

REVIEW

Open Access



Biomolecular engineering for nanobio/bionanotechnology

Teruyuki Nagamune* 

Abstract

Biomolecular engineering can be used to purposefully manipulate biomolecules, such as peptides, proteins, nucleic acids and lipids, within the framework of the relations among their structures, functions and properties, as well as their applicability to such areas as developing novel biomaterials, biosensing, bioimaging, and clinical diagnostics and therapeutics. Nanotechnology can also be used to design and tune the sizes, shapes, properties and functionality of nanomaterials. As such, there are considerable overlaps between nanotechnology and biomolecular engineering, in that both are concerned with the structure and behavior of materials on the nanometer scale or smaller. Therefore, in combination with nanotechnology, biomolecular engineering is expected to open up new fields of nanobio/bionanotechnology and to contribute to the development of novel nanobiomaterials, nanobiodevices and nanobiosystems. This review highlights recent studies using engineered biological molecules (e.g., oligonucleotides, peptides, proteins, enzymes, polysaccharides, lipids, biological cofactors and ligands) combined with functional nanomaterials in nanobio/bionanotechnology applications, including therapeutics, diagnostics, biosensing, bioanalysis and biocatalysts. Furthermore, this review focuses on five areas of recent advances in biomolecular engineering: (a) nucleic acid engineering, (b) gene engineering, (c) protein engineering, (d) chemical and enzymatic conjugation technologies, and (e) linker engineering. Precisely engineered nanobiomaterials, nanobiodevices and nanobiosystems are anticipated to emerge as next-generation platforms for bioelectronics, biosensors, biocatalysts, molecular imaging modalities, biological actuators, and biomedical applications.

Keywords: Engineered biological molecules, Therapy, Diagnosis, Biosensing, Bioanalysis, Biocatalyst, Nucleic acid engineering, Gene engineering, Protein engineering, Conjugation technologies

1 Introduction

Nanotechnology is the creation and utilization of materials, devices, and systems through controlling matter on the nanometer scale, and it is the key technology of the twenty-first century. The ability to exploit the structures, functions and processes of biological molecules, complexes and nanosystems to produce novel functional nanostructured biological materials has created the rapidly growing fields of nanobiotechnology and bionanotechnology, which are fusion research fields of nanotechnology and biotechnology [1]. Although these words are often used interchangeably, in this review, they are utilized in terminologically different ways, as follows.

Nanobiotechnology is used in relation to the ways in which nanotechnology is used to create materials, devices and systems for studying biological systems and developing new biological assay, diagnostic, therapeutic, information storage and computing systems, among others. These systems use nanotechnology to advance the goals of biological fields. Some nanobiotechnologies scale from the top down, such as from microfluidics to nanofluidic biochips (e.g., lab-on-a-chip for continuous-flow separation and the detection of such macromolecules as DNA and proteins [2], point-of-care biosensors for detecting biomarkers and clinical diagnosis [3–7], and solid-state nanopore sensors for DNA sequencing [8]). Other nanobiotechnologies scale from the bottom up for the fabrication of nanoscale hybrid materials, such as complexes consisting of nanoparticles (NPs) (e.g., magnetic NPs, AuNPs and AgNPs, silica NPs, quantum dots

*Correspondence: nagamune@bio.t.u-tokyo.ac.jp
Department of Chemistry and Biotechnology, Graduate School of Engineering, The University of Tokyo, Tokyo, Japan

(QDs), polymeric micelles, liposomes, dendrimers, and fullerenes) and biological molecules, which are highly useful for biosensing, bioimaging, diagnostic and therapeutic applications in healthcare [9–15].

On the other hand, bionanotechnology refers to the ways in which biotechnology is used to improve existing or create new nanotechnologies through the study of how biological systems work and the applications of biological molecules and systems to nanotechnology. DNA and RNA nanotechnologies, the utilization of the base-pairing and molecular self-assembly properties of nucleic acids to create useful materials, such as DNA origami, DNA nanomachines, DNA scaffolds for electronics, photonics and protein arrays, and DNA and RNA aptamers, ribozymes and riboswitches, are important examples of bionanotechnology [16, 17]. Another important area of research involves taking advantage of the self-assembly properties of peptides, proteins and lipids to generate well-defined 3D structures, functional protein complexes, nanofilms and other nanostructures, such as micelles, reverse micelles and liposomes, which could be used as novel approaches for the large-scale production of programmable nanomaterials [18–20]. The application of carbohydrate polymers combined with nanotechnology in tissue engineering and medicine are also potential research fields for the development of novel biomaterials for biosensing, bioimaging, diagnostic and drug-delivery systems [21].

With either nanobiotechnology or bionanotechnology, biological molecules are indispensable building blocks for fabricating functional nanomaterials, nanodevices and nanosystems. However, from the viewpoint of applying biological materials to nanotechnology, biological materials found in nature always have sufficient functions and properties. Recent advances in biomolecular engineering, such as genetic engineering, DNA and RNA engineering, protein engineering, site-specific chemical and enzymatic conjugation technologies, self-assembly technology and massive high-throughput screening (HTS) methods, have enabled us to improve, stabilize, integrate and alter the functions and properties of biological materials. Thus, it is possible to create engineered biological materials with functions and properties that are optimized for various uses in the fields of bioelectronics, biosensors, biocatalysis, molecular imaging, biological actuators, drug delivery systems, biomaterials for tissue engineering and regenerative medicine.

In this review, recent studies applying engineered biological materials to nanobio/bionanotechnology are discussed, and various biomolecular engineering technologies are highlighted.

2 Application of engineered biological molecules to nanobio/bionanotechnology

Nanobio/bionanotechnology has created new opportunities for advances in diverse fields, including life science, medicine, electronics, engineering, and biotechnology. Nanoscale materials [e.g., NPs, nanowires, nanofibers, and nanotubes (NTs)] combined with various engineered biological molecules (e.g., proteins, enzymes, oligonucleotides, polysaccharides, lipids, biological cofactors and ligands) have been explored in many biological applications (e.g., therapy, diagnosis, bioimaging, biosensing, bioanalysis, biocatalysis, cell and organ chips, bioelectronic devices, and biological separation) (Fig. 1). Their novel and unique properties and functions, such as high volume-to-surface ratio, improved solubility, quantum size, macroscopic quantum tunnel and multifunctionality, result in nanobiomaterials that are drastically different from their corresponding bulk materials.

The current review is focused on advances in the development of nanobiomaterials for applications in therapy, diagnosis, biosensing, bioanalysis and biocatalysis because nanobiomaterials for cell and organ chips [22–25], bioelectronic devices [26, 27] and biological separation [28] have recently been reviewed in this journal.

2.1 Nanobiomaterials for therapy and diagnosis

Smart therapeutic and diagnostic or bioimaging NPs carrying cargo materials, such as drugs, DNAs, RNAs, proteins, and imaging reagents, have been widely developed [11, 13, 29–33]. To achieve intracellular NP and drug delivery, many strategies for overcoming various biological barriers are needed, including the following: (i) preventing removal from the circulation by cells of the reticuloendothelial system; (ii) targeting specific cells; (iii)

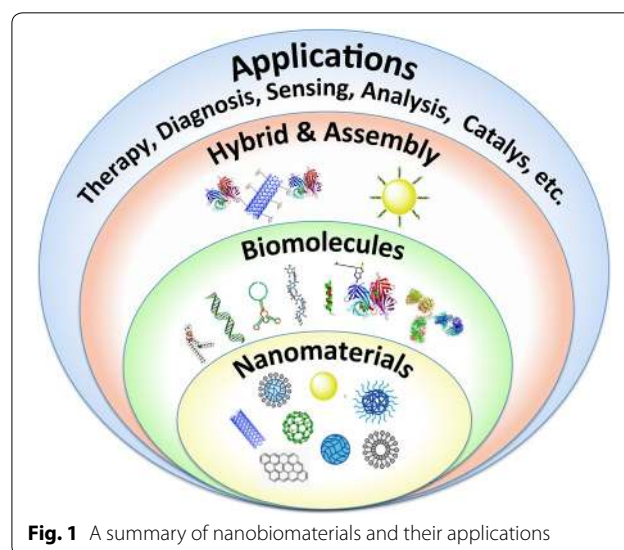


Fig. 1 A summary of nanobiomaterials and their applications

internalization into cells; (iv) escaping from endosomes; (v) trafficking to specific organelles; and (vi) controlling the release of payloads (e.g., drugs, DNAs or RNAs).

2.1.1 Preventing removal from the circulation

NPs made of hydrophobic synthetic polymers, metals or inorganic materials are usually not blood compatible. Their injection into the body can provoke a coagulation response and activate the complement cascade; subsequently, they can be recognized by phagocytes and macrophages, rendering them useless or harmful. The surface modification of NPs with hydrophilic synthetic or biological polymers, such as polyethylene glycol (PEG) [34], heparin [35] or dextran [36], forms a steric brush that imparts resistance to protein adsorption. This type of surface modification shows increased intrinsic anticoagulant and anti-complement properties, as well as other biological activities; in addition, it extends the circulation half-life and reduces the immunogenicity of NPs in the human body. The conformation of polymer chains on the surface also influences the pharmacokinetics and biodistribution of NPs.

2.1.2 Targeting specific cells

The surface modification of NPs with biological ligands, such as folate, arginine-glycine-aspartate (RGD) peptides, aptamers, transferrin, antibodies or small antibody fragments, facilitates NP targeting, imaging and internalization into specific cells, e.g., cancer cells, and tumor tissues.

Folate is a well-known small molecule frequently used as a cancer cell-targeting ligand that binds to folate receptors with high affinity. The chemical conjugation of folate onto the surface of NPs can significantly promote their targeted delivery into cancer cells that overexpress folate receptors [37].

Proliferating tumors are known to generate new blood vessels. This process is an important feature of tumor development characterized by the unique overexpression of the integrins $\alpha_v\beta_3$ and $\alpha_v\beta_5$ by nascent endothelial cells during angiogenesis in various tumors, but not by ordinary endothelial cells. Peptides possessing the RGD sequence bind the integrins $\alpha_v\beta_3$ and $\alpha_v\beta_5$ with high affinity. Cyclic RGD peptides show higher affinity and stability than do linear RGD peptides, which allows their use for developing integrin-selective, targeting NPs [38].

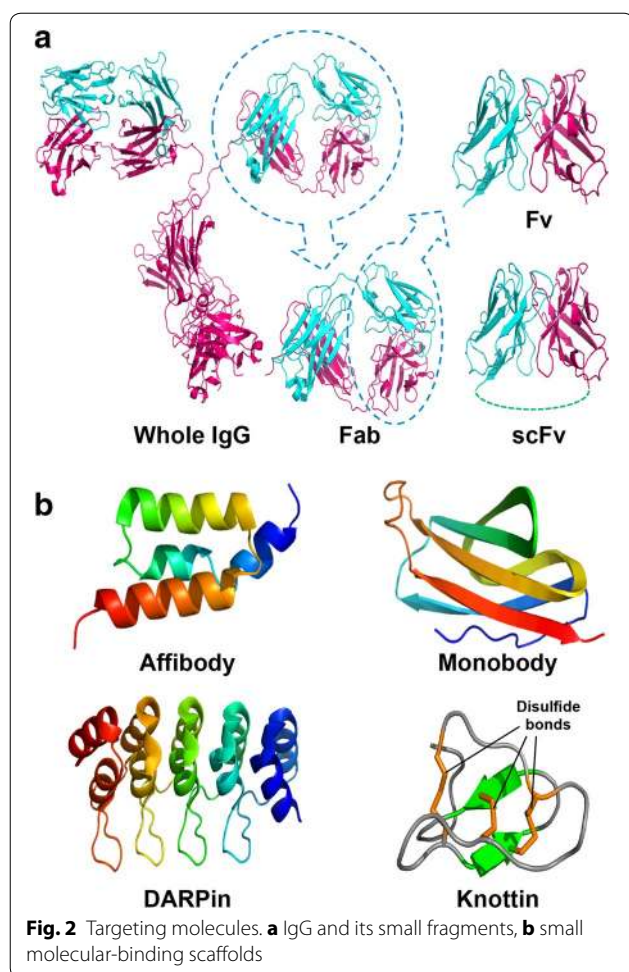
Aptamers are short, single-stranded RNA or DNA oligonucleotides (15–40 bases) that can bind to target molecules with high affinity and specificity due to the ability of the molecules to fold into unique conformations with three-dimensional (3D) structures. A large number of aptamers have been screened against aberrantly activated proteins in cancer cells, such as vascular endothelial

growth factor, platelet-derived growth factor, and nuclear factor kappa-light-chain-enhancer of activated B cells. Specific aptamers for targets can be selected from a large number of random sequences (libraries of 10^{15} random oligonucleotides) via the systematic evolution of ligands by exponential enrichment (SELEX) [39]. Aptamers generally have less immunogenicity, which can lead to improved biodistribution in the human body. NP surfaces can easily be conjugated with aptamers, and the conjugates show efficient cancer cell targeting and internalization [40]. Small molecules, peptides and aptamers are preferred for targeting and imaging ligands because they can be simply conjugated to NPs via facile chemical conjugation methods.

Transferrin (Tf) is a monomeric glycoprotein that can transport iron atoms into cells. Upon the binding of Tf to the Tf receptor (TfR), the Tf/TfR complex is internalized by cells through receptor-mediated endocytosis. TfR has been explored as a target for delivering anti-cancer drugs into cancer cells due to its overexpression by malignant tumor cells. TfR can be targeted by direct interaction with Tf displayed on the surface of NPs [41].

Monoclonal IgG antibodies (mAbs) have been the preferred targeting molecules for receptors, membrane proteins and glyco-antigens on the surface of cancer cells. Because many breast cancer cells overexpress human epidermal growth factor receptor-2 (HER-2), NPs coated with anti-HER-2 antibodies can target breast cancer cells with high specificity. Similarly, epidermal growth factor receptor (EGFR) can be targeted by anti-EGFR antibodies. Despite the immense efforts directed toward their development, mAb-conjugated NPs still encounter many challenges and limitations, such as the difficulty or cost of manufacturing, immunogenicity, and penetration into tumor tissues, as mAbs are very large (150–170 kDa, 15–20 nm in diameter) and complex molecules. Alternatively, after proper engineering, small antibody fragments [e.g., antigen-binding fragment (Fab: ~55 kDa) and variable fragment (Fv: ~27 kDa)] can be used as they can retain the targeting affinity and specificity of the original whole antibody (Fig. 2a). For example, the single-chain variable fragment (scFv: ~28 kDa) that consists of variable heavy- and light-chain domains connected with a flexible peptide linker can be used to target cells with high binding affinity and specificity.

Additionally, many alternative molecular scaffolds to mAbs have been investigated and developed in recent years, largely by the pursuit of much smaller (<20 kDa) targeting molecules with their putatively superior transport properties (Fig. 2b) [42]. These scaffolds include affibodies (~8 kDa) with three-helix bundles structure derived from the Z domain of protein A, DARPIn with three or more repeated small domains (~6 kDa)



consisting of two α -helices separated by a β -turn derived from ankyrin repeat proteins, and monobody with seven β -sheets forming a β -sandwich and three exposed loops from the 10th human fibronectin extracellular type III domain (~ 10 kDa). These scaffolds are lacking disulfide bonds that make it possible to produce functional scaffolds regardless of the redox potential of the cellular environment, including the reducing environment of the cytoplasm and nucleus. Another scaffold is knottins (~ 3.5 kDa) comprising a family of exceptionally small and highly stable proteins found in many species with structural homology involving a triple-disulfide stabilized knot motif. The randomization of loops or surfaces in conjunction with phage, ribosome or cell surface display technologies is used to engineer these molecular scaffolds and select binders to target molecules from many random libraries.

2.1.3 Internalization into cells

The surface modification of NPs with cell-penetrating peptides (CPPs) [43], such as the Tat peptide, penetratin,

pVEC, transportan, MPG, Pep-1 and polyarginines, could facilitate the internalization of NPs into cells through either direct entry into the cytosol or endosomal pathways. The Tat peptide, penetratin and pVEC are short peptides (~ 20 -mers) derived from the basic domain of the HIV-1 trans-activator of transcription (Tat) protein, the third helix of the Antennapedia homeodomain and cadherin, respectively. Transportan, MPG and Pep-1 are chimeric peptides (~ 30 -mers) that are formed by the fusion of two natural sequences derived from galanin/mastoparan, HIV-gp41/SV40 T-antigen and HIV-reverse transcriptase/SV40 T-antigen, respectively. These CPPs mostly bear a net positive charge and consist of amino acid (AA) sequences with repeated basic AA units and hydrophobic or aromatic AAs. The repeated basic AA units might contribute to not only the binding of CPPs to the negatively charged cell surface but also the endosomal escape of CPPs via conformational change under the acidic pH conditions of late endosomes.

2.1.4 Endosomal escape

The endosomal-escape ability of NPs is indispensable for the delivery of NPs into the cytosol and to organelles within the cell. Peptide-based endosomal-escape agents have been developed, and these are derived from the small-peptide domains of several viral, bacterial and human sources [44]. For example, the HA2 subunit of the *Haemophilus influenzae* hemagglutinin (HA) protein of the influenza virus with a short chain of an N-terminal anionic peptide has shown fusogenic activity. At a low pH, the protonation of the glutamate (Glu) and the aspartate (Asp) causes a conformational change of this peptide from a random coil into an amphiphilic α -helical structure. This change allows the amphiphilic α -helical peptide to bind to the endosomal membrane, causing membrane disruption. A pH-sensitive peptide GALA with repeating glutamate-alanine-leucine-alanine (Glu-Ala-Leu-Ala) units could disturb the lipid bilayer by the same mechanism and facilitate the endosomal escape of GALA-modified NPs at acidic pH values. Arginin (Arg)-rich peptides and cationic peptides, also derived from viral proteins, could mimic the endosomal-disruptive properties of viral particles [45]. Several chemical polymers, such as polyethylenimine- and imidazole-containing polymers, with endosomal-disruptive properties have been reported. These polymers have a buffering capacity ranging from pH 5.0–7.2 and can promote endosome osmotic swelling and disruption via the proton sponge effect [46]. Recently, a conformation-switchable synthetic lipid consisting of two alkyl chains on a di(methoxyphenyl)pyridine (pH-switchable unit) and a polar head group at the para position to the pyridine N atom was reported; upon protonation, hydrogen bonding induced a relative

orientation change of the two alkyl chains, which disturbed the lipid packing of the membranes and conferred endosomal-escape properties [47].

2.1.5 Trafficking to specific organelles

In eukaryotic cells, proteins are specifically sorted during or after translation and delivered from the cytosol to target organelles, such as the nucleus, endoplasmic reticulum, peroxisomes and mitochondria. These proteins contain organelle-targeting peptide signals often found at the N-terminal extension consisting of a short, positively charged stretch of basic AAs and a long α -helical stretch of hydrophobic AAs [48, 49], and a database of protein localization signals has been constructed based on experimental protein localization [50]. Gene delivery systems for the gene therapy of chromosomal and mitochondrial DNA have been developed by chemically conjugating nuclear and mitochondrial targeting signal peptides to NPs consisting of therapeutic DNAs [51].

2.1.6 Controlling payload release

In many cases, NPs in the endosomes or the cytoplasm must collapse to allow the release of their payloads. Several strategies using stimulus-responsive moieties built into NPs have been utilized to improve the efficiency of controlled release [31]. These include pH-sensitive and thermal-sensitive polymers, which control interactions between payloads and NPs [52], and external stimulus-sensitive crosslinkers, which conjugate payloads with NPs [53], such as pH-labile linkers, photosensitive- and enzyme-cleavable linkers, and disulfide crosslinkers that are sensitive to a reducing intracellular environment.

The difference in pH values existing between healthy tissues (pH 7.4) and the extracellular environment of solid tumors (pH 6.5–6.8), as well as between the cytosol (pH 7.4) and endosomes (pH 5–6), has been extensively utilized to trigger the release of drugs into a specific organ or intracellular compartment. Polymers with functional groups that can alter the structure and hydrophobicity of NPs as a result of protonation or deprotonation in response to pH variation can be utilized in pH-sensitive polymeric NPs. Notable examples of pH-sensitive polymers include poly(acryl amide) (PAAm), poly(acrylic acid) (PAA), poly(methacrylic acid) (PMAA), poly(methyl acrylate) (PMA), poly(diethylaminoethyl methacrylate) (PDEAEMA), poly(diallyl dimethylammonium chloride) (PDDA) and poly(dimethyl aminoethyl methacrylate) (PDMAEMA).

Temperature-sensitive polymers and hydrogels exhibit a volume phase transition at a certain temperature, which causes a dramatic change in the hydration state. This phase transition reflects competing hydrogen-bonding properties, where intra- and intermolecular hydrogen

bonding of the polymer molecules are favorable compared to the solubilization of the polymers by water. Examples of thermo-sensitive polymers are poly(*N*-isopropyl acrylamide) (PNIPAAm), poly(*N,N*-diethyl acrylamide) (PDEAAm), poly(methyl vinyl ether) (PMVE), poly(*N*-vinyl caprolactam) (PVCL), and poly(ethylene oxide)-poly(propylene oxide)-poly(ethylene oxide) (PEO-PPO-PEO).

In the case of polymer–drug conjugates, pH-sensitive linkages, such as oxime (pH < 5), hydrazone (pH < 5), hydrazide (pH < 5) and acetal (pH < 4–5), have been used to directly attach drug molecules to polymers. The use of light as a stimulus to trigger drug release has been actively explored owing to its high spatiotemporal resolution. Photosensitivity is often introduced to NPs through functional groups that can change their conformations and structures (e.g., azobenzene, pyrene, nitrobenzene and spirobenzopyran groups) or break their chemical bonds (e.g., arylcarbonylmethyl, nitroaryl, arylmethyl and coumarin-4-ylmethyl groups) upon irradiation [54, 55].

Enzymes perform a vast array of important functions inside our body. For example, hydrolytic enzymes overexpressed in cancer cells and tumor tissue can break certain bonds (e.g., ester, amide, glucuronide and phosphodiester bonds) within biopolymers, causing polymer structure disassembly or destruction. Notable examples of these enzymes are esterase, matrix metalloproteinase, β -glucuronidase and alkaline phosphatase. These enzymatic reactions can be utilized to trigger drug release [56].

2.1.7 Recent advances in targeted drug delivery and bioimaging

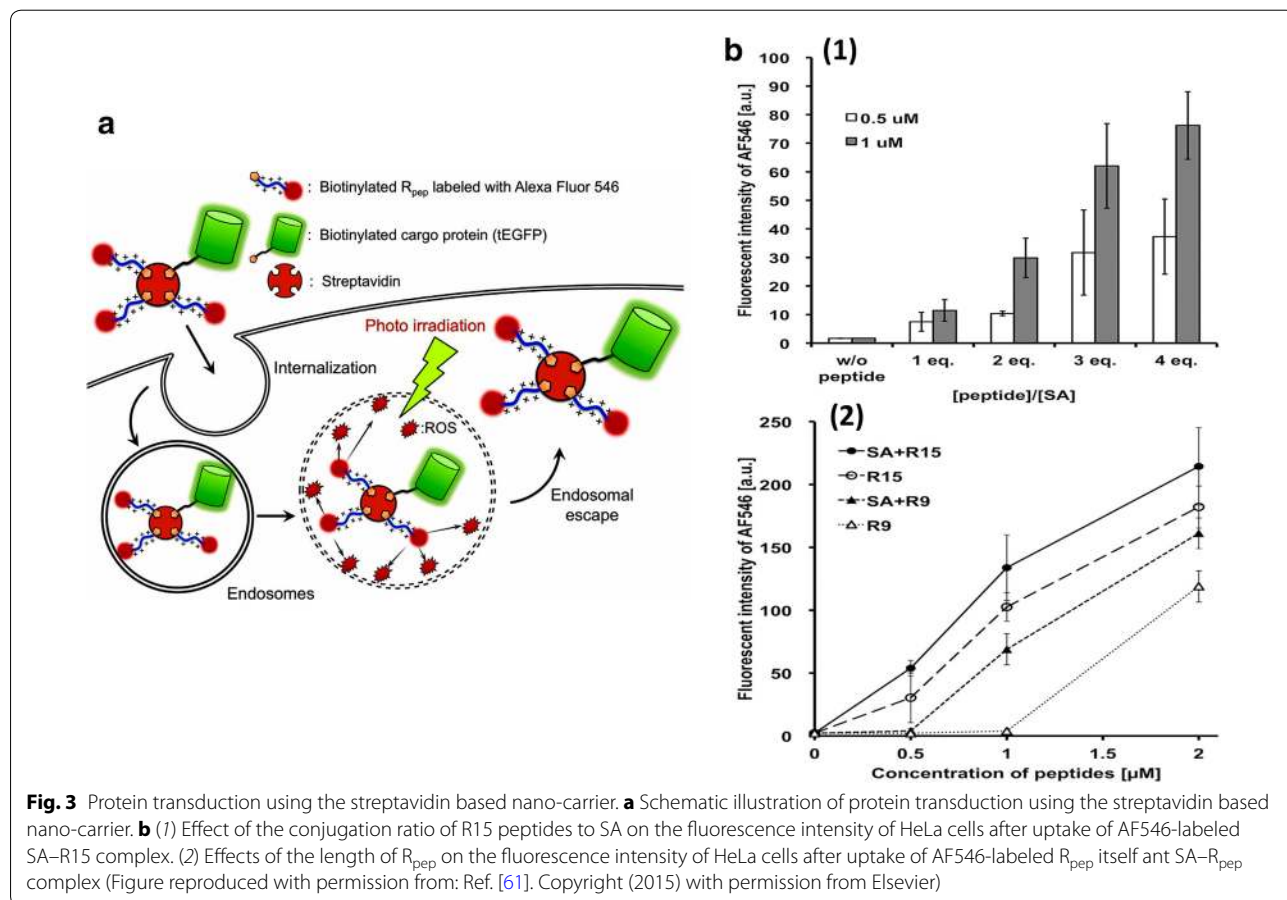
A major challenge of targeted drug delivery and bioimaging in therapeutics and diagnostics is the fabrication of NPs modified with various functional biomolecules for overcoming the above-mentioned biological barriers with a triggered cargo release system. Pluronic polymer-based micelles, to which folic acid (FA), redox-sensitive thiol groups and the anti-cancer drug doxorubicin (DOX) are chemically conjugated with pH-sensitive linkers, could be successfully delivered into multidrug-resistant (MDR) tumors in mice and exerted high cytotoxicity in the DOX-resistant MDR tumors by bypassing MDR efflux [57]. The carboxylate graphene oxide (GO)-based nanocarrier was multifunctionalized by poly(ethylene glycol) (PEG) terminated with an amino group and an FA group (FA-PEG-NH₂) via the amidation reaction. The GO-based nanocarrier could adsorb large amounts of DOX on the GO surface via π - π stacking interactions at a neutral pH but release it at an acidic pH. The DOX-loaded FA-PEG-modified GO-based nanocarrier not only showed stable dispersibility and targetability to

cancer cells with high FA receptor expression levels but also exhibited the low pH-activated controlled release of DOX in the endosomes of cells [58].

Nanohydrogels composed of filamentous bacteriophages and AuNPs, which were self-assembled via electrostatic interactions between the phage-capsid proteins and imidazole-modified AuNPs, have been developed and utilized for noninvasive imaging and targeted drug delivery in preclinical mouse models of breast and prostate cancer. The phage-based nanohydrogels could be multifunctionalized by fusing peptides, e.g., tumor-targeting ligands and CPPs, to phage-capsid proteins and by incorporating temperature-sensitive liposomes or mesoporous silica NPs containing imaging reagents and drugs. Because AuNPs packed densely within the nanohydrogel, their surface plasmon resonance shifted to the near-infrared (NIR) range, thereby allowing the NIR laser-mediated spatiotemporal photothermal release of cargo from temperature-sensitive liposomes [59]. Multifunctionalized AuNPs are generally constructed by the covalent assembly of an Au core with thiolated ligands. Novel multifunctionalized AuNPs have been assembled in one step by the nucleic acid hybridization of thiolated

oligodeoxynucleotide-modified AuNPs with a library of functional molecule-conjugated complementary peptide nucleic acids (PNAs). The PNAs were functionalized by conjugation with 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid for chelating ⁶⁴Cu for positron emission tomography imaging, PEG for conferring stealth properties, and Cy5 for fluorescent imaging. These NPs demonstrated good stability in vivo by showing biodistribution behavior in mice [60].

Recently, streptavidin (SA)-containing multifunctionalized NPs for carrying various biotinylated functional biomolecules have been reported. SA is a homo-tetramer protein, and each subunit can tightly bind to biotin molecule. We developed an SA-based cell-permeable nano-carrier equipped with photosensitizers as a versatile vehicle for spatiotemporally controlled cargo protein delivery into the cytosol (Fig. 3a) [61]. These nanocarriers can be prepared by attaching photosensitizer (Alexa Fluor 546: AF546)-modified biotinylated CPPs (oligo-arginine peptide R9 or R15) to a few biotin-binding sites of SA. Furthermore, a biotinylated target cargo protein is also loaded onto this carrier complex by using the remaining biotin-binding site of SA. Conjugation with



more than three CPPs per SA significantly raised the cell-permeability of the SA–CPP complexes into HeLa cells (Fig. 3b). Under optimized conditions, the SA–CPP (R15) complex could be delivered into cells with both high efficiency and low cytotoxicity. Furthermore, the internalized AF546-modified SA complex could spatiotemporally escape from the endosome in a light-irradiated area.

