systems could achieve comparable modular and efficient reactivity. I describe here the concept of click biology as a set of reactions derived from the regular building blocks of living cells, rapidly forming covalent bonds to specific partners under cell-friendly conditions (Fig. 1).


The initial formulation of click chemistry in 2001 (ref. 1) was influential on how researchers thought about the importance of different chemical reactions and the design of useful syntheses³. Click chemistry referred to the development of a few reactions with highly favorable characteristics that enabled their efficient use in diverse situations. Three widely used click chemistry reactions are shown in Fig. 1: copper-catalyzed azide–alkyne cycloaddition (CuAAC), strain-promoted azide–alkyne cycloaddition and inverse electron-demand Diels–Alder⁴.

Table 1 describes the initial criteria defining a click chemistry reaction¹ and then additional criteria as the field matured². Click chemistry

has been criticized for introducing undesired scars from the ligation. However, a key aspect of the click chemistry philosophy was highlighting the importance of ease and simplicity in joining molecules to make a dent in the extraordinary breadth of possible chemical space. Small-molecule space (<500 Da) has $\sim 10^{63}$ diversity⁵; macromolecular space will be much larger, so new approaches are required to accelerate exploration. As George S. Hammond said in his 1968 Norris Award lecture, “The most fundamental and lasting objective of synthesis is not production of new compounds, but production of properties”.

Click chemistry expanded to have great importance in biological research and has reached clinical trials^{6,7}. For developments in this field, Carolyn Bertozzi, Morten Meldal and K. Barry Sharpless received the 2022 Nobel Prize in Chemistry³. Diverse approaches have been employed in chemical modification of biomolecules, so concepts of click chemistry provide valuable insights into how such approaches may be advanced through click biology.

Biological systems can achieve characteristics of a genetically encoded click chemistry, while employing functional groups that are naturally available and possess only moderate intrinsic reactivity (for example, amines, carboxylic acids, hydroxyls, thiols)⁸. In contrast to click chemistry, functional groups for click biology are not intrinsically

¹Department of Pharmacology, University of Cambridge, Cambridge, UK. ²Engineering Biology Interdisciplinary Research Centre, University of Cambridge, Cambridge, UK.  e-mail: mh2186@cam.ac.uk

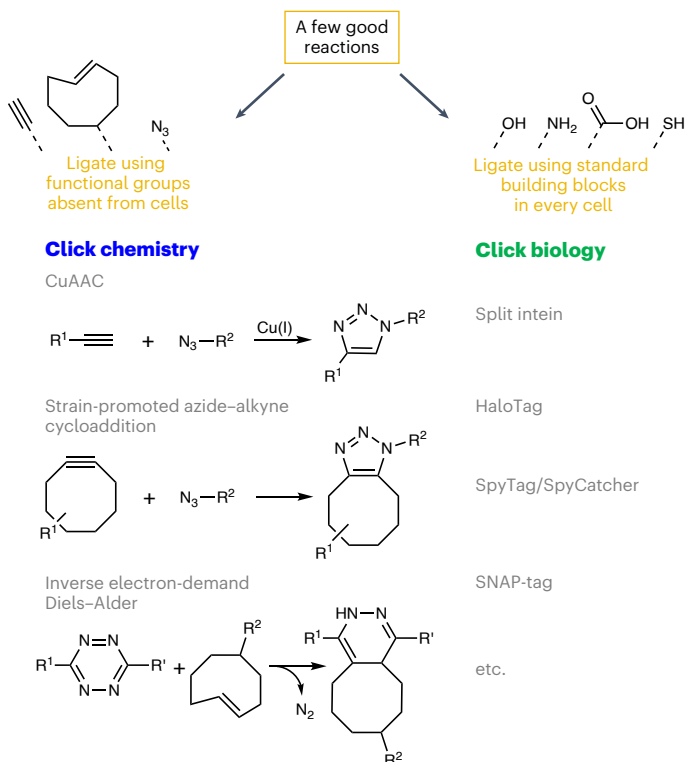


Fig. 1 | Overview of the comparison of click chemistry and click biology.

Click biology uses canonical building blocks, for example, combinations of the regular 20 amino acids, to achieve favorable reactions in a cellular context. Widely used click chemistry reactions are shown next to examples of click biology reactions.

spring-loaded for reaction, but macromolecular organization generates a precise three-dimensional orientation and tuning of the local environment for efficient reactivity (Fig. 1). The extended interaction surface of biological macromolecules facilitates high specificity as well as high-affinity binding for coupling at low reactant concentrations, important in complex biological systems. These features are akin to the advantages of enzymes over small-molecule catalysts⁹, but in click biology the reactants are trapped and not turned over. This Perspective introduces prototypical click biology reactions and considers how they match up to criteria of click chemistry. I will explore unique opportunities that arise from rapid and efficient covalent reactivity, employing building blocks naturally present in living systems. Finally, I will consider directions to enhance the power of the click biology toolbox.

Criteria for a click biology reaction

The ‘stringent criteria’ for click chemistry¹ are compared to criteria for click biology in Table 1. The overriding feature of a click reaction is robust efficiency: robust to what is attached to reacting groups and robust to different conditions. Click biology may be employed on isolated components, but it is central to consider applicability in living systems; hence, the criteria differ from those of click chemistry.

Click biology criterion 1: production from regular biological building blocks

Click biology uses building blocks universal through the biosphere, which facilitates use without optimization in different cell types and species as well as environmentally friendly large-scale production.

Click biology criterion 2: fusion tolerance is wide

The dream of biological modularity is that each module in combination retains all the activity of the module in isolation. Cells express

thousands of different proteins of diverse size and biophysical characteristics. Modules for click biology would ideally retain full activity when linked at either terminus of different biomolecules of interest or ideally also when inserted at central locations.

Click biology criterion 3: reaction yield is high

In the test tube, it may be easy to provide an excess of one component to drive a reaction; in living systems excess reactant is not so easily removed. Click biology reactions ideally give high yield with a 1:1 reactant ratio, generating a product that is stable despite challenges of time, high temperature or force.

Click biology criterion 4: rate constant is high

Biologists often employ components at nanomolar concentrations, and biological events occur in less than a second. High rate constants enable researchers to achieve deeper understanding and increased control of biological processes.

Click biology criterion 5: condition tolerance is broad

Reactions would ideally retain high rate constants despite the diversity through biology in temperature, pH, redox and ionic composition.

Click biology criterion 6: specificity in diverse cellular contexts

The initial formulation of click chemistry focused on selective reaction within a particular small molecule¹; later click chemistry aimed for selective reaction in complex biological environments (bioorthogonality)² (Table 1). Click biology should allow reaction at a specific site on a purified macromolecule but would ideally allow a selective reaction in living cells, multicellular organisms or even in human clinical use.

These click biology criteria are illustrated schematically in Fig. 2a.

Selected reactions for click biology

Here, I highlight widely used reactions that might be considered as click biology (Fig. 2). The focus is on bimolecular reactions in which at least one component contains only naturally occurring biological building blocks and the reaction shows selectivity amid cellular complexity. The description is not exhaustive: the reviews cited below provide details of related methods.

Split intein formation

Inteins are protein elements that splice themselves out of a polypeptide, linking two exteins¹⁰. The first split intein was found in 1998 (ref. 11), and their potential was quickly realized (Fig. 2b). After noncovalent interaction between Int^N and Int^C, *N*-*X* acyl shift generates an ester or a thioester (Supplementary Fig. 1a). Next, *trans*-(thio)esterification, asparagine cyclization and *X*-*N* acyl shift create the final spliced protein^{12,13}. Given these intricate steps, competing side reactions of *N*-terminal cleavage and/or *C*-terminal cleavage occur to some degree¹⁴ (Supplementary Fig. 1a).

Initial split inteins took hours and showed moderate yield¹⁰. Engineering the split intein Gp41-1 enabled a splicing half-time of ~5 s at 37 °C, with a dissociation constant for the splicing-inactive fragments of 9 nM¹⁵. The extensive contact surface between Int^N and Int^C lends itself to multiple orthogonal split inteins¹⁶. The separation site is often reengineered, and split intein elements have been decreased to six residues¹⁷. Commonly, researchers include three amino acid residues from the original extein to promote reaction¹⁴, while certain flanking sequences accelerate splicing¹⁸. The size of all tags is compared in Supplementary Table 1.

Widely used Npu and Gp41-1 split inteins use reactive cysteines and require reducing agent. The cysteine-less (CL) intein has serine at autocatalytic positions and was fully active without reducing agent¹⁹. The 26-amino-acid CL Int^N was used for site-specific labeling at the cell surface. The reaction was typically conducted over 6 h and was active from 0 to 37 °C. The VidaL split intein from metagenomic sequencing in

Table 1 | Criteria for click chemistry reactions (left column) compared to click biology reactions (right column)

Click chemistry	Click biology
Initial criteria (Kolb et al.)	Central criteria
<ul style="list-style-type: none"> • Readily available starting materials and reagents 	<ol style="list-style-type: none"> 1. Production from regular biological building blocks <p>One part genetically encodable using regular amino acids, nucleic acid bases or sugars. No need for UV or toxic metal ions (Ni²⁺, Cr³⁺)</p>
<ul style="list-style-type: none"> • Modular and wide scope 	<ol style="list-style-type: none"> 2. Fusion tolerance is wide. <p>Active when fused to diverse partners, with tolerance to steric crowding and good solubility</p>
<ul style="list-style-type: none"> • High yields • Product stable under physiological conditions 	<ol style="list-style-type: none"> 3. Reaction yield is high. <p>No large excess of one reactant; few side reactions. No reversal under physiological conditions and preferably stable to boiling or lyophilization</p>
<ul style="list-style-type: none"> • Proceeds rapidly to completion 	<ol style="list-style-type: none"> 4. Rate constant is high. <p>Second-order rate constant not far from diffusion limit, so rapid reactivity with low concentration of reactant</p>
<ul style="list-style-type: none"> • Reaction conditions simple (insensitive to O₂+H₂O) • No solvent or solvent benign (H₂O) or easily removed 	<ol style="list-style-type: none"> 5. Condition tolerance is broad. <p>4–37°C and pH 6–8, 0–2 mM Ca²⁺</p> <p>Insensitive to oxidizing or reducing conditions</p>
<ul style="list-style-type: none"> • Selective for single product; stereospecific 	<ol style="list-style-type: none"> 6. Specificity in diverse cellular contexts <p>Selective for single site on target and amid diverse components in multiple organisms</p>
<ul style="list-style-type: none"> • Byproducts inoffensive 	<ul style="list-style-type: none"> • Byproducts often H₂O, Cl⁻, NH₃
<ul style="list-style-type: none"> • Simple product isolation 	
Later criteria	Desirable criteria
<ul style="list-style-type: none"> • Bioorthogonality 	
<ul style="list-style-type: none"> • Orthogonal to other click chemistry 	<ul style="list-style-type: none"> • Orthogonal to other click biology • Extensive infrastructure of parts to conjugate

The overriding feature of a click reaction is robust efficiency: readily available starting materials reacting with broad scope, high yield, high rate constant, tolerance to diverse conditions and selectivity. UV, ultraviolet.

Lake Vida in Antarctica is active from 4 to 37 °C and even in 4 M NaCl^{20,21}. VidaL Int^N is 16 residues with a reaction half-time of 1 min. Vidal enabled traceless in vitro assembly of nucleosomes bearing macroH2A and even splicing of a posttranslationally modified histone tail in cells²⁰. Overall, split inteins perform molecular surgery, achieving high-yield linkage of protein building blocks in vitro or in diverse cellular contexts.

SpyTag and SpyCatcher

Spontaneous formation of isopeptide bonds was first discovered in the capsid of the bacteriophage HK97. An isopeptide bond is an amide bond involving one or more protein side chains. The capsid of HK97 achieves extreme stability through interlocking protein rings like chain mail^{22,23}. A range of proteins on the surface of Gram-positive bacteria also perform spontaneous isopeptide bond formation, either as an amidation or a transamidation, with an apposed glutamic acid facilitating the reaction through proton shuffling²⁴. To harness this chemistry, various bacterial domains were split into peptide and protein partners that react together irreversibly²⁵. The first pair with fast reactivity was SpyTag and SpyCatcher. Here, an aspartic acid on the peptide partner (SpyTag) undergoes spontaneous amidation with a lysine on the protein partner (SpyCatcher) simply upon mixing (half-time of

74 s with the Tag and Catcher components each at 10 μM)²⁶ (Fig. 2c). Through phage selection²⁷ and computational stabilization, the reaction rate was engineered to approach the diffusion limit for SpyTag003 and SpyCatcher003 (ref. 28). SpyTag003 (16 amino acids, 1.9 kDa) and SpyCatcher003 (113 amino acids, 12 kDa) do not contain any cysteines, so the reaction is the same in reducing or oxidizing conditions²⁸. The reaction is efficient from 4 to 37 °C, from pH 5 to 8 and in the presence of diverse buffers, with or without various detergents^{28,29}.

Tag–Catcher pairs orthogonal to SpyTag–SpyCatcher include SnoopTag–SnoopCatcher³⁰, DogTag–DogCatcher³¹ and SilkTag–SilkCatcher³². DogTag is loopophilic: well tolerated in various tight turns of proteins^{31,33}. As in click chemistry, where CuAAC can be reused through sequential deprotections³⁴, distinct SpyTag003 ligands can be plugged into a scaffold containing twin SpyCatcher003 using the SpyMask strategy. SpyTag003 ligand A reacts at the open SpyCatcher003, and then tobacco etch virus protease uncages a masking peptide to allow SpyTag003 ligand B to couple at the second SpyCatcher003 for generation of conformationally constrained bispecific antibodies³⁵.