Photolytic protein aggregates (P-Aggs) for light-controllable nanocarriers have also been developed using SA [62]. Submicron-scaled P-Aggs were constructed by mixing SA and cargo proteins labeled with a biotinylated caging reagent (BCR) and were utilized as a facile and versatile platform for the light-induced release of cargo proteins (Fig. 4). The size of P-Aggs could be controlled either by adding an excess of biotin to the above mixture to stop the increase in P-Agg size or by conducting a mixing reaction in a water pool of reverse micelles and adding biotinylated-PEG to stop the increase in P-Agg size. For example, P-Aggs were prepared by mixing SA, a BCR-caged transferrin-doxorubicin conjugate (Tf-DOX)

and biotinylated AF647. These P-Aggs multifunctionalized with Tf, Alexa Fluor 647 and DOX were introduced into human colon cancer cells by endocytosis via TfR, followed by the selective release of DOX from the P-Aggs in light-irradiated cells, resulting in the spatiotemporal induction of target cancer cell apoptosis (Fig. 5).

We also developed a method for preparing SA-immobilized redox-sensitive nanohydrogels via peptide tag-induced disulfide formation mediated by horseradish peroxidase (HRP) (Fig. 6a) [63]. In this system, the peptides with sequences of HHHHHHC (C-tag) and GGGGY (Y-tag) were genetically fused to the N- and C-termini of SA (C-SA-Y), respectively. Here, H, C, G and Y denote histidine, cystein, glycine and tyrosine, respectively. The C-SA-Y was mixed with HRP- and thiol-functionalized 4-arm PEG to yield a C-SA-Y-immobilized hydrogel (C-SA-Y gel) crosslinked with redox-sensitive disulfide bonds. The C-SA-Y immobilized in the hydrogel retained its affinity for biotin, allowing the incorporation of any biotinylated functional biomolecules or synthetic chemical

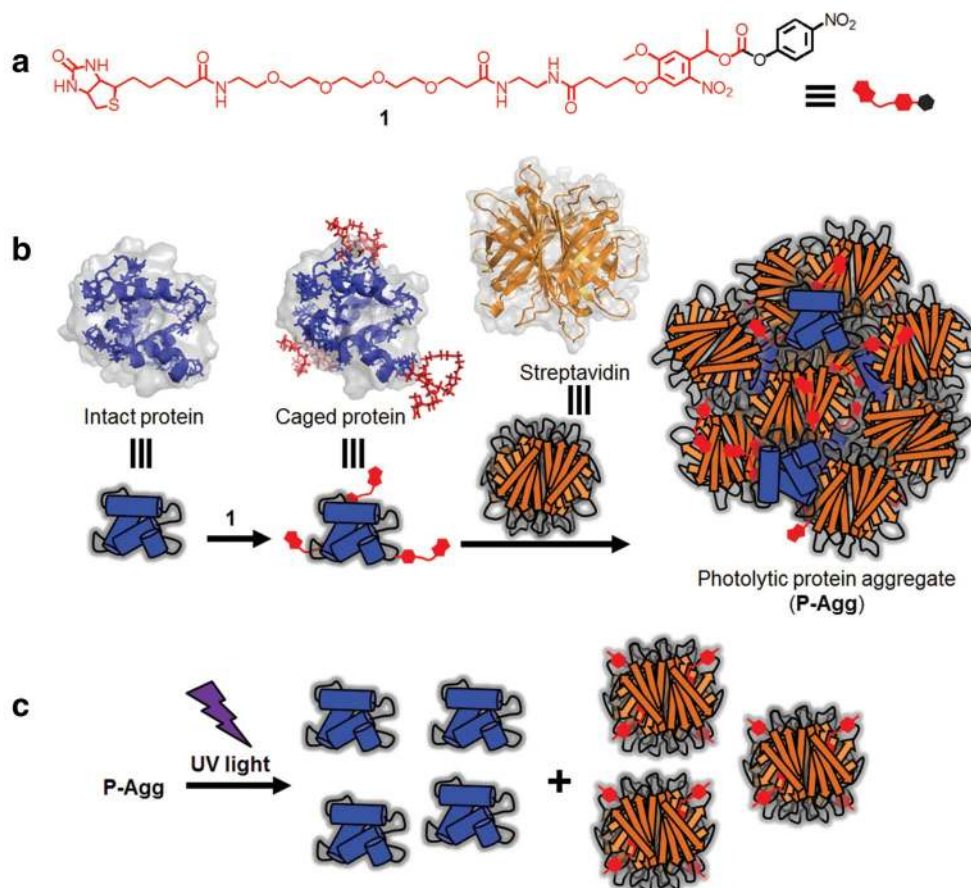
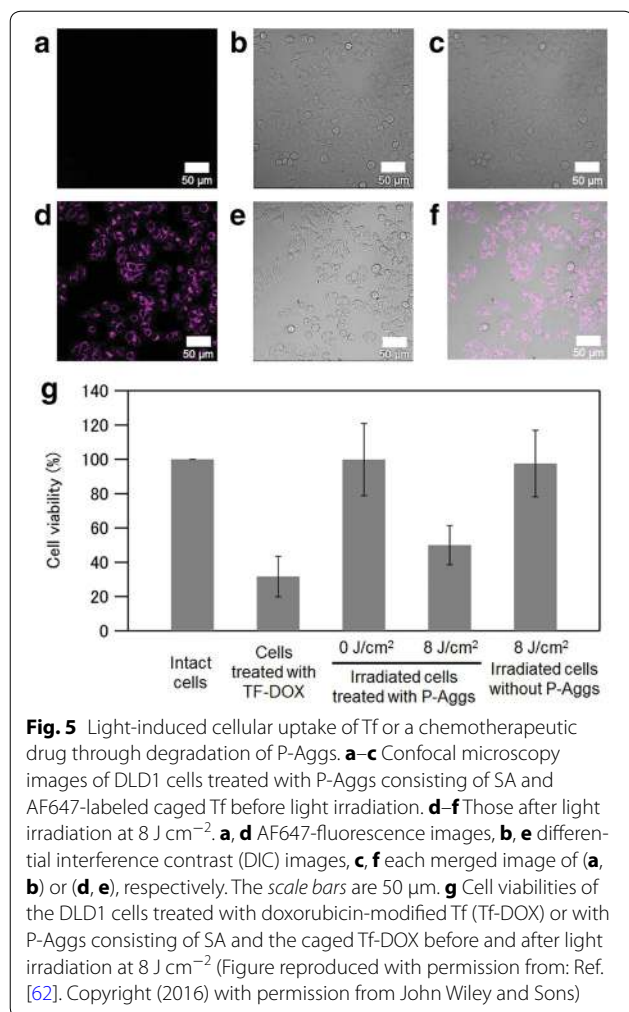


Fig. 4 Schematic illustration of photolytic P-Aggs formation and light-induced release of active proteins. **a** The chemical structure of BCR 1 consisting of a biotinylated photo-cleavable protection group (red) and an amino-reactive group (black). **b** Schemes of P-Aggs formation. **c** Protein photo-liberation from P-Aggs (Figure reproduced with permission from: Ref. [62]. Copyright (2016) with permission from John Wiley and Sons)



agents into the hydrogel via biotin-SA interaction. The C-SA-Y gel was further prepared within a reverse micelle system to yield a nanosized hydrogel, rendering it a potential drug delivery carrier. A C-SA-Y nanogel functionalized with biotinylated CPP (biotin-G3R15GYC-Alexa Fluor 546) and Alexa Fluor 488-labeled saporin was prepared, and we investigated its efficacy as a drug delivery system for human colon adenocarcinoma DLD1 cells (Fig. 6a). The cell viability assay results revealed that the C-SA-Y nanogel prepared without CPP or saporin hardly affected the viability of DLD1 cells. In contrast, treatment of the human colon cancer cells with the C-SA-Y gel functionalized with both CPP and saporin resulted in a marked decrease in cell proliferation (Fig. 6b). These results indicated that the C-SA-Y nanogel had been internalized into the cells via the CPP, reducing cell variability by cytotoxicity of saporin. The internalization ability of CPP and cytotoxicity of saporin were therefore successfully integrated in the C-SA-Y nanogel, with both properties working cooperatively to yield a cytotoxic C-SA-Y nanogel.

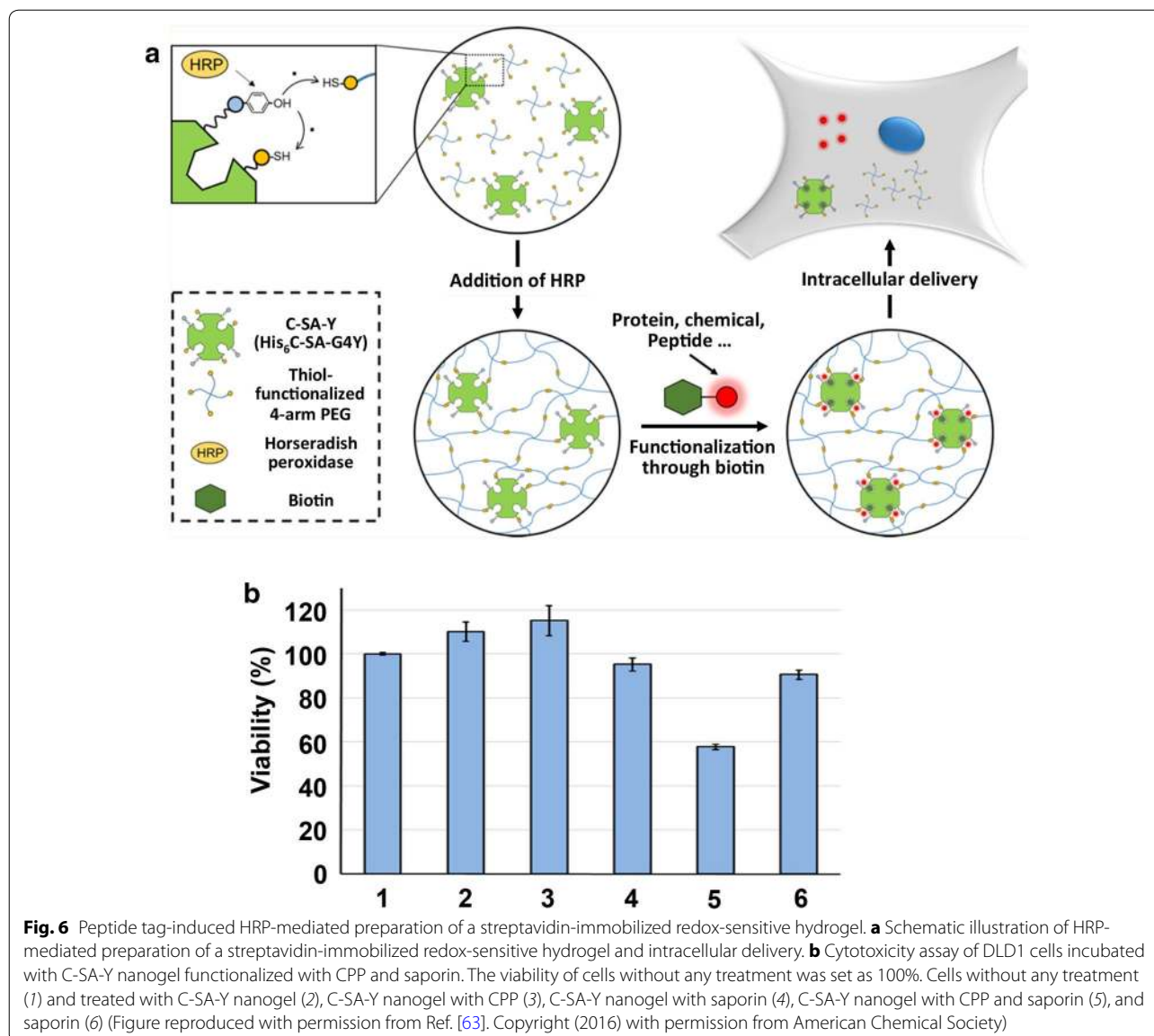
2.2 Nanobiomaterials for biosensing and bioanalysis

Biosensing and bioanalysis based on new nanomaterials and nanotechnology in the areas of nanoelectronics, nanooptics, nanopatterns and nanofabrication have a wide range of promising applications in point-of-care diagnostics, earlier disease diagnosis, pathological testing, food testing, environmental monitoring, drug discovery, genomics and proteomics. The rapid development of nanotechnology has resulted in the successful synthesis and characterization of a variety of nanomaterials, making them ideal candidates for signal generation and transduction in sensing. In other words, the unique properties and functionalization of biomaterial-conjugated nanostructures make them very useful for signal amplification in assays, other biomolecular recognition events and fabricating functional nanostructured biointerfaces [64, 65]. Therefore, nanomaterials and nanofabrication technologies play significant roles in fabricating biosensors and biodevices (e.g., colorimetric, fluorescent, electrochemical, surface-enhanced Raman scattering, localized surface plasmon resonance, quartz crystal microbalance and magnetic resonance imaging (MRI)), including implantable devices [66] for the detection of a broad range of biomarkers with ultrahigh sensitivity and selectivity and rapid responses.

2.2.1 Nanomaterials for enhancing sensitivity of biosensing and bioanalysis

During the last decade, several promising nanomaterials (e.g., QDs, NPs, carbon nanotubes (CNTs) and graphene) with biomolecule-modified surfaces have been widely used in the fields of biosensing, bioanalysis and diagnostics [67–70]. For example, one of the first nanomaterials to have an impact on amperometric biosensors was CNTs, which have such advantages as a small size with a large surface area, an excellent electron transfer ability, and easy biomolecule immobilization. CNT-modified electrodes improved current densities and enhanced the reactivity of biomolecules, redox cofactors and redox enzymes. In addition, aligned CNT forests facilitated direct electron transfer with the redox centers of enzymes, resulting in improved overall performance of enzyme electrodes and enzyme-labeled immunosensors [71].

Several nanomaterials have shown great promise in imaging due to their intrinsic imaging characteristics, such as their brightness, sharp bandwidth and long-term stability (e.g., fluorescent agents, such as QDs [72], magnetic NPs in MRI [73] and colloidal AuNPs [74]). For imaging, nanomaterials can be targeted to specific disease sites within the body by conjugating the materials to biomarker-specific biomolecules. These biomaterial-based imaging agents can also provide information in



addition to anatomical data, e.g., information relating to physiology and function, which enables more accurate and early disease diagnosis, such as the highly sensitive detection of early-stage cancer [75].

2.2.2 Nanofabrication technologies for biosensing and bioanalysis

Microarrays [76] and microfluidic [77, 78] platforms coupled with biomolecule-conjugated nanomaterials (e.g., QDs, NPs, or CNTs conjugated with enzymes, antibodies, DNAs, or aptamers) have enabled the simultaneous multiplex detection of many disease biomarkers for cancer, infectious diseases, diabetes, cardiovascular diseases and Alzheimer's disease. For example, novel electrochemiluminescence (ECL) microwell array [79] and microfluidic [80]

immunoassay devices equipped with capture-antibody-decorated single-walled carbon nanotube (SWCNT) forests on pyrolytic graphite chips have been developed. The $[\text{Ru}(\text{bpy})_3]^{2+}$ -doped silica NPs covered with thin hydrophilic polymer films prepared by the sequential layer-by-layer deposition of positively charged PDDA and negatively charged PAA were used as ECL labels in these systems for highly sensitive two-analyte detection. Antibodies to prostate specific antigen (PSA) and interleukin (IL)-6 were chemically conjugated to either SWCNTs or polymer-coated RuBPY-silica-Ab₂ NPs via amidization with 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC) and *N*-hydroxysulfosuccinimide (NHSS). The microfluidic immunoassay device provided the simultaneous detection of the biomarker proteins PSA and IL-6 in

serum, demonstrating high sensitivity and detection limits in the low femtogram per milliliter range (10^{-21} M range) (Fig. 7) [80]. These platforms explored the detection of ultra-low concentrations of target biomarkers and have realized rapid, ultrasensitive and cost-effective bioassays requiring minimum sample volumes, which will enable primary care physicians and patients to perform assays in their respective settings, using so-called point-of-care diagnostics. The detection of cancer biomarkers by immunoassays and sensors using these engineered nanomaterials could also enable the diagnosis of cancer at very early stages [81, 82].

Fabrication must employ strategies to control chemistry to ensure not only that patterns and structures are generated at the desired location and within an appropriate time frame but also that undesired side reactions are prevented. Bionanofabrication, the use of biological materials and mechanisms for the construction of nanodevices for biosensing and bioanalysis, offers convergent approaches for building nanointerfaces between biomolecules and devices by either enzymatic assembly or self-assembly. For example, film-forming pH-sensitive chitosan directly assembles on electrodes under physiological conditions in response to electrode-imposed voltages (i.e., electrodeposition). Through recombinant technology, biomolecular engineering allows target proteins to be endowed with peptide tags [e.g., a Glutamine (Gln)-tag for transglutaminase-mediated crosslinking between the side chains of Gln and Lysine (Lys) residues] for assembly, which enables fabrication and controls

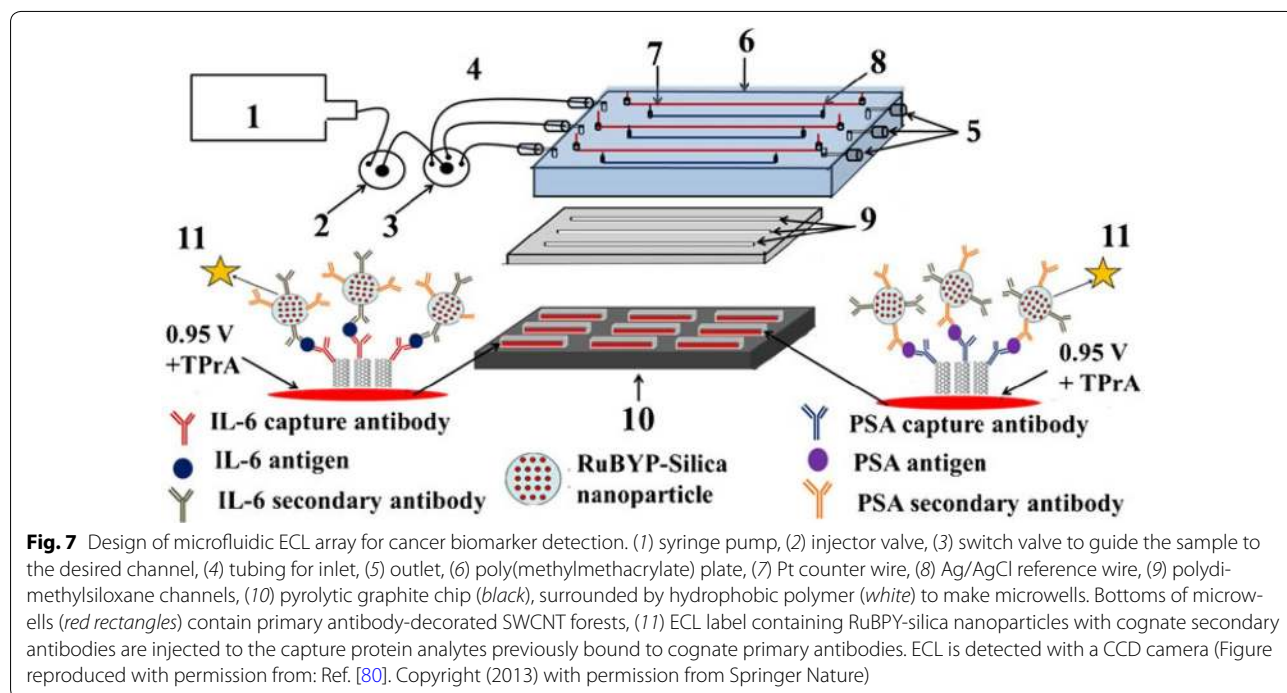
bioconjugation chemistry through molecular recognition for the enzymatic generation of covalent bonds (Fig. 8) [83]. These self-assembly and enzymatic assembly methods also provide mechanisms for construction over a hierarchy of length scales. Bionanofabrication will enable the effective interfacing of biomolecules with nanomaterials to create implantable devices.

2.3 Nanobiomaterials for biocatalysis

The use of nanomaterials for enzyme immobilization and stabilization is highly effective not only in stabilizing the enzyme activity but also in developing other advantageous properties, including high enzyme loading and activity, an improved electron transfer rate, low mass transfer resistance, high resistance to proteolytic digestion and the easy separation and reuse of biocatalysts by magnetic force [84]. The immobilization or entrapment of enzymes on the surface or interior of nanocarriers has been accomplished using various nanomaterials, such as polymer NPs (e.g., polylactic acid, polystyrene, polyvinyl alcohol, and chitosan), magnetic and superparamagnetic NPs, polymer nanofibers (e.g., nylon, polyurethane, polycarbonate, polyvinyl alcohol, polylactic acid, polystyrene, and carbon), CNTs, GO nanosheets, porous silica NPs, sol-gel NPs and viral NPs [85–87].

2.3.1 Enzyme immobilization

There are considerable advantages of effectively immobilizing enzymes for modifying nanomaterial surface



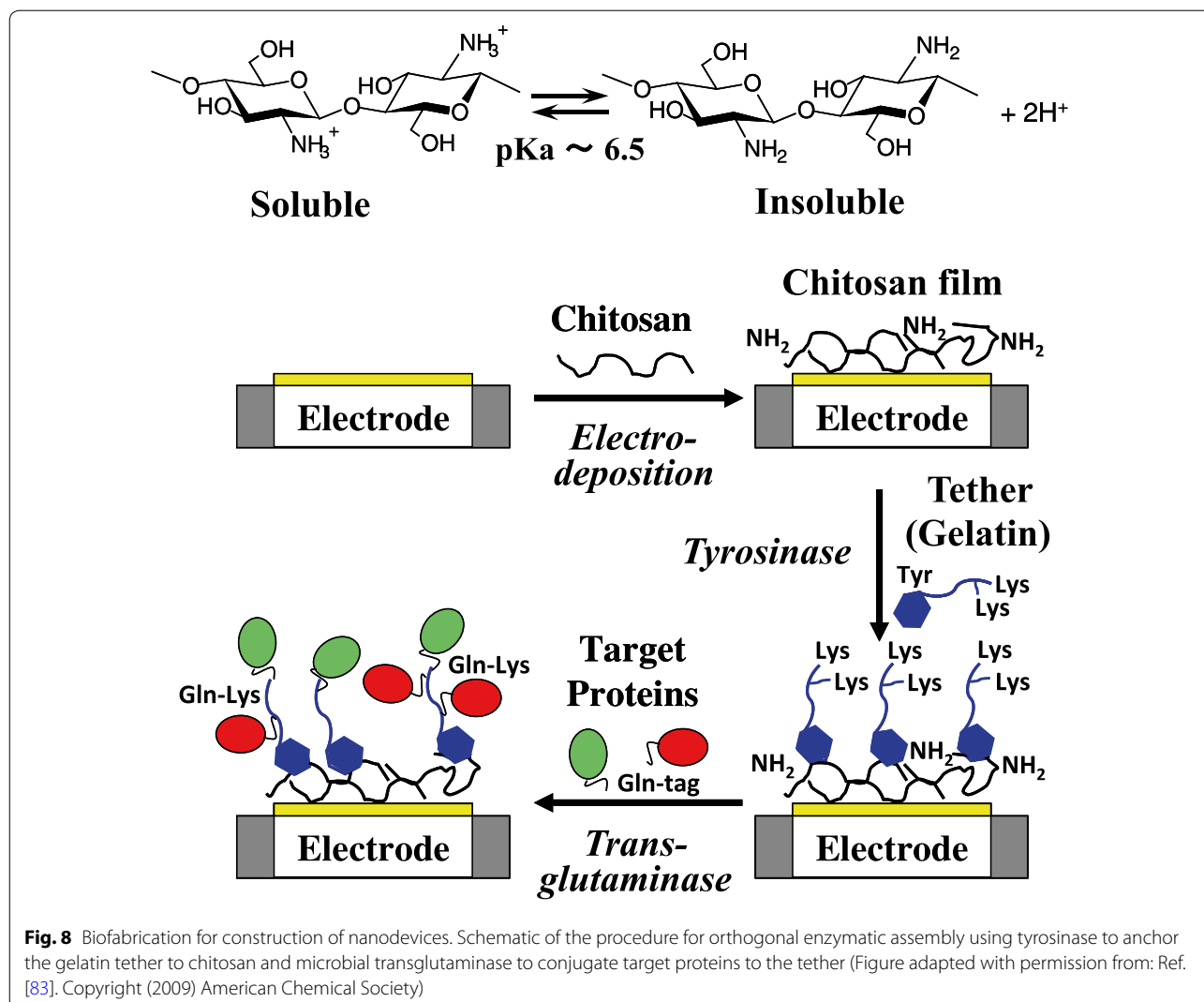


Fig. 8 Biofabrication for construction of nanodevices. Schematic of the procedure for orthogonal enzymatic assembly using tyrosinase to anchor the gelatin tether to chitosan and microbial transglutaminase to conjugate target proteins to the tether (Figure adapted with permission from: Ref. [83]. Copyright (2009) American Chemical Society)

properties and grafting desirable functional groups onto their surface through chemical functionalization techniques. The surface chemistry of a functionalized nanomaterial can affect its dispersibility and interactions with enzymes, thus altering the catalytic activity of the immobilized enzyme in a significant manner. Toward this end, much effort has been exerted to develop strategies for immobilizing enzymes that remain functional and stable on nanomaterial surfaces; various methods including, physical and/or chemical attachment, entrapment, and crosslinking, have been employed [86, 88, 89]. In certain cases, a combination of two physical and chemical immobilization methods has been employed for stable immobilization. For example, the enzyme can first be immobilized by physical adsorption onto nanomaterials followed by crosslinking to avoid enzyme leaching. Both glutaraldehyde and carbodiimide chemistry, such

as dicyclohexylcarbodiimide/*N*-hydroxysuccinimide (NHS) and EDC/NHS, have been commonly utilized for crosslinking. However, in some cases, enzymes dramatically lose their activities because many conventional enzyme immobilization approaches, which rely on the nonspecific absorption of enzymes to solid supports or the chemical coupling of reactive groups within enzymes, have inherent difficulties, such as protein denaturation, poor stability due to nonspecific absorption, variations in the spatial distances between enzymes and between the enzymes and the surface, decreases in conformational enzyme flexibility and the inability to control enzyme orientation.

To overcome these problems, many strategies for enzyme immobilization have been developed. One approach is known as ‘single-enzyme nanoparticles’ (SENs), in which an organic–inorganic hybrid polymer

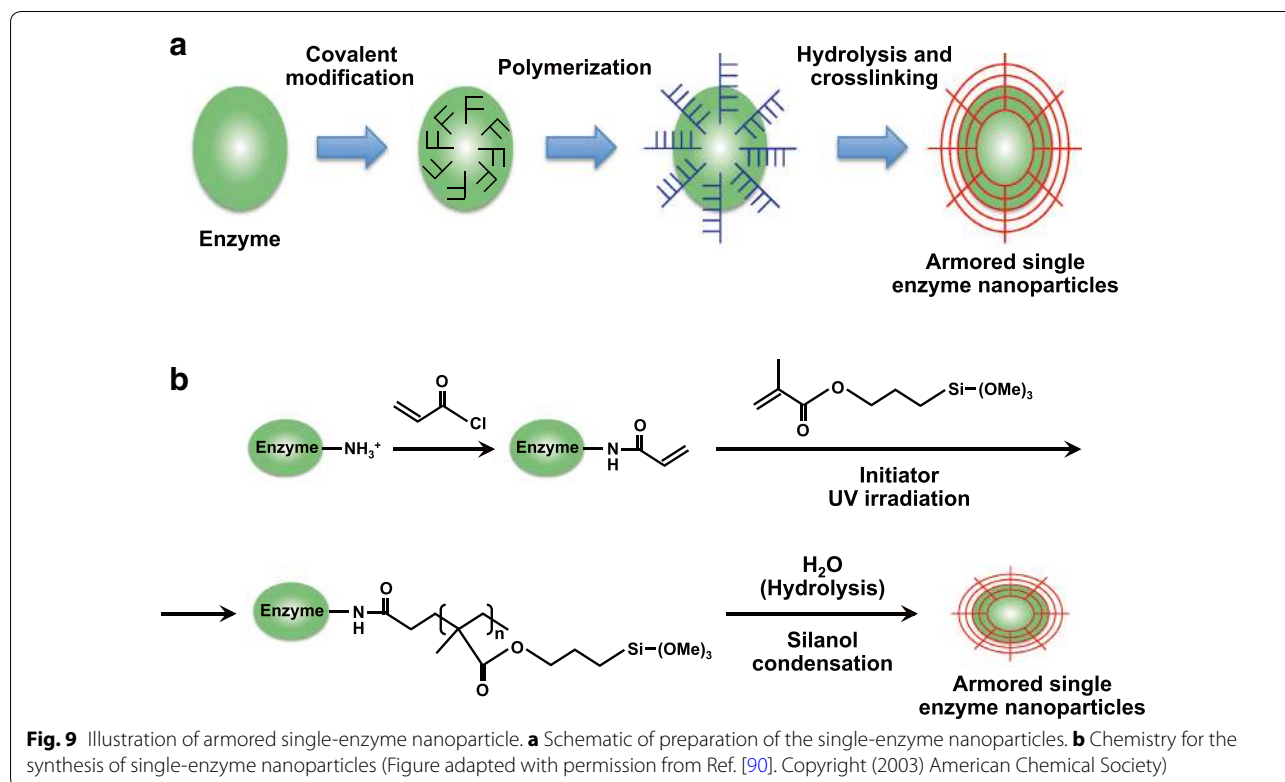
network less than a few nanometers in thickness is built up from the surface of an enzyme. The synthesis of SENs involves three reactions: first, amino groups on the enzyme surface react with acryloyl chloride to yield surface vinyl groups; then, free-radicals initiate vinyl polymerization from the enzyme surface using a vinyl monomer and pendant trimethoxy-silane groups; finally, orthogonal polymerization occurs via silanol condensation reactions to crosslink the attached polymer chains into a network (Fig. 9). It was demonstrated that SENs can be immobilized in mesoporous silica; additionally, this method of immobilization was shown to provide a much more stable immobilized enzyme system than that of native enzymes immobilized by either adsorption or covalent bonding in the same material [90]. Another approach is to introduce molecular interfaces between a solid surface and enzymes. Several methods based on this approach have been reported, such as the surface modification of solid supports with hydrophilic synthetic polymers [91, 92] and peptides [93] with specificities and affinities toward enzymes, and the fusion of enzymes with peptide tags [94] or anchor proteins [95, 96].

Peptides with an affinity for nanomaterials have been identified from a combinatorial peptide library, and these peptides are promising tools for bottom-up fabrication technology in the field of bionanotechnology. Through the use of these peptides, enzymes can be

directly immobilized on a substrate surface with desired orientations and without the need for substrate surface modification or complicated conjugation processes. For example, an Au-binding peptide was applied to direct the self-assembly of organophosphorus hydrolase onto an AuNP-coated graphene chemosensor. This electrochemical biosensor system could detect pesticides with a fast response time, low detection limit, better operating stability and high sensitivity [97].

The amphiphilic protein HFBI (7.5 kDa), class II hydrophobin, that is produced by *Trichoderma reesei* adheres to solid surfaces and exhibits self-organization at water–solid interfaces. A fusion protein between HFBI and glucose oxidase (GOx-HFBI) with a 21-AA flexible linker (linker sequence: SGSVTSTSKTTATASKTSTST) was constructed. This fusion protein exhibited the highest levels of both protein adsorption and high GOx activity owing to the presence of the HFBI spacer and flexible linker, which forms a self-organized protein layer on solid surface and enables the GOx component in the fusion protein to be highly mobile, respectively [95].

The crystalline bacterial cell surface layer (S-layer) proteins of prokaryotic organisms constitute a unique self-assembly system that can be employed as a patterning element for various biological molecules, e.g., glycans, polysaccharides, nucleic acids, and lipids. One of the most excellent properties of S-layer proteins is their



capability to self-assemble into monomolecular protein lattices on artificial surfaces (e.g., plastics, noble metals or silicon wafers) or on Langmuir lipid films or liposomes. A fusion protein between the S-layer protein SbpA from *Bacillus sphaericus* CCM 2177 and the enzyme laminarinase (LamA) from *Pyrococcus furiosus* fully retained the self-assembly capability of the S-layer moiety, and the catalytic domain of LamA was exposed at the outer surface of the formed protein lattice. The enzyme activity of the S-layer fusion protein monolayer on silicon wafers, glass slides and different types of polymer membranes was compared with that of only LamA immobilized with conventional techniques. LamA aligned within the S-layer fusion protein lattice catalyzed two-fold higher glucose release from the laminarin polysaccharide substrate compared with the randomly immobilized enzyme. Thus, S-layer proteins can be utilised as building blocks and templates for generating functional nanostructures at the meso- and macroscopic scales [98].

2.3.2 Multienzyme complex systems

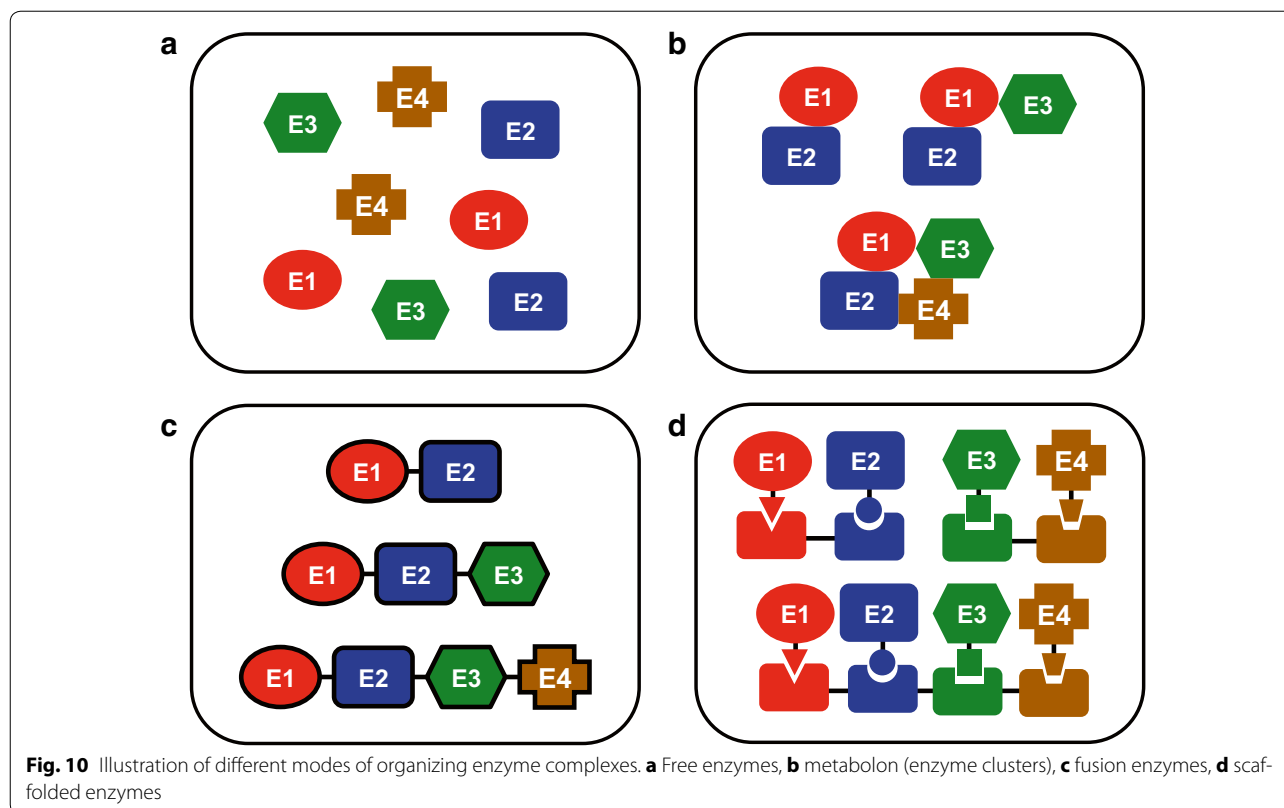
In nature, the macromolecular organization of multi-enzyme complexes has important implications for the specificity, controllability, and throughput of multi-step biochemical reaction cascades. This nanoscale macromolecular organization has been shown to increase the local concentrations of enzymes and their substrates, to enhance intermediate channeling between consecutive enzymes and to prevent competition with other intracellular metabolites. The immobilization of an artificial multienzyme system on a nanomaterial to mimic natural multienzyme organization could lead to promising biocatalysts. However, the above-mentioned immobilization methods for one type of enzyme on nanomaterials cannot always be applied to multienzyme systems in a straightforward manner because it is very difficult to control the precise spatial placement and the molecular ratio of each component of a multienzyme system using these methods. Therefore, strategies have been developed for the fabrication of multienzyme reaction systems [99, 100], such as genetic fusion [101], encapsulation [102] in reverse micelles, liposomes, nano/mesoporous silica or porous polymersomes, scaffold-mediated co-localization [103], and scaffold-free, site-specific, chemical and enzymatic conjugation [104, 105].

In many organisms, complex enzyme architectures are assembled either by simple genetic fusion or enzyme clustering, as in the case of metabolons, or by cooperative and spatial organization using biomolecular scaffolds, and these enzyme structures enhance the overall biological pathway performance (Fig. 10) [103, 106, 107]. In metabolons, such as nonribosomal peptide synthase, polyketide synthase, fatty acid synthase and acetyl-CoA

carboxylase, reaction intermediates are covalently attached to functional domains or subunits and transferred between domains or subunits. Alternatively, substrate channeling in such multienzyme complexes as metabolons, including by glycolysis, the Calvin and Krebs cycles, tryptophan synthase, carbamoyl phosphate synthetase, and dhurrin synthesis, is utilized to prevent the loss of low-abundance intermediates, to protect unstable intermediates from interacting with solvents and to increase the effective concentration of reactants. Additionally, scaffold proteins are involved in many enzymatic cascades in signaling pathways (e.g., the MAPK scaffold in the MAPK phosphorylation cascade pathway) and metabolic processes (e.g., cellulosomes from *Clostridium thermocellum*). From a practical point of view, there are several obstacles for the genetic fusion of over three enzymes to construct multienzyme complexes. First, large recombinant fusion proteins are easily misfolded and subsequently are either proteolyzed or form inactive inclusion bodies in *E. coli*. Furthermore, the optimum refolding conditions of each enzyme motif in fusion proteins are not always identical. Last, rational design methods for peptide linkers between enzymes that enable control of linker spatial arrangement and orientation have not yet been developed [106]. Additionally, engineering the required interfacial interactions for efficient enzyme clustering is extremely challenging. Therefore, flexible post-translational methods using enzymatic site-specific protein–protein conjugation and synthetic scaffolds by employing orthogonal interaction domains for assembly have been particularly attractive because of the modular nature of biomolecular design [103].

2.3.2.1 Post-translational enzymatic modification-based multienzyme complexes Many proteins are subjected to post-translational enzymatic modifications in nature. The natural post-translational processing of proteins is generally efficient and site-specific under physiological conditions. Therefore, in vitro and in vivo enzymatic protein modifications have been developed for site-specific protein–protein conjugation. The applications of enzymatic modifications are limited to recombinant proteins harboring additional protein/peptide tags. However, protein assembly using enzymatic modifications (e.g., inteins, sortase A, and transglutaminase) is a promising method because it is achieved simply by mixing proteins without special techniques [106].

Recently, we demonstrated a covalently fused multienzyme complex with a “branched structure” using microbial transglutaminase (MTGase) from *Streptomyces mobaraensis*, which catalyzes the formation of an ϵ -(γ -glutamyl) lysine isopeptide bond between the side chains of Gln and Lys residues. A cytochrome P450 enzyme

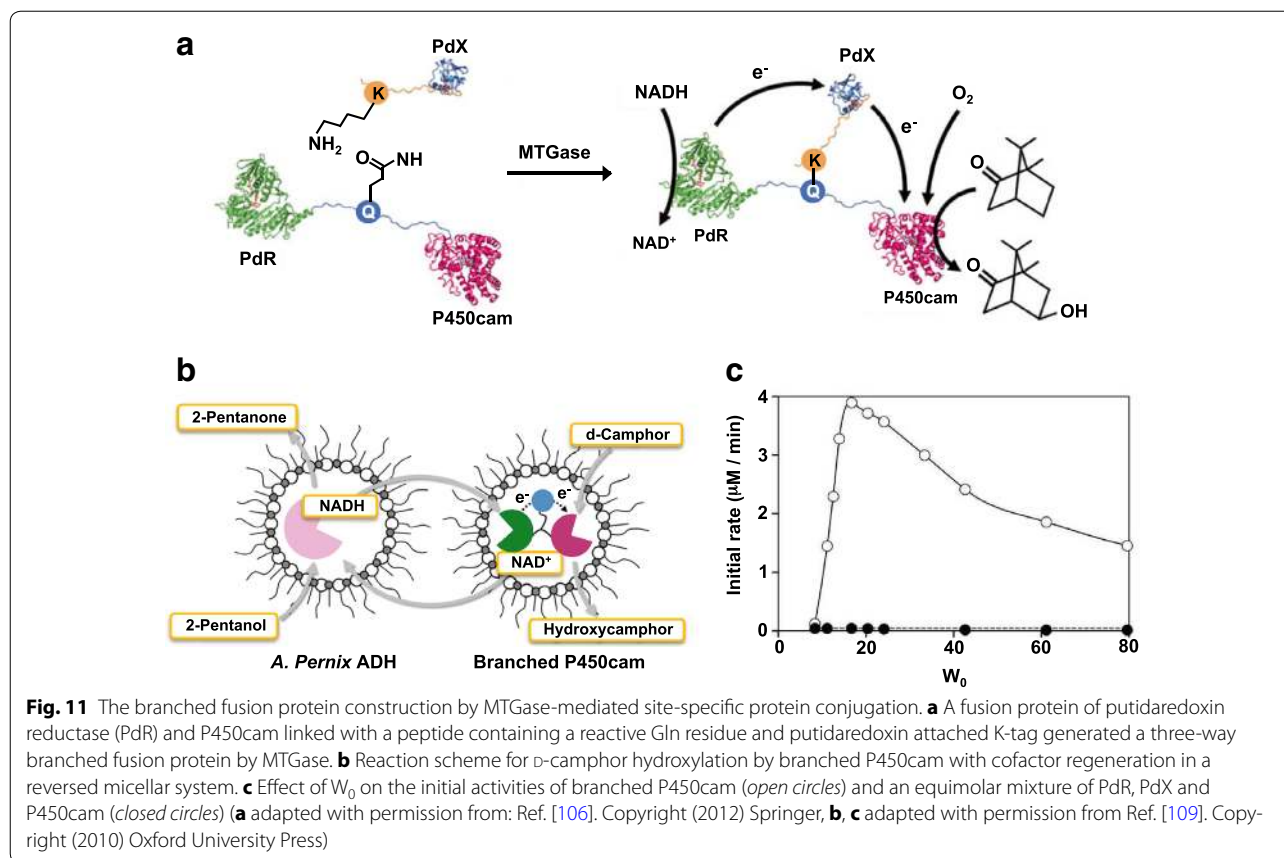


from *Pseudomonas putida* (P450cam) requires two soluble redox proteins, putidaredoxin (PdX) and putidaredoxin reductase (PdR), to receive electrons from NADH for its catalytic cycle, in which PdX reduced by PdR with NADH activates P450cam. Therefore, it has been suggested that the complex formation of P450cam with PdX and PdR can enhance the electron transfer from PdR to PdX and from PdX to P450cam. This unique multienzyme complex with a branched structure that has never been obtained by genetic fusion showed a much higher activity than that of tandem linear fusion P450cam genetically fused with PdX and PdR (Fig. 11a) [108]. This multienzyme complex with a branched structure was further applied to a reverse micelle system. When the solubility of substrate is quite low in an aqueous solution, the reverse micelle system is often adopted for simple, one-step enzymatic reactions because the substrate can be solubilized at a high concentration in an organic solvent, subsequently accelerating the reaction rate. In the case of a multienzyme system, especially systems including electron transfer processes, such as the P450cam system, the reverse micelle system is difficult to apply because each component is usually distributed into different micelles and because the incorporation of all components into the same aqueous pool of micelles is very difficult. Unlike the natural P450cam system, all components of the branched

P450cam system were incorporated into the same aqueous pool of micelles at a 1:1:1 ratio (Fig. 11b) and enabled both extremely high local protein concentrations and efficient electron transfer to P450cam, resulting in a reaction activity higher than that of a reverse micelle system composed of an equimolar mixture of PdR, PdX and P450cam (Fig. 11c) [109].

2.3.2.2 Scaffold protein-based multienzyme complexes Scaffold proteins enable the precise spatial placement of the components of a multienzymatic reaction cascade at the nanometer scale. Scaffolds are involved in many enzymatic reaction cascades in signaling pathways and metabolic processes [110], and they can provide advantages over reactions catalyzed by freely diffusing enzymes by segregating reactions, increasing throughput and providing modularity for the construction of novel reaction networks. Recently, various multienzyme systems have been developed using natural scaffold proteins [111] and synthetic scaffolds [112] composed of elements of natural scaffold proteins, such as cellulosomes [113] and signal transduction scaffolds [114].

Proliferating cell nuclear antigen (PCNA) is a DNA-sliding clamp that forms a symmetrical ring-shaped structure encircling double-stranded DNA (dsDNA) and acts as a scaffold for DNA-related enzymes, such as

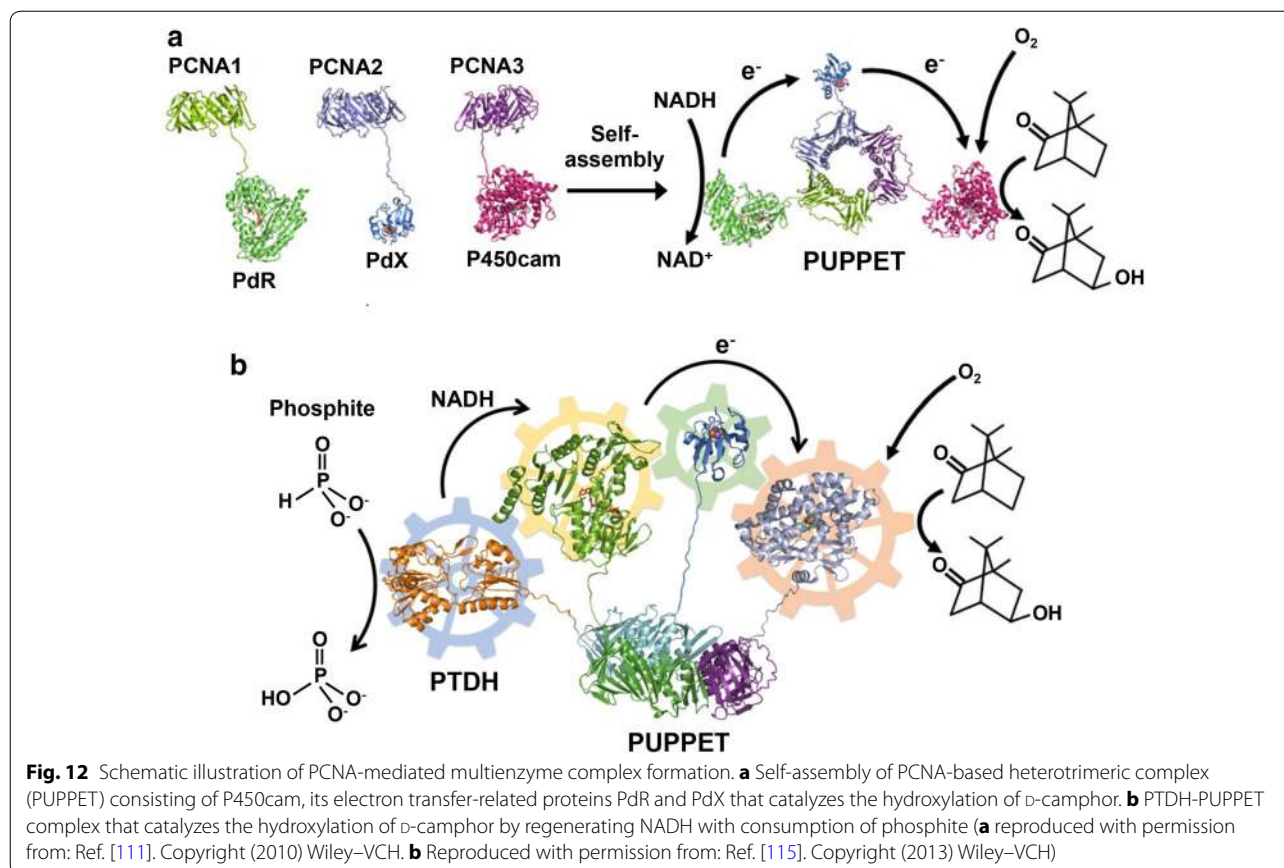


DNA polymerase and helicase. The archaeon *Sulfolobus solfataricus* has three distinct PCNA genes with the three expressed PCNA proteins, PCNA1, PCNA2 and PCNA3, which form a heterotrimeric complex. These three PCNAs were fused to the three component proteins (i.e., PdR, PdX, and P450cam) composing the *P. putida* P450 system (Fig. 12a). The resulting fusion proteins, PCNA1-PdR, PCNA2-PdX and PCNA3-P450cam, completely retained the functions of the component proteins, including the heterotrimerization of the PCNAs, the catalytic activities of PdR and P450cam, and the electron transfer function of PdX. The three fusion proteins immediately formed a heterotrimeric complex in vitro by mixing. Compared to an equimolar mixture of PdR, PdX and P450cam, the complex showed a 52-fold enhancement in the monooxygenase activity of P450cam because of efficient electron transfer within the complex from PdR to PdX and from PdX to P450cam [111]. This system based on the PCNA scaffold was further extended to a phosphite-driven self-sufficient P450cam system in vitro by incorporating phosphite dehydrogenase (PTDH) for cofactor NADH regeneration (Fig. 12b) [115]. The K_m value of PTDH-incorporated PUPPET (PTDH-PUPPET) for NAD^+ ($51.0 \pm 2.7 \mu M$) in the presence of D-camphor

and phosphite was slightly smaller than that of an equimolar mixture of PUPPET and PTDH ($69.7 \pm 4.8 \mu M$). This result indicates that the oxidation of NADH by the PdR domain in PTDH-PUPPET might increase the effective local concentration of NAD^+ around the PTDH domain and that this proximity effect on cofactor channeling could potentially be improved by optimizing the arrangement of PTDH and PdR on the PCNA scaffold.

Designer cellulosomes containing four different enzymes (two cellulases and two xylanases) from *Thermobifida fusca* have been reported, where four dockerin-fused cellulolytic enzymes were incorporated into specific locations on an artificial, chimeric scaffold containing four cohesins corresponding to each dockerin. As expected, compared to their free enzyme mixture system without the chimeric scaffolding, the resulting multi-enzyme complexes exhibited enhanced activity (~2.4-fold) on wheat straw as a complex cellulosic substrate [116].

Recently, Deuber et al. demonstrated in vivo multi-enzyme complex formation in *E. coli* cells via synthetic protein scaffold expression. Protein scaffolds with various arrangements of fusion domains were built from the interaction domains of signaling proteins, the mouse SH3 and PDZ domains and the rat GTPase protein-binding



domain (GBD). The three enzymes acetoacetyl-CoA thiolase, hydroxymethylglutaryl-CoA synthase and hydroxymethylglutaryl-CoA reductase, which catalyze a cascade reaction from acetyl-CoA to mevalonate, were genetically tagged with their cognate peptidyl ligands. These protein scaffolds and enzymes with peptidyl ligands were co-expressed in *E. coli* cells. A significant 77-fold increase in mevalonate production was achieved by the expression of the optimized scaffold: $(\text{GBD})_1\text{-(SH3)}_2\text{-(PDZ)}_2$ [114].

2.3.2.3 Oligonucleotide scaffold-based multienzyme complexes DNA has numerous attractive features as a scaffold for multienzyme complexes. Its properties, such as high rigidity, programmability, complexity and assembly through complementary hybridization, allow DNA to form excellent scaffolds with linear, two-dimensional (2D) and 3D structures (e.g., simple dsDNA helices, Holliday junctions, DNA tiles, and DNA origami) for arranging multiple enzymes with controlled spacing in linear, 2D or 3D geometric patterns and for constructing interactive multienzyme complexes and networks [117–120]. DNA-protein conjugates are necessary to achieve DNA-directed protein assembly for the fabrication of multienzyme complexes on DNA scaffolds. However, this requirement

makes it difficult to utilize this assembly method *in vivo*. Currently, there are several methodologies for conjugating proteins with DNA [117]. Proteins have been assembled onto DNA scaffolds through intervening adapter molecules, such as biotin–streptavidin, Ni–NTA–hexahistidine, antibodies–haptens and aptamers. Alternatively, direct covalent conjugation with DNA can be achieved by modifying cysteine (Cys) or Lys residues via disulfide or maleimide coupling, as well as by bioorthogonal chemistry, such as expressed protein ligation, Staudinger ligation and Huisgen cycloaddition.

By utilizing DNA nanostructures as assembly scaffolds, it has become feasible to organize the DNA-directed assembly of artificial multienzyme complexes. DNA-mediated assembly was employed to control the activity of a multidomain enzyme. Cytochrome P450 BM3 (P450 BM3) is composed of two domains, a flavin adenine dinucleotide and flavin mononucleotide-containing reductase domain (BMR) and a heme-containing monooxygenase domain (BMP). P450 BM3 shows monooxygenase activity by transferring electrons to BMP from NADPH through BMR. Both subdomains were genetically fused to the HaloTag protein, a self-labeling enzyme, enabling bioconjugation with chloroalkane-modified DNAs and

subsequently reconstituting BM3 activity by DNA-mediated assembly. The arrangement of the two domains on a DNA scaffold can control the distance between them. The distance-dependent activity of multidomain P450 BM3 complexes was investigated by varying the length of spacing scaffolds between the BMR and BMP domains. The resulting changes in distance between the redox centers of the two domains regulated the efficiency of electron transfer and thus the enzymatic activity of the reconstituted P450 BM3 [121].

2D DNA nanostructures provide an even greater opportunity to organize multienzyme systems into more complicated geometric patterns. Thiolated nucleic acids were covalently linked to glucose oxidase (GOx) and horseradish peroxidase (HRP) by using *N*-[(1-maleimidopropyl)sulphosuccinimide ester] as a bifunctional crosslinker. The GOx/HRP enzyme cascade was organized on 2D hexagonal DNA strips via self-assembly. The distance between two enzymes was controlled by varying the positions of two free DNA tethers on the hexagonal DNA strips. The complementary DNA-conjugated enzymes organized on the two-hexagon strips (shorter distances) showed 1.2-fold higher activity than the four-hexagon strips. With shorter distances, intermediate (H_2O_2) diffusion was more efficient, which therefore resulted in increased cascade reaction efficiency. However, the enzyme cascade was not activated in the absence of the DNA scaffolds or in the presence of foreign DNA [122]. These observations indicate that spatial arrangement at the nanometer scale using a 2D nanostructure comprising a rigid DNA duplex could control the flux of an intermediate from a primary enzyme to a secondary enzyme and that the flux control dominated the multienzyme cascade reaction rate.

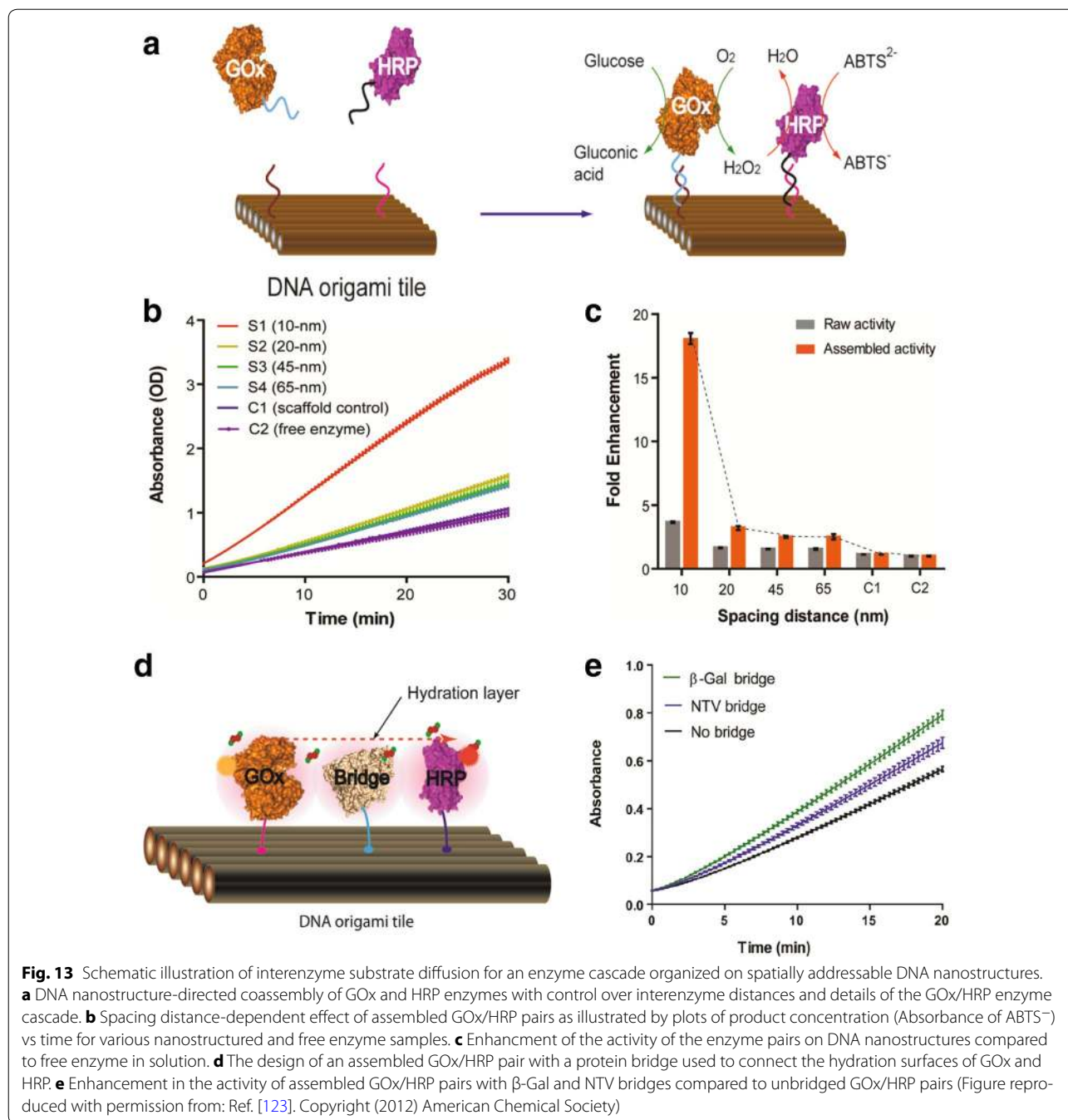
More accurate distance control of the GOx/HRP enzyme cascade was realized using DNA origami tiles as a scaffold. The distance between enzymes was systematically varied from 10–65 nm, and the corresponding activities were evaluated. The study revealed the existence of two different distance-dependent kinetic processes associated with the assembled enzyme pairs. Strongly enhanced activity was observed when the enzymes were closely spaced, while the activity decreased drastically for enzymes as little as 20 nm apart. Increasing the spacing further showed much weaker distance dependence (Fig. 13a–c). This study revealed that intermediate transfer between enzymes might occur at the connected hydration shells for closely spaced enzymes. This mechanism was verified by constructing different sizes of non-catalytic protein bridges (β -galactosidase (β -Gal) and NeutrAvidin (NTV)) between GOx and HRP to facilitate intermediate transfer across protein surfaces. The bridging protein changed the Brownian diffusion, resulting

in the restricted diffusion of H_2O_2 along the hydration layer of the contacted protein surfaces and enhancing the enzyme cascade reaction activity (Fig. 13d, e) [123].