HaloTag

HaloTag is a 33-kDa protein engineered from haloalkane dehalogenase³⁶. The first step for haloalkane dehalogenase is nucleophilic attack on an alkyl halide by aspartate 106 to form an ester intermediate. Histidine 272 then activates water to hydrolyze the ester and release an alcohol, regenerating the enzyme. In HaloTag, mutation of histidine 272 to phenylalanine causes the ester to be trapped, creating a suicide enzyme³⁶ (Fig. 2d). This mutant had a second-order rate constant for the chloroalkyl-tetramethylrhodamine ligand of 67 M⁻¹ s⁻¹, improved to 2.7 × 10⁶ M⁻¹ s⁻¹ by engineering³⁶. A few years later, the faster HaloTag7 was engineered, bearing mutations around the ligand-binding tunnel, with disruption of surface positive patches and modification to the C terminus. Exploring linkers for fusion identified sequences that minimized degradation and improved expression, indicating that default glycine–serine-rich linkers are not always ideal³⁷. HaloTag has continued to be refined by removing a disulfide³⁸ and variants allowing multiplexed imaging via different fluorescence lifetimes³⁹.

HaloTag has most commonly been used for fluorescent labeling, anchoring dyes with greater brightness and photostability than fluorescent proteins^{36,40}. Lipoic acid ligase couples alkyl halides site specifically to a peptide tag, enabling HaloTag to link covalently to another protein⁴¹. HaloTag can tether molecules inducibly at specific sites in cells. Tethering hydrophobic ligands allows rapid induction of protein degradation, including in zebrafish embryos and a mouse tumor model⁴². Recruiting caged electrophiles to specific sites in cells using HaloTag has uncovered signaling responses that reactive electrophiles activate⁴³. Overall, HaloTag provides an optimized protein fusion for rapid coupling of alkyl halide ligands with specificity in diverse biological contexts.

SNAP-tag

SNAP-tag is a 19-kDa protein that was engineered from the human DNA repair enzyme *O*-alkylguanine alkyltransferase⁴⁴. Alkyltransferase repair proteins are naturally suicide enzymes⁴⁴. Cysteine 145 is activated by proximity to histidine 146 and is poised for S_N2 reaction with alkylated guanines. SNAP-tag forms an irreversible thioether adduct to diverse benzylguanine-linked reactants (Fig. 2e). Mutation of glycine 160 to tryptophan in SNAP-tag facilitates reaction to benzylguanines⁴⁴. SNAP-tag has been principally used for fluorescence imaging⁴⁵. To minimize the need for washing away uncoupled dye to visualize SNAP-tag location, fluorogenic substrates have been developed^{45,46}. CLIP-tag (19 kDa) was similarly engineered from *O*-alkylguanine alkyltransferase, reacting selectively with O²-benzylcytosine conjugates (Supplementary Fig. 1b), providing an efficient orthogonal pair⁴⁷.

SNAP-tag can interface with other click biology: a bifunctional ligand in cells inducibly bridges a SNAP-tag fusion to a HaloTag fusion,

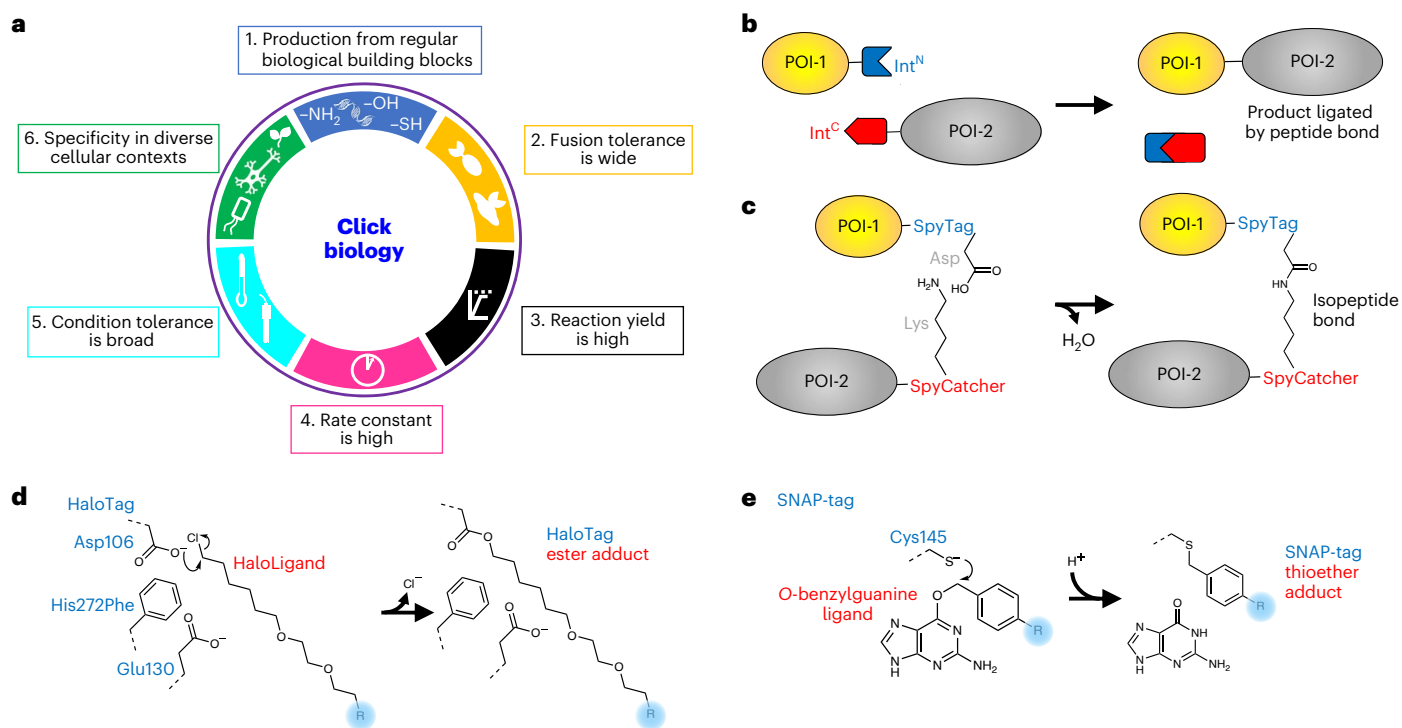


Fig. 2 | Click biology criteria and example reactions. **a**, Schematic of criteria for click biology. **b–e**, Examples of click biology reactions: split intein reconstitution (**b**), SpyTag and SpyCatcher (**c**), HaloTag (**d**) and SNAP-tag (**e**). POI, protein of interest; R, probe of interest.

and such linkage can be photocleavable⁴⁸. Overall, SNAP-tag provides a robust protein module for rapid irreversible coupling of benzylguanine building blocks, showing specificity in diverse cellular contexts.

Other valuable biological couplings

Here I highlight other important reactions to protein components that are highly effective in certain situations but differ in certain characteristics compared to the above reactions.

Coupling of surface-exposed cysteines

The most common covalent reaction to link proteins is formation of a disulfide bond. Disulfide bonds can be engineered to connect two proteins of interest^{49,50} but with limitations:

- Promiscuous reaction, as many proteins have exposed cysteines, while reduced glutathione and other free thiols are abundant in biological systems⁵¹.
- Instability of disulfides, particular in reducing environments such as the cytosol and the nucleus. Even in oxidizing environments (for example, endoplasmic reticulum, Golgi apparatus and outside the cell), disulfide bonds can rearrange⁵².
- Moderate yields for joining component A to component B, because of competition from homodimerization to A2 and B2.
- Side products include cysteine oxidation to sulfenic acid, sulfinic acid or sulfonic acid⁵³ and reaction with nitric oxide to make nitrosothiols⁵⁴.

FIAsH, in which a bisarsenical reacts with a tetracysteine motif, represents an ingenious approach to site-specific reaction⁵⁵. FIAsH relies on a subtle balance of the chelate effect to reverse promiscuous reaction with other closely positioned cysteines in cells. Hence, FIAsH has faced issues in specificity compared to HaloTag and SNAP-tag⁵⁵.

TMP-tag is based on engineering of dihydrofolate reductase to tether acrylamide-linked probes for proximity-based ligation to a surface cysteine. TMP-tag is an exciting approach for covalent labeling,

with especially good qualities for insertion in loops⁵⁶, although relying on exposed cysteine may face difficulties in oxidizing compartments.

Streptavidin and biotin

The interaction of biotin with streptavidin or similar avidin and NeutrAvidin proteins, abbreviated as (strept)avidin, may be considered a precursor of click biology. (Strept)avidin has been used for decades when biologists or even material scientists want a stable and fast-binding anchor^{57,58}. AviTag peptide enabled biotin to be ligated site specifically by the BirA enzyme to specific proteins in complex environments^{59,60}. However, (strept)avidin–biotin has attendant limitations:

- Dissociation of biotinylated ligand becomes substantial in response to time, force or harsh conditions⁶¹.
- Biotin and biotinylated protein are present in all organisms, leading to background binding^{58,59}. Depleting biotin or overexpressing biotin-binding proteins is toxic⁶².
- (Strept)avidin is a tetramer, while mutations to make monomeric versions decrease affinity^{63,64}, impeding use as a fusion partner.

Nonetheless, there is extensive infrastructure from the availability of thousands of reagents and surfaces linked to biotin or (strept)avidin. (Strept)avidin–biotin can be combined with click biology through SpyAvidin⁶⁵.

Enzymatic ligations

Various enzymes have been elegantly engineered for specific coupling to peptides or proteins (for example, butelase, omniligase, subtiligase, microbial transglutaminase and sortase)⁶⁶. Sortase has been most widely employed, particularly for coupling small molecules to C termini⁶⁷. Phosphopantetheinyl transferase enables efficient coupling of coenzyme A-linked probes to acyl carrier proteins, peptide carrier proteins or short peptides, such as the 11-residue ybbR tag⁶⁸. Formylglycine-generating enzyme (FGE) recognizes a cysteine in a motif as short as six residues, incorporating an aldehyde that allows

coupling of hydrazide or aminoxy probes⁶⁹ or more stable Pictet–Spengler ligation⁷⁰.

Enzymatic ligation requires three components (the catalyst and two reactants) and so has different features from two-component click biology reactions. Simply in terms of collisions, it is less likely for three components to come together than two components. The above enzymatic ligations usually require one component in high micromolar concentrations, which is fine for purified components but may reduce generality for cellular application, given the limited solubility or abundance of many substrates^{66,67,71}. Horseradish peroxidase (HRP), APEX2, BioID, miniSOG or PafA generate broadly reactive species for proximity labeling and so are optimized for low specificity⁷², in contrast to the click biology criterion of specificity in diverse cellular contexts.

Incorporation of click chemistry building blocks into biological macromolecules

Bioorthogonal click chemistry provided huge opportunity for modification of biological macromolecules². One can exploit the promiscuity of natural enzymes to incorporate modified substrates. Such metabolic labeling has been employed for sugars, proteins, lipids, nucleic acids and core metabolism^{2,34,73,74}. Because the same building blocks are often incorporated into multiple sequences (for example, azidohomoalanine in place of methionines in different proteins), this approach is powerful for broad profiling of cellular processes^{34,75}.

Monosaccharides bearing click chemistry probes can be taken up by cells, activated to the sugar nucleotide and then incorporated by glycosyltransferases⁷⁶. Labeled molecules traffic within cells, including to the cell surface, allowing subsequent reaction with click chemistry probes. Metabolic glycan labeling was a breakthrough for imaging glycoprotein trafficking, even in multicellular organisms, and for enriching specific glycoprotein classes³.

The nucleoside 5-ethynyl-2'-deoxyuridine (EdU) is taken up by cells, activated to the triphosphate and incorporated into DNA in place of thymidine by DNA polymerases, allowing detection of the terminal alkyne with azide probes. EdU provides a powerful tool to monitor DNA synthesis and cell proliferation without denaturing the double helix^{34,77}.

Enzymes may also be engineered for selective recognition of substrates containing click chemistry functional groups. This approach has been creatively advanced for incorporation of noncanonical amino acids by the ribosome, using a specific transfer RNA (tRNA) (often a suppressor tRNA recognizing the amber stop codon) and an engineered aminoacyl tRNA synthetase charging suppressor tRNA with the noncanonical amino acid⁷⁸. Amino acids bearing azide, alkyne, tetrazine, cyclopropene and *trans*-cyclooctene functionality have been incorporated, enabling a range of fundamental biological advances⁷⁸. Genetic code expansion has unique precision in which any residue may be targeted for modification. The next section considers how simplicity and scale differ for such an approach compared to click biology.

Other self-labeling enzyme probes

Important work has engineered other self-labeling enzymes, with some of the most established being cutinase, PYP-tag and BL-tag. Cutinase is a serine esterase that forms stable adducts upon reaction with phosphonate esters⁷⁹. Cutinase is 22 kDa and monomeric and tolerates fusion at either terminus; its two disulfide bonds may be a challenge for certain fusions⁷⁹. Cutinase has been especially powerful for monolayer attachment and sequential protein assembly⁷⁹. PYP-tag is based on the 14-kDa photoactive yellow protein, the cysteine of which couples with 7-hydroxycoumarin-3-carboxylic acid-linked fluorophores⁸⁰. PYP-tag reacts with thioester ligands and generates a thioester product⁸⁰. BL-tag was engineered from a 29-kDa TEM-1 β -lactamase in which a serine forms an ester adduct with β -lactam variants, but mutation of

glutamate 166 to asparagine causes this adduct to remain trapped⁸¹. BL-tag allows labeling with diverse β -lactam fluorophores in living cells, including with fluorogenic probes⁸¹.

The fit of current reactions to ideal click biology criteria

Here, I explore why click biology criteria matter for biological research as well as how current reactions match up to ideal goals.

Click biology criterion 1: production from regular biological building blocks

Mathematicians and physicists put a high value on simplicity, chemists and biologists not always so much⁸². Click biology employs building blocks universal through the biosphere, which facilitates:

- Simple generalization to different systems.

Click biology can be used without optimization in different species, simply requiring delivery of the right nucleic acid sequence. SpyTag–SpyCatcher has been employed without alteration in phage, bacteria, archaea, plants, fungi, *Caenorhabditis elegans*, *Drosophila melanogaster* and mammals⁸³. HaloTag and SNAP-Tag have also been successful in diverse species^{40,45}.