An enzyme cascade nanoreactor was constructed by coupling GOx and HRP using both a planar rectangular orientation and short DNA origami NTs. Biotinylated GOx and HRP were positioned on the streptavidin-decorated planar rectangular DNA sheet via the biotin–avidin interaction with a specific interenzyme distance (i.e., the distance between GOx and HRP) of 15 nm. This DNA sheet equipped with GOx and HRP was then rolled into a confined NT, resulting in the encapsulation of the enzymes in a nanoreactor. Remarkably, the enzymatic coupling efficiency of this enzyme cascade within short DNA NTs was significantly higher than that on the planar rectangular DNA sheet alone. When both enzymes were confined within the DNA NTs, H_2O_2 could not diffuse out of the diffusion layer, which was much thicker than the diameter of the DNA NTs (20 nm), resulting in a high coupling of the reaction intermediate H_2O_2 between the enzymes [124].

A similar modular type of enzyme cascade nanoreactor was constructed using 3D DNA origami building blocks. Each of the DNA origami units contained three biotin-conjugated strands protruding from the inner surface of the tubular structure. The deglycosylated avidin and NTV were immobilized on the inner surface of the units via the biotin–avidin interaction to facilitate the further binding of biotinylated enzymes. Biotinylated GOx and HRP were anchored inside the origami compartment with the help of NTV. The resulting GOx- and HRP-immobilized tubular DNA origami structures were connected together by hybridizing 32 short (3–6 bases) sequences. The GOx/HRP cascade reaction of the assembled dimer nanoreactor showed significantly higher activity than that without a DNA scaffold [125].

Engineered RNA modules were assembled into discrete (0D), one-dimensional (1D) and 2D scaffolds with distinct protein-docking sites (duplexes with aptamer sites) and used to control the spatial organization of a hydrogen-producing pathway in bacteria. The 0D, 1D and 2D RNA scaffolds were assembled in vivo through the incorporation of two orthogonal aptamers for capturing the target phage-coat proteins MS2 and PP7. Cells expressing the designed RNA scaffold modules and both ferredoxin/MS2 (F_M) and [FeFe]-hydrogenase/PP7 (H_P) fusion proteins showed remarkable increases in hydrogen production. Namely, 4-, 11- and 48-fold enhancements in hydrogen production compared with that of control cells were observed from the RNA-templated hydrogenase and ferredoxin cascade reactions in cells expressing 0D, 1D and 2D RNA scaffolds, respectively. This study suggests that a metabolic engineering approach can be used



to introduce structural nucleic acid nanostructures inside cells for the organization of multienzyme reaction pathways [126].

3 Biomolecular engineering for nanobio/bionanotechnology

Biomolecular engineering addresses the manipulation of many biomolecules, such as nucleic acids, peptides, proteins, carbohydrates, and lipids. These molecules are

the basic building blocks of biological systems, and there are many new advantages available to nanotechnology by manipulating their structures, functions and properties. Since every biomolecule is different, there are a number of technologies used to manipulate each one individually.

Biomolecules have various outstanding functions, such as molecular recognition, molecular binding, self-assembly, catalysis, molecular transport, signal transduction, energy transfer, electron transfer, and luminescence.

These functions of biomolecules, especially nucleic acids and proteins, can be manipulated by nucleic acid (DNA/RNA) engineering, gene engineering, protein engineering, chemical and enzymatic conjugation technologies and linker engineering. Subsequently, engineered biomolecules can be applied to various fields, such as therapy, diagnosis, biosensing, bioanalysis, bioimaging, and biocatalysis (Fig. 14).

3.1 Nucleic acid engineering

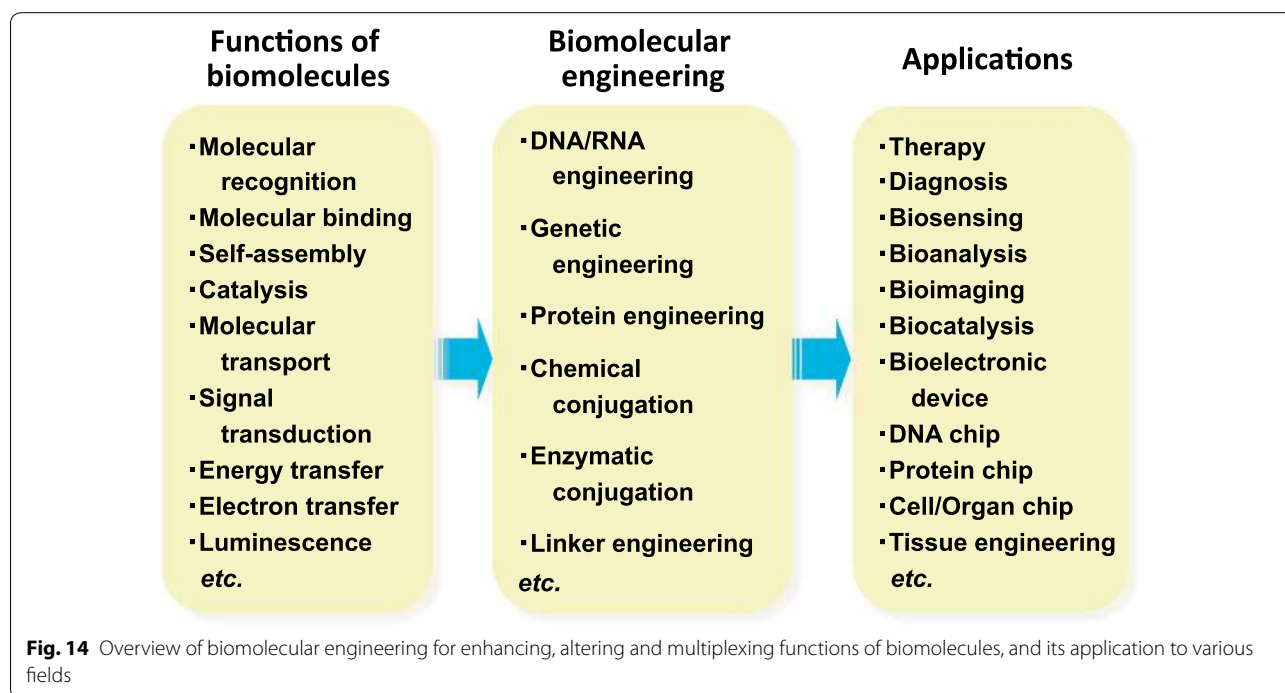
Nucleic acids, such as DNA and RNA, exhibit a wide range of biochemical functions, including the storage and transfer of genetic information, the regulation of gene expression, molecular recognition and catalysis. Nucleic acid engineering based on the base-pairing and self-assembly characteristics of nucleic acids is key for DNA/RNA nanotechnologies, such as those involving DNA/RNA origami, aptamers, and ribozymes [16, 17, 127].

3.1.1 DNA/RNA origami

DNA/RNA origami, a new programmed nucleic acid assembly system, uses the nature of nucleic acid complementarity (i.e., the specificity of Watson–Crick base pairing) for the construction of nanostructures by means of the intermolecular interactions of DNA/RNA strands. 2D and 3D DNA/RNA nanostructures with a wide variety of shapes and defined sizes have been created with precise control over their geometries, periodicities and topologies [16, 128, 129]. Rothemund developed a versatile

and simple ‘one-pot’ 2D DNA origami method named ‘scaffolded DNA origami,’ which involves the folding of a long single strand of viral DNA into a DNA scaffold of a desired shape, such as a square, rectangle, triangle, five-pointed star, and even a smiley face using multiple short ‘staple’ strands [130]. To fabricate and stabilize various shapes of DNA tiles, crossover motifs have been designed through the reciprocal exchange of DNA backbones. Branched DNA tiles have also been constructed using sticky ends and crossover junction motifs, such as tensegrity triangles (rigid structures in a periodic-array form) and algorithmic self-assembled Sierpinski triangles (a fractal with the overall shape of an equilateral triangle). These DNA tiles can further self-assemble into NTs, helix bundles and complex DNA motifs and arrays [17]. 3D DNA origami structures can be designed by extending the 2D DNA origami system, e.g., by bundling dsDNAs, where the relative positioning of adjacent dsDNAs is controlled by crossovers or by folding 2D origami domains into 3D structures using interconnection strands [131]. 3D DNA networks with such topologies as cubes, polyhedrons, prisms and buckyballs have also been fabricated using a minimal set of DNA strands based on junction flexibility and edge rigidity [17].

Because the folding properties of RNA and DNA are not exactly the same, the assembly of RNA was generally developed under a slightly different perspective due to the secondary interactions in an RNA strand. For this reason, RNA tectonics based on tertiary interactions

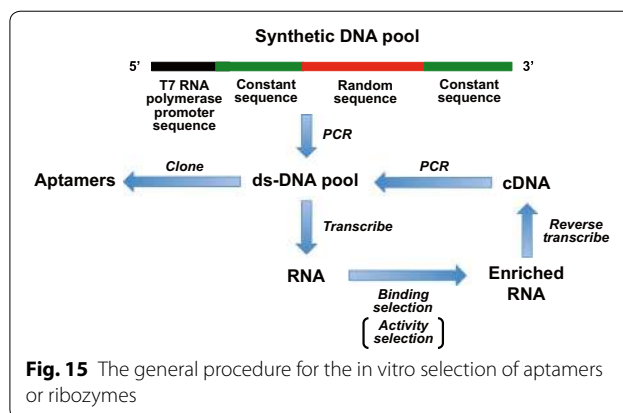


have been introduced for the self-assembly of RNA. In particular, hairpin–hairpin or hairpin–receptor interactions have been widely used to construct RNA structures [16]. However, the fundamental principles of DNA origami are applicable to RNA origami. For example, the use of three- and four-way junctions to build new and diverse RNA architectures is very similar to the branching approaches used for DNA. Both RNA and DNA can form jigsaw puzzles and be developed into bundles [17].

One of the most important features of DNA/RNA origami is that each individual position of the 2D structure contains different sequence information. This means that the functional molecules and particles that are attached to the staple strands can be placed at desired positions on the 2D structure. For example, NPs, proteins or dyes were selectively positioned on 2D structures with precise control by conjugating ligands and aptamers to the staple strands. These DNA/RNA origami scaffolds could be applied to selective biomolecular functionalization, single-molecule imaging, DNA nanorobot, and molecular machine design [131]. The potential use of DNA/RNA nanostructures as scaffolds for X-ray crystallography and nanomaterials for nanomechanical devices, biosensors, biomimetic systems for energy transfer and photonics, and clinical diagnostics and therapeutics have been thoroughly reviewed elsewhere [16, 17, 127–129]; readers are referred to these studies for more detailed information.

3.1.2 Aptamers

Aptamers are single-stranded nucleic acids (RNA, DNA, and modified RNA or DNA) that bind to their targets with high selectivity and affinity because of their 3D shape. They are isolated from 10^{12} to 10^{15} combinatorial oligonucleotide libraries chemically synthesized by in vitro selection [132]. Many protocols, including high-throughput next-generation sequencing and bioinformatics for the in vitro selection of aptamers, have been developed and have demonstrated the capacity of aptamers to bind to a wide variety of target molecules, ranging from small metal ions, organic molecules, drugs, and peptides to large proteins and even complex cells or tissues [39, 133–136]. The general in vitro selection procedure for an aptamer, SELEX (Fig. 15), is as follows: a synthetic DNA pool is prepared by chemical synthesis. DNAs consist of a random or mutagenized sequence region flanked on each end by a constant sequence and with a T7 RNA polymerase promoter at the 5' end. This DNA is amplified by a few cycles of polymerase chain reaction (PCR) and subsequently transcribed in vitro to make the RNA pool. The RNA molecules are then selected based on their binding affinity to the target molecule, for example, by passing them through a target-immobilized affinity column. The retained RNAs are eluted, reverse



transcribed, amplified by PCR, and transcribed; then, the entire cycle is repeated. After multiple rounds of selection (generally 6–18 rounds), quite large populations ($>10^{13}$ different sequences) can be sieved, the ratio of active-to-inactive RNA sequences increases and finally the pool becomes dominated by molecules that can bind the target molecule.

Chemically modified nucleotides provide several advantages, such as enhanced nuclease resistance, an improved binding affinity, increased oligonucleotide pool diversity and improved success rate of selection. Therefore a modified oligonucleotide pool is becoming more popular for aptamer selection. Although chemically modified nucleotides and deoxynucleotide triphosphates cannot be recognized by wild-type T7 RNA polymerases and A-type DNA polymerases, such as *Taq* polymerase, fortunately, modified nucleotide triphosphates (2'-fluoro pyrimidines, 2'-O-methyl nucleotides) and functionalized 2'-deoxynucleotide triphosphates with AA-like residues (e.g., indole, benzyl, or alkyne moieties) can be recognized by some mutant RNA polymerases [137] and B-type polymerases, and *Pwo* and *Vent* (exo-) DNA polymerases [138], respectively.

Specific aptamers against diverse targets have been developed and aptamer-conjugated nanomaterials such as drug-encapsulated polymer NPs, CNTs, AuNPs, QDs and DNA origami demonstrated potential in applications ranging from therapy, targeted drug delivery, sensors and diagnostic reagents to aptamer-directed protein arrays on DNA nanostructures. The details of such applications will not be covered in this review; readers are referred to several recently published reviews [29–31, 40, 64, 68, 132, 139, 140].

3.1.3 Ribozymes

Natural ribozymes are RNA molecules that have enzymatic activity for cleaving phosphodiester linkages. Therefore, ribozymes have significant potential for use

in cancer, genetic disease, and viral therapeutics by specifically inhibiting gene expression through cleaving RNA substrates, such as mRNA, with the viral genome of RNA containing a sequence complementary to the catalytic center of the ribozymes [141].

Natural ribozymes bind to substrate RNAs through Watson–Crick base pairing, which offers the sequence-specific cleavage of substrate RNAs. Two ribozymes, the ‘hammerhead’ ribozyme and the ‘hairpin’ ribozyme, have been extensively studied [142]. The catalytic motif of a ribozyme is surrounded by a flanking sequence that is responsible for ‘guiding’ the ribozyme to its target RNA and giving stability to the structure. With the hammerhead ribozyme, cleavage is dependent on divalent metal ions, such as magnesium, and can occur after any NUH triplet (where N = any nucleotide and H = A, C or U) within the target RNA sequence. The kinetics of the reaction can vary significantly (up to one or more orders of magnitude) with different triplet-flanking sequence combinations; thus, the choice of an appropriate ribozyme cleavage site is the first and most important step in hammerhead ribozyme design [143].

Artificial ribozymes with catalytic properties have been isolated by *in vitro* selection from random or combinatorial nucleic acid libraries. Variations of the aptamer selection strategies can be used to isolate catalytic nucleic acid sequences by changing the binding selection step of the aptamer selection process to an activity selection step (Fig. 15). Such approaches have been used to change the function of known ribozymes and to create completely new ones from a random or combinatorial nucleic acid pool [144]. A broad range of chemical reactions could be catalyzed, such as the formation, cleavage and rearrangement of various types of covalent bonds. Examples including not only the cleavage or ligation of RNA substrates by phosphoester transfer at the phosphorus center [144, 145] but also Diels–Alder reactions, N-glycosidic bond formation, alkylations, acylations, and amide bond formations at the carbon centers [144, 146] have been reviewed. The catalytic performance, nuclease resistance and diversity of the oligonucleotide pools of ribozymes could also be enhanced by the incorporation of chemically modified nucleotides, as utilized in aptamer selection protocols [146].

Ribozymes can be expressed from a vector, which offers the advantage of the continued intracellular production of these molecules. However, the turnover rates of ribozymes are rather low in some cases, since dissociation from the cleavage product is the rate-limiting step that controls their usefulness. Furthermore, some ribozymes require high divalent metal ion concentrations for efficient substrate cleavage, which may limit their use in intracellular environments. All of these concerns, as

well as off-target activity, resistance to serum and cellular nucleases, and cell-specific, targeted delivery, need to be addressed and overcome in order to utilize ribozymes in therapies. Ribozymes can be hardly incorporated into cells in their naked forms and often required a vehicle for efficient delivery. Many classes of nanomaterials including cationic liposomes, cationic polymer micelles [147] and spherical nucleic acids composed of inorganic core and densely packed, highly oriented nucleic acid shell [148] have been used as delivery vehicles to prevent nuclease-dependent degradation and to enhance cell-targeting and intracellular transduction [143].

3.2 Gene engineering

Gene engineering is a powerful tool for creating artificial genes for proteins and enzymes with desired, improved and multiple properties such as molecular recognition, molecular binding, self-assembly, catalysis, molecular transport, signal transduction, energy transfer, electron transfer, and luminescence, which contribute to develop novel nanobiomaterials, nanobiodevices and nanobiosystems. This technology has been employed to evolve genes *in vitro* through an iterative process consisting of recombinant generation. Coupled with the powerful HTS or selection methods, gene engineering has been widely applied to solve problems in protein engineering. This technology includes technologies for direct gene manipulation, such as gene mutagenesis, DNA sequence amplification [e.g., PCR and rolling circle amplification (RCA)], DNA shuffling and gene fusion.

There are many methods to generate genetic diversity and to create combinatorial libraries. For example, a variety of *in vitro* gene manipulation techniques have been developed over the past decade that allow various types of directed changes in a gene by modifying (inserting, deleting or replacing) one or more codons (gene mutagenesis), swapping domains between related functional gene sequences (DNA shuffling) and fusing domains from different functional gene sequences (gene fusion), resulting in the creation of diverse collections of mutant gene clones. There are two main types of mutagenesis, i.e., random and site-directed mutagenesis.

3.2.1 Random mutagenesis

With random mutagenesis, point mutations are introduced at random positions in a gene of interest, typically through error-prone PCR mutagenesis, in which $MnCl_2$ is added to the reaction mixture to cause a reduction in the fidelity of the DNA amplification [149]. The modified error-prone PCR method, which achieves higher frequencies of base substitutions and both transition and transversion mutations, was developed using mixtures of triphosphate derivatives of nucleoside analogs

[150, 151]. An error-prone RCA method, which is an isothermal DNA amplification method with the addition of $MnCl_2$ to the reaction mixture, was also developed for random mutagenesis [152]. Different in vitro chemical mutagenesis methods have also been used to introduce random mutations into a gene of interest. In these methods, bases of DNA are modified by chemical mutagens, such as nitrous acid, bisulfate, hydroxylamine and ethyl methane sulfonate, and these methods have less bias than does mutagenesis using PCR-based methods [153]. Randomized sequences are then cloned into a suitable expression vector, and the resulting mutant libraries can be screened to identify mutants with altered or improved properties.

3.2.2 Site-directed mutagenesis

Site-directed mutagenesis is a method for altering a gene sequence at a selected location by using overlapping extension PCR. Point mutations, insertions, or deletions are introduced by incorporating DNA primers containing the desired modification with a DNA polymerase in an amplification reaction. Site-saturation mutagenesis further allows the substitution of predetermined protein sites against all twenty possible AAs at once by employing degenerate primers in which the three bases of the targeted codon are replaced by mixtures, most commonly NNN or NNK (N = A, C, G or T; K = G or T). A completely randomized codon, NNN, results in a library size of 64 different sequences encoding all 20 AAs and 3 stop codons. On the other hand, NNK codons reduce the library size by half, still encoding 20 AAs, with the advantage of having only one stop codon. In this configuration, the AAs W, F, I, Y, Q, N, H, K, D, E, M and C are encoded by a single codon, while G, A, V, P, and T, and L, S, and R are encoded by two and three codons, respectively [154].

3.2.3 DNA shuffling

DNA shuffling is a method for the in vitro recombination of homologous genes to quickly generate a large library of chimeric progeny genes incorporating sequence fragments from a number of parent genes by random fragmentation through DNase I and PCR extension without primers for reassembly; this process is followed by PCR amplification with primers to generate full-length chimeras suitable for cloning into an expression vector (Fig. 16a) [155]. One significant drawback of this DNA-shuffling method is the low frequency of chimeric genes in the shuffled library, which may be due to the homo-duplex formation of DNA fragments derived from the same parental genes at the annealing step, the probability of which is much higher than that of hetero-duplex formation. To address this problem, a modified

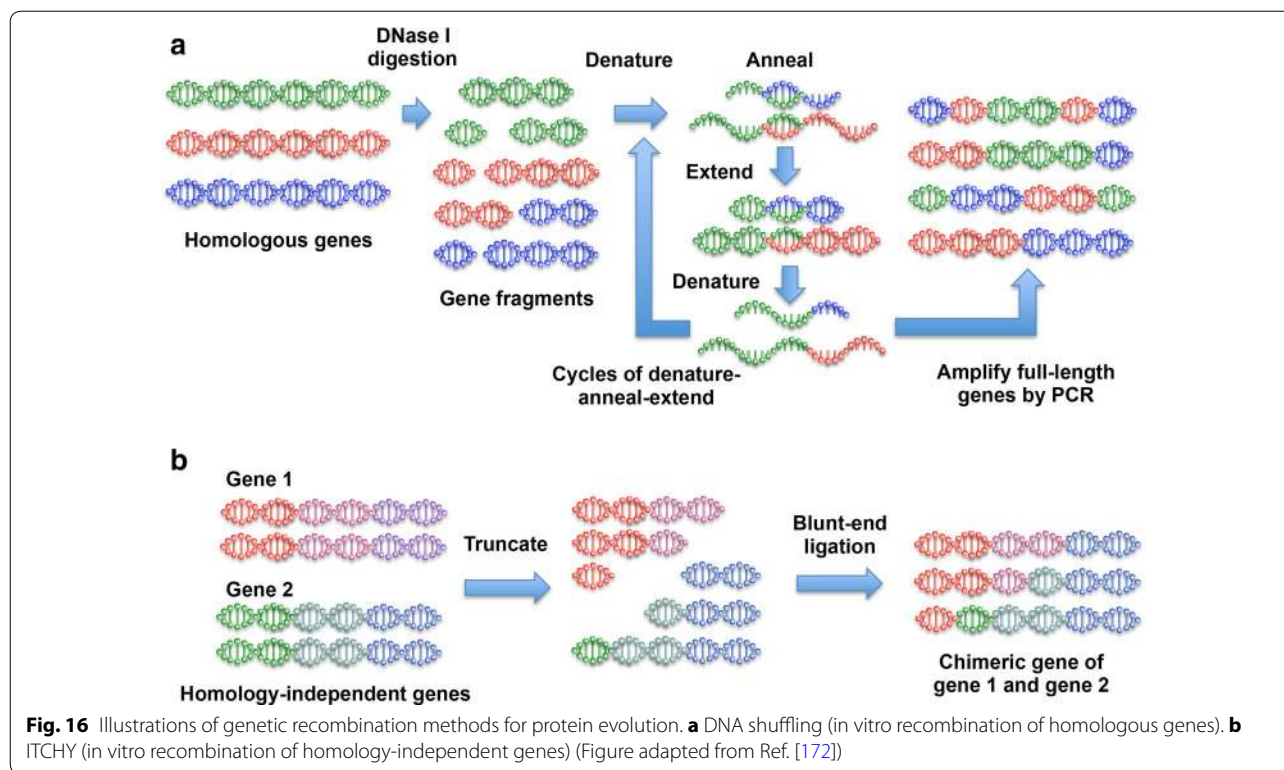
DNA-shuffling method can be used; this method involves the fragmentation of the parental genes using restriction enzymes rather than DNase I [156] or uses single-stranded DNA (ssDNA) templates rather than dsDNA templates for DNase I fragmentation [157]. Since the use of ssDNA as templates will decrease the probability of homo-duplex formation, the percentage of the parental genes in the shuffled library should be significantly reduced.

DNA shuffling has been extended to distantly or completely unrelated gene families, which require methods that do not rely on homologous recombination because of the degree of sequence divergence. Sequence homology-independent protein recombination [158] and incremental truncation for the creation of hybrid enzymes lead to the formation of chimeric genes (Fig. 16b) [159]. The rearrangement of these chimeras by shuffling yields functional hybrids [160]. The main advantage of these methods is that knowledge about detailed protein structure is not required [161].

Exon shuffling is a natural molecular mechanism for the formation of new eukaryotic genes. New exon combinations can be generated by recombination within the intervening intron sequences, yielding new rearranged genes with altered functions. The natural process of exon shuffling can be mimicked in vitro by generating libraries of exon-shuffled genes and subsequently screening target DNA from libraries [162]. In this method, exons or combinations of exons that encode protein domains are amplified by PCR using mixtures of chimeric oligonucleotides that determine which exons are spliced together. By means of a self-priming overlap polymerase reaction, mixtures of these PCR fragments are combinatorially assembled into full-length genes. Recombination is performed by connecting an exon from one gene to an exon from a different gene. In this way, two or more exons from different genes can be combined together ectopically, or the same exon can be duplicated, to create a new exon–intron structure.

3.2.4 Gene fusion

Fusion genes are created by genetically fusing the open reading frames of two or more genes in-frame through ligation or overlap extension PCR. To construct such fusion genes, two types of connection are possible. One is 'end-to-end' fusion, in which the 5' end of one gene is linked to the 3' end of the other gene. The second is insertional fusion, in which one gene is inserted in-frame into the middle of the other parent gene [163]. These methods offer various advantages for producing fusion genes with high throughput in different orientations and including linker sequences to maximize the performance of fusion partners [164].



3.3 Protein engineering

The field of protein engineering has always played a central role in biological science, biomedical research, and biotechnology. Protein engineering is also indispensable technology to design useful and valuable building blocks for nanobio/bionanotechnology to fabricate a variety of artificial self-assembled protein systems with nanoscale structures [18, 19], proteins with tagged peptides for immobilization on NPs [94] and engineered proteins for applications to bioelectronic devices [23, 26, 27], therapy [42, 44, 45, 67, 165], bio-imaging [67, 166], biosensing [83, 97, 167], and biocatalysis [87, 89, 95, 98, 101, 103, 108, 110–116]. There are two general strategies for protein engineering, i.e., rational protein design and directed evolution (high-throughput library screening- or selection-based approaches) (Fig. 17).

3.3.1 Rational protein design

In rational protein design (Fig. 17, the left panel), detailed knowledge of the structure and function of a protein is used to make desired changes to the protein. In general, this approach has the advantage of creating functionally improved proteins easily and inexpensively, since site-directed mutagenesis techniques allow precise changes in AA sequences, loops and even domains in proteins

[161]. However, the major drawback of protein redesign is that detailed structural knowledge of a protein is often unavailable, and, even when it is available, substitutions at sites buried inside proteins are more likely to break their structures and functions. Therefore, it is still very difficult to predict the effects of various mutations on the structural and functional properties of the mutated protein, although many studies have been done to predict the effects of AA substitutions on protein functions [168]. Another rational protein design method is computational protein design, which aims to design new protein molecules with a target folding protein structure, novel function and/or behavior. In this approach, proteins can be designed by transcendently setting AA sequences compatible with existing or postulated template backbone structures (de novo design) or by making calculated variations to a known protein structure and its sequence (protein redesign) [169]. Rational protein design approaches make predicted AA sequences of protein that will fold into specific 3D structures. Subsequently, these predicted sequences should be validated experimentally through the chemical synthesis of an artificial gene, followed by protein expression and purification. The details of computational protein design methods will not be covered in this review; readers are referred to several recently published reviews [170, 171].