Proteins containing site-specific noncanonical amino acids are not so readily available as starting materials, requiring expression of three components: the gene of interest with a site for noncanonical amino acid incorporation, the noncanonical tRNA and the aminoacyl tRNA synthetase⁷⁸. Only one component needs to be encoded for a click biology building block, which simplifies delivery (gene therapy vectors have limited capacity)⁸⁴ and decreases metabolic demands⁸⁵. Noncanonical amino acid mutagenesis usually requires delivery of the amino acid to the cell and then washing out the amino acid before the other reacting part can be added⁸⁶: labeling is inhomogeneous (Fig. 3). For split intein reconstitution and SpyTag–SpyCatcher, the reaction is homogeneous: as soon as components are synthesized, they can react (Fig. 3). Homogeneous reaction may simplify use, particularly in living organisms, where washing away unreacted building blocks is not straightforward.

- Environmentally friendly large-scale production.

Avoiding the need to synthesize new building blocks chemically should facilitate cheap scalable production with low environmental impact⁸⁷. A fascinating feature of the original click chemistry review was discussion on mitigating the explosion risk of azides¹. Low environmental impact (for example, no organic solvent, using only renewable resources) is more true of split inteins and SpyTag–SpyCatcher, where both reactive parts can be encoded by the cell, rather than HaloTag and SNAP-tag, where one reactive part requires chemical synthesis. Cheap production for click biology may provide opportunities in low-price markets, for example, enzymes in chicken feed or plant-based meat replacements^{88,89}. Fitting the click chemistry criterion (Table 1), side products from click biology are inoffensive (Cl⁻ for HaloTag, H₂O for SpyTag–SpyCatcher, guanine for SNAP-tag).

Click biology criterion 2: fusion tolerance is wide

Modularity is central to enhance the ability to engineer biological systems⁹⁰. Because of the complexity of protein folding and the role of terminal groups in the function of certain proteins, no genetic fusion tag will be universally successful. Thousands of successful fusions have been made to click biology components^{14,40,45,83}, but it is important to understand the limitations. For split inteins, the reaction must occur at a specific terminus¹⁴. HaloTag, SNAP-tag, SpyTag003 and SpyCatcher003 react via side chains and function at either the N or the C terminus^{40,45,91}. Tolerance of fusion to either terminus enables the generation of nonlinear protein architectures as well as allowing a single virus-like particle (VLP) to display antigens that prefer fusion at

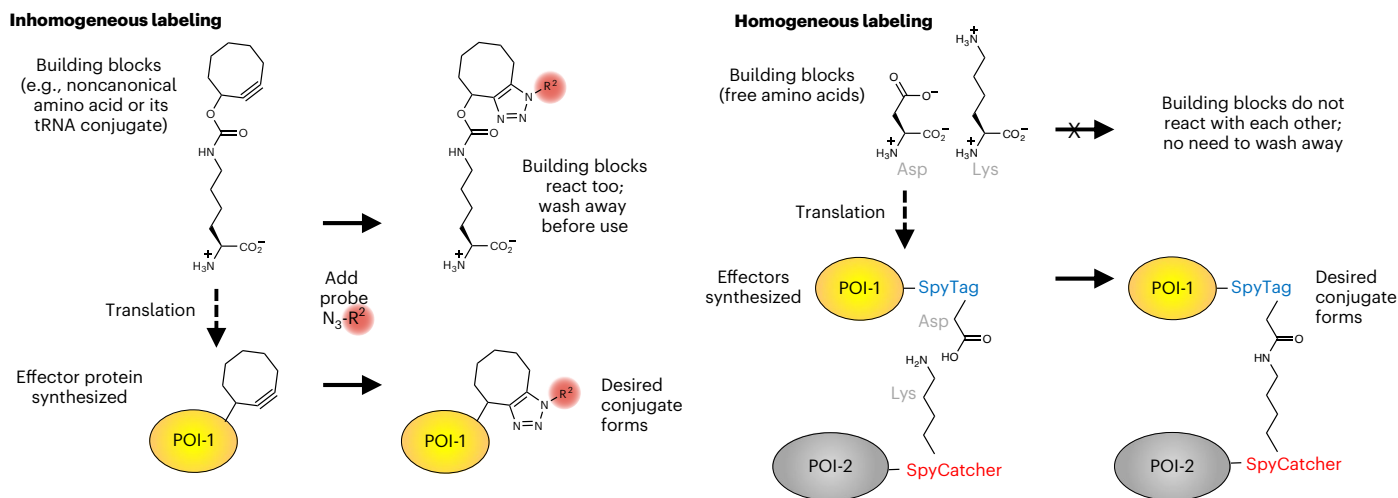


Fig. 3 | Inhomogeneous compared to homogeneous protein labeling. For click chemistry, the building block usually has comparable reactivity to the same functional group in a macromolecule, so labeling is inhomogeneous. For click biology, there is only reactivity in the folded environment, so side reactions from building blocks are not a concern.

their N terminus (for example, influenza neuraminidase) or C terminus (for example, influenza hemagglutinin)⁹².

The number of residues in a tag does not correlate with the chance of disruption. Folded domains (usually >50 residues) are able to bury hydrophobic residues in the protein core. Shorter tags may have exposed hydrophobic residues that promote promiscuous interaction or recognition by quality control machinery⁹³. Such quality control may alter trafficking (for example, retention in the endoplasmic reticulum) or promote degradation or stress responses⁹⁴. High charge-density promotes nonspecific interactions: high positive charge favors binding to nucleic acids or phospholipid bilayers⁹⁵. For a fast rate constant, one needs high-affinity initial noncovalent interaction, which correlates with interaction surface area⁹⁶. Hence, there are biophysical reasons why it is unlikely for a five-amino acid peptide to have a high-affinity specific interaction that leads to an efficient reaction⁹⁶.

Click biology criterion 3: reaction yield is high

Reaction yield and stability of the ligated product for weeks to months are especially important for devices in synthetic biology. For vaccines and diagnostics in global health, there is urgent need to maintain long-lasting activity without a cold chain⁹⁷. High-affinity noncovalent interactions may apparently maintain stability by dissociating but rebinding⁶¹. However, if such interactions are multiplexed on a protein chip, binding or unbinding will mix up reagents. For surface engineering of cancer-targeting cell therapy⁹⁸ or vaccine particles that stimulate the immune system over several days, coupling stability is crucial⁹⁹.

It is unusual to see equilibrium constants measured for click reactions, partly because of the difficulty in measuring low concentrations of unconverted reactant for efficient reactions. An operational definition for conversion of one reactant is more common: for reaction $A + B$ to AB , then $[AB]/[A]$ in the presence of excess B after reaching equilibrium. An operational definition may be more appropriate in click biology because living systems are not at equilibrium: biomolecules are being synthesized, trafficking and being degraded, not to mention posttranslational modification or changes in folding. If the reaction changes the molecular weight substantially (for example, for split inteins or SpyTag–SpyCatcher), western blot provides a convenient way to quantify conversion²⁸.

In some situations, high yield is not essential and there may be a more practical criterion: that is, when imaging my HaloTag fusion in a cell, is there sufficient fluorescent signal to see that protein move? In other situations, high yield is key: that is, if one saturates

the HaloTag fusion with the blocking 7-bromoheptanol, any newly coupled dye represents newly synthesized HaloTag fusion¹⁰⁰. Fluorescent pulse labeling avoids more toxic approaches to track protein lifetime, such as protein synthesis inhibitors. Pulse labeling was taken to another level through HaloTag's irreversible reaction within protein fibers, recording fluorescent memory like the ticker tape of a cell's transcriptional activation¹⁰¹. Using click biology to link DNA-binding elements to protein-based transcriptional activators in the nucleus, there is no way to wash away excess components, so high yield is essential¹⁰². For multistep ligation, high yield minimizes truncated products. Nine steps each at 95% efficiency give an overall yield of 63%; nine steps each at 50% give an overall yield of 0.2%. In the presence of excess partner, reaction with SpyTag–SpyCatcher or SnoopTag–SnoopCatcher gives >95% conversion, allowing multiple sequential reactions³⁰.

For the highest yield, one needs a highly stable product. Like triazoles, isopeptide bonds from Tag–Catcher have chemical, biological and mechanical stability^{1,24,26}. Esters from the HaloTag reaction have lower chemical stability. Yield also depends sensitively on the likelihood of side reactions. Split intein yield can be limited by competition with hydrolysis or other trap states¹⁴ (Supplementary Fig. 1a), but, for the latest split inteins, such a side reaction is only a few percent²⁰.

Subverting click biology for reversibility can also bring benefits. Nonreactive variants of SpyTag–SpyCatcher allow analysis of cell adhesion dynamics²⁸ or switchable protein elution¹⁰³, while reversible HaloTag allows dye exchange to resist photobleaching¹⁰⁴.

Click biology criterion 4: rate constant is high

Why do biologists have a need for speed? There are three main routes to high conversion: reactants at high concentration, a high rate constant or waiting a long time. For small molecules in isolation, consistent with the original click chemistry, reactants can indeed be in the millimolar range. However, many proteins will not be soluble above the low micromolar range, while the median concentration of a protein in eukaryotic cells is in the low nanomolar range⁷¹. Therefore, high rate constants are crucial to wide use of click biology. For bioconjugation on isolated components, it may be feasible to be patient and wait days, while the biomolecules may avoid degradation. However, many biological processes occur during the blink of an eye: membrane fusion, ion channel opening, action potential and muscle contraction¹⁰⁵. Therefore, high rate constants enable deeper understanding and increased control of biological processes. Raising concentrations of reactive dye leads

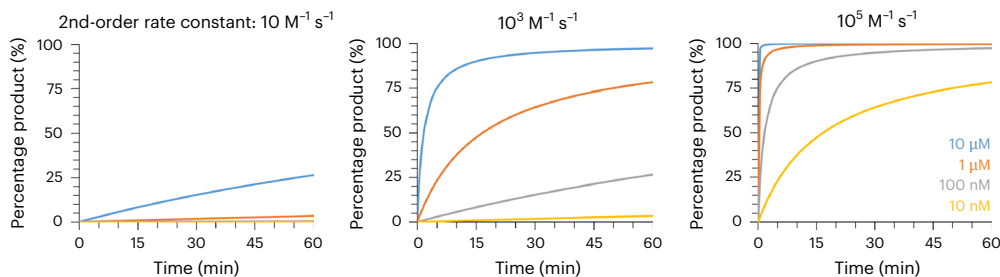


Fig. 4 | Consequence of click reaction speed. Species A reacts with species B to give AB, with the reaction first order in A and first order in B and second order overall, assuming no reversibility. Percentage product formation over time with

both reactants initially at 10 nM (yellow), 100 nM (gray), 1 μ M (orange) or 10 μ M (blue), with a second-order rate constant of 10 (left), 10^3 (middle) or 10^5 $M^{-1} s^{-1}$ (right).

to high background from nonspecific dye interaction and fluid-phase uptake into cellular vesicles. The dramatic consequences for reaction progress in different concentration ranges with different rate constants are illustrated in Fig. 4.

For CuAAC, rate constants of 10 to 100 $M^{-1} s^{-1}$ are typical, which are good for the reaction of highly soluble isolated components⁴. For the copper-free reaction in strain-promoted azide-alkyne cycloaddition, rate constants of 10^{-2} to $1 M^{-1} s^{-1}$ are commonly observed⁴. The inverse electron-demand Diels-Alder reaction of *trans*-cyclooctene with tetrazines has rate constants of $1 M^{-1} s^{-1}$ up to an impressive $10^6 M^{-1} s^{-1}$, although some pairs lacked storage stability and had increased side reactions⁴. For SpyTag003-SpyCatcher003, the rate constant of $5.5 \times 10^5 M^{-1} s^{-1}$ is close to the diffusion limit given the size of reactants²⁸. HaloTag7 rate constants are almost diffusion limited for particular rhodamine ligands (for the dye Live580, $1.4 \times 10^8 M^{-1} s^{-1}$), while rate constants were two to five orders of magnitude slower for nonfluorophore ligands¹⁰⁶. SNAP-tag's rate constant is 10^4 – $10^5 M^{-1} s^{-1}$ for a range of reactants¹⁰⁶. SNAPf represents a second-generation SNAP-tag, displaying faster reactions with Alexa Fluor 488, Atto 549 and Alexa Fluor 647 probes⁴⁶. For split inteins, researchers report half-times with each reactant in the low micromolar range (for example, 5 s for Gp41-1 or 1 min for VidaL)^{15,20}, so the fit to this speed criterion is less clear.

Click biology criterion 5: condition tolerance is broad

In the initial formulation of click chemistry, requirements on reaction tolerance would be undemanding for most biological components: tolerance to water and oxygen¹. Subsequently, Barry Sharpless said, "The CuAAC reaction cannot be stopped. It'll go 100% every time. You can have it in urine. You can have it in minestrone soup"¹⁰⁷. CuAAC can sometimes form reactive oxygen species and other radicals with consequent nucleic acid damage, so control of catalyst loading and inert atmosphere have occasionally been required³⁴. Challenges for the robustness of click chemistry also come from unintended reaction of strained alkynes with thiols¹⁰⁸, reduction of azides within cells¹⁰⁹, breakdown of tetrazines¹¹⁰ and isomerization or polymerization of some *trans*-cyclooctenes¹¹¹.

When describing click biology reactions above, a single rate constant was presented. However, these rate constants often represent best-case scenarios. Users are often surprised when reactions appear less efficient in the rough-and-tumble of real application. Small changes in conditions or fusion partners may have large effects on rate. Rate constants can be orders of magnitude slower upon fusion to large species, notably nanoparticles, as well as being slowed by steric hindrance (as for dense coupling on a VLP)^{99,112}. It is not easy to please all of the people all of the time: biochemists with labile complexes want reactions rapid at 4 $^{\circ}C$, while cell biologists want rapid labeling at 37 $^{\circ}C$. As catalysts need suitable flexibility¹¹³, even for single-turnover enzymes, it is not trivial to obtain high activities at both temperatures. Researchers focused on extracellular biology want reaction in an oxidizing environment with high Na^+ and millimolar Ca^{2+} ; researchers working on

the cytosol want reaction in a reducing environment with high K^+ and nanomolar Ca^{2+} . Organisms are active from sea ice (-20 $^{\circ}C$) to hydrothermal vents (121 $^{\circ}C$)^{114,115}. Reaction at pH 7–8 fits most applications, but pH ranges from 1 to 3 in the human stomach and 4.5 in lysosomes to pH 11 in parts of the mosquito midgut¹¹⁶.