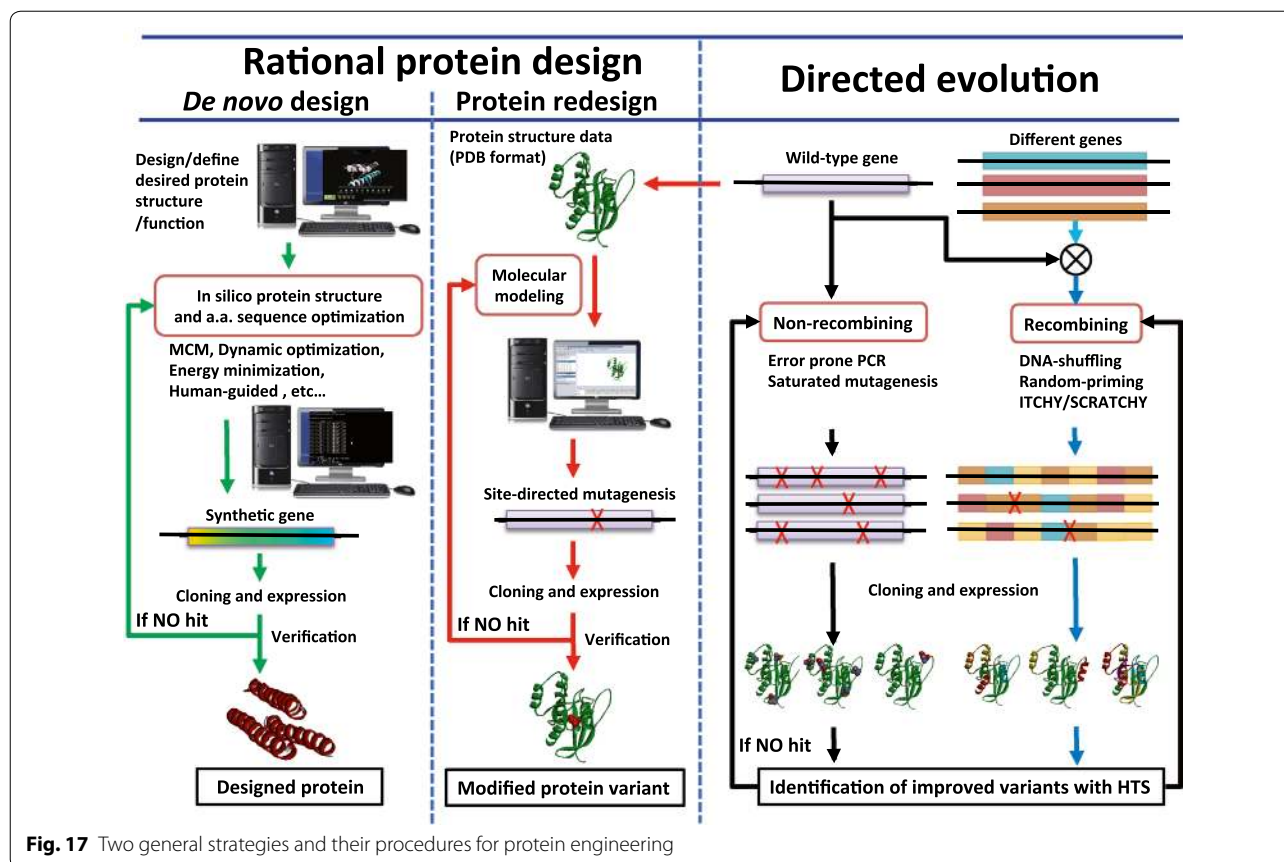


Fig. 17 Two general strategies and their procedures for protein engineering

3.3.2 Directed evolution (protein engineering based on high-throughput library screening or selection)

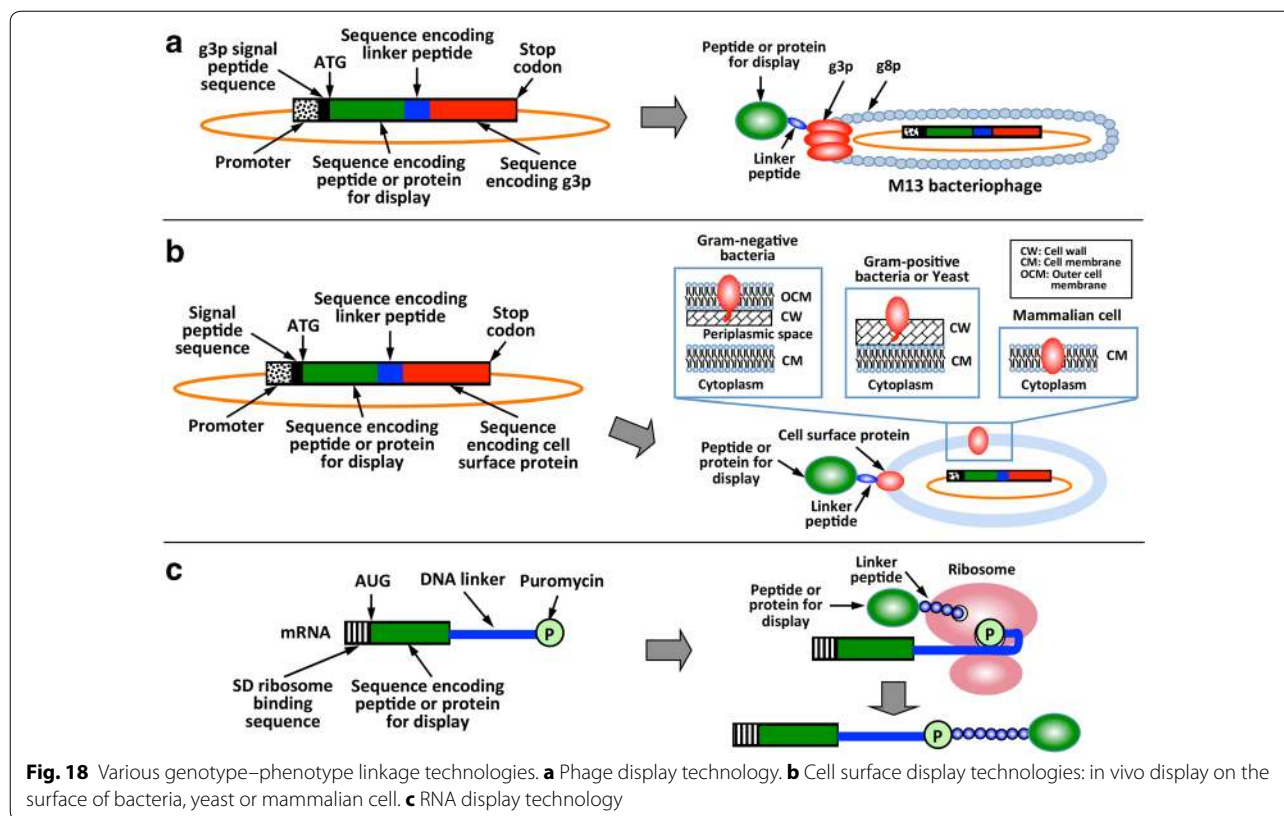
The directed evolution approach (Fig. 17, the right panel) involves many technologies, such as gene library diversification, genotype–phenotype linkage technologies, display technologies, cell-free protein synthesis (CFPS) technologies, and phenotype detection and evaluation technologies [172]. This approach mimics the process of natural selection (Darwinian evolution) to evolve proteins toward a target goal. It involves subjecting a gene to iterative rounds of mutagenesis (creating a molecular library with sufficient diversity for the altered function), selection (expressing the variants and isolating members with the desired function), and amplification (generating a template for the next round). This process can be performed *in vivo* (in living cells), or *in vitro* (free in solutions or microdroplets). Molecular diversity is typically created by various random mutagenesis and/or *in vitro* gene recombination methods, as described in “Gene engineering”.

Functionally improved variants are identified by an HTS or selection method and then used as the parents for the next round of evolution. The success of directed evolution depends on the choices of both

diversity-generation methods and HTS/selection methods. The key technology of HTS/selection methods is the linkage of the genotype (the nucleic acid that can be replicated) and the phenotype (the functional trait, such as binding or catalytic activity). Aptamer and ribozyme selection from nucleic acid libraries can be performed much faster than those of functional proteins because the nucleic acids themselves have binding or catalytic activities (i.e., selectable phenotypes), such that the genotype and phenotype are identical. However, since proteins cannot be amplified, it is necessary to have a linkage between the phenotype exhibited by the protein and the genotype (mRNA or DNA) encoding it to evolve proteins.

Many genotype–phenotype linkage technologies have been developed; these link proteins to their corresponding genes (Fig. 18) [172–174]. Genotype–phenotype linkage technologies can be divided into *in vivo* and *in vitro* display technologies. *In vitro* display technologies can be further classified into RNA display and DNA display technologies.

In vivo display technology includes phage display [175] and baculovirus display [176], in which a protein gene designated for evolution is fused to a coat protein gene and expressed as a fusion protein on the surface of phage



and virus particles. Cell surface display technologies are also in vivo display technologies and use bacteria [177, 178], yeast [179, 180] and mammalian cells [181] as host cells, in which the fusion gene resulting from a protein gene and a partial (or full) endogenous cell surface protein gene is expressed and displayed on the cell surface. These in vivo display technologies can indirectly link a protein designated for evolution and its gene through the display of the protein on biological particles or cells. However, the library sizes of in vivo display technologies are usually restricted to the 10^8 – 10^{11} size range by the efficiency of the transformation and transduction steps of their encoding plasmids.

In vitro display technologies are based on CFPS systems. Recent advances in CFPS technologies and applications have been reviewed elsewhere [182–185]. RNA display technology includes mRNA display and ribosome display [186]. mRNA display covalently links a protein to its coding mRNA through a puromycin linker that is covalently attached to the protein via ribosome-catalyzed peptide bond formation. Ribosome display noncovalently links a protein to its coding mRNA genetically fused to a spacer sequence lacking a stop codon through a ribosome because the nascent protein does not dissociate from the ribosome. Such display technologies using in vitro translation reactions can screen proteins that would be

toxic to cells and can cover quite large libraries (10^{15}) by bypassing the restricted library size bottleneck of in vivo display technologies (Table 1).

There are several in vitro DNA display technologies, such as CIS display [187], *M. Hae* III display [188], STABLE display [189], microbead display [190] and in vitro compartmentalization (IVC) [191]. CIS display noncovalently links RepA (DNA-binding protein) fusion protein and its coding DNA template through the interaction between RepA and the CIS element of the DNA template. For *M. Hae* III display, the DNA methyltransferase *M. Hae* III covalently links a protein and its DNA template. IVC technology uses the aqueous droplets in water–oil emulsions to compartmentalize individual genes and gene products. STABLE display and microbead display technologies utilize noncovalent biotin–streptavidin binding to link biotin-labeled DNA templates and streptavidin-fused proteins.

The details of HTS and selection methods, such as fluorescence-activated cell sorting-based phenotype detection and evaluation technologies coupled with these display technologies [172, 192–195], as well as the applications of the directed evolution of enzymes, antibodies, receptors and other proteins in such areas as environmental issues, catalysis, gene therapy, and therapeutic protein and vaccine development will not be covered in

Table 1 Various display technologies

Technology (typical number of sequences screened per library)	Description	Strengths or weaknesses
Bacterial cell display (10^8 – 10^9)	Fusion gene libraries of the target proteins and bacterial surface proteins Fusion proteins are displayed on bacterial cell surface	Selects proteins displayed on bacterial cell surfaces Flow cytometry allows multiparameter, quantitative screening <i>Smaller library size</i> <i>Cannot screen proteins that would be toxic to cells</i>
Yeast or mammalian cell display (10^8 – 10^{10})	Fusion gene libraries of the target protein and cell surface proteins of yeast or mammalian cells Fusion proteins are displayed on cell surface	Selects proteins displayed on eukaryotic cell surfaces Flow cytometry allows multiparameter, quantitative screening <i>Smaller library sizes</i> <i>Cannot screen proteins that would be toxic to cells</i>
Phage or baculovirus display (10^{11})	Fusion gene libraries of the target protein and phage or virus coat proteins Infected bacteria produces phage or virus particles displaying fusion protein libraries on the surface	Robust and quick <i>Cannot screen proteins that would be toxic to cells</i>
Ribosome display (10^{15})	mRNA-target protein complexes are displayed on stalled ribosomes in cell free protein synthesis system Reverse-transcription PCR allows amplification after rounds of selections	Large library size Can screen proteins that would be toxic to cells <i>Requires stringent conditions and stable proteins</i>
mRNA display (10^{15})	mRNA-target protein fusions are synthesized in cell free protein synthesis system by conjugating them through a puromycin linker Reverse-transcription PCR allows amplification after rounds of selections	Large library size Can screen proteins that would be toxic to cells <i>Works well with small proteins but not large ones</i> <i>Requires stringent conditions</i>

this review; readers are referred to several recently published articles and reviews [42, 172, 175, 177, 180, 181, 186, 187, 193–208].

Recent, significant advances in protein engineering have come through computational methods, such as SCHEMA, ProSAR, and ROSETTA. Computational design based on these methods greatly decreases the need for probing randomized sequence space, rendering the route to novel biocatalysts much more efficient [195, 209–212]. Therefore, in the future, more detailed knowledge about the relationship between protein structures and functions, as well as advancements in high-throughput technology, may greatly expand the capabilities of protein engineering.

3.4 Chemical and enzymatic conjugation technologies

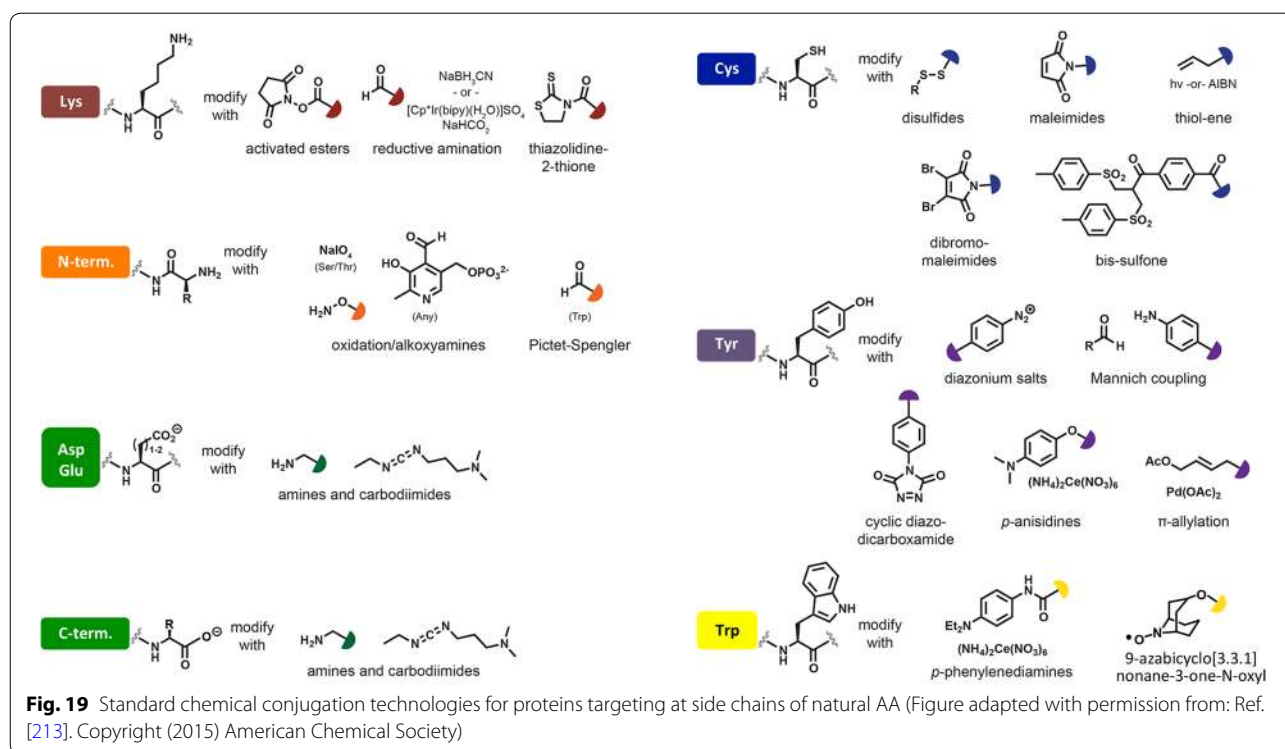
In the current postgenomic era, many studies require chemically modified proteins or protein bioconjugates that are impossible to prepare via standard ribosomal synthesis. Conjugation technologies to site-specifically modify proteins with diverse natural and unnatural functionalities have been developed in the last two decades. These technologies have been widely utilized to fabricate hybrid biomolecular material, such as protein/protein, protein/peptide, protein/nucleic acid, protein/lipid, protein/oligosaccharide, and protein/ligand hybrids, and hybrid materials comprising biomolecules and inorganic/

organic materials for use in nanobio/bionanotechnology. These technologies range from classical chemical bioconjugation technologies targeting natural AAs to more sophisticated approaches, such as unnatural AA (UAA) incorporation based on amber stop codon suppression, bioorthogonal chemical conjugations, protein chemical ligations and enzymatic conjugations.

3.4.1 Chemical conjugation technologies targeting natural AAs

Standard chemical conjugation technologies for proteins target the side chains of natural AAs, such as the primary amine groups (R–NH₂) of Lys residue and the N-terminus, the carboxylic acid groups (R–COOH) of Asp, Glu and the C-terminus, the thiol group (R–SH) of Cys, the phenyl ring of tyrosine (Tyr) and the indole ring of tryptophan (Trp) (Fig. 19) [213].

Lys is one of the most common AA residues in proteins with an average abundance of approximately 6% and is often surface-exposed due to its hydrophilicity; therefore, it is an excellent target site for conjugation. On the other hand, the N-terminus provides a more site-selective location but is not always surface-exposed. The primary amine of Lys has been predominantly functionalized with *N*-hydroxysuccinimidyl-esters (NHS-esters), NHS-ester sulfates or isothiocyanates. In these electrophilic reagents, NHS-esters are highly used for primary



amine-targeted functionalization because of the reaction simplicity. A limitation of NHS-esters is a side reaction of hydrolysis in water (<5 h half-life), which accelerates as the pH increases above 7. This hydrolysis competes with desired reactions and reduces reaction efficiency [214].

The N-terminus can be selectively targeted for modification when it is sufficiently accessible and not post-translationally modified. The transamination reaction mediated by pyridoxal-5-phosphate can be applied to the modification of the N-terminal residue without the presence of toxic Cu(II) or denaturing organic cosolvents, although proteins possessing N-terminal serine (Ser), threonine (Thr), Cys, or Trp residues will be incompatible with this method because of known side reactions with aldehydes [215].

Asp and Glu are also the most common AA residues in naturally occurring proteins; they have an average abundance of approximately 12%, are often surface-exposed and are excellent target conjugation sites. The carboxylic acid side chains of Asp, Glu and the C-terminus can be functionalized by carbodiimide chemistry, typically using EDC, which has been widely used for covalently crosslinking a carboxylic acid and amine. However, the relatively high abundance of Lys, Asp and Glu and the high solvent accessibility of their side chains make it impossible to modify a single site on the protein surface using these methods.

Cys is not definitively hydrophilic or hydrophobic, and it is an attractive residue site for directed target

conjugation because its average abundance in naturally occurring proteins is estimated to be approximately 1%. The relatively low abundance of Cys facilitates the genetic modification of the protein sequence to introduce a unique Cys. The nucleophilic side chain of Cys can be site-selectively targeted to create a well-defined conjugate. At slightly basic pH levels, the thiolate moiety can be modified with disulfides, maleimides, thiol-ene, dibromo-maleimides or bis-sulfone. Modification with disulfide (under mild oxidative condition) and maleimide (Michael addition) reagents produces disulfide and thiosuccinimide bond linkages that are not stable in the presence of free thiols, such as reduced glutathione (GSH) abundant in the cytoplasm of cells [213]. This GSH-sensitive conjugation property has been positively utilized for the release of drug delivery system payloads in the cytoplasm. In contrast, the ring-opening hydrolysis of thiosuccinimide using maleimide derivative incorporating a basic amino group adjacent to the maleimide, positioned to provide intramolecular catalysis of thiosuccinimide ring hydrolysis, yields a stable conjugate (e.g., an antibody–drug conjugate) [216].

Methods for the conjugation of Tyr, which has an average abundance of 3% in proteins, have also been developed. In the presence of strong oxidizing agents (e.g., H₂O₂) and appropriate catalysts, the phenolic side chain of the Tyr residue can crosslink with other phenolic compounds. The oxidizing agents required to catalyze these

reactions are not discerning, and there is concern over causing undesired side reactions to other portions of proteins. To overcome this problem, a Tyr coupling reaction has been developed; it involves an electrophilic reagent, imines formed in situ from aldehydes and electron-rich anilines. This three-component Mannich-type coupling reaction is highly selective for Tyr and proceeds under mild conditions [217].

Conventional methods for the conjugation of Trp, which has an average abundance of approximately 1%, require toxic heavy metals or biochemically incompatible conditions. Some of these methods also exhibit cross reactivity with other AAs (particularly Tyr), thus limiting the range of applications. Recently, a transition metal-free method using 9-azabicyclo[3.3.1]nonane-3-one-*N*-oxyl (keto-ABNO) for the conjugation of Trp was reported. This new method showed novel features, such as high Trp selectivity, the formation of single conjugates with high homogeneity, facile conjugation at an ambient temperature and nearly neutral pH and a short reaction time [218].

3.4.2 Chemical conjugation technologies targeting UAAs

Although the site-specific substitution of AAs with rare AAs, such as Cys or Tyr, provides multiple options for the covalent functionalization of proteins, it is possible to damage the folding or function of the proteins when the genetic modifications for such chemical conjugations become too extensive. The incorporation of UAAs into proteins allows for greater flexibility in protein modifications. Unlike the natural AA residues, these UAAs can contain entirely unique reactive moieties, such as azide, cyano, iodo, bromo, boc, and dansyl groups, and thereby afford completely bioorthogonal reactivity that can occur inside of living systems without interfering with native biochemical processes. The site-specific incorporation (SSI) of UAAs utilizes nonsense codons to incorporate one or a few UAAs into a single or multiple defined locations in a protein. The amber stop codon suppression is commonly employed, in which the target-location codon is mutated to an amber stop codon [219]. To incorporate the UAA, a tRNA and aminoacyl tRNA-synthetase (aaRS) are engineered to pair with the UAA and recognize the amber codon as a sense codon. SSI substitutes a natural AA with an unnatural analog, where the endogenous aminoacyl tRNA-synthetase (aaRS) accepts the UAA and aminoacylates the appropriate tRNA with the UAA. Incubating an auxotrophic expression host in culture media containing the analog UAA instead of the specific natural AA leads to UAA incorporation at nearly every location where the specific natural AA would have been incorporated [220]. The SSI of UAAs has been further extended to CFPS systems [219, 221].

The incorporation of multiple different UAAs has been achieved by the extension of codon-anticodon pairs using a different four-base codon for each tRNA [222]. Technology using acylating ribozyme (flexizyme) instead of ssRS has been developed for in vitro semi-enzymatic synthesis and acylation [223]. Therefore, SSI is minimally invasive and allows the incorporation of any UAA into a specific site of a protein with minor effects.

3.4.3 Bioorthogonal chemical conjugation technologies

UAA-incorporated proteins or peptides can be chemoselectively conjugated with other biomolecules and synthetic inorganic/organic materials bearing bioorthogonal functional groups. Other biomolecules, such as nucleic acids, lipids, and synthetic inorganic/organic materials modified with bioorthogonal functional groups, can also be subjected chemoselectively to a bioconjugation reaction. This reaction should occur in an aqueous solution at near physiological pH, have rapid kinetics even with submillimolar concentrations of reactants, and occur at physiological temperatures. The chemical reactions that satisfy these requirements include the ketone/hydroxyamine condensations, Huisgen [3 + 2] cycloaddition, the Staudinger ligation, the Diels–Alder cycloadditions and a photo-Click cycloadditions (Fig. 20) [224, 225].

3.4.3.1 Conjugation reactions through aldehydes and ketones The carbonyl groups of aldehydes and ketones can specifically react with strong α -effect nucleophiles, such as hydrazides ($R-N-NH_2$) and alkoxyamines ($R-O-NH_2$), under acidic conditions (pH 4–6) to produce stable hydrazones and oximes, respectively. Since these ketone/aldehyde condensations show rather slow kinetics with second-order rate constants, large excesses of the conjugation reagent are necessary in order to achieve good labeling efficiency. In general, ketone/aldehyde condensations are best suited for conjugation reactions in vitro or at the cell surface because the reaction requires an acidic pH and high concentrations of the labeling reagent, which are problematic in terms of cell toxicity [224].

3.4.3.2 Conjugation reactions through azides The azide group is truly orthogonal in its reactivity to the majority of biological functionalities. The Huisgen 1,3-dipolar cycloaddition of alkynes and azides has been widely adopted since this reaction is both selective and high yielding when catalyzed by Cu(I) for Cu-chelating propylene derivatives or by strain release for strained cycloheptyne derivatives. Another cycloaddition reaction, the inverse-electron-demand Diels–Alder reaction between tetrazines and strained alkenes or alkynes, yields dihydropyridazines or pyridazines with nitrogen gas as the only byproduct. These reactions have recently been explored

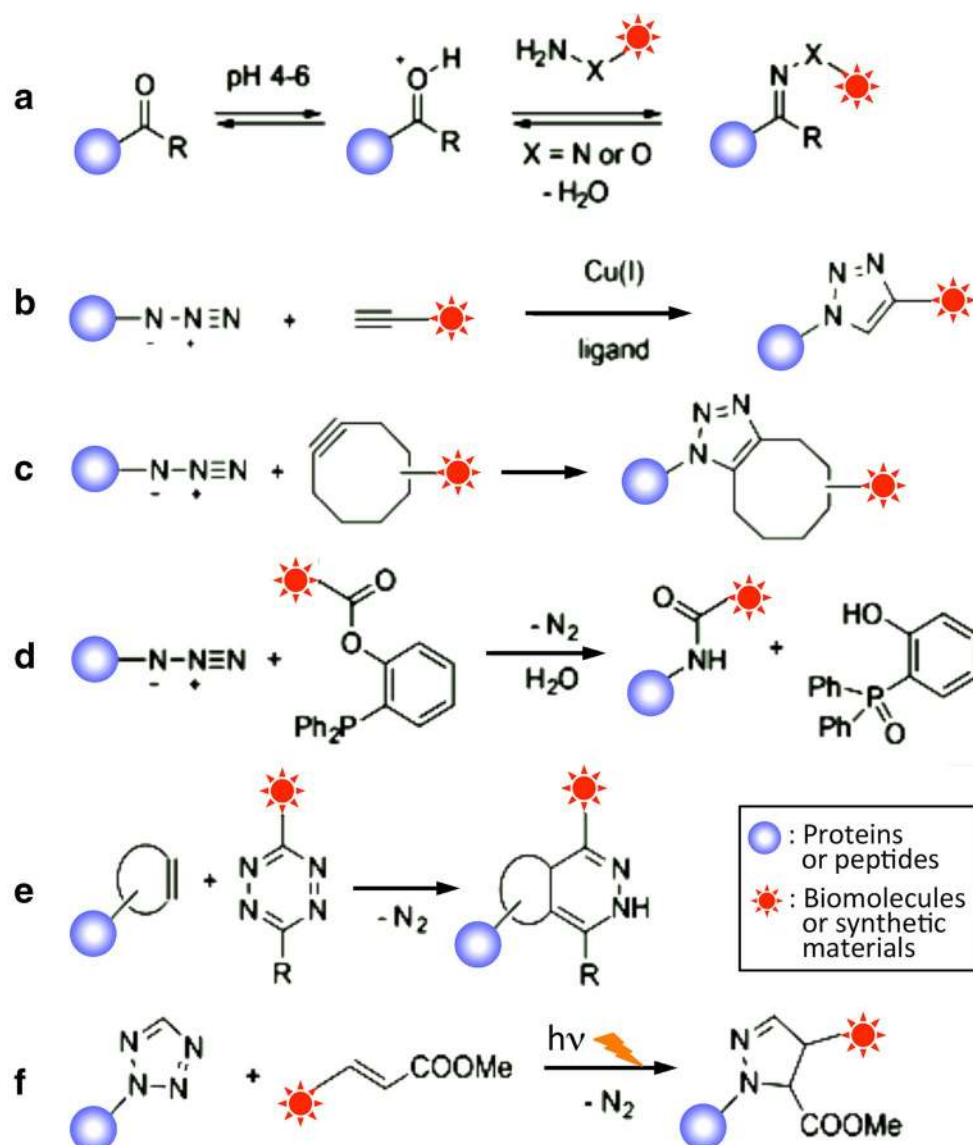


Fig. 20 Chemoselective bioconjugation reactions. **a** Ketone/hydroxylamine condensations. **b** Copper-catalyzed alkyne-azide Huisgen cycloadditions. **c** Strain-promoted alkyne-azide cycloadditions. **d** Staudinger ligation. **e** Diels-Alder cycloadditions. **f** Photo-click cycloadditions (Figure adapted with permission from: Ref. [224]. Copyright (2014) American Chemical Society)

as chemoselective reactions for labeling and manipulating biomolecules in their native states. The reactions are extraordinarily fast (up to $10^5 \text{ M}^{-1} \text{ s}^{-1}$) and have improved second-order kinetics relative to the chelating Cu(I)-catalyzed azide-alkyne cycloaddition ($10\text{--}200 \text{ M}^{-1} \text{ s}^{-1}$) [224]. The 1,2,3-triazole linkage formed in the cycloaddition reaction (click reaction) is thermodynamically and hydrolytically stable. This reaction is insensitive to aqueous conditions and pH levels ranging from 4 to 12, succeeds over a broad temperature range, and tolerates a wide variety of functional groups. Pure products can be easily isolated by

simple filtration or extraction without chromatography or recrystallization. Many other bioorthogonal conjugation reactions have been reported; readers can refer to recent reviews [224, 225].