Hyperthermophilic archaea have provided a split intein that splices well at 80 $^{\circ}C$ ¹¹⁷ and a DNA alkyltransferase that couples at 70 $^{\circ}C$ and from pH 5 to 10 (ref. 118). Intein splicing that is activated at 70–80 $^{\circ}C$ enabled hot-start PCR for reduced background DNA amplification¹¹⁹. SpyTag-SpyCatcher regained reactivity after unfolding by heat, drying or sodium dodecyl sulfate (SDS), as applied to nanofibrous protein films or making enzymes resilient to boiling^{120,121}.

Click biology criterion 6: specificity in diverse cellular contexts

Achieving reaction at a single site on a specific target, despite the presence of thousands of biomolecules composed from the same building blocks, is a key goal for click biology. Nonetheless, perfect selectivity is impossible to achieve and is not necessary; many clinically effective pharmaceuticals interact with unintended targets¹²². HaloTag and SpyTag003-SpyCatcher003 have shown good specificity in western blotting on various cell types^{28,123}. Further mass spectrometry would give deeper insight into any unintended reaction products. Possible noncovalent interactions of click biology components may have cellular relevance. Toxicity represents a central indicator for lack of specificity. The key operational definition of specificity comes from negative controls looking for the relevant signal, apparently from ligated product AB, in the absence of component A or the absence of component B. For example, what is the fluorescence image from adding the chloroalkyl-fluorophore to control cells that do not express HaloTag? Different organisms and compartments may show different specificity. Labeling of reactant A by reactant B may be apparently selective when reactant A is expressed at high levels (that is, a million copies per cell). When reactant A is expressed at the level of a low-abundance protein (thousands of copies per mammalian cell)⁷¹, competing processes may dominate. Overexpression of biomolecules can distort function, so further advancement of click biology will be required to ensure suitable application across the range of expression levels.

Ideal tags and probes have low intrinsic reactivity, minimizing off-target reaction. It is worth exploring how any undesired reactions of the more electrophilic phosphonate esters (for cutinase), thioesters (PYP-tag) or β -lactams (BL-tag) compare to side reactions from alkyl halides for HaloTag and O-alkylated probes for SNAP-tag and CLIP-tag^{79,80}. For spontaneous isopeptide bond formation, there is low intrinsic chemical reactivity (carboxylic acid on SpyTag and amine on SpyCatcher); the SpyTag-SpyCatcher reaction occurs in animal models, including redirecting chimeric antigen receptor (CAR) T cell killing for combating tumor evasion^{98,124}.

Overall, it is important that specificity in click chemistry and click biology is not considered as binary (specific or nonspecific) but a continuum that varies according to biological circumstance.

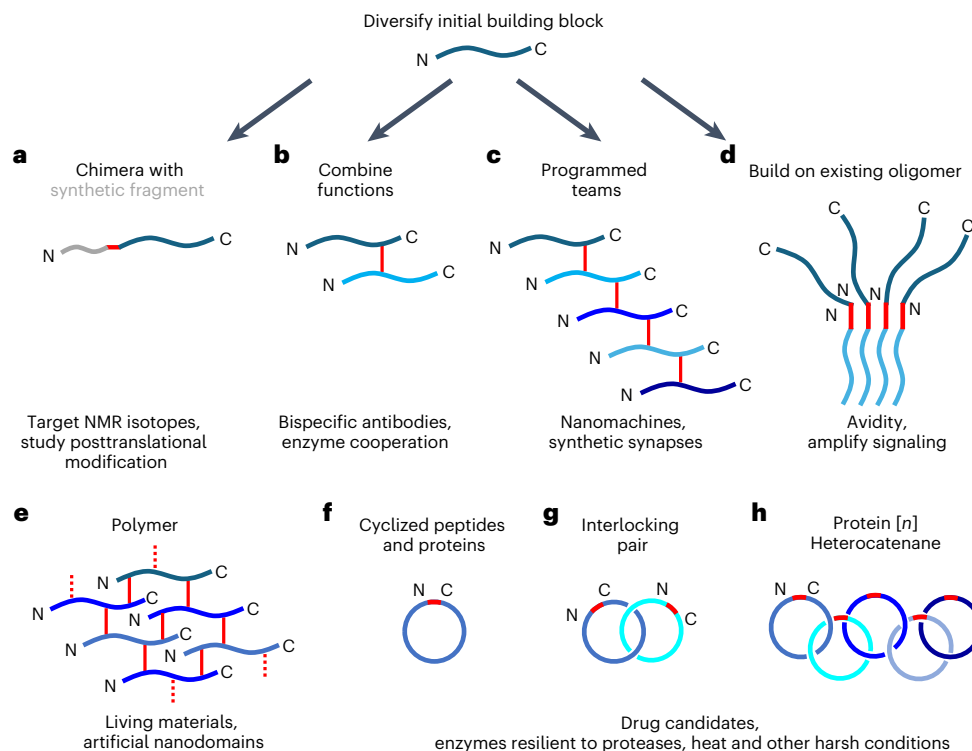


Fig. 5 | Noncanonical architectures from click biology. At the top, the linear structure in blue represents the connectivity from the N terminus to the C terminus for a protein of interest fused to a click biology partner. Ligation is shown generically in red. **a**, Linkage to a fragment synthesized in vitro that contains specific isotopes or posttranslational modification. NMR, nuclear magnetic resonance. **b**, Linkage to another effector biomolecule. **c**, Programmed

linkage to multiple different effector biomolecules. **d**, Linkage to a scaffold to enhance activity through oligomerization. **e**, Multiple site-specific reactions on each chain to generate a hydrogel. **f**, Self-reaction through a peptide or isopeptide bond. **g**, Reaction to generate an interlocking pair of chains. **h**, Reactions to generate multiple interlocking chains.

Opportunities from click biology

The robust reactions of click biology generate unique opportunities from basic research to disease prevention.

Construction of noncanonical architectures

Click biology provides a simple way to generate novel robust biomolecular architectures (Fig. 5). Noncanonical linkages enable exploration of new functional space, particularly for cooperation between functional units (Fig. 5a,b). Natural biomaterials may employ cross-links for mechanical performance, including disulfides, transglutaminase-catalyzed isopeptide bonds and sulfilimine cross-links in keratin of hair, horn or collagen^{125,126}. Natural biomaterials are difficult to recapitulate synthetically, depending on the interplay between many structural biomolecules, enzymes and cell types¹²⁷. Click biology helps to abstract features of biomaterials programmably.

Split inteins are amenable to solid-phase peptide synthesis, enabling linkage of ribosome-expressed proteins to isotopically labeled or nongenetically encoded probes, to enhance the size limit for nuclear magnetic resonance structure determination^{14,128} (Fig. 5a). High yield enables multiple sequential reactions, analogous to programmed in vitro synthesis of peptide or oligonucleotide chains. ‘Megamolecules’ were generated through small-molecule bridges linking HaloTag, SNAP-tag and cutinase¹²⁹ (Fig. 5c). ‘Polyproteams’ used sequential SpyCatcher–SnoopCatcher bridges to extend chains of building blocks bearing a SpyTag and a SnoopTag to find synergy in death receptor and receptor tyrosine kinase signaling³⁰ (Fig. 5c). Gigavalent display employed polyacrylamide beads that covalently assemble 10^9 protein molecules through SNAP-tag, HaloTag, SpyTag–SpyCatcher and SnoopTag–SnoopCatcher bridging¹³⁰. Biological signaling is often driven by clustering, so modular ligation of a protein

of interest, through mixing with a panel of SpyCatcher-linked oligomers, allows rapid exploration of the optimal cluster size¹³¹ (Fig. 5d). SpyTag–SpyCatcher can also drive polymerization to form hydrogels tailored with motifs for cellular interaction, degradability¹³² and light responsiveness¹³³ (Fig. 5e).

Cyclization can occur spontaneously from a single polypeptide. Split intein circular ligation of peptides and proteins ingeniously encodes Int^C at the N terminus and Int^N at the C terminus, so splicing generates a cyclic product with therapeutic potential¹³⁴ (Fig. 5f). Libraries of up to 10^8 different cyclic peptides allow in-cell screens in *Escherichia coli*, yeast and human cells¹³⁵. Enzymes with SpyTag at one terminus and SpyCatcher at the other terminus undergo an intramolecular reaction¹³⁶. Such SpyRing enzymes often gain resilience to harsh conditions, such as during industrial processing¹²¹, facilitated by the SpyTag–SpyCatcher moiety as well as chain topology (Fig. 5f). Resilience depends partly on decreased unfolding but more on increased refolding¹²¹. Nanodiscs cyclized with SpyTag–SpyCatcher enhance purification of membrane complexes from lipid bilayers²⁹.

Click biology allows a range of tadpole-shaped or H-shaped protein architectures, including knots enhancing stability^{137,138} (Fig. 5g) and even multiple interlocking rings toward the synthesis of five chains interlocked like Olympic rings¹³⁹ (Fig. 5h).

Bridging biological and nonbiological building blocks

Through rapid, selective and high-yield reactions, click biology can impact the labeling of biomolecules in isolation and especially in cells or organisms (Fig. 6). HaloTag and SNAP-tag are widely used for fluorescent tracking of protein dynamics in microorganisms, animals and plants^{40,45} (Fig. 6d). HaloTag has been used as a reporter gene for targeting positron emission tomography probes¹⁴⁰ or for magnetic

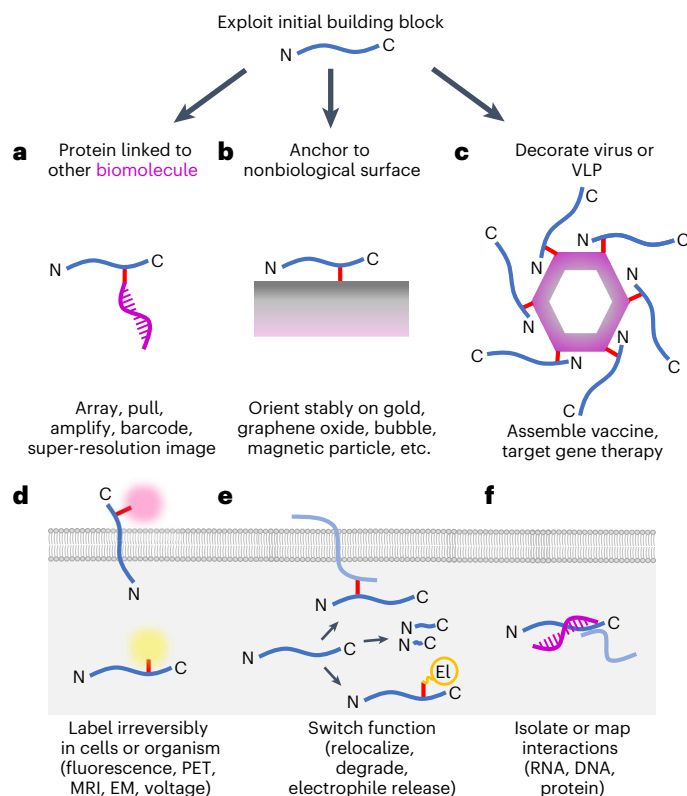


Fig. 6 | Extending molecular function through click biology. At the top, the linear structure in blue represents the connectivity from the N terminus to the C terminus for a protein of interest fused to a click biology partner. Ligation is shown generically in red. **a–c**, This polypeptide can be linked to other biomolecules or nonbiological probes (**a**), surfaces or nanoparticles (**b**) or viruses or VLPs (**c**). Bottom, cellular applications of click biology. **d**, Irreversible labeling at the plasma membrane or inside the cell. EM, electron microscopy; MRI, magnetic resonance imaging; PET, positron emission tomography. **e**, Ligation for switching protein function, such as through relocalization, degradation or inducible delivery of a caged electrophile (El). **f**, Irreversible capture of a protein of interest after lysis through click biology enhances mapping of cellular interactions, such as by RNA sequencing or mass spectrometry.

resonance imaging¹⁴¹. SNAP-tag enables subcellularly targeted anchoring of probes to sense viscosity, Ca^{2+} , Zn^{2+} , NO, H_2O_2 or protein aggregation (via solvatochromic dyes)⁴⁵. Click biology can functionalize protein polymers (including amyloids extending from bacteria) with gold nanoparticles, quantum dots and carbon nanotubes¹⁴². SpyTag–SpyCatcher bridges proteins stably and site specifically to diverse surfaces (magnetic beads, atomic force microscopy tips, bubbles)^{91,143,144} (Fig. 6b). SpyTag captures proteins from cells on graphene oxide with minimal sample processing to enhance structure determination by cryoelectron microscopy^{145,146}. The membrane permeability of probes for HaloTag and SNAP-tag provides a clear advantage over SpyTag–SpyCatcher and split inteins in targeting synthetic probes for intracellular applications (Fig. 6d).

Engineering cellular intelligence

A major focus of synthetic biology has been the programming of tuneable and logic-gated cell behavior through transcriptional circuits⁹⁰. However, there is a substantial time delay before transcription changes cell behavior as well as the difficulty of nucleic acid delivery for therapeutics. Hence, it is important to develop improved routes for responsiveness and logic gating at the protein level. Unraveling the logic of immune synapses is facilitated by titratable and combinatorial display

of ligands, using ‘CombiCells’ displaying SpyCatcher in different formats at the plasma membrane¹⁴⁷ (Fig. 6d).

Split inteins are especially important in logic-gated responses¹⁴. To increase the barrier to the spread of herbicide resistance from a genetically modified plant, two segments of a herbicide resistance protein were fused to each part of a split intein, with genes inserted on different chromosomes¹⁴⁸. For inducible activation of therapeutics, nanobodies were linked to split barnase and a split intein. Binding of the nanobodies to a cancer cell target enables intein splicing, barnase reconstitution and activation of cytotoxicity¹⁴⁹. Uncaging split inteins by proteases allows detection or response to viral infection¹⁵⁰. Splicing of address tags through split inteins provides control of protein translocation to specific organelles¹⁴ (Fig. 6e).

Fundamental cell biology: surviving tough conditions and traceability

Covalent reactions have a great advantage in ease of quantitation. To study whether a noncovalent binding protein has encountered its protein target inside a cell is not easy. Pulldown interaction depends sensitively on isolation conditions. Fluorescence resonance energy transfer detection requires expensive equipment and careful calibration. However, if two proteins of interest are equipped with a split intein or SpyTag–SpyCatcher, interaction is easily quantified by gel shift assays (after boiling in sodium dodecyl sulfate to break noncovalent interactions)^{14,28}.

A key way to learn about biological function is by identifying which components a biomolecule interacts with. A few candidate contacts can be probed inside living cells by fluorescence microscopy, but broad analysis requires breaking open cells and high-throughput analysis (by sequencing nucleic acids or protein mass spectrometry). A common concern is avoiding weak promiscuous binding interactions, which is best done by intense washing. Cross-linking and immunoprecipitation is a key method to determine protein interactions with RNA but suffers from frequent false positives. Linking a candidate RNA-binding protein to SpyTag or HaloTag allows covalent anchoring to resin before a denaturing wash to remove noncovalent nonspecific complexes, helping to map cell-relevant protein–RNA interactions^{151,152} (Fig. 6f).

The response of cells and biomolecules to mechanical force has historically been underappreciated¹²⁷. It is important to have a stable grip on what one is testing mechanically, so click biology facilitates simple, precise and strong immobilization (Fig. 6b). There has been wide use of HaloTag and SpyTag–SpyCatcher for single-molecule force spectroscopy through atomic force microscopy or magnetic tweezers^{26,153}. Such covalent anchoring enabled the study of force in cell-based systems, uncovering mechanical responsiveness of the giant protein titin¹⁵⁴ or myosin assemblies¹⁵⁵.

Enhanced viral vector and VLP function

Viral vectors and VLPs present the most immediate prospect of medical impact for click biology (Fig. 6c). Adeno-associated virus (AAV) vectors are one of the most successful gene therapy platforms but are held back by the size of their genetic cargo. Multiple AAV vectors, each encoding one protein fragment of the final product, employ split intein splicing for successful gene therapy in the retina¹⁵⁶. Click biology ligation of targeting ligands to the surface of viral vectors has shown promise to address inefficient delivery to the tissue of interest to avoid the majority going to the liver^{33,157,158}.

Individual foreign proteins induce weak and transient immune responses when injected into an animal. Decoration of a foreign protein in a repeating virus-sized organization induces larger and longer-lived immune responses, characteristic of an effective vaccine^{159,160}. Hence, there has been interest in standardized scaffolds to which diverse antigens can be coupled. Conventional cross-linkers to bridge amines or thiols have limitations in uniformity of antigen display⁹⁹. Click chemistry

has been employed for vaccine candidates, with cowpea mosaic virus chemically conjugated to an azide¹⁶¹. Challenges with cost and scale of production are especially important for vaccines. The potential toxicity of copper has also been a concern for regulators¹⁶². Split inteins and SpyTag–SpyCatcher have been employed to generate a range of vaccine candidates, including bacterial outer membrane vesicles¹⁶³, viral vectors³³ and VLPs^{164,165}. Modular assembly allows generality in decorating VLPs using antigens with symmetry (for example, cyclic dimers, trimers or tetramers) that may not fit the symmetry for genetic fusion to a particular capsid⁹².

Click chemistry has reached clinical trials in therapeutics and diagnostics. Tetrazine-containing hyaluronic acid is injected into the tumor, followed by intravenous injections of caged doxorubicin linked to *trans*-cyclooctene; reaction at the tumor triggers localized release of the toxic doxorubicin (ClinicalTrials.gov ID [NCT04106492](#))⁷. Also in clinical trials is click chemistry-based labeling of an anti-tumor antibody with a radioisotope ([NCT04692831](#)) as well as pretargeting to enhance positron emission tomography imaging of cancer ([NCT05737615](#))⁶.

Click biology has reached clinical trials in the field of vaccines. VLPs assembled through SpyTag–SpyCatcher are in phase 1 against malaria ([NCT05357560](#))¹⁶⁶ and cytomegalovirus ([NCT06145178](#)), and a related Tag–Catcher VLP completed a phase 3 trial against SARS-CoV-2, showing non-inferior results to the approved messenger RNA vaccine Comirnaty ([NCT05329220](#))¹⁶⁷.

Conclusions and future directions

The concept of click biology can illuminate opportunities for exploring and enhancing biological function. Click biology has many similarities to click chemistry and highlights strengths and weaknesses of an approach centered on exotic functional groups, rather than on reactivity through ubiquitous biological building blocks. Both approaches simplify access to new combinations, focusing on the ease more than the perfection of the join. Noncanonical amino acid mutagenesis with click chemistry reactants provides extraordinary precision in targeted protein ligation. By contrast, click biology provides simple generalization to different cellular systems and should facilitate cheap scalable production of reactive modules. In click chemistry, the unassembled building block (for example, azide-containing amino acid) has similar reactivity to the reactant incorporated in the macromolecule. Click biology facilitates homogeneous (wash-free) labeling, as reactivity only emerges when the macromolecule is expressed and folded.

To summarize opportunities arising from click biology, these reactions provide a simple route to noncanonical biomolecular architectures, providing a dimension to protein engineering beyond evolving or designing the best primary structure¹⁶⁸. Through efficient coupling of nonbiological components in living systems, click biology supercharges the application of ingenious small-molecule and nanotechnology reporters for maximal impact on biological understanding. Through split protein reconstitution, click biology provides logic-gated behavior that has previously centered at the nucleic acid level. The resilience to harsh conditions from click biology empowers biochemical and mechanochemical exploration of cellular complexity. Finally, simple modular decoration of viral vectors and VLPs may lead to the first clinical applications of click biology. There are no perfect click biology reactions, but there are useful ones. Priorities to advance click biology include:

- Increase infrastructure of relevant tools.

When the first click chemistry review was published¹, there were minimal alkyne or azide reagents commercially available. Now diverse toolboxes for click chemistry can be purchased, related to cross-linkers, nucleic acids, peptides and drug discovery³⁴. Look in the supermarket for pre-sliced apples and pre-grated cheese; apparently simple steps are obstacles for the consumer. It is important not to underestimate

the importance of obtaining resources off the shelf for wide use of a method. Hence, more commercially available beads, columns, dyes and linkers for click biology will have major impact.

- Advance click biology beyond linking proteins to other proteins or to small molecules.

There are wide gaps in reactions meeting click biology criteria for linkage to DNA, RNA, lipids, phospholipids and carbohydrates (Fig. 6a). Linking proteins to DNA is best studied, because of the great opportunities such as barcoding and super-resolution fluorescence microscopy¹⁶⁹. Oligonucleotides can be conjugated chemically with ligands for coupling to HaloTag or SNAP-tag, allowing integration with DNA nanotechnology¹⁷⁰. Uracil in DNA allows coupling to the uracil-DNA glycosylase UdgX¹⁷¹. To form covalent bonds to natural DNA bases, a key advance was gene A*-tag, where polarization by a divalent metal ion helps tyrosine nucleophilic attack on the phosphodiester backbone¹⁷². The minimal VirD2 protein forms a covalent bond with single-stranded DNA bearing a specific recognition sequence, although only 40% of protein gets labeled¹⁷³. Orthogonal HUH endonucleases allow up to 95% coupling to DNA in 5 min, although they face challenges with reversibility and the requirement for millimolar Mn²⁺ (refs. [169,174](#)).

Ribozymes or DNAzymes for diverse reactions have been generated through systematic evolution of ligands by exponential enrichment^{175,176}. DNAzymes can ligate RNA, coupling 2'- or 3'-hydroxyl groups on one RNA with another 5'-triphosphate-containing RNA oligonucleotide, generating linear or branched RNA chains^{175,177}. Challenges for nucleic acid-only click biology include the slowness of many nucleic acid-directed reactions¹⁷⁵ as well as receptors detecting unnatural nucleic acid cellular location or modification, which may initiate immune activation¹⁷⁸.

- Quantify and enhance tolerance of click biology fusions on a proteome-wide scale.

When a new click biology reaction is generated, it is relatively easy to quantify the fit to criteria on yield, rate and condition tolerance. However, evaluating fit to criterion 2 (broad fusion tolerance) and criterion 6 (specificity in diverse cellular contexts) has relied upon uncoordinated work over years. Screens on AAV capsid showed a range of tolerated insertion sites for SNAP-tag and SpyCatcher, even in an assembly that progresses through a full virus life cycle; in some cases, insertion enhanced fitness¹⁷⁹. Proteome-wide libraries in *Saccharomyces cerevisiae* have been powerful in exploring fusion tolerance of fluorescent protein tags^{180,181}. Such libraries of click biology fusions will enable new biological exploration and will also establish principles for maximizing fusion tolerance. By focusing on the cell's 'awkward customers' who do not put up with fusions well tolerated elsewhere, iterating screens with module variants would lead to the most general fusion tolerance.

Rapidly growing metagenomic libraries increase the discovery of catalytic activities that can be engineered into self-labeling enzymes¹⁸². Machine learning web servers are increasingly successful at predicting the activity and substrates of enzymes from sequence alone¹⁸³. AlphaFold, ESMFold and RoseTTAFold allow effective three-dimensional structure prediction on large datasets as well as providing insight into protein dynamics^{183,184}. Training of large language models allows the sequence space of functional protein folds to be predicted, accelerating design of novel properties¹⁸⁴. Recent focus has been on plastic-degrading enzymes^{182,185}, but similar approaches will be applicable for click biology.

Improved directed evolution, including continuous evolution, will accelerate refinement of click biology for new targets and challenging environments¹⁸⁶. Despite recent leaps in computational design, directed evolution still has an important role¹⁸⁷: to progress promising leads to meet stringent click biology criteria.

- Enhance clinical applicability of click biology.

It normally takes many years for new reactions to move from the laboratory into commercial or clinical application. Unknowns need to be addressed in product-market fit, scalability, intellectual property and the understanding of consumers and regulators. Click biology has recently entered clinical trials, and it will be important to gain further understanding of the safety profile and scalability. Reduction in immunogenicity would also enhance construction of modular therapeutics using click biology⁹². The immune system can detect even small molecular changes. A big decrease in anti-cancer antibody response was seen when immunizing against tumor-associated carbohydrates displayed on VLPs through triazole linkage compared to amide linkage¹⁸⁸. Also, little is known about the long-term cellular fate of unusual cross-links from click biology (for example, lysine–aspartate isopeptides from SpyTag–SpyCatcher or thioethers from SNAP-tag). Even for click chemistry, many aspects are unexplored. For example, terminal alkynes may be converted by cytochrome P450 to ketenes, which can lead to enzyme inactivation¹⁸⁹. In this newly emerging field, there is still much to learn in pharmacodynamics and pharmacokinetics. Simplicity is the ultimate sophistication and when click biology reaches a high-enough standard, the chemistry and biological application will be so straightforward, fast and selective that no one will even think about the underlying engineering anymore.