3.4.4 Chemical ligation technologies

Native chemical ligation (NCL) has become the most general and robust method for the conjugation of protein-peptide, protein-protein, protein-DNA, and protein-NP materials because it is simple, general, and has a high yield efficiency [226]. NCL is a chemoselective coupling reaction

that generates a native peptide bond by a reversible transthioesterification between a peptide fragment containing an N-terminal Cys residue (α -Cys) and another peptide fragment bearing a C-terminal α -thioester group, followed by an irreversible intramolecular N-S acyl shift (Fig. 21). This reaction proceeds efficiently under physiological conditions and is compatible with all natural AA side chains. Therefore, through the recombinant preparation of proteins having an α -Cys residue, NCL can be used to generate proteins containing modifications at their N-termini. It is advantageous to conduct NCL in an aqueous solution at a neutral pH even though a C-terminal α -thioester derivative can be competitively hydrolyzed. Recent extensions of NCL, such as ligation rate acceleration, chemoselective post-ligation modifications, and the streamlined ligation of multiple peptide fragments, have been reviewed [227].

Expressed protein ligation (EPL) and protein trans-splicing (PTS) are both intein-based chemical conjugation technologies that permit the assembly of a protein from smaller synthetic and/or recombinant unprotected polypeptide building blocks (Fig. 22). An intein is an internal protein domain that can autocatalytically excise itself from a precursor protein. The cis-splicing of intein by the addition of high concentrations of thiol derivatives

can be used to generate a C-terminal α -thioester of a protein from protein-intein fusion. In EPL, one or more of the peptides is of recombinant origin, but the actual ligation step is still a chemical process and can be performed under a wide range of reactions to introduce a variety of functional materials, such as fluorophores, UAAs, isotopic labels, and post-translational modifications, into a large number of proteins [228]. By contrast, PTS post-translationally links two recombinant protein fragments. An intein domain is split into two fragments (split intein or trans-splicing intein), Int^N and Int^C , which are fused to the flanking polypeptides, termed the N and C exteins (Ex^N and Ex^C). The ligation step in PTS must be performed under conditions compatible with protein folding because the process involves the functional reconstitution of a split intein. In this step, Ex^N - Int^N and Int^C - Ex^C associate, fold to form a functional intein, restore autocatalytic protein splicing activity to excise the Int^N - Int^C , and ligate the flanking Ex^N and Ex^C with a peptide bond of Cys.

Although the advances in NCL, EPL and PTS made it possible to precisely introduce a variety of functional materials into peptides and proteins, these technologies also have some drawbacks, as follows. (1) The

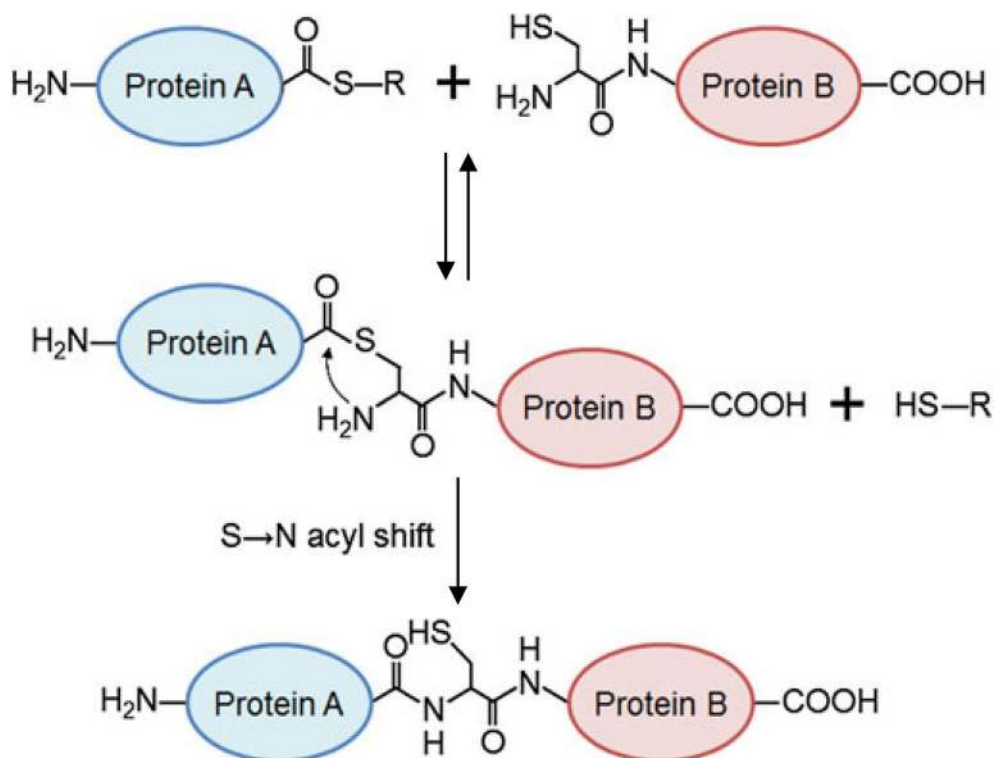


Fig. 21 Native chemical ligation. Native chemical ligation (NCL) is a chemoselective coupling reaction that links a peptide fragment containing an N-terminal Cys (α -Cys) residue and another peptide fragment bearing a C-terminal α -thioester group by a native peptide bond (Figure reproduced with permission from: Ref. [106]. Copyright (2012) Springer)

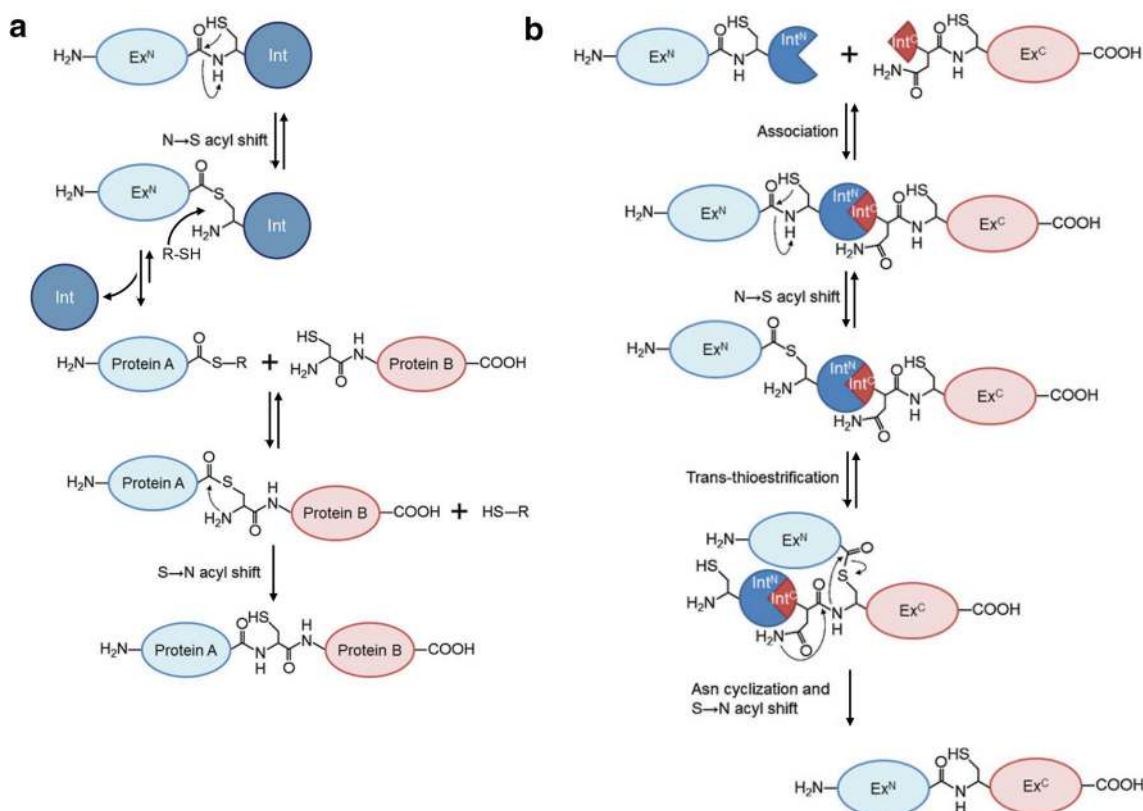


Fig. 22 Intein-based chemical conjugation. **a** Expressed protein ligation (EPL) is a semisynthetic version of NCL in which synthetic and recombinant polypeptides are chemically ligated together. Proteins (A) expressed as intein fusions can be cleaved from the intein with a variety of thiols to give the corresponding α -thioester derivative. Proteins (B) containing N-terminal Cys can be made recombinantly by masking the Cys with a protease tag that can be later removed. **b** Protein trans-splicing (PTS) post-translationally links two protein fragments. An intein domain is split into two fragments, Int^N and Int^C , which are fused to the flanking exons, Ex^N and Ex^C . Ex^N - Int^N and Int^C - Ex^C associate and fold to form a functional intein. This functional intein can restore protein splicing activity to excise itself, and to conjugate Ex^N and Ex^C with a peptide bond (Figures adapted with permission from: Ref. [106]. Copyright (2012) Springer)

preparation of synthetic peptide α -thioesters is still technically difficult. (2) Since the ligation process is a chemical reaction, the higher concentrations of both or either of the reactants are required. (3) The application of EPL to many disulfide bond-containing proteins is restricted or complicated since the use of high concentrations (usually more than several tens of mM) of thiol derivatives is needed to induce thiolysis of the protein-intein fusions. (4) The expression of intein-based fusion proteins often results in the formation of inclusion bodies due to the large protein sizes and poor solubility, which requires additional refolding steps.

3.4.5 Enzymatic conjugation technologies

In nature, numerous proteins are post-translationally modified by enzymes and play important roles in controlling cellular processes, such as metabolism, signal transduction, gene expression, and cell differentiation. These enzymes participating in post-translational modifications

catalyze the covalent addition of some chemical groups (e.g., phosphate, acetate, amide, and methyl groups and biotin, flavins, carbohydrates and lipids) to the N- or C-terminus or a side chain of an AA residue at specific site in a protein; these enzymes can also catalyze the cleavage and ligation of peptide backbones in proteins. Natural post-translational modifications of proteins are generally efficient under physiological conditions and site-specific. Therefore, a variety of transferase or ligase enzymes have been repurposed for site-specific protein modification. Typically, a small tag peptide sequence incorporated into the target protein is recognized by the post-translational modification enzyme as a substrate and then transfers functional moieties from an analog of its natural substrate onto the tag (Fig. 23). Examples include formylglycine-generating enzyme (FGE), protein farnesyltransferase (PFTase), *N*-myristoyltransferase (NMTase), biotin ligase (BirA), lipoic acid ligase (LAL), microbial transglutaminase (MTGase), sortase A (SrtA),

glutathione *S*-transferase (GST), SpyLigase, and several engineered self-labeling protein tags. Except for self-labeling protein tags, a primary benefit of this approach is the small size of the peptide tag that must be incorporated into proteins, which ranges from 3 to 15 residues. Some enzymes only recognize the tag peptide at a specific position in the primary sequence of the protein (often the N- or C-terminus), while others are not inherently limited by tag position.

Enzymatic protein conjugation technologies, including non-site-specific crosslinking by such oxidoreductases as peroxidase, laccase, tyrosinase, lysyl oxidase, and amine oxidase, are reviewed elsewhere [105]. Here, we briefly review recent enzymatic conjugation technologies for site-specific protein conjugation and crosslinking of biomolecules and synthetic materials. The applications of enzymatic conjugations and modifications of proteins with other biomolecules and synthetic materials are

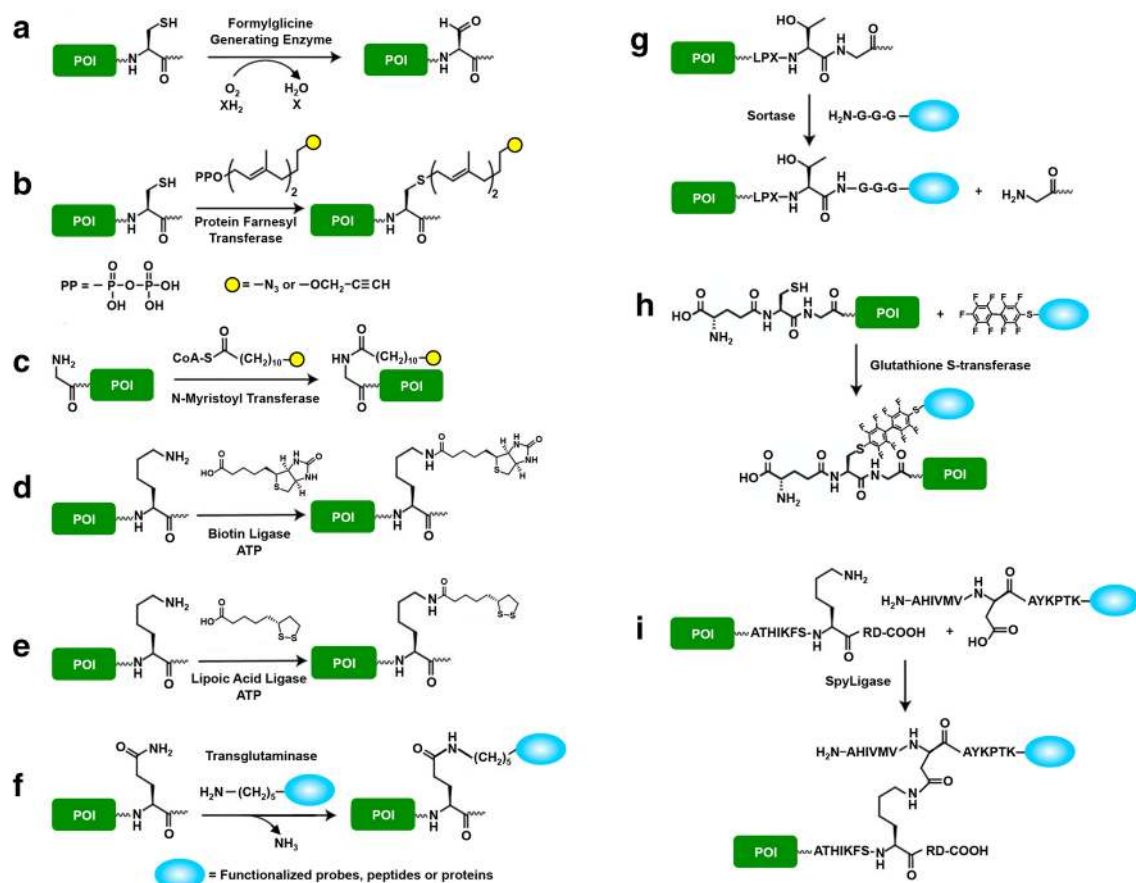


Fig. 23 Chemoenzymatic labeling strategies of the protein of interest (POI) using post-translational modification enzymes. **a** Formylglycine generating enzyme (FGE) recognizes LCXPXR peptide motif and converts the side chain of Cys residue into an aldehyde group. The POI fused to the aldehyde tag can be further functionalized with aminoxy or hydrazide probes. **b** Farnesyltransferase (FTase) recognizes the four AAs sequence $\underline{C}A_1A_2X$ (A_1 and A_2 are non-charged aliphatic AAs and X is C-terminal Met, Ser or Phe) at the C-terminus and catalyzes the attachment of the farnesyl isoprenoid group to the Cys residue. The POI can be further labeled by bioorthogonal chemical conjugation of the farnesyl moiety functionalized with azide or alkyne. **c** *N*-Myristoyl transferase (NMT) recognizes the $\underline{G}XXXS$ peptide motif at the N-terminus and attaches a myristate group to an N-terminal Gly residue. The POI can be further labeled by bioorthogonal chemical conjugation of myristate moiety functionalized with azide or alkyne. **d** Biotin ligase recognizes the GGLNDIFEAQKIEWH peptide motif derived from biotin carboxyl carrier protein and catalyzes the transfer of biotin from an ATP intermediate (biotinyl *S*'-adenylate) to Lys residue. Biotinylated POI can then be labeled with streptavidin conjugated with a variety of chemical probes. **e** Lipoic acid ligase recognizes the GFEEIDKVVYDLDA peptide motif and catalyzes the attachment of lipoic acid or its derivatives to Lys residue in the motif. **f** Transglutaminase (TGase) catalyzes the transamination reaction and forms an iso-peptide bond between Gln in POI and Lys residue-functionalized small molecule probes, peptides or proteins. **g** Sortase cleaves LPXTG peptide tag fused to POI between Thr and Gly residue and conjugates oligo Gly-functionalized small molecule probes, peptides or proteins to POI by forming a peptide bond between Thr and Gly residues. **h** GST catalyzes Cys arylation and conjugates probes bearing a 4-mercaptoperfluorobiphenyl moiety to the N-terminal γ -Glu-Cys-Gly sequence of POI. **i** SpyLigase catalyzes the generation of an isopeptide bond between Lys residue in KTag and Asp residue in SpyTag

limited to recombinant proteins harboring additional protein/peptide tags. However, protein functionalization using enzymatic conjugations is a promising method because it is achieved simply by mixing proteins without special techniques. The details of enzymatic conjugation technology applications will not be covered in this review; readers are referred to several recently published reviews [229–232].

3.4.5.1 FGE The FGE oxidizes Cys or Ser residue to formylglycine (FGly) within a conserved 13-AA consensus sequence found in prokaryotic Type I sulfatases. The modification is thought to occur co-translationally, before protein folding. The consensus sequence can be incorporated into heterologous proteins expressed in *E. coli*, where it is modified efficiently by a co-expressed bacterial FGE. Furthermore, the minimized core motif sequence CX(P/A)XR or SXPXR, derived from the most highly conserved portion of the FGE recognition site, directed the efficient conversion of Cys or Ser to FGly. The aldehyde-bearing residue FGly can be subsequently used for covalent conjugation using complementary aminoxy- or hydrazide-functionalized moieties by ketone-reactive chemistries (Fig. 23a) [233].

3.4.5.2 PFTase PFTase is an α/β heterodimer enzyme that catalyzes the transfer of a farnesyl isoprenoid group from farnesyl pyrophosphate (FPP) through a thioether bond to a sulfur atom on a Cys in a tetrapeptide sequence (denoted as a CA₁A₂X-box, here C is Cys, A₁ and A₂ are aliphatic AAs, and X is one of a variety of AAs) four residues from the C-terminus (Fig. 23b). Since PFTase can tolerate many simple modifications to the aldehyde-containing isoprenoid substrate, it can be used to introduce a diverse range of functionalities into proteins containing a CA₁A₂X-box positioned at the C-terminus. Subsequent chemoselective reactions with the resulting protein can then be used for a wide range of applications. The catalytic activity of PFTase toward various FPP analogs has been greatly improved by site-directed mutagenesis around the substrate-binding pocket of PFTase [234].

3.4.5.3 NMTase NMTase from *Candida albicans* catalyzes the acyl transfer of myristic acid from myristoyl-CoA to the amino group of an N-terminal glycine (Gly) residue of a protein to form an amide bond. NMTase recognizes the sequence GXXX(S/T), where X can be any AA (Fig. 23c). This enzyme can successfully transfer alkyne- and azide-containing myristic acid analogs that incorporated the bioorthogonal groups at the distal end of the lipid to the N-terminal Gly residue of recombinant proteins containing an N-terminal myristoylation motif. This method provides a convenient and potentially gen-

eral method for N-terminal-specific recombinant protein labeling [235].

3.4.5.4 BirA BirA from *E. coli* catalyzes the adenosine triphosphate (ATP)-dependent amide bond formation between the carboxylic group of biotin and the ϵ -amino group of a Lys in an acceptor peptide sequence (23 AA residues) (Fig. 23d). This acceptor sequence was further optimized to a 15-AA acceptor peptide sequence (GLNDIFEAKQKIEWHE) [236]. BirA can be used to site-specifically conjugate a biotin moiety to recombinant proteins by the genetic fusion of the BirA recognition acceptor peptide sequence with the target protein. The enzymatic biotin labeling to a protein allows the subsequent formation of very strong noncovalent conjugate with avidin due to the low dissociation constant between biotin and avidin ($\sim 10^{-15}$ M). Another orthogonal acceptor sequence for yeast BirA has been further developed to enable two-color imaging [237]. The substrate tolerance of BirA was also expanded to biotin analogs, including ketone, azide, and alkyne groups, which contain alternative functionalities suitable for bioorthogonal reactions [238].

3.4.5.5 LAL LAL from *E. coli* catalyzes the ATP-dependent amide bond formation between the carboxylic group of lipoic acid and the ϵ -amino group of a lysine in an optimized 13-AA recognition acceptor sequence (GFEIDKLVWYDLDA) (Fig. 23e). The Trp 37 residue at the lipoic acid-binding pocket of LAL was substituted with small AA residues to accept a wider range of lipoic acid analogs containing an aliphatic azide, aryl-aldehyde, or aryl-hydrazine moiety [239]. These lipoic acid analogs are attached to a Lys residue in the acceptor sequence of a protein and are then used to conjugate diverse functional molecules by bioorthogonal reactions.

3.4.5.6 MTGase Transglutaminase is a unique enzyme that catalyzes the acyl-transfer reaction between the γ -carboxamide group of a Gln residue in proteins and a wide variety of unbranched primary amines, commonly the ϵ -amino group of a Lys residue, and forms an isopeptide bond between the side chain of Gln residues and primary amines (Fig. 23f). Because this conjugation reaction is irreversible, involves the release of ammonia and proceeds quickly even under low temperature conditions (~ 15 °C), the conjugation product is stable, and a high yield can be obtained. MTGase is isolated from *Streptomyces mobaraensis*, which is widely used in the food industry, and recognizes various peptide sequences consisting of Gln residues. A notable correlation was observed between the polypeptide chain regions of high temperature factor (B-factor) determined crystallographically and the MTGase attacking sites, thus indicating

the role of polypeptide chain mobility or local unfolding in dictating site-specific enzymatic modifications [240]. Consequently, enhanced MTGase polypeptide chain flexibility limits the enzymatic reaction with Gln residues on rigid polypeptide in globular proteins. Therefore, it is possible to predict the site(s) of Gln residue modifications by MTGase on the basis of local structure and dynamics of polypeptide chain containing Gln residue.

Because of its openness with regard to the primary amine substrate, MTGase is an attractive catalyst for generating protein conjugates with small functional molecules, lipids, nucleic acids, synthetic polymers, e.g., PEG, peptides and even other proteins. Although the substrate specificity of MTGase toward the polypeptide sequence containing a Gln residue (Q-tag) has not yet been clarified, the Q-tag derived from the polypeptides of globular proteins, the ribonuclease S-peptide (KETAAKFERQHMS and its Lys to Ala-substituted peptide AETAAAKFERQHMS), the F-helix peptide of horse heart myoglobin (PLAQSH) or the designed N-terminal oligo-Gly tag (N-Gly₅), which are recognized as a Gln-substrate by MTGase, can be utilized as Q-tag substrates [108, 241–244]. For protein modification by MTGase, these Q-tags are incorporated at the N- or C-terminus or inside the loop region of proteins by genetic means. Subsequently, MTGase can site-specifically conjugate the Q-tag in the protein with a primary amine-containing short synthetic linker or a Lys residue-containing polypeptide tag (KTag) harboring a functional moiety. However, one of the drawbacks of conjugating proteins possessing many Lys and Gln residues is that the activity of MTGase toward Gln and Lys residues makes it difficult to control the site(s) of modification.

3.4.5.7 SrtA SrtAs are cell envelope-bound housekeeping transpeptidases from gram-positive bacteria. SrtA attaches surface proteins, such as virulence factors, to the penta-Gly motif of branched lipid II, the peptidoglycan precursor. SrtA recognizes the peptide sequence (LPXTG) and catalyzes the cleavage of the amide bond between the Thr and Gly residues by means of an active site Cys residue (Cys184) (Fig. 23g). This process generates a covalent acyl-enzyme intermediate. The carboxyl group of the Thr of the thioester intermediate then undergoes nucleophilic attack by an amino group of the oligo-Gly substrates, producing ligated products and altering the primary structure. Recent reports have demonstrated that the ϵ -amino group of Lys residues can also act as a nucleophile instead of the ϵ -amino group of oligo-Gly [245]. Since both of the optimized recognition peptide sequences, LPETGG [246] and oligo-Gly with more than two repeats [247], for SrtA-

mediated transpeptidation are very short, these motifs can be easily incorporated into proteins or polypeptides either by standard genetic means or chemical peptide synthesis. Benefiting from its simplicity and specificity, a soluble truncated *Staphylococcus aureus* SrtA that lacks the N-terminal membrane-anchoring motif has begun to be applied for a wide variety of protein engineering and bioconjugation purposes, including the in situ site-specific fluorescent labeling of membrane proteins [248–252] and the fabrication of an electrochemically active protein bilayer on electrodes [253].

Unfortunately, since this conjugation reaction is reversible and the acyl-enzyme intermediate is hydrolyzed by water even in the presence of enough oligo-Gly nucleophiles, the conjugation reaction does not proceed to completion. However, we have overcome this limitation by introducing a β -hairpin structure around the ligation site of products and preventing substrate recognition by SrtA, thereby successfully stabilizing conjugation products and providing a high yield [254]. *S. aureus* SrtA needs Ca²⁺ for stabilizing the active site conformation, and its strong Ca²⁺ dependency makes *S. aureus* SrtA difficult for use under low Ca²⁺ concentrations and in the presence of Ca²⁺-binding substances. To overcome this problem, we designed an *S. aureus* SrtA heptamutant (P94R/E105K/E108A/D16N/D165A/K190E/K196T) that exhibited a high Ca²⁺-independent catalytic activity and successfully catalyzed a selective protein–protein ligation in living cells, which usually retain low Ca²⁺ concentrations [255]. These recent advances in *S. aureus* SrtA-mediated ligation will contribute to the development and design of many other protein conjugates and multienzyme complexes both in vitro and in vivo.

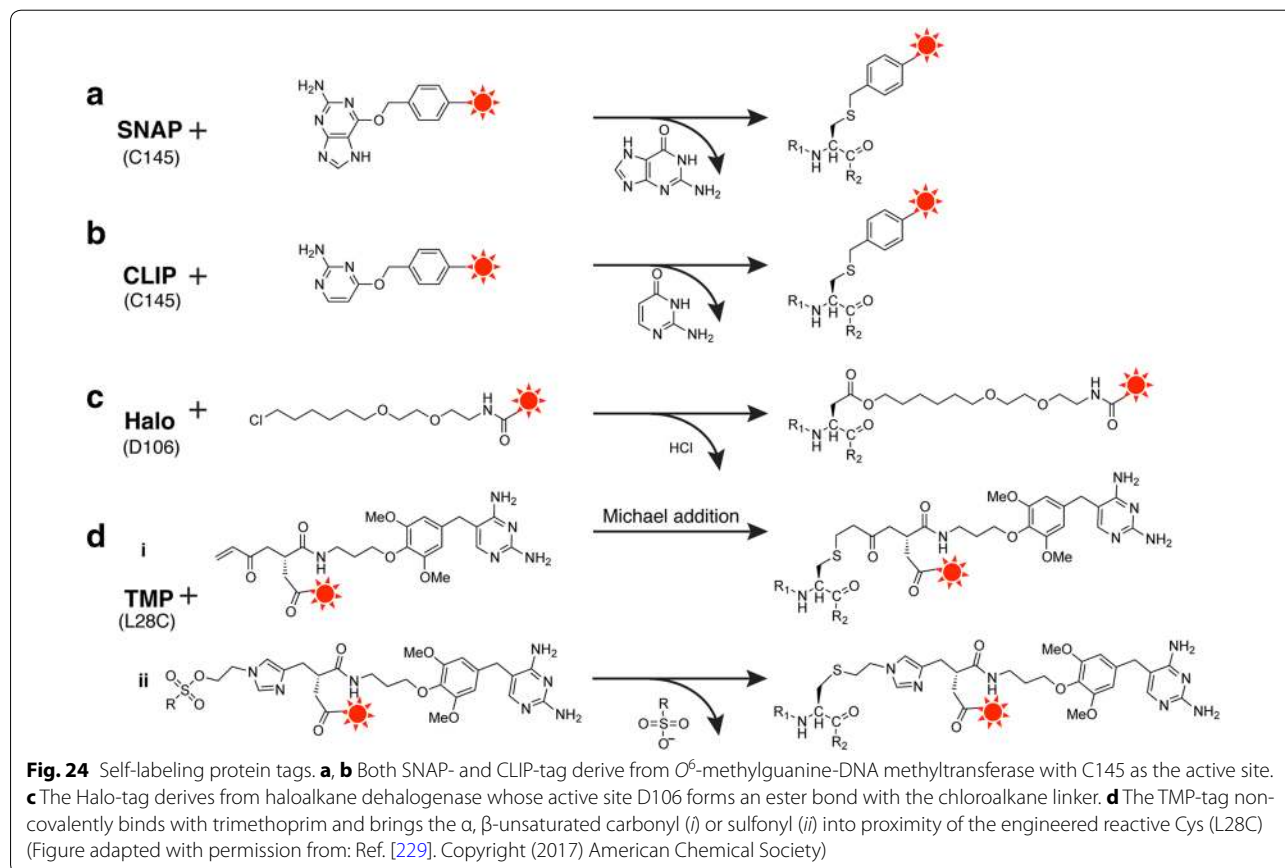
3.4.5.8 GST GST catalyzes conjugation reactions between the Cys residue of glutathione (GSH, γ -Glu-Cys-Gly) and various electrophiles and allows the cell to detoxify xenobiotics in vivo (Fig. 23h). The ubiquitous nature of GST facilitates this bioconjugation with polypeptides bearing an N-terminal GSH in aqueous media and enables the chemo- and regioselective functionalization of a single Cys thiol group of GSH based on a nucleophilic aromatic substitution reaction between Cys residues and perfluoroarenes, even in the presence of other unprotected Cys residues and reactive functional groups on the same polypeptide chain. This conjugation reaction can be carried out over a wide range of temperatures (4–60 °C) and in co-solvent system with the addition of organic solvents (up to 20%) [256]. However, this technology is currently limited to peptide-based couplings due to the requirement for both an N-terminal γ -Glu-Cys-Gly sequence and a perfluoraryl reaction partner.