References

- Kolb, H. C., Finn, M. G. & Sharpless, K. B. Click chemistry: diverse chemical function from a few good reactions. *Angew. Chem. Int. Ed. Engl.* **40**, 2004–2021 (2001).
The pioneering treatise setting out philosophy, practice and potential of click chemistry.
- Patterson, D. M., Nazarova, L. A. & Prescher, J. A. Finding the right (bioorthogonal) chemistry. *ACS Chem. Biol.* **9**, 592–605 (2014).
- Bertozzi, C. A special virtual issue celebrating the 2022 Nobel Prize in Chemistry for the development of click chemistry and bioorthogonal chemistry. *ACS Cent. Sci.* **9**, 558–559 (2023).
- Oliveira, B. L., Guo, Z. & Bernardes, G. J. L. Inverse electron demand Diels–Alder reactions in chemical biology. *Chem. Soc. Rev.* **46**, 4895–4950 (2017).
- Bohacek, R. S., McMartin, C. & Guida, W. C. The art and practice of structure-based drug design: a molecular modeling perspective. *Med. Res. Rev.* **16**, 3–50 (1996).
- Bauer, D., Cornejo, M. A., Hoang, T. T., Lewis, J. S. & Zeglis, B. M. Click chemistry and radiochemistry: an update. *Bioconjug. Chem.* **34**, 1925–1950 (2023).
- Srinivasan, S. et al. SQ3370, the first clinical click chemistry-activated cancer therapeutic, shows safety in humans and translatability across species. Preprint at *bioRxiv* <https://doi.org/10.1101/2023.03.28.534654> (2023).
- Sun, F. & Zhang, W.-B. Genetically encoded click chemistry. *Chin. J. Chem.* **38**, 894–896 (2020).
- Kraut, D. A., Carroll, K. S. & Herschlag, D. Challenges in enzyme mechanism and energetics. *Annu. Rev. Biochem.* **72**, 517–571 (2003).
- Shah, N. H. & Muir, T. W. Inteins: nature's gift to protein chemists. *Chem. Sci.* **5**, 446–461 (2014).
- Wu, H., Hu, Z. & Liu, X.-Q. Protein trans-splicing by a split intein encoded in a split DnaE gene of *Synechocystis* sp. PCC6803. *Proc. Natl Acad. Sci. USA* **95**, 9226–9231 (1998).
The discovery of how a DNA polymerase component from a cyanobacterium is expressed in two fragments, which reconstitute and splice together the functional protein, leading on to the future identification of many split inteins and powerful tools for click biology.
- Eryilmaz, E., Shah, N. H., Muir, T. W. & Cowburn, D. Structural and dynamical features of inteins and implications on protein splicing. *J. Biol. Chem.* **289**, 14506–14511 (2014).
- Mills, K. V., Johnson, M. A. & Perler, F. B. Protein splicing: how inteins escape from precursor proteins. *J. Biol. Chem.* **289**, 14498–14505 (2014).
- Anastassov, S., Filo, M. & Khammash, M. Inteins: a Swiss army knife for synthetic biology. *Biotechnol. Adv.* **73**, 108349 (2024).
- Bhagawati, M. et al. In cellulo protein semi-synthesis from endogenous and exogenous fragments using the ultra-fast split Gp41-1 intein. *Angew. Chem. Int. Ed. Engl.* **59**, 21007–21015 (2020).
- Pinto, F., Thornton, E. L. & Wang, B. An expanded library of orthogonal split inteins enables modular multi-peptide assemblies. *Nat. Commun.* **11**, 1529 (2020).
- Carvajal-Vallejos, P., Pallisse, R., Mootz, H. D. & Schmidt, S. R. Unprecedented rates and efficiencies revealed for new natural split inteins from metagenomic sources. *J. Biol. Chem.* **287**, 28686–28696 (2012).
- Gramespacher, J. A., Stevens, A. J., Thompson, R. E. & Muir, T. W. Improved protein splicing using embedded split inteins. *Protein Sci.* **27**, 614–619 (2018).
- Bhagawati, M. et al. A mesophilic cysteine-less split intein for protein trans-splicing applications under oxidizing conditions. *Proc. Natl Acad. Sci. USA* **116**, 22164–22172 (2019).
This study overcomes the obstacle of many split inteins requiring reducing conditions, optimizing a bacteriophage split intein that contains key reactive serines instead of cysteines, allowing molecular surgery at the cell surface.
- Burton, A. J., Haugbro, M., Parisi, E. & Muir, T. W. Live-cell protein engineering with an ultra-short split intein. *Proc. Natl Acad. Sci. USA* **117**, 12041–12049 (2020).
Lake Vida has ice all year round and contains water with exceptional salinity. From this frigid landscape, a split intein platform was shown to react with outstanding speed and efficiency, allowing precision editing on histones within living cells.
- Neugebauer, M., Böcker, J. K., Matern, J. C. J., Pietrokovski, S. & Mootz, H. D. Development of a screening system for inteins active in protein splicing based on intein insertion into the LacZ α -peptide. *Biol. Chem.* **398**, 57–67 (2017).
- Popa, M. P., McKelvey, T. A., Hempel, J. & Hendrix, R. W. Bacteriophage HK97 structure: wholesale covalent cross-linking between the major head shell subunits. *J. Virol.* **65**, 3227–3237 (1991).
- Wikoff, W. R. et al. Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* **289**, 2129–2133 (2000).
- Kang, H. J. & Baker, E. N. Intramolecular isopeptide bonds: protein crosslinks built for stress? *Trends Biochem. Sci.* **36**, 229–237 (2011).
- Zakeri, B. & Howarth, M. Spontaneous intermolecular amide bond formation between side chains for irreversible peptide targeting. *J. Am. Chem. Soc.* **132**, 4526–4527 (2010).
- Zakeri, B. et al. Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc. Natl Acad. Sci. USA* **109**, E690–E697 (2012).
- Keeble, A. H. et al. Evolving accelerated amidation by SpyTag/SpyCatcher to analyze membrane dynamics. *Angew. Chem. Int. Ed. Engl.* **56**, 16521–16525 (2017).
- Keeble, A. H. et al. Approaching infinite affinity through engineering of peptide–protein interaction. *Proc. Natl Acad. Sci. USA* **116**, 26523–26533 (2019).
- Zhang, S. et al. One-step construction of circularized nanodiscs using SpyCatcher–SpyTag. *Nat. Commun.* **12**, 5451 (2021).
- Veggiani, G. et al. Programmable polyproteins built using twin peptide superglues. *Proc. Natl Acad. Sci. USA* **113**, 1202–1207 (2016).
- Keeble, A. H. et al. DogCatcher allows loop-friendly protein–protein ligation. *Cell Chem. Biol.* **29**, 339–350 (2022).
- Fan, R. & Aranko, A. S. Catcher/Tag toolbox: biomolecular click-reactions for protein engineering beyond genetics. *ChemBioChem* **25**, e202300600 (2024).

33. Dicks, M. D. J. et al. Modular capsid decoration boosts adenovirus vaccine-induced humoral immunity against SARS-CoV-2. *Mol. Ther.* **30**, 3639–3657 (2022).
34. Fantoni, N. Z., El-Sagheer, A. H. & Brown, T. A hitchhiker's guide to click-chemistry with nucleic acids. *Chem. Rev.* **121**, 7122–7154 (2021).
35. Driscoll, C. L., Keeble, A. H. & Howarth, M. R. SpyMask enables combinatorial assembly of bispecific binders. *Nat. Commun.* **15**, 2403 (2024).
36. Los, G. V. et al. HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem. Biol.* **3**, 373–382 (2008).
- A tour de force in enzymology converts a bacterial dehalogenase into a self-labeling enzyme, including extensive semi-rational optimization and demonstration of the power of the covalent labeling across a range of methods.**
37. Encell, L. P. et al. Development of a dehalogenase-based protein fusion tag capable of rapid, selective and covalent attachment to customizable ligands. *Curr. Chem. Genomics* **6**, 55–71 (2012).
38. Deprey, K. & Kritzer, J. A. HaloTag forms an intramolecular disulfide. *Bioconjug. Chem.* **32**, 964–970 (2021).
39. Frei, M. S. et al. Engineered HaloTag variants for fluorescence lifetime multiplexing. *Nat. Methods* **19**, 65–70 (2022).
40. Chen, W., Younis, M. H., Zhao, Z. & Cai, W. Recent biomedical advances enabled by HaloTag technology. *Biocell* **46**, 1789–1801 (2022).
41. Liu, D. S., Phipps, W. S., Loh, K. H., Howarth, M. & Ting, A. Y. Quantum dot targeting with lipoic acid ligase and HaloTag for single-molecule imaging on living cells. *ACS Nano* **6**, 11080–11087 (2012).
42. Neklesa, T. K. et al. Small-molecule hydrophobic tagging-induced degradation of HaloTag fusion proteins. *Nat. Chem. Biol.* **7**, 538–543 (2011).
43. Liu, X., Long, M. J. C. & Aye, Y. Proteomics and beyond: cell decision-making shaped by reactive electrophiles. *Trends Biochem. Sci.* **44**, 75–89 (2019).
44. Keppler, A. et al. A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol.* **21**, 86–89 (2003).
- Conversion of a DNA repair enzyme into a powerful approach for cellular labeling: the establishment of SNAP-tag chemistry.**
45. Dreyer, R., Pfkwa, R., Barth, S., Hunter, R. & Klumperman, B. The evolution of SNAP-tag labels. *Biomacromolecules* **24**, 517–530 (2023).
46. Sun, X. et al. Development of SNAP-tag fluorogenic probes for wash-free fluorescence imaging. *ChemBioChem* **12**, 2217–2226 (2011).
47. Gautier, A. et al. An engineered protein tag for multiprotein labeling in living cells. *Chem. Biol.* **15**, 128–136 (2008).
48. Zimmermann, M. et al. Cell-permeant and photocleavable chemical inducer of dimerization. *Angew. Chem. Int. Ed. Engl.* **53**, 4717–4720 (2014).
49. Nakamura, H., Yoshikawa, M., Oda-Ueda, N., Ueda, T. & Ohkuri, T. A comprehensive analysis of novel disulfide bond introduction site into the constant domain of human Fab. *Sci. Rep.* **11**, 12937 (2021).
50. Rossi, E. A., Goldenberg, D. M. & Chang, C. H. Complex and defined biostructures with the dock-and-lock method. *Trends Pharmacol. Sci.* **33**, 474–481 (2012).
51. Liu, X. et al. Structure-based reactivity profiles of reactive metabolites with glutathione. *Chem. Res. Toxicol.* **33**, 1579–1593 (2020).
52. Chiu, J. & Hogg, P. J. Allosteric disulfides: sophisticated molecular structures enabling flexible protein regulation. *J. Biol. Chem.* **294**, 2949–2960 (2019).
53. Garrido Ruiz, D., Sandoval-Perez, A., Rangarajan, A. V., Gunderson, E. L. & Jacobson, M. P. Cysteine oxidation in proteins: structure, biophysics, and simulation. *Biochemistry* **61**, 2165–2176 (2022).
54. Dent, M. R. & DeMartino, A. W. Nitric oxide and thiols: chemical biology, signalling paradigms and vascular therapeutic potential. *Br. J. Pharmacol.* <https://doi.org/10.1111/bph.16274> (2023).
55. Griffin, B. A., Adams, S. R. & Tsien, R. Y. How FLAsH got its sparkle: historical recollections of the biarsenical-tetracysteine tag. *Methods Mol. Biol.* **1266**, 1–6 (2015).
56. Mo, J. et al. Third-generation covalent TMP-tag for fast labeling and multiplexed imaging of cellular proteins. *Angew. Chem. Int. Ed. Engl.* **61**, e202207905 (2022).
57. Liu, W., Samanta, S. K., Smith, B. D. & Isaacs, L. Synthetic mimics of biotin/(strept)avidin. *Chem. Soc. Rev.* **46**, 2391–2403 (2017).
58. Green, N. M. Avidin and streptavidin. *Methods Enzymol.* **184**, 51–67 (1990).
59. Beckett, D., Kovaleva, E. & Schatz, P. J. A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Sci.* **8**, 921–929 (1999).
60. Howarth, M., Takao, K., Hayashi, Y. & Ting, A. Y. Targeting quantum dots to surface proteins in living cells with biotin ligase. *Proc. Natl Acad. Sci. USA* **102**, 7583–7588 (2005).
61. Chivers, C. E. et al. A streptavidin variant with slower biotin dissociation and increased mechanostability. *Nat. Methods* **7**, 391–393 (2010).
62. Christeller, J. T., Markwick, N. P., Burgess, E. P. J. & Malone, L. A. The use of biotin-binding proteins for insect control. *J. Econ. Entomol.* **103**, 497–508 (2010).
63. Howarth, M. et al. A monovalent streptavidin with a single femtomolar biotin binding site. *Nat. Methods* **3**, 267–273 (2006).
64. Demonte, D., Drake, E. J., Lim, K. H., Gulick, A. M. & Park, S. Structure-based engineering of streptavidin monomer with a reduced biotin dissociation rate. *Proteins* **81**, 1621–1633 (2013).
65. Fairhead, M. et al. SpyAvidin hubs enable precise and ultrastable orthogonal nanoassembly. *J. Am. Chem. Soc.* **136**, 12355–12363 (2014).
66. Narayanan, K. B. & Han, S. S. Peptide ligases: a novel and potential enzyme toolbox for catalytic cross-linking of protein/peptide-based biomaterial scaffolds for tissue engineering. *Enzyme Microb. Technol.* **155**, 109990 (2022).
67. Pishesha, N., Ingram, J. R. & Ploegh, H. L. Sortase A: a model for transpeptidation and its biological applications. *Annu. Rev. Cell Dev. Biol.* **34**, 163–188 (2018).
68. Yin, J. et al. Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. *Proc. Natl Acad. Sci. USA* **102**, 15815–15820 (2005).
69. Wu, P. et al. Site-specific chemical modification of recombinant proteins produced in mammalian cells by using the genetically encoded aldehyde tag. *Proc. Natl Acad. Sci. USA* **106**, 3000–3005 (2009).
70. Agarwal, P., van der Weijden, J., Sletten, E. M., Rabuka, D. & Bertozzi, C. R. A Pictet–Spengler ligation for protein chemical modification. *Proc. Natl Acad. Sci. USA* **110**, 46–51 (2013).
71. Sun, B. & Kumar, S. Protein adsorption loss—the bottleneck of single-cell proteomics. *J. Proteome Res.* **21**, 1808–1815 (2022).
72. Guo, J. et al. The development of proximity labeling technology and its applications in mammals, plants, and microorganisms. *Cell Commun. Signal.* **21**, 269 (2023).
73. Seckute, J. & Devaraj, N. K. Expanding room for tetrazine ligations in the in vivo chemistry toolbox. *Curr. Opin. Chem. Biol.* **17**, 761–767 (2013).
74. Uematsu, M. & Baskin, J. M. Chemical approaches for measuring and manipulating lipids at the organelle level. *Cold Spring Harb. Perspect. Biol.* **15**, a041407 (2023).
75. Johnson, J. A., Lu, Y. Y., Van Deventer, J. A. & Tirrell, D. A. Residue-specific incorporation of non-canonical amino acids into proteins: recent developments and applications. *Curr. Opin. Chem. Biol.* **14**, 774–780 (2010).