3.4.5.9 SpyLigase SpyLigase is an artificial ligase obtained by engineering a domain (CnaB2) from the fibronectin adhesion protein FbaB of *Streptococcus pyogenes* (Spy), which is essential for the bacteria to invade human cells. Within CnaB2, there is a post-translational modification to form an isopeptide bond between Lys31 and Asp117 residues, which is catalyzed by an apposed Glu77 residue. Based on the 3D structure and isopeptide bond formation mechanism of CnaB2, the domain was rationally split into three parts, SpyTag (AHIVMVDAYKPTK), KTag (ATHIKFSKRD) and SpyLigase (11 kDa, containing the catalytic Glu77 residue). SpyLigase was derived from CnaB2 first by the removal of SpyTag and KTag, and then by circular permutation via replacing residues from the C-terminus of CnaB2 with a Gly/Ser linker, followed by N-terminal CnaB2 residues. SpyLigase not only can ligate KTag and SpyTag fused at the C- or N-terminus of peptides but can also direct the ligation of KTag to SpyTag inserted in the middle of a protein (Fig. 23i). The yield of conjugation products decreased from approximately 50–10% by elevating the reaction temperature from 4 to 37 °C, likely due to a dynamic change in the secondary structure of SpyLigase [257].

3.4.6 Self-labeling protein tag-based chemoenzymatic conjugation technologies

Chemoenzymatic labeling methods exploit the exquisite molecular recognition mechanism between substrates/inhibitors and enzymes to create a new specific covalent linkage between them by engineering enzymes (Fig. 24) [229].

3.4.6.1 SNAP-tag SNAP-tag (20 kDa) was derived from the human DNA repair protein *O*⁶-alkylguanine-DNA alkyl-transferase (AGT). The normal function of AGT is to repair *O*⁶-alkylated guanine in DNA by transferring the alkyl group in an SN2 reaction to a reactive Cys145 residue in AGT. The repair mechanism is unusual because the protein is irreversibly inactivated. Consequently, the reaction of AGT-fusion proteins with *O*⁶-benzylguanine (BG) derivatives harboring functional moieties leads to the irreversible and covalent labeling of the fusion proteins since the functional moieties on BG are transferred along with the benzyl group of BG to the reactive Cys, creating a stable thioether covalent bond. The SNAP-tag-mediated labeling of proteins in bacteria and yeast is specific, since the respective endogenous AGTs do not accept



BG as substrates, whereas AGT-deficient cell lines should be used for labeling in mammalian cells [258].

3.4.6.2 CLIP-tag Subsequently, AGT mutant-based CLIP-tag, which reacts specifically with O^2 -benzylcytosine (BC) derivatives, was developed by directed evolution. To generate a mutant library of AGT, AA residues at positions with indirect proximity to BG bound in the active site were chosen with the aid of the crystal structure of wild-type AGT. After two-step library screenings using yeast and phage display, CLIP-tag, the eight-point mutant of AGT (Met60Ileu, Tyr114Glu, Ala121Val, Lys131Asn, Ser135Asp, Leu153Ser, Gly157Pro, Glu159Leu) was selected. CLIP-tag with potent catalytic activity exhibited a 105-fold change in substrate specificity and a 100-fold greater preference for BC over BG [259]. The mutual orthogonality of the SNAP- and CLIP-tags enables the simultaneous labeling of multiple proteins in the same cellular context.

3.4.6.3 HaloTag *Rhodococcus* haloalkane dehalogenase (DhaA) removes halides from aliphatic hydrocarbons by a nucleophilic displacement mechanism. A covalent ester bond is formed during catalysis between an Asp106 residue in the enzyme and the hydrocarbon substrate. The base-catalyzed hydrolysis of this covalent intermediate subsequently releases the hydrocarbon as an alcohol and regenerates the Asp106 nucleophile for additional rounds of catalysis. The based-catalyzed cleavage is mediated by a conserved His272 residue located near the Asp106 nucleophile. HaloTag (33 kDa) was derived from a mutant DhaA, whose catalytic His272 residue is substituted with a Phe residue and does not exhibit the enzymatic activity of intermediate cleavage. However, the apparent binding rates of haloalkanes to this mutant are low compared to those of common affinity-based interactions, such as biotin–streptavidin, potentially hampering the practical utility of this mutant as a protein tag. To overcome this issue, several variants with dramatically improved binding rates were identified using a semi-rational strategy, protein–ligand binding complex modeling, site-saturation mutagenesis, and HTS for faster binding kinetics. A mutant with three point substitutions, Lys175Met/Cys-176Gly/Tyr273Leu, i.e., HaloTag, has a high apparent second-order rate constant, thus allowing the labeling reaction to reach completion even under low haloalkane ligand concentrations [260]. Covalent bond formation between the HaloTag and chloroalkane linker (14 atoms long with 6 carbon atoms proximal to the terminal chlorine) functionalized with small synthetic molecules is highly specific, occurs rapidly under physiological conditions and is essentially irreversible. Therefore, the HaloTag-fused pro-

tein can be covalently labeled with a variety of functional group-modified chloroalkane linkers and can be applied to a wide range of fluorescent labels, affinity handles, or solid supports.

3.4.6.4 Trimethoprim (TMP)-tag TMP-tag (18 kDa) was derived from *E. coli* dihydrofolate reductase (eDHFR), which binds the small-molecule inhibitor TMP with high affinity (~ 1 nM K_D) and selectivity (affinities for mammalian DHFRs are $K_D > 1$ μ M). The first-generation TMP-tag harnessed the high-affinity interaction between eDHFR and TMP to form long-duration and yet reversible binding without covalent bond formation. The second-generation, engineered, self-labeling TMP-tag (Leu28Cys) exploited a proximity-induced Michael addition reactivity between a Cys28 residue engineered on the eDHFR surface near the TMP binding site and a mild electrophile, such as an α , β -unsaturated carbonyl moiety, e.g., the β -carbon of acrylamide, or a sulfonyl group installed on the TMP derivatives. To optimize the positioning of the Cys residue nucleophile and the acrylamide electrophile of the TMP derivatives, the site of point mutation on the eDHFR surface and the atom length of the spacer between the 4'-OH group of the TMP and the reactive β -carbon of the acrylamide functional group were investigated based on the molecular modeling of the eDHFR and TMP derivative complexes. After subsequent combinatorial screening in vitro, the combination of the TMP-tag (Leu28Cys) and the TMP derivatives with a 10-atom spacer was selected and exhibited superior specificity and efficiency in protein labeling with fluorophores for live cell imaging [261]. Since the covalent TMP-tag is based on a modular organic reaction rather than a specific enzyme modification, it is easier to build additional features into the covalent TMP-tag.

Self-labeling protein tags, such as SNAP-, CLIP-, Halo- and TMP-tags, feature exquisite specificity and broad applicability to the areas of subcellular protein imaging in live cells, the fabrication of protein–DNA, protein–peptide and protein–protein complexes, and protein immobilization on solid materials, but they are limited by their large molecular size (20–30 kDa) and expensive substrate derivatives, except for HaloTag.

3.5 Linker engineering

Linker engineering is also an important technology for controlling the distances, orientations and interactions among functional components crosslinked in conjugates. Linkers are indispensable units for the fabrication of multidimensional biomaterials or complexes of bio/organic/inorganic materials. Such linkers can be classified as chemical or biological linkers, such as oligonucleotides or polypeptides.

3.5.1 Chemical linkers

Chemical linkers have been widely used to modify or crosslink biomolecules, such as proteins, peptides, nucleic acids and drugs, synthetic polymers and solid surfaces with functional molecules and materials. Chemical linkers can be characterized by the following properties: chemical specificity, reactive groups, spacer arm length, water solubility, cell membrane permeability, spontaneously reactive or photoreactive groups, and cleavability by such stimuli as pH, redox, and light. Particularly, spacer arm length and water solubility are important parameters for protein modifications and crosslinking using chemical linkers. For example, when biomolecules are functionalized with small molecules, such as fluorophores or bioorthogonal functional groups, rigid, short methylene arms are utilized as spacers. Various photocleavable, short chemical linkers were also developed to control the functions of crosslinked biomolecules [54, 262, 263]. In contrast, when proteins are functionalized with hydrophobic or large materials, hydrophilic, flexible, long spacer arms formed from PEG chains are often utilized to increase the water solubility of functionalized chemical linkers and to avoid steric hindrance between proteins and functionalized materials. We utilized PEG chains as chemical linkers to prepare a Fab'-green fluorescent protein (GFP) immunoconjugate for a homogeneous immunoassay [264], an enzyme-streptavidin conjugate for enzyme activity control [265, 266], and a *Synechocystis* sp. DnaB intein-TMP conjugate for in vitro protein ligation [267], and the results showed that the length of the PEG chemical linkers affected both the conjugation efficiency and the controllability of protein function. We also produced antibody-lipid and peptide-lipid conjugates for cell surface display [268–270] using PEG chain linkers. Although there are enormous bioconjugation applications for biomolecules using chemical linkers, the details of recent applications are reviewed elsewhere [271–279].

3.5.2 Biological linkers

3.5.2.1 Oligonucleotide linkers In the bottom-up fabrication of nanoscale systems, synthetic DNA oligonucleotides are extraordinarily useful as a construction unit. The extremely high specificity of Watson–Crick base pairing allows one to readily design DNA linkers by using the predictable adenine–thymine (A–T) and guanine–cytosine (G–C) hydrogen-bonding interaction between complementary nucleic acids. In practice, short DNA oligomers with approximately 10–30 nucleotides (mostly 21 nucleotides forming a 7-nm long base pair segment) have been utilized as linkers to noncovalently conjugate complementary oligonucleotide-modified materials by hybridization and facilitate the fabrication of a wide variety of

programed structures [117–120, 280]. These DNA linkers have been utilized to immobilize functional materials (e.g., DNA, aptamers, peptides, proteins, antibodies, enzymes, and NPs) on complementary DNA-modified solid supports for bioanalysis [117, 281], to fabricate multifunctional NPs for biosensing and bioimaging [65, 68, 77, 79], for DNA origami, and for placing cascading multienzyme complexes on DNA scaffolds [120, 122–125]. Although short DNA linkers display a relatively high physicochemical stability in vitro, some approaches, such as the utilization of unnatural base DNA or PNA, are required for in vivo applications to prevent degradation by nucleases.

PNA is a DNA analog with a noncyclic, peptide-like backbone (Fig. 25). Owing to its flexible and neutral backbone instead of a negatively charged deoxyribose phosphate backbone, PNA exhibits very good hybridization properties with DNA, RNA, PNA, and DNA duplexes at low and even high ion concentrations, as well as a higher temperature stability than the corresponding pure nucleic acid complexes. Therefore, PNA can highly discriminate mismatched DNA and has a stronger binding affinity for complementary DNA than does its DNA counterpart. PNA also displays a very high stability against enzymatic degradation due to its peptide-like backbone [282]. Applications of PNA linkers in the fields of therapy, diagnosis, and biosensing have been reviewed [282–284]. For example, coupling a radioactively labeled PNA to a TfR

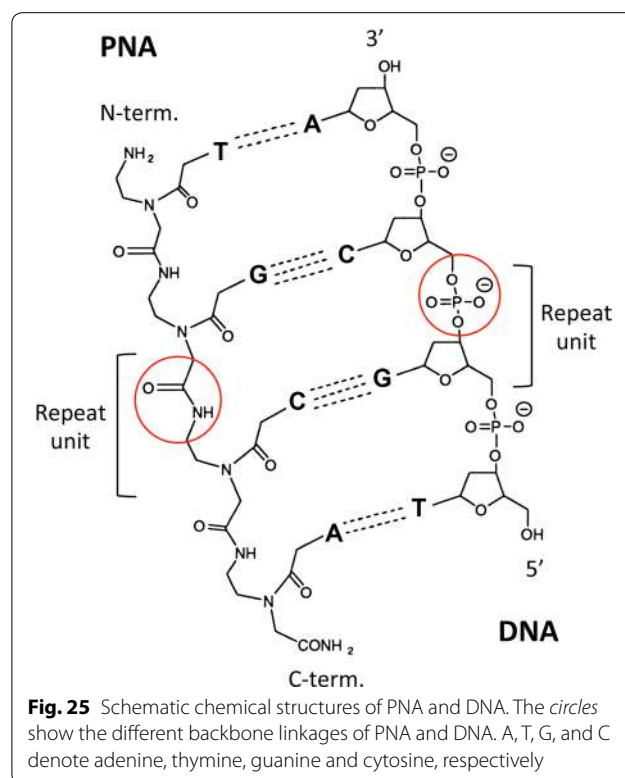


Fig. 25 Schematic chemical structures of PNA and DNA. The circles show the different backbone linkages of PNA and DNA. A, T, G, and C denote adenine, thymine, guanine and cytosine, respectively

mAb rendered the antisense agent transportable through the blood–brain barrier [285]. The identification and differentiation of tuberculous and nontuberculous mycobacteria in liquid cultures were clinically evaluated by a fluorescence hybridization assay using PNA [286]. The detection of a complementary oligonucleotide at a femtomolar ($\sim 10^{-15}$ M) level was accomplished based on the ion-channel sensor technique using Au electrodes modified with self-assembled monolayers of the PNA probe and 8-amino-1-octanethiol [287].

3.5.2.2 Three classes of peptide linkers The concepts of the protein domains and modules were first proposed in 1973 by Wetlaufer [288] and 1981 by Go [289], respectively. These concepts gave insights into domains and modules as the basic structural, functional or evolutionary units of proteins. A wide variety of naturally occurring multidomain fusion proteins with different architectures have been generated through evolution and characterized to meet the functional requirements of living organisms at the molecular level [290]. The strategies used by nature to evolve fusion proteins have been mimicked by the construction of hybrid or chimeric proteins using molecular biology techniques. Inspired by natural fusion proteins, synthetic fusion proteins have been designed to achieve synergistically improved bioactivities or to generate novel functional combinations derived from each of their component moieties, which are integrated into one molecule by peptide linkers. The fusion proteins have been widely applied in various areas, including recombinant protein production by the tag-mediated enhancement of protein expression, solubility and high-throughput purification [291, 292], fluorescent protein-mediated molecular imaging [293], advanced biocatalysis [101, 108, 111, 115, 164, 290, 294–297], biosensing and bioelectronic materials [290, 298–300], pharmaceuticals, diagnostics and therapeutics [208, 290, 301, 302], reporter protein-mediated immunoassays [303–310], the chimeric receptor-mediated control of cell fate, e.g., growth, death, migration or differentiation [311–319], the library selection of antibodies [203, 320, 321] and antibody-mediated drug delivery [218, 322, 323].

Genetic fusion and enzymatic conjugation technologies have been commonly adopted for the construction of fusion proteins. Among them, an end-to-end genetic fusion is the simplest method for constructing a fusion protein, where the coding genes of functional units are combined together and expressed in a suitable host organism. Direct tandem genetic fusion through restriction enzyme sites is simple; the flexible and unstructured N- or C-terminal regions of the component proteins and additional short peptides derived from restriction enzyme sites act as a peptide linker to provide enough

space between the functional units of a fusion protein for correct folding. However, if the N- or C-terminus is not flexible or not long enough to prevent steric hindrance, this effect will reduce the degrees of freedom of units in fusion protein dynamics and may cause unfavorable results, such as inclusion body formation derived by protein misfolding, a loss of function and a low yield of functional fusion proteins. For this reason, longer peptide linkers are generally inserted between functional units [290].

Peptide linkers are generally classified into three groups according to their structures: flexible linkers, rigid linkers, and site-specific linkers cleavable by proteolytic enzyme digestion. In addition to the basic role of linking functional units together or releasing functional units (e.g., toxin release in drug delivery systems, affinity tag cleavage from tag-fused recombinant pharmaceutical proteins in the purification process), peptide linkers may offer many other advantages for the production of fusion proteins, such as improving biological activity and structural stability and achieving desirable biopharmaceutical pharmacokinetic profiles [324]. Therefore, peptide linkers play a variety of structural and functional roles in fusion proteins.

3.5.2.3 Flexible peptide linkers Flexible linkers are frequently adopted as natural inter-domain peptide linkers in multidomain proteins when the joined domains require a certain degree of movement or interaction. Based on the analysis of AA preferences for residues contained in these natural flexible linkers, it has been revealed that they are generally composed of small, nonpolar (e.g., Gly) or polar (e.g., Ser, Thr) residues [325]. The small size of these AA residues provides flexibility and enables the mobility of the connected functional units. The incorporation of Ser or Thr can maintain the stability of the peptide linker in aqueous solutions by forming hydrogen bonds with water molecules, thereby reducing unfavorable interactions between the linker and protein moieties. The most widely used synthetic flexible linker is the G_4S -linker, $(G_4S)_n$, where n indicates the number of G_4S motif repeats. By changing the repeat number “ n ,” the length of this G_4S linker can be adjusted to achieve appropriate functional unit separation or to maintain necessary interactions among units, thus allowing proper folding or achieving optimal biological activity [324]. Poly-Gly (G_n) linkers also form an elongated structure similar to that of the unstable 3_{10} -helix conformation. Since Gly has the greatest freedom in backbone dihedral angles among the natural AAs, G_n linkers can be assumed to be the most “flexible” polypeptide linkers [326]. In addition to the G_4S linkers and poly-Gly linkers, many other flexible linkers, such as KESGSVSSEQLAQFRSLD and EGKSSGSGSESKST

for the construction of a single-chain variable fragment (scFv), have been designed by searching libraries of 3D peptide structures derived from protein data banks for crosslinking peptides with proper V_H and V_L molecular dimensions [327]. These flexible linkers are also rich in small or polar AAs, such as Gly, Ser, and Thr, and they contain additional AAs, such as Ala, to maintain flexibility, as well as large polar AAs, such as Glu and Lys, to increase the solubility of fusion proteins.

3.5.2.4 Rigid peptide linkers Rigid linkers act as stiff spacers between the functional units of fusion proteins to maintain their independent functions. The typical rigid linkers are helix-forming peptide linkers, such as the polyproline (Pro) helix (P_n), poly-Ala helix (A_n) and α -helix-forming Ala-rich peptide (EA_3K) $_n$, which are stabilized by the salt bridges between Glu^- and Lys^+ within the motifs [328]. Fusion proteins with helical linker peptides are more thermally stable than are those with flexible linkers. This property was attributed to the rigid structure of the α -helical linker, which might decrease interference among the linked moieties, suggesting that changes in linker structure and length could affect the stability and bioactivity of functional moieties.

The Pro-rich peptide (XP) $_n$, with X designating any AA, preferably Ala, Lys, or Glu, can also constrain the linker to an extended conformation with relatively limited flexibility. The Pro residue is a very unique AA; it is a cyclic AA, and its side chain cyclizes back to the amide on the backbone, which restricts the conformation of its backbone to a small range of backbone angles. Since the Pro residue has no amide hydrogen to form a hydrogen bond with other AAs, it can avoid ordered structures and prevent interactions between the linkers and neighboring domains [324, 329]. Therefore, Pro residues in (XP) $_n$ linkers can increase the linker stiffness and effectively separate neighboring domains.

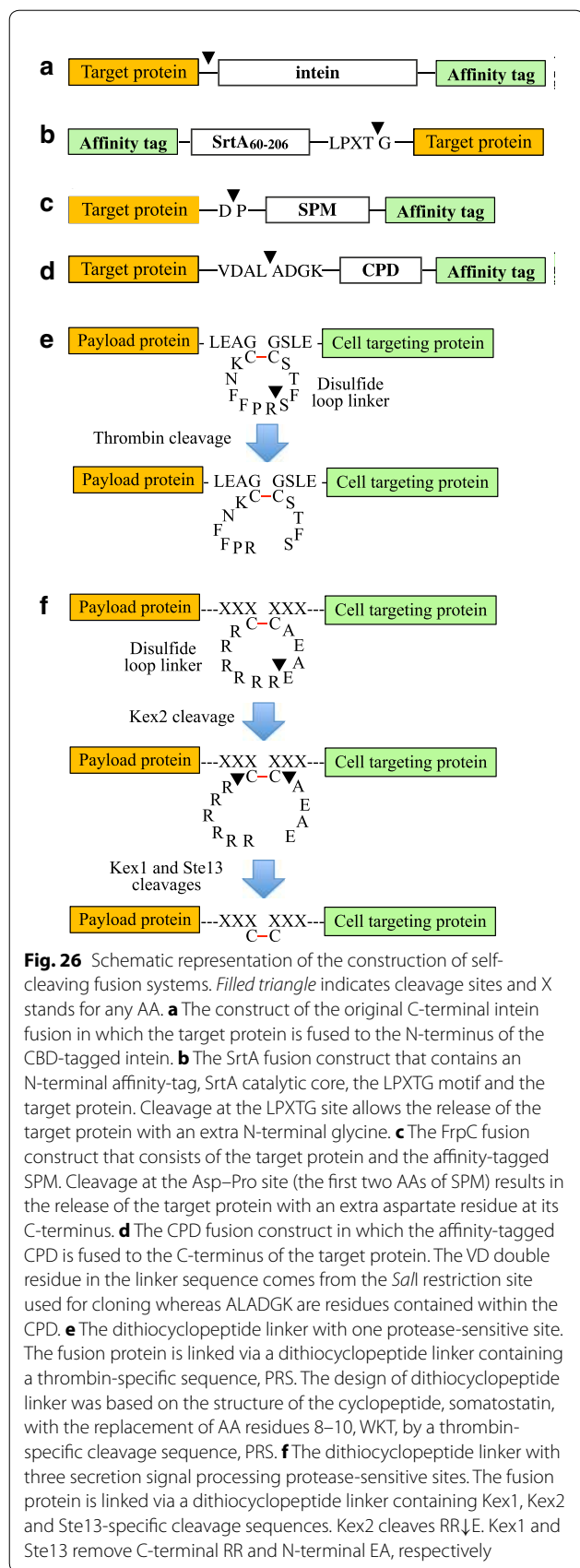
3.5.2.5 Site-specific, cleavable peptide linkers Genetic fusion technology provides an effective means for recombinant protein expression and purification. A comprehensive review of affinity tags can be found elsewhere [292, 330]. Examples of affinity tags include poly-His, FLAG, HA, strep II, the calmodulin-binding peptide and the chitin-binding domain. These tags specifically interact with their partner molecules and allow the fused protein to be captured by corresponding partner molecule-modified matrices. In most cases, the tags are removed from the fusion proteins after an affinity tag-assisted purification process to obtain the final product consisting of pure target protein. This is usually achieved by enzymatic or chemical cleavage at the junction between the tag and the target protein. Endoproteases commonly used to cleave

fusion tags include factor Xa (I(E/D)GR↓X), enterokinase (DDDDK↓X), thrombin (LVPR↓GS), tobacco etch virus protease (ENLYFQ↓(G/S)) and a genetically engineered derivative of human rhinovirus 3C protease, PreScission™ (LEVLQ↓GP) [329, 330]. Chemicals that are specific and efficient for the chemical cleavage of proteins in solution are CNBr (Met↓), 2-(2'-nitrophenylsulfonyl)-3-methyl-3-bromoindolenine (Trp↓), 2-nitro-5-thiocyanobenzoic acid (Cys↓), formic acid (Asp↓Pro) and hydroxylamine (Asn↓Gly) [331]. Here, the down arrow and X in parenthesis indicate the cleavage site of the recognition site and any AA, respectively. In general, enzymatic cleavage is site-specific and can be carried out under mild conditions. However, cleavage efficiency may vary with different fusion proteins. Steric hindrance or the presence of unfavorable residues around the cleavage site could result in inefficient processing. In contrast to enzymatic cleavage, chemical cleavage offers a less expensive alternative but requires harsh conditions that may cause side-chain modifications. Furthermore, since chemical cleavage usually targets specific residues or dipeptide linkages, the frequent presence of the single- or double-residue site recognized by these chemicals within the AA sequence of the target protein limits its use [332].

Self-cleaving tags are a special group of fusion tags that are based on protein modules (e.g., intein, SrtA, the FrpC module, and the Cys protease domain) and possess inducible proteolytic activities. Fusion proteins containing them can be site-specifically self-cleaved by the trigger of a low molecular weight compound or a change in its conformation. Combined with appropriate affinity tags, self-cleaving tags enable fusion protein purification, cleavage and target protein separation to be achieved in a single step [332].

In the case of Intein-tag, the target protein is fused to the N-terminus of intein [e.g., VMA intein from *Saccharomyces cerevisiae* (51 kDa) or DnaB intein from *Synechocystis* sp. strain PCC6803 (17 kDa)], whose C-terminus is conjugated with an affinity tag (Fig. 26a). Intein-mediated site-specific cleavage can be triggered by thiol reagents, such as dithiothreitol or β -mercaptoethanol.

As for SrtA-tag, the fusion protein consists of an N-terminal affinity tag, a SrtA catalytic core, the LPXTG motif and the target protein (Fig. 26b). On-resin cleavage can be induced by incubation in a Ca^{2+} ion-containing buffer, and the released target protein, with an extra Gly residue at its N-terminus, can then be collected. However, this system has a potential drawback. Although the activity of SrtA from *S. aureus* is inducible by Ca^{2+} ions and moderate conditions, it is not completely suppressed during protein expression because abundant soluble Mg^{2+} ions (10^3 - to 10^4 -fold greater in concentration than Ca^{2+} ions) in the cytosol can partly replace Ca^{2+} ions in function



[333], which causes unwanted fusion cleavage at an early stage.

The FrpC module is an iron-regulated protein produced by the gram-negative bacterium *Neisseria meningitidis*. The fusion construct contains the target protein, which is at the N-terminal moiety, and the affinity-tagged self-processing module (SPM) (Fig. 26c). The DNA coding sequence for the first four AAs of the SPM, which are Asp-Pro-Leu-Ala, contains an *NheI* restriction site that can be used for cloning. The Ca²⁺ ion-addition induces SPM-mediated cleavage, resulting in the release of the target protein with an extra Asp residue at the C-terminus.

Vibrio cholerae secretes a toxin with large, multifunctional, auto-processing repeats; this toxin undergoes proteolytic cleavage during translocation into host cells. The proteolysis of the toxin is mediated by a conserved internal Cys protease domain (CPD), which is activated upon the binding of the small molecule inositol polyphosphate (IP6). Affinity-tagged CPD can be fused to the C-terminus of the target protein (Fig. 26d). The IP6-addition triggers CPD-mediated cleavage, which allows the target protein to be released. Depending on the cloning site used, one or more additional residues may be appended to the C-terminus of the target protein.

Other applications of cleavable linkers are drug delivery systems to release free functional units of fusion proteins in vivo. These linkers are designed to cleave under specific conditions, such as the presence of reducing reagents or proteases. This linker system enables fusion proteins to reduce steric hindrance and improve both the independent actions and bioactivities of individual functional units after in vivo cleavage. The reduction of disulfide bonds in vivo has been widely applied for the release of payloads from drug delivery systems fabricated by chemical conjugation technologies. Similarly, disulfide linkers cleavable in vivo were designed for recombinant fusion proteins [334, 335]. One such disulfide linker (LEAGCKNFFPR↓SFTSCGSLE) is based on a dithiocyclopeptide containing an intramolecular disulfide bond formed between two Cys residues on the linker, as well as a thrombin recognition sequence (PRS) between the two Cys residues (Fig. 26e). Another disulfide linker (CRRRRRREAEAC) also contains an intramolecular disulfide bond and a peptide sequence sensitive to the secretion signal-processing proteases of the yeast secretory pathway. During protein expression, this linker is first cleaved by the protease Kex2 at CRRRRR↓EAEAC, followed by the removal of the dipeptides RR and EA by the secretion signal-processing proteases Kex1 and Ste13 (C↓RR↓RR↓RR, EA↓EA↓C), respectively (Fig. 26f). As a result, the AAs between the two Cys residues in the linker were completely removed during secretion, and

the disulfide linked fusion protein was directly expressed by *Pichia pastoris*.