76. Dube, D. H. & Bertozzi, C. R. Metabolic oligosaccharide engineering as a tool for glycobiology. *Curr. Opin. Chem. Biol.* **7**, 616–625 (2003).
77. Salic, A. & Mitchison, T. J. A chemical method for fast and sensitive detection of DNA synthesis in vivo. *Proc. Natl Acad. Sci. USA* **105**, 2415–2420 (2008).
78. de la Torre, D. & Chin, J. W. Reprogramming the genetic code. *Nat. Rev. Genet.* **22**, 169–184 (2021).
79. Hodneland, C. D., Lee, Y.-S., Min, D.-H. & Mrksich, M. Selective immobilization of proteins to self-assembled monolayers presenting active site-directed capture ligands. *Proc. Natl Acad. Sci. USA* **99**, 5048–5052 (2002).
80. Mizukami, S., Hori, Y. & Kikuchi, K. Small-molecule-based protein-labeling technology in live cell studies: probe-design concepts and applications. *Acc. Chem. Res.* **47**, 247–256 (2014).
81. Minoshima, M. et al. Development of a versatile protein labeling tool for live-cell imaging using fluorescent β -lactamase inhibitors. *Angew. Chem. Int. Ed. Engl.* **62**, e202301704 (2023).
82. Whitesides, G. M. Cool, or simple and cheap? Why not both? *Lab Chip* **13**, 11–13 (2013).
- An important philosophical perspective on the value of simplicity, in contrast to the frequent academic tendency to promote expensive and intricate solutions.**
83. Keeble, A. H. & Howarth, M. Insider information on successful covalent protein coupling with help from SpyBank. *Methods Enzymol.* **617**, 443–461 (2019).
84. McLachlan, G. et al. Progress in respiratory gene therapy. *Hum. Gene Ther.* **33**, 893–912 (2022).
85. Borkowski, O. et al. Cell-free prediction of protein expression costs for growing cells. *Nat. Commun.* **9**, 1457 (2018).
86. Nikić, I., Kang, J. H., Girona, G. E., Aramburu, I. V. & Lemke, E. A. Labeling proteins on live mammalian cells using click chemistry. *Nat. Protoc.* **10**, 780–791 (2015).
87. Zimmerman, J. B., Anastas, P. T., Erythropel, H. C. & Leitner, W. Designing for a green chemistry future. *Science* **367**, 397–400 (2020).
88. Paloheimo, M., Haarmann, T., Mäkinen, S. & Vehmaanperä, J. Production of industrial enzymes in *Trichoderma reesei*. In *Gene Expression Systems in Fungi: Advancements and Applications* (eds Schmoll, M. & Dattenböck, C.) 23–57 (Springer Nature, 2016).
89. Tyndall, S. M., Maloney, G. R., Cole, M. B., Hazell, N. G. & Augustin, M. A. Critical food and nutrition science challenges for plant-based meat alternative products. *Crit. Rev. Food Sci. Nutr.* **64**, 638–653 (2024).
90. Way, J. C., Collins, J. J., Keasling, J. D. & Silver, P. A. Integrating biological redesign: where synthetic biology came from and where it needs to go. *Cell* **157**, 151–161 (2014).
91. Keeble, A. H. & Howarth, M. Power to the protein: enhancing and combining activities using the Spy toolbox. *Chem. Sci.* **11**, 7281–7291 (2020).
92. Rahikainen, R. et al. Overcoming symmetry mismatch in vaccine nanoassembly through spontaneous amidation. *Angew. Chem. Int. Ed. Engl.* **60**, 321–330 (2021).
93. Balchin, D., Hayer-Hartl, M. & Hartl, F. U. Recent advances in understanding catalysis of protein folding by molecular chaperones. *FEBS Lett.* **594**, 2770–2781 (2020).
94. Radanović, T. & Ernst, R. The unfolded protein response as a guardian of the secretory pathway. *Cells* **10**, 2965 (2021).
95. Jain, T. et al. Biophysical properties of the clinical-stage antibody landscape. *Proc. Natl Acad. Sci. USA* **114**, 944–949 (2017).
96. Houk, K. N., Leach, A. G., Kim, S. P. & Zhang, X. Binding affinities of host–guest, protein–ligand, and protein–transition-state complexes. *Angew. Chem. Int. Ed. Engl.* **42**, 4872–4897 (2003).
97. Yen, C. et al. The development of global vaccine stockpiles. *Lancet Infect. Dis.* **15**, 340–347 (2015).
98. Minutolo, N. G. et al. Quantitative control of gene-engineered T-cell activity through the covalent attachment of targeting ligands to a universal immune receptor. *J. Am. Chem. Soc.* **142**, 6554–6568 (2020).
99. Brune, K. D. & Howarth, M. New routes and opportunities for modular construction of particulate vaccines: stick, click, and glue. *Front. Immunol.* **9**, 1432 (2018).
100. Merrill, R. A. et al. A robust and economical pulse-chase protocol to measure the turnover of HaloTag fusion proteins. *J. Biol. Chem.* **294**, 16164–16171 (2019).
101. Lin, D. et al. Time-tagged ticker tapes for intracellular recordings. *Nat. Biotechnol.* **41**, 631–639 (2023).
102. Donovan, D. A. et al. Engineered chromatin remodeling proteins for precise nucleosome positioning. *Cell Rep.* **29**, 2520–2535 (2019).
103. Vester, S. K. et al. SpySwitch enables pH- or heat-responsive capture and release for plug-and-display nanoassembly. *Nat. Commun.* **13**, 3714 (2022).
104. Kompa, J. et al. Exchangeable HaloTag ligands for super-resolution fluorescence microscopy. *J. Am. Chem. Soc.* **145**, 3075–3083 (2023).
105. Moran, U., Phillips, R. & Milo, R. SnapShot: key numbers in biology. *Cell* **141**, 1262 (2010).
106. Wilhelm, J. et al. Kinetic and structural characterization of the self-labeling protein tags HaloTag7, SNAP-tag, and CLIP-tag. *Biochemistry* **60**, 2560–2575 (2021).
107. Krämer, K. How click conquered chemistry. *Chem World* www.chemistryworld.com/features/how-click-conquered-chemistry/4016366.article (2022).
108. van den Bosch, S. M. et al. Evaluation of strained alkynes for Cu-free click reaction in live mice. *Nucl. Med. Biol.* **40**, 415–423 (2013).
109. Sasmal, P. K. et al. Catalytic azide reduction in biological environments. *ChemBioChem* **13**, 1116–1120 (2012).
110. Versteegen, R. M., Rossin, R., ten Hoeve, W., Janssen, H. M. & Robillard, M. S. Click to release: instantaneous doxorubicin elimination upon tetrazine ligation. *Angew. Chem. Int. Ed. Engl.* **52**, 14112–14116 (2013).
111. Darko, A. et al. Conformationally strained *trans*-cyclooctene with improved stability and excellent reactivity in tetrazine ligation. *Chem. Sci.* **5**, 3770–3776 (2014).
112. Swift, J. L., Heuff, R. & Cramb, D. T. A two-photon excitation fluorescence cross-correlation assay for a model ligand–receptor binding system using quantum dots. *Biophys. J.* **90**, 1396–1410 (2006).
113. Pabis, A., Rizzo, V. A., Sanchez-Ruiz, J. M. & Kamerlin, S. C. Cooperativity and flexibility in enzyme evolution. *Curr. Opin. Struct. Biol.* **48**, 83–92 (2018).
114. Clarke, A. et al. A low temperature limit for life on earth. *PLoS ONE* **8**, e66207 (2013).
115. Kashafi, K. & Lovley, D. R. Extending the upper temperature limit for life. *Science* **301**, 934 (2003).
116. Linser, P. J., Smith, K. E., Seron, T. J. & Neira Oviedo, M. Carbonic anhydrases and anion transport in mosquito midgut pH regulation. *J. Exp. Biol.* **212**, 1662–1671 (2009).
117. Choi, J. J. et al. Protein *trans*-splicing and characterization of a split family B-type DNA polymerase from the hyperthermophilic archaeal parasite *Nanoarchaeum equitans*. *J. Mol. Biol.* **356**, 1093–1106 (2006).
118. Perugino, G. et al. Activity and regulation of archaeal DNA alkyltransferase. *J. Biol. Chem.* **287**, 4222–4231 (2012).
119. Wang, Y., Shi, Y., Hellinga, H. W. & Beese, L. S. Thermally controlled intein splicing of engineered DNA polymerases provides a robust and generalizable solution for accurate and sensitive molecular diagnostics. *Nucleic Acids Res.* **51**, 5883–5894 (2023).

120. Courchesne, N. M. D., Duraj-Thatte, A., Tay, P. K. R., Nguyen, P. Q. & Joshi, N. S. Scalable production of genetically engineered nanofibrous macroscopic materials via filtration. *ACS Biomater. Sci. Eng.* **3**, 733–741 (2017).
121. Schoene, C., Bennett, S. P. & Howarth, M. SpyRing interrogation: analyzing how enzyme resilience can be achieved with phytase and distinct cyclization chemistries. *Sci. Rep.* **6**, 21151 (2016).
122. Mouysset, B., Le Grand, M., Camoin, L. & Pasquier, E. Poly-pharmacology of existing drugs: how to crack the code? *Cancer Lett.* **588**, 216800 (2024).
123. England, C. G., Luo, H. & Cai, W. HaloTag technology: a versatile platform for biomedical applications. *Bioconjug. Chem.* **26**, 975–986 (2015).
124. Liu, X. et al. Split chimeric antigen receptor-modified T cells targeting glypican-3 suppress hepatocellular carcinoma growth with reduced cytokine release. *Ther. Adv. Med. Oncol.* **12**, 1758835920910347 (2020).
125. Roy, A. & Gauld, J. W. Sulfilimine bond formation in collagen IV. *Chem. Commun.* **60**, 646–657 (2024).
126. Alibardi, L. Vertebrate keratinization evolved into cornification mainly due to transglutaminase and sulfhydryl oxidase activities on epidermal proteins: an immunohistochemical survey. *Anat. Rec.* **305**, 333–358 (2022).
127. Rajendran, A. K. et al. Trends in mechanobiology guided tissue engineering and tools to study cell–substrate interactions: a brief review. *Biomater. Res.* **27**, 55 (2023).
128. Otomo, T., Teruya, K., Uegaki, K., Yamazaki, T. & Kyogoku, Y. Improved segmental isotope labeling of proteins and application to a larger protein. *J. Biomol. NMR* **14**, 105–114 (1999).
129. Modica, J. A., Skarpathiotis, S. & Mrksich, M. Modular assembly of protein building blocks to create precisely defined megamolecules. *ChemBioChem* **13**, 2331–2334 (2012).
Demonstration of the power of high-yield irreversible covalent coupling with HaloTag and cutinase for generating novel classes of protein architecture.
130. Fryer, T. et al. Gigavalent display of proteins on monodisperse polyacrylamide hydrogels as a versatile modular platform for functional assays and protein engineering. *ACS Cent. Sci.* **8**, 1182–1195 (2022).
131. Khairil Anuar, I. N. A. et al. Spy&Go purification of SpyTag-proteins using pseudo-SpyCatcher to access an oligomerization toolbox. *Nat. Commun.* **10**, 1734 (2019).
132. Sun, F., Zhang, W. B., Mahdavi, A., Arnold, F. H. & Tirrell, D. A. Synthesis of bioactive protein hydrogels by genetically encoded SpyTag–SpyCatcher chemistry. *Proc. Natl Acad. Sci. USA* **111**, 11269–11274 (2014).
133. Wang, R., Yang, Z., Luo, J., Hsing, I. M. & Sun, F. B₁₂-dependent photoresponsive protein hydrogels for controlled stem cell/protein release. *Proc. Natl Acad. Sci. USA* **114**, 5912–5917 (2017).
134. Scott, C. P., Abel-Santos, E., Jones, A. D. & Benkovic, S. J. Structural requirements for the biosynthesis of backbone cyclic peptide libraries. *Chem. Biol.* **8**, 801–815 (2001).
Given the power of cyclization to reduce peptide flexibility and degradation, showing how split intein cyclization allows selection of potential drug leads.
135. Sohrabi, C., Foster, A. & Tavassoli, A. Methods for generating and screening libraries of genetically encoded cyclic peptides in drug discovery. *Nat. Rev. Chem.* **4**, 90–101 (2020).
136. Schoene, C., Bennett, S. P. & Howarth, M. SpyRings declassified: a blueprint for using isopeptide-mediated cyclization to enhance enzyme thermal resilience. *Methods Enzymol.* **580**, 149–167 (2016).
137. Zhang, W. B., Sun, F., Tirrell, D. A. & Arnold, F. H. Controlling macromolecular topology with genetically encoded SpyTag–SpyCatcher chemistry. *J. Am. Chem. Soc.* **135**, 13988–13997 (2013).
Establishing a simple route to prepare a panel of programmed nonlinear protein topologies through efficient isopeptide bond formation.
138. Li, T., Zhang, F., Fang, J., Liu, Y. & Zhang, W.-B. Rational design and cellular synthesis of proteins with unconventional chemical topology. *Chin. J. Chem.* **41**, 2873–2880 (2023).
139. Zhang, F., Liu, Y., Da, X.-D. & Zhang, W.-B. Toward selective synthesis of protein olympiades via orthogonal active templates in one step. *CCS Chem.* **6**, 1047–1059 (2023).
Integrating supramolecular chemistry with protein engineering, this paper harnesses multiple genetically directed couplings to achieve intricate mechanically interlocked architectures.
140. Hong, H. et al. HaloTag: a novel reporter gene for positron emission tomography. *Am. J. Transl. Res.* **3**, 392–403 (2011).
141. Strauch, R. C. et al. Reporter protein-targeted probes for magnetic resonance imaging. *J. Am. Chem. Soc.* **133**, 16346–16349 (2011).
142. Li, Y. et al. Conformable self-assembling amyloid protein coatings with genetically programmable functionality. *Sci. Adv.* **6**, eaba1425 (2020).
143. Williams, D. M. et al. Facile protein immobilization using engineered surface-active biofilm proteins. *ACS Appl. Nano Mater.* **1**, 2483–2488 (2018).
144. Ma, W. et al. Modular assembly of proteins on nanoparticles. *Nat. Commun.* **9**, 1489 (2018).
145. Wang, F. et al. General and robust covalently linked graphene oxide affinity grids for high-resolution cryo-EM. *Proc. Natl Acad. Sci. USA* **117**, 24269–24273 (2020).
146. Ramlaul, K. et al. A 3D-printed flow-cell for on-grid purification of electron microscopy samples directly from lysate. *J. Struct. Biol.* **215**, 107999 (2023).
147. Patel, A. et al. Using CombiCells, a platform for titration and combinatorial display of cell surface ligands, to study T-cell antigen sensitivity modulation by accessory receptors. *EMBO J.* **43**, 132–150 (2024).
148. Wang, X.-J. et al. Gene-splitting technology: a novel approach for the containment of transgene flow in *Nicotiana tabacum*. *PLoS ONE* **9**, e99651 (2014).
149. Free, K., Nakanishi, H. & Itaka, K. Development of synthetic mRNAs encoding split cytotoxic proteins for selective cell elimination based on specific protein detection. *Pharmaceutics* **15**, 213 (2023).
150. Gramespacher, J. A., Stevens, A. J., Nguyen, D. P., Chin, J. W. & Muir, T. W. Intein zymogens: conditional assembly and splicing of split inteins via targeted proteolysis. *J. Am. Chem. Soc.* **139**, 8074–8077 (2017).
151. Zhao, Y. et al. SpyCLIP: an easy-to-use and high-throughput compatible CLIP platform for the characterization of protein–RNA interactions with high accuracy. *Nucleic Acids Res.* **47**, e33 (2019).
152. Gu, J. et al. GoldCLIP: gel-omitted ligation-dependent CLIP. *Genomics Proteomics Bioinformatics* **16**, 136–143 (2018).
153. Taniguchi, Y. & Kawakami, M. Application of HaloTag protein to covalent immobilization of recombinant proteins for single molecule force spectroscopy. *Langmuir* **26**, 10433–10436 (2010).
154. Rivas-Pardo, J. A. et al. A HaloTag-TEV genetic cassette for mechanical phenotyping of proteins from tissues. *Nat. Commun.* **11**, 2060 (2020).
155. Napierski, N. C. et al. A novel ‘cut and paste’ method for in situ replacement of cMyBP-C reveals a new role for cMyBP-C in the regulation of contractile oscillations. *Circ. Res.* **126**, 737–749 (2020).
156. Tornabene, P. et al. Intein-mediated protein *trans*-splicing expands adeno-associated virus transfer capacity in the retina. *Sci. Transl. Med.* **11**, eaav4523 (2019).

157. Muik, A. et al. Covalent coupling of high-affinity ligands to the surface of viral vector particles by protein *trans*-splicing mediates cell type-specific gene transfer. *Biomaterials* **144**, 84–94 (2017).
158. Sabin, L. et al. Viral particles retargeted to skeletal muscle. US Patent WO/2023/081850 (2023).
159. Bachmann, M. F. & Jennings, G. T. Vaccine delivery: a matter of size, geometry, kinetics and molecular patterns. *Nat. Rev. Immunol.* **10**, 787–796 (2010).
160. Hills, R. A. & Howarth, M. Virus-like particles against infectious disease and cancer: guidance for the nano-architect. *Curr. Opin. Biotechnol.* **73**, 346–354 (2022).
161. Wang, Q. et al. Bioconjugation by copper(I)-catalyzed azide–alkyne [3+2] cycloaddition. *J. Am. Chem. Soc.* **125**, 3192–3193 (2003).
162. Li, X. & Xiong, Y. Application of ‘click’ chemistry in biomedical hydrogels. *ACS Omega* **7**, 36918–36928 (2022).
163. van den Berg van Saparoea, H. B., Houben, D., de Jonge, M. I., Jong, W. S. P. & Luirink, J. Display of recombinant proteins on bacterial outer membrane vesicles by using protein ligation. *Appl. Environ. Microbiol.* **84**, e02567-17 (2018).
164. Brune, K. D. et al. Plug-and-Display: decoration of virus-like particles via isopeptide bonds for modular immunization. *Sci. Rep.* **6**, 19234 (2016).
165. Thrane, S. et al. Bacterial superglue enables easy development of efficient virus-like particle based vaccines. *J. Nanobiotechnology* **14**, 30 (2016).
166. King, L. D. W. et al. Preclinical development of a stabilized RH5 virus-like particle vaccine that induces improved anti-malarial antibodies. *Cell Rep. Med.* **5**, 101654 (2024).
167. Smit, M. J. et al. First-in-human use of a modular capsid virus-like vaccine platform: an open-label, non-randomised, phase 1 clinical trial of the SARS-CoV-2 vaccine ABNCoV2. *Lancet Microbe* **4**, e140–e148 (2023).
- Entry of click biology into human clinical trials through modular decoration of VLPs.**
168. Baker, D. What has de novo protein design taught us about protein folding and biophysics? *Protein Sci.* **28**, 678–683 (2019).
169. Du, J., Kong, Y., Wen, Y., Shen, E. & Xing, H. HUH endonuclease: a sequence-specific fusion protein tag for precise DNA–protein conjugation. *Bioorg. Chem.* **144**, 107118 (2024).
170. Sacca, B. et al. Orthogonal protein decoration of DNA origami. *Angew. Chem. Int. Ed. Engl.* **49**, 9378–9383 (2010).
171. Komiya, E. et al. Exploration and application of DNA-binding proteins to make a versatile DNA–protein covalent-linking patch (D-Pclip): the case of a biosensing element. *J. Am. Chem. Soc.* **146**, 4087–4097 (2024).
172. Mashimo, Y., Maeda, H., Mie, M. & Kobatake, E. Construction of semisynthetic DNA–protein conjugates with Phi X174 Gene-A* protein. *Bioconjug. Chem.* **23**, 1349–1355 (2012).
- Pioneering the site-specific covalent linkage of protein to DNA through engineering of a phage ligation system.**
173. Bernardinelli, G. & Högberg, B. Entirely enzymatic nanofabrication of DNA–protein conjugates. *Nucleic Acids Res.* **45**, e160 (2017).
174. Lovendahl, K. N., Hayward, A. N. & Gordon, W. R. Sequence-directed covalent protein–DNA linkages in a single step using HUH-tags. *J. Am. Chem. Soc.* **139**, 7030–7035 (2017).
- Establishment of a panel of efficient modules, engineered from viruses and bacteria, to empower spontaneous covalent coupling of DNA to protein.**
175. Silverman, S. K. Catalytic DNA: scope, applications, and biochemistry of deoxyribozymes. *Trends Biochem. Sci.* **41**, 595–609 (2016).
176. Sun, P., Gou, H., Che, X., Chen, G. & Feng, C. Recent advances in DNAzymes for bioimaging, biosensing and cancer therapy. *Chem. Commun.* **60**, 10805–10821 (2024).
177. Baum, D. A. & Silverman, S. K. Deoxyribozyme-catalyzed labeling of RNA. *Angew. Chem. Int. Ed. Engl.* **46**, 3502–3504 (2007).
178. Kong, L.-Z. et al. Understanding nucleic acid sensing and its therapeutic applications. *Exp. Mol. Med.* **55**, 2320–2331 (2023).
179. Hoffmann, M. D. et al. Multiparametric domain insertional profiling of adeno-associated virus VP1. *Mol. Ther. Methods Clin. Dev.* **31**, 101143 (2023).
180. Huh, W. K. et al. Global analysis of protein localization in budding yeast. *Nature* **425**, 686–691 (2003).
181. Weill, U. et al. Genome-wide SWAp-Tag yeast libraries for proteome exploration. *Nat. Methods* **15**, 617–622 (2018).
182. Eiamthong, B. et al. Discovery and genetic code expansion of a polyethylene terephthalate (PET) hydrolase from the human saliva metagenome for the degradation and bio-functionalization of PET. *Angew. Chem. Int. Ed. Engl.* **61**, e202203061 (2022).
183. Mican, J. et al. Exploring new galaxies: perspectives on the discovery of novel PET-degrading enzymes. *Appl. Catal. B Environ.* **342**, 123404 (2024).
184. Lin, Z. et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* **379**, 1123–1130 (2023).
185. Lu, H. et al. Machine learning-aided engineering of hydrolases for PET depolymerization. *Nature* **604**, 662–667 (2022).
186. Esvelt, K. M., Carlson, J. C. & Liu, D. R. A system for the continuous directed evolution of biomolecules. *Nature* **472**, 499–503 (2011).
187. Lovelock, S. L. et al. The road to fully programmable protein catalysis. *Nature* **606**, 49–58 (2022).
188. Yin, Z. et al. Significant impact of immunogen design on the diversity of antibodies generated by carbohydrate-based anticancer vaccine. *ACS Chem. Biol.* **10**, 2364–2372 (2015).
189. Ortiz de Montellano, P. R. Acetylenes: cytochrome P450 oxidation and mechanism-based enzyme inactivation. *Drug Metab. Rev.* **51**, 162–177 (2019).

Acknowledgements

Funding was provided by the Engineering and Physical Sciences Research Council (EPSRC EP/T030704/1).

Competing interests

M.R.H. is an inventor on a patent on traptavidin (UK Intellectual Property Office 0919102.4), a SpyBiotech cofounder and shareholder, and an inventor on patents on spontaneous amide bond formation (UK Intellectual Property Office 2117283.8, 2104999.4, 1915905.2, 1903479.2, 1819850.7, 1706430.4, 1705750.6, 1509782.7, 1002362.0).

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41589-025-01944-x>.

Correspondence should be addressed to Mark R. Howarth.

Peer review information *Nature Chemical Biology* thanks Ali Tavassoli and the other, anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

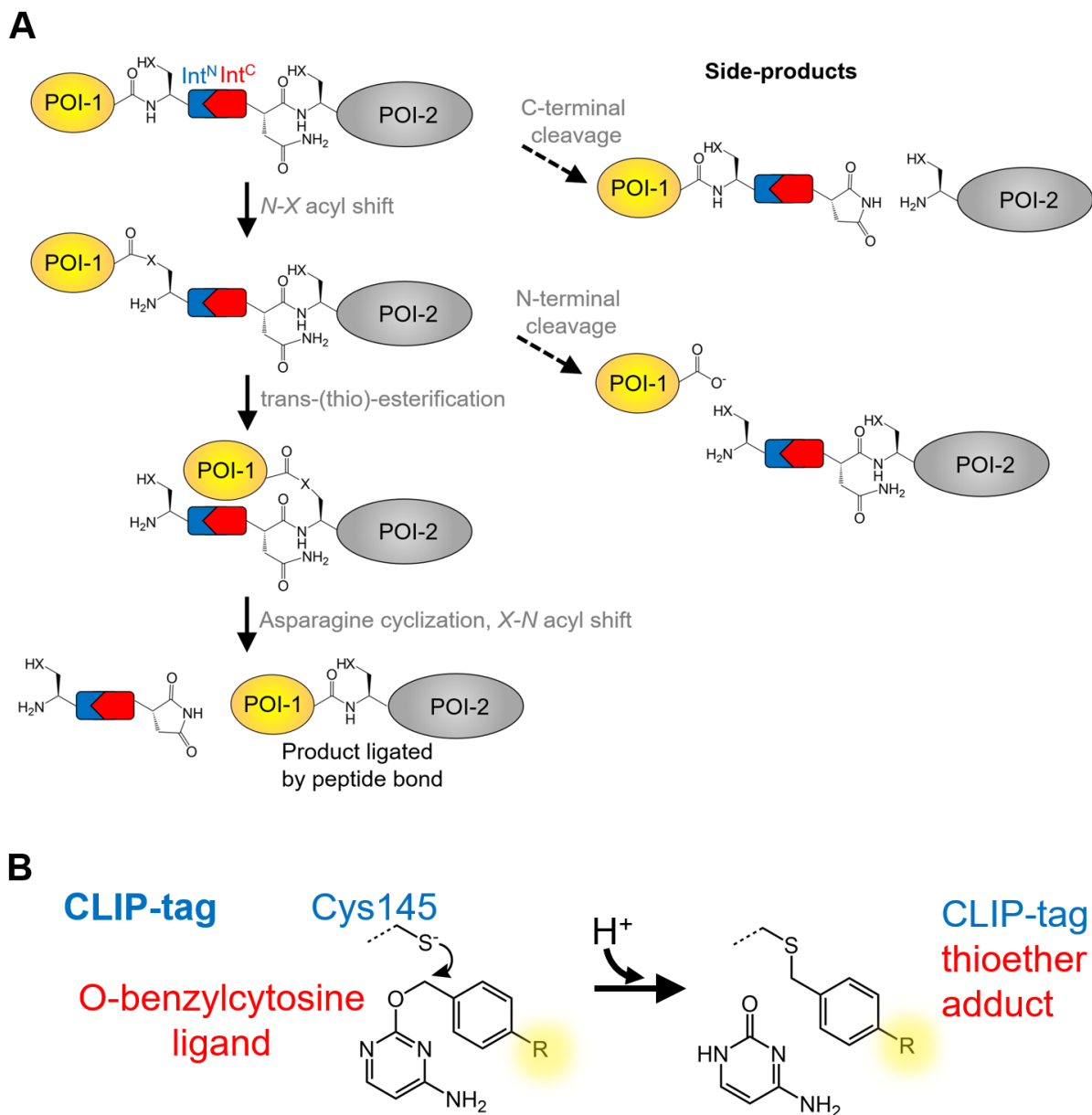
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© Springer Nature America, Inc. 2025

Click biology highlights the opportunities from reliable biological reactions

In the format provided by the authors and unedited



Supplementary Figure 1. (A) Split intein reaction mechanism. POI = Protein of interest. Potential side-reactions are shown on the right. Unintended C-terminal cleavage may occur at various points in the splicing process. (B) CLIP-tag reaction. R- represents the probe of interest, often a fluorophore.

Ligation system	Tag size (amino acid residues)
Split inteins	<i>tags released after reaction</i>
Npu	Int ^N 102, Int ^C 35
Gp41-1	Int ^N 88, Int ^C 37
CL	Int ^N 26, Int ^C 126
VidaL	Int ^N 16, Int ^C 124
Tag/Catchers	
SpyTag003/SpyCatcher003	16 / 113
SnoopTag/SnoopCatcher	12 / 112
DogTag/DogCatcher	23 / 104
SilkTag/SilkCatcher	20 / 79
Self-labeling enzymes	
HaloTag	297
SNAP-tag	182
CLIP-tag	182
TMP-tag	159
Cutinase	214
PYP-tag	125
BL-tag	263
Enzymatic ligations	
Butelase	3
Omniligase	4 (Ct), 2 (Nt)
Subtligase	Most unblocked N-termini
Microbial transglutaminase	1 or 4
ybbR	11
Formylglycine generating enzyme	6 or 13
Sortase	5 (Ct), 1-5 (Nt)
Nucleic acid coupling	
UdgX	209
A*-tag	341
HUH-tag mMobA	186
Other	
Non-canonical amino acid	1
FIAsh	15

Supplementary Table 1: *The size of ligation systems described in this Perspective, in terms of the tag size in amino acids. Ct refers to tags placed at the C-terminus. Nt refers to tags placed at the N-terminus. Spacers or fusions for purification (e.g. His-tags) are excluded from the calculation. Apart from split inteins, tags are retained on protein constructs post-coupling.*