3.5.2.6 The effect of linker composition, flexibility/rigidity and length on the functions and conformations of fusion proteins The folding, stability, proteolytic sensitivity and function of fusion proteins might be affected by the AA composition and the flexibility/rigidity and length of the peptide linkers. For example, fusion proteins consisting of a cellulose-binding domain of *Neocallimastix patriciarum* cellulase A (Cel6A) and lipase B from *Candida antarctica* were constructed by connecting two functional units with different linker peptides (4–44 AA residues, different Asn residue numbers and positions for potential N-glycosylation sites) derived from the natural peptide linker contained in Cel6A. Analyses of linker stability toward proteolysis and the cellulose-binding activity and lipase activity of the fusion proteins were conducted; the results revealed that fusion proteins with shorter linkers (4–16 AA residues) were more stable against proteolysis but had slightly lower cellulose-binding capacities than those containing longer linkers. However, all fusion proteins retained the lipase-specific activity of the wild-type protein [336].

Bifunctional fusion proteins composed of the catalytic domains of endoglucanase (Endo5A) and β -glucosidase (Gluc1C) from a *Paenibacillus* strain were constructed by changing the connection order of two domains and linking them with flexible peptide linkers of different lengths $(G_4S)_n$ ($n = 0-3$). The results indicated that the substrate affinity K_m and catalytic efficiency k_{cat}/K_m of Gluc1C were sensitive to its position, as it showed a decline in both affinity and catalytic efficiency when Gluc1C was placed at the N-terminus of the fusion protein. However, there was no direct relationship of linker length with either Endo5A or Gluc1C activity [337].

Tandem fusion proteins of human serum albumin and onconase (ONC) with flexible linkers $(G_4S)_n$ ($n = 0-3$) were constructed and expressed in *P. pastoris*. The expression level of the fusion proteins had no relationship with the linker length. However, while the ONC moiety of the fusion protein without a linker ($n = 0$) showed no cytotoxicity toward tumor cells, this gradually improved with increasing linker length [338].

For the development of a bifunctional immunoreagent, the B1 domain of *Streptococcal* protein G (SpG), which binds to the Fc region and CH1 domain of IgG, was fused with luciferase from *Vargula hilgendorffii* (Vluc) using flexible peptide linkers $(G_4S)_n$ ($n = 0-1$). The resulting fusion protein, SpG- $(G_4S)_n$ -Vluc, retained the bioluminescence activity of the Vluc moiety but lost the binding affinity of SpG to IgG. However, inserting the three α -helices bundle D domain of protein A from *S. aureus*

(SpA) between the SpG and the (G_4S) linker successfully recovered the binding affinity of SpG to the CH1 domain of IgG [339].

Fusion protein pairs for noncompetitive and homogeneous immunoassays were developed by optimizing the flexible G_4S linker length of each fusion protein. This assay system is based on the antigen-dependent reassociation of antibody variable regions (V_H , V_L) and the subsequent complementation of the β -Gal domains $\Delta\alpha$ and $\Delta\omega$. The best pair was found to be V_H - $(G_4S)_2$ - $\Delta\alpha$ and V_L - $(G_4S)_1$ - $\Delta\omega$, which, at its optimal concentration, showed a 2.5-fold increase in β -Gal activity upon antigen addition [340].

Chimeric receptors (chimeras of anti-fluorescein (FL) scFv and an engineered c-Mpl receptor possessing only signaling mediator STAT3-binding motifs) were designed by changing the peptide linker length between the binding motifs of JAK and STAT3 using flexible linkers $(G_4S)_n$ ($n = 0, 3, 6, 9$). The activation level of STAT3 was quantitatively evaluated by detecting the level of phosphorylated STAT3 after the stimulation of chimeric receptor-expressing cells with FL-labeled bovine serum albumin (BSA-FL). The results showed that the STAT3 activation levels were 0.8-, 1.5- and 1.4-fold greater with $(G_4S)_3$, $(G_4S)_6$ and $(G_4S)_9$, respectively, than without a linker. Therefore, changes in the distance from the JAK-binding domain to the STAT3-binding motif exerted relatively minor effects on the phosphorylation level of STAT3 [341].

Helical poly-Ala linkers $(Ala)_n$ ($n = 0-4$) were inserted between the transmembrane and intracellular domains of a chimeric receptor (a tandem fusion protein of anti-FL scFv/intracellular domain-truncated EpoR/gp130 intracellular domain), and the effect of linker length on cell proliferation was investigated by stimulating chimeric receptor-expressing cells with BSA-FL. A periodic enhancement in cell proliferation was induced by the insertion of one to four Ala residues. The chimeric receptors with linkers $(Ala)_n$ ($n = 0, 1$) transduced a growth signal, while growth activity was lost when $(Ala)_n$ ($n = 2-4$) linkers were inserted. Furthermore, the extracellular EpoR D1 domain-truncated chimeric receptor showed different patterns in the periodic enhancement of cell proliferation by the insertion of one to four Ala residues. In this case, the chimeric receptors with linkers $(Ala)_n$ ($n = 0, 3, 4$) failed to transduce a growth signal, whereas growth activity was restored when one or two Ala residues were inserted. These results clearly demonstrate the importance of intracellular domain orientation for the activation of chimeric receptors, which is readily controlled by the 109° rotation of the α -helix Ala linker with each increment of one Ala residue [342].

To construct a ligand-inducible scFv dimer, anti-ErbB2 scFv was fused with FKBP_{F36V}, which is a mutant of FK-binding protein 12 that can be dimerized by the synthetic homodimeric ligand AP20187. The three kind of linkers, i.e., flexible (G₄S)₃, rigid α -helix (EA₃K)₃ and DKTHCP(G₄S)₂, derived from the hinge region of IgG were inserted between scFv and FKBP_{F36V}, and the effect of linker properties on the activity of the fusion protein dimer, which can dimerize the artificial chimeric receptor ErbB2-gp130 expressed on the cell surface and induce cell proliferation signaling from the dimerized chimeric receptor, were investigated. The results showed that the fusion protein with the hinge linker was the best for activating ErbB2-gp130 chimera-induced cell proliferation [320].

It has been demonstrated that the selective complex formation of P450cam with its redox partner proteins, PdX and PdR, can be achieved by fusing each component to the C-terminus of a different subunit of the

heterotrimer PCNA from *Sulfolobus solfataricus* to form a self-assembling scaffold [111]. To enhance the activity of this self-assembled multienzyme complex, the peptide linker connecting PdX with PCN2 was optimized using various peptide linkers, such as flexible linkers (G₄S)_n (n = 1–6), helical and rigid Pro-rich linkers (G₄S-(P₅)_n-G₄S) (n = 1–5) and other linkers (G₄S-LVPRGS-G₄S). Although the activity was affected by the lengths of both the rigid Pro-rich linkers and the flexible linkers, the Pro-rich linkers provided the greatest activity enhancement. The optimized Pro-rich linker (G₄S-(P₅)₄-G₄S) enhanced the activity by 1.9-fold compared with the G₄S-LVPRGS-G₄S linker, while the (G₄S)_n (n = 1–6) linker did not yield activity higher than the maximum activity of the optimized Pro-rich linker. Both peptide linker rigidity/flexibility and length were found to be important for enhancing overall multienzyme complex activity (Fig. 27) [343].

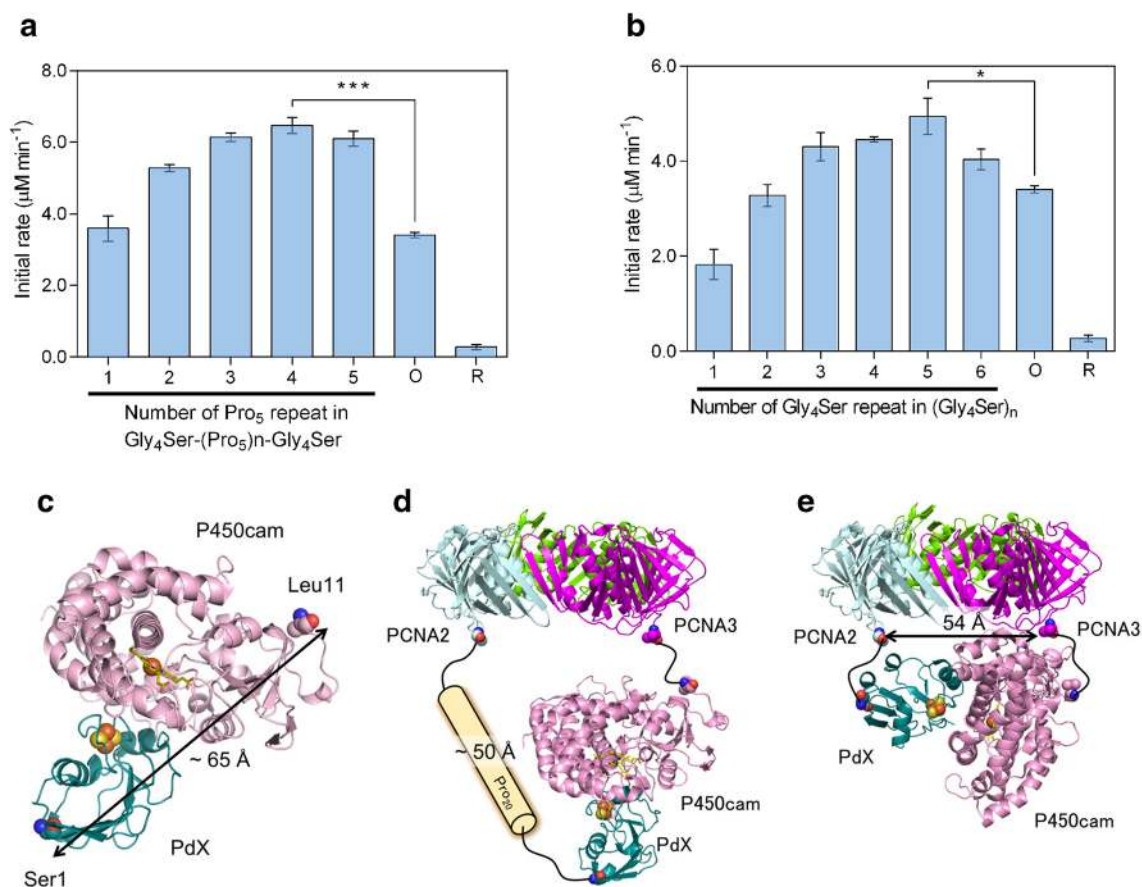
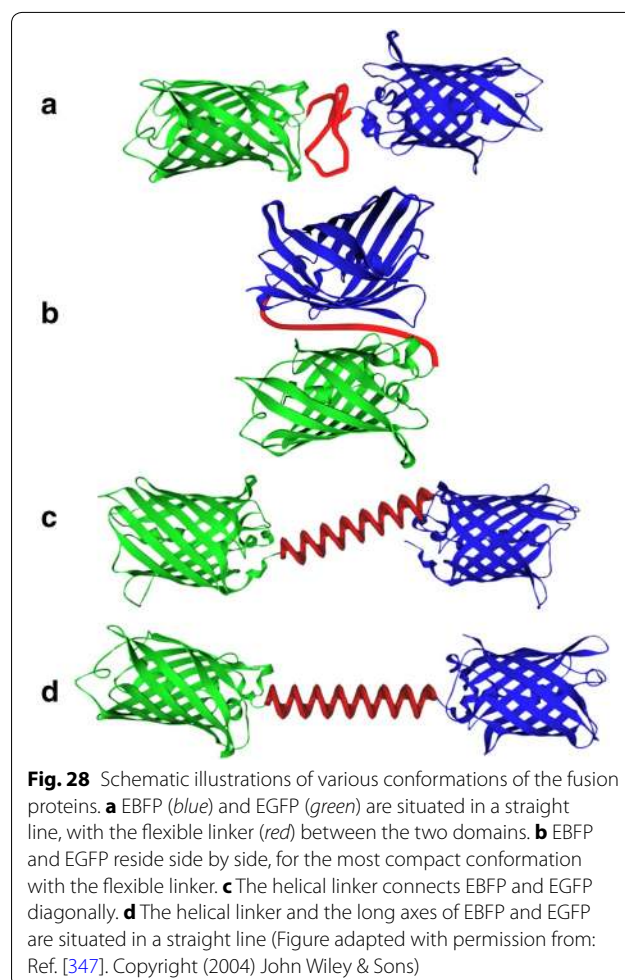


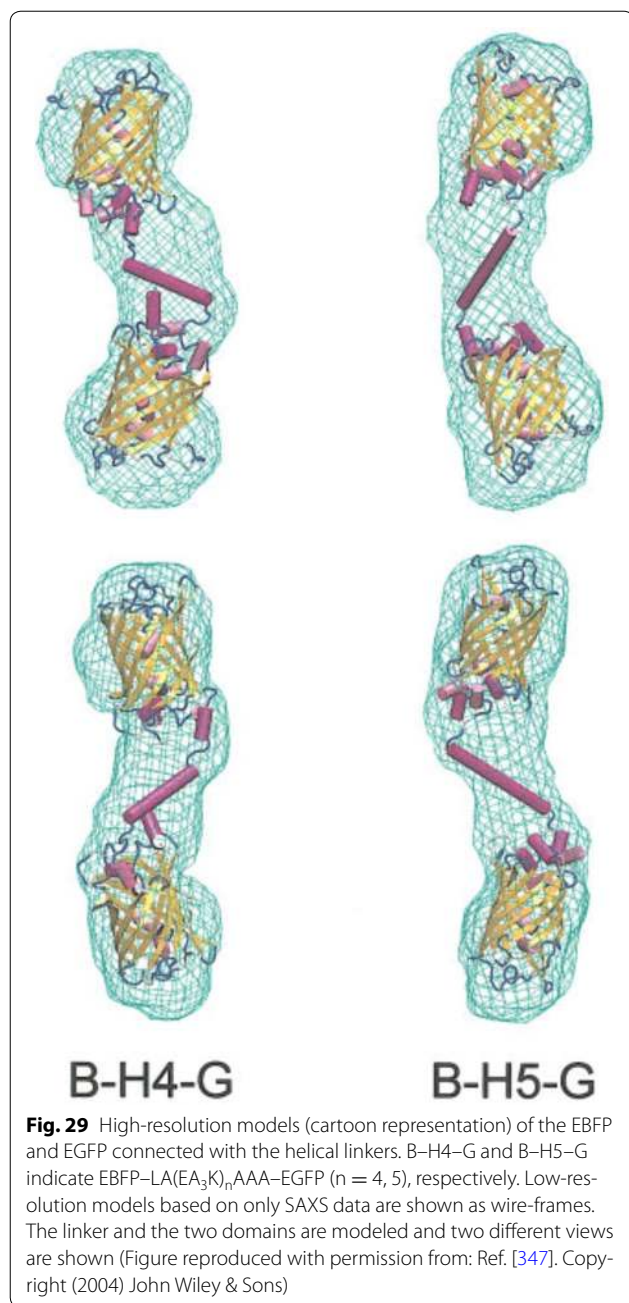
Fig. 27 Optimization of the PCNA2-PdX fusion protein linker in PUPPET. **a** P450cam oxidation activities of the PUPPET linker variants, PUPPET-Pn (n = 1–5). **b** P450cam oxidation activities of the PUPPET linker variants, PUPPET-Gn (n = 1–6). **c** A docking model of P450cam and PdX. **d** Spatial arrangement of P450cam and the PCNA ring when the PdX-binding site of P450cam faces in the same direction to the PCNA ring. **e** Spatial arrangement of P450cam and the PCNA ring when the PdX-binding site of P450cam faces in a perpendicular direction to the PCNA ring (Figures reproduced from Ref. [343])

The tandem fusion proteins β -glucanase (Gluc)/xylanase (Xyl) were constructed using peptide linkers, such as flexible linkers $(G_4S)_n$ ($n = 0-3$), α -helical linkers $(EA_3K)_n$ ($n = 0-3$) and others (MGSSSN designed using the software of the web server LINKER [344], and TGSRYMELGATQGMGEALTRGM derived from the two α -helix bundle of *Humicola insolens* endocellulase). The effects of the linkers on the thermal stability and catalytic efficiency of both enzymes were analyzed. The Gluc moieties of most fusion constructs showed greater stability at 40–60 °C than did the parental Gluc and the linker-free fusion protein. All the Xyl moieties showed thermal stabilities similar to that of the parental Xyl, at ≥ 60 °C. It was also revealed that the catalytic efficiencies of the Gluc and Xyl moieties of all the fusion proteins were 3.04- to 4.26-fold and 0.82- to 1.43-fold those of the parental moieties, respectively. The flexible linker $(G_4S)_2$ resulted in the best fusion proteins, whose catalytic efficiencies were increased by 4.26-fold for the Gluc moiety and by 1.43-fold for the Xyl moiety. The Gluc and Xyl moieties of the fusion protein with the rigid linker $(EA_3K)_3$ also showed 3.62- and 1.31-fold increases in catalytic efficiency [345].

Aiming to clarify the criteria for designing peptide linkers for the effective separation of the domains in a bifunctional fusion protein, a systematic investigation was carried out. As a model, the fusion proteins of two *Aequorea* GFP variants, enhanced GFP (EGFP) and enhanced blue fluorescent protein (EBFP), were employed. The secondary structure of the linker and the relative distance between EBFP and EGFP were examined using circular dichroism (CD) spectra and fluorescent resonance energy transfer (FRET), respectively. The following AA sequences were designed and utilized as peptide linkers: a short linker (SL); LAAA (4 AAs) (derived from the cleavage sites for *Hind*III and *Not*I); flexible linkers $(G_4S)_n$ AAA ($n = 3, 4$); α -helical linkers $LA(EA_3K)_n$ AAA ($n = 3-5$); and a three α -helix bundle from the B domain of SpA (LFNKEQQNAFYELHLPNLNEEQRNGFIQSLKDDPSQSANLLAEAKKLNDQA AAA). The differential CD spectra analysis suggested that the $LA(EA_3K)_n$ AAA linkers formed an α -helix and that the α -helical contents increased as the number of the linker residues increased. In contrast, the flexible linkers formed a random, coiled conformation. The FRET from EBFP to EGFP decreased as the length of the helical linkers increased, indicating that distances increased in proportion to the length of the linkers. The results showed that the helical linkers could effectively separate the neighboring domains of the fusion protein. In the case of the fusion proteins with the flexible linkers, the FRET efficiency was not sensitive to linker length and was highly comparable to that of the fusion proteins with the SL, although the flexible linkers were much longer

than the SL, again indicating that the flexible linkers had a random, coiled conformation [346]. The real in situ conformations of these fusion proteins and structures of the linkers were further analyzed using synchrotron X-ray small-angle scattering (SAXS). The SAXS experiments indicated that the fusion proteins with flexible linkers assume an elongated conformation (Fig. 28a) rather than the most compact conformation (Fig. 28b) and that the distance between EBFP and EGFP was not regulated by the linker length. On the other hand, fusion proteins with helical linkers $[LA(EA_3K)_n$ AAA $n = 4, 5]$ were more elongated than were those with flexible linkers, and the high-resolution models (Fig. 29) showed that the helical linkers connected the EBFP and EGFP domains diagonally (Fig. 28c) rather than longitudinally (Fig. 28d). However, in the case of the shorter helical linkers ($n = 2, 3$, especially $n = 2$), fusion protein multimerization was observed. Since most residues of the short helical linkers are situated closer to the two domains of the fusion protein, the charged residues, Glu and Lys in the (EA_3K) unit are likely to form ion pairs with the oppositely charged





residues on the top surfaces of EBFP and EGFP. Consequently, this ion pairs formation causes destabilization of the short helix and melted helix linkers may act as attractants for the attachment of neighboring molecules due to their charges and hydrophobicity, thereby causing multimerization of fusion protein. On the other hand, in the case of the fusion protein with the longer helical linkers (n = 4, 5), the linkers retained the α -helix structure and could solvate monomeric fusion proteins. These results clearly suggested the outstanding ability of the

rigid helical linkers to control the distance and reduce the interference between the domains [347]. This study is the first example of modeling in situ fusion protein conformations and linker structures by combining SAXS data of fusion proteins, structural information of the functional units from the Brookhaven Protein Data Bank (PDB), and molecular dynamics calculations of peptide linker structures. Recently, this modeling method was applied to evaluate the in situ conformations and structures of fusion proteins composed of a de novo two-helix bundle protein and a single trimeric foldon domain of fibrin from the bacteriophage T4 connected by a short peptide linker (KLAAA). Size exclusion chromatography, multi-angle light scattering, analytical ultracentrifugation, and SAXS analyses indicated that the small (S form), middle (M form), and large (L form) forms of the fusion protein oligomers exist as 6-, 12-, and 18-mers, respectively. The SAXS data further suggested that the S and M forms have barrel- and tetrahedron-like shapes, respectively [348].

The selection of a suitable peptide linker, which allows a desirable conformation and interaction among functional units in fusion proteins, is key to the successful design of fusion proteins. Generally, rigid linkers exhibit relatively stiff structures by adopting α -helical structures or by containing multiple Pro residues with the *cis* isomer of the peptide bond. Under many circumstances, they can separate the functional domains in fusion protein more efficiently than do flexible linkers. The length of the linkers can be easily adjusted by changing the linker-unit repeat-number, such as (EA₃K), to achieve an optimal distance between functional units. As a result, when the spatial separation of the functional units is critical to avoid steric hindrance and to preserve the folding, stability and activity of each unit in the fusion proteins, rigid linkers would be selected. However, there are other types of fusion proteins, in which functional units are required to have a certain degree of movement/interaction or a precise proximal spatial arrangement and orientation to form complexes. In such cases, flexible linkers are often chosen because they can serve as a passive linker to maintain a distance or to adjust the proximal spatial arrangement and orientation of functional units. However, optimizing the peptide linker sequence and predicting the spatial linker arrangement and orientation are more difficult for flexible linkers than for rigid linkers. Current strategies are mostly empirical and intuitive and have a high uncertainty. Therefore, computational simulation technologies for predicting fusion protein conformations and linker structures would potentially encourage rational flexible linker design with improved success rates.

3.5.2.7 Rational algorithms and software for designing linker sequences and structures The rational design of

fusion proteins with desired conformations, properties and functions is a challenging issue. Most current approaches to linker selection and design processes for fusion proteins are still largely dependent on experience and intuition; such selection processes often involve great uncertainty, particularly in the case of longer flexible linker selection, and many unintended consequences, such as the misfolding, low yield and reduced functional activity of fusion proteins may occur. This is mostly because of our limited understanding of the sequence–structure–function relationships in these fusion proteins. To overcome this problem, the computational prediction of fusion protein conformation and linker structure can be considered a cost-effective alternative to experimental trial-and-error linker selection. Based on the structural information of individual functional units and linkers (either from the PDB or homology modeling), considerable progress has been made in predicting fusion protein conformations and linker structures [290]. Approaches for the design or selection of flexible linker sequences to connect two functional units can be categorized into two groups. The first group comprises library selection-based approaches, in which a candidate linker sequence is selected from a loop sequence library without consideration of the conformation or placement of functional units in the fusion proteins. The second group comprises modeling-based approaches, in which functional unit conformation and placement and linker structure and AA composition would be optimized by simulation.

Regarding the first approach, a computer program called LINKER was developed. This web-based program (<http://astro.temple.edu/feng/Servers/Bioinformatic-Servers.htm>) automatically generated a set of peptide sequences based on the assumption that the observed loop sequences in the X-ray crystal structures or the nuclear magnetic resonance structures were likely to adopt an extended conformation as linkers in a fusion protein. Loop linker sequences of various lengths were extracted from the PDB, which contains both globular and membrane proteins, by removing short loop sequences less than four residues and redundant sequences. LINKER searched its database of loop linker sequences with user-specified inputs and outputted several candidate linker sequences that meet the criteria. The basic input to the program was the desired length of the linker, expressed as either the number of residues or a distance in angstroms. Additional input parameters included potential cleavage sites for restriction endonucleases or proteases to avoid such that the selected linkers would be resistant against the restriction enzymes and the specified protease during the DNA cloning and protein purification process, respectively. The users could also include AA composition preferences (e.g., eliminate

bulky hydrophobic residues) to further select their linkers of interest. The output of LINKER included a list of peptide sequences with the specified lengths, sequence characteristics and chemical features of every linker sequence shown by hydrophobicity plots [344, 349]. However, although the PDB database has expanded tremendously during the last decade, no further updates or improvements were made to the LINKER website since it was created, and it is no longer accessible.

The web-based program LinkerDB (<http://www.ibi.vu.nl/programs/linkerdbwww/>) also provides a database containing linker sequences with various confirmations and a search engine. The search algorithm accepts several query types (e.g., PDB code, PDB header, linker length, secondary structure, sequence or solvent accessibility). The program can provide the linker sequences fitting the searching criteria as well as other information, such as the PDB code and a brief description of the source protein, the linker's position within the source protein, linker length, secondary structure, and solvent accessibility. Users can search for sequences with desired properties and obtain candidate sequences from natural multidomain proteins [329].

Another server website for facilitating linker selection and fusion protein modeling is SynLinker (<http://bioinfo.bti.a-star.edu.sg/linkerdb>). It contains information regarding 2260 linkers, consisting of natural linkers extracted from multidomain proteins in the latest PDB, as well as artificial and empirical linkers collected from the literature and patents. A user may specify multiple query criteria to search SynLinker, such as the PDB ID of the source proteins, protein names, the number of AA residues in a linker, and/or the end-to-end distance of a linker conformation in Angstroms (Å). Additionally, the user can choose a linker starting residue, ending residue, AA enrichment, AA depletion and/or protease sensitivity as a desired linker property in the recombinant fusion protein. Once a query is submitted, both the natural and artificial/empirical linkers in SynLinker are searched simultaneously, yielding a list of potential linker candidates satisfying the desired selection criteria together with information about the AA composition radar chart and the conformation of the selected linker, as well as the fusion protein structure and hydropathicity plot [350].

As for modeling-based approaches, the conformation and placement of functional units in fusion proteins, of which 3D structures are available from the PDB or homology modeling, can be predicted by computer-aided modeling. A modeling tool known as FPMOD was developed and can generate fusion protein models by connecting functional units with flexible linkers of proper lengths, defining regions of flexible linkers, treating the structures of all functional units as rigid bodies and

rotating each of them around their flexible linker to produce random structures. This tool can extensively test the conformational space of fusion proteins and finally generate plausible models [351]. This tool has been applied to designing FRET-based protein biosensors for Ca^{2+} ion by qualitatively predicting their FRET efficiencies, and the predictions strongly agreed with the experimental results [352].

A similar modeling tool was developed for assembling structures of isolated functional units to constitute multidomain fusion proteins. However, this approach of assembling functional units is different from the method of testing conformational space. In this method, an ab initio protein-modeling method is utilized to predict the tertiary structure of fusion proteins, the conformation and placement of functional units and the linker structure. This method samples the degrees of freedom of the linker (in other words, domain assembly as a linker-folding problem) rather than those of the rigid bodies, as adopted in FPMOD. The method consists of an initial low-resolution search, in which the conformational space of the linker is explored using the Rosetta de novo structure prediction method. This is followed by a high-resolution search, in which all atoms are treated explicitly, and backbone and side chain degrees of freedom are simultaneously optimized. The obtained models with the lowest energy are often very close to the correct structures of existing multidomain proteins with very high accuracy [353].

A method called pyDockTET (tethered-docking) uses rigid-body docking to generate domain–domain complexes that are scored by the electrostatic and desolvation energy terms, as well as a pseudo-energy term reflecting restraints from linker end-to-end distances; in this manner, near-native pair-wise domain poses are selected. The optimal linker sequence length (in the number of residues) with the linker ends (defined as the distance between the $\text{C}\alpha$ atoms of the two ends of a linker) is selected from a flexible linker database, which consists of 542 linkers with sequence lengths ranging from 2 to 29 AAs derived from the inter-domain linkers of multidomain structures in the PDB [354].

A fusion protein consisting of a protein called cell-traversal protein for ookinetes and sporozoites (CelTOS) antigen from *Plasmodium falciparum* (the deadliest of malaria species) and human IL-2 as an adjuvant was designed to develop a candidate vaccine against malaria. CelTOS and IL-2 were linked together directly or by using different flexible linkers, including $(\text{G})_8$, (G_4S) and $(\text{G}_4\text{S})_3$. Since the N-terminus of IL-2 and the C-terminus of CelTOS are critical to preserve their stability and bioactivity, the fusion protein was designed by linking the C-terminus

of IL-2 with the N-terminus of CelTOS. The tertiary structures of the fusion proteins were predicted in silico by the I-TASSER online server (<http://zhanglab.ccmb.med.umich.edu/I-TASSER/>) [355]. The model with the highest confidence score (C-score: a scoring function based on the relative clustering structural density and the consensus significance score of multiple threading templates) was considered as the best model. The selected structures of the fusion proteins with different linkers were then validated and analyzed using a Ramachandran plot assessment [356]. All the results verified the $(\text{G}_4\text{S})_3$ linker as the most suitable for separating these proteins [357].

The important issue to be addressed in structure prediction is the method of searching the large and complex conformational space to rapidly reach at the minimum energy structure, which is presumed to be the native fold. The genetic algorithm combined with an extremely fast technique to search the conformation space exhaustively and build a library of possible low-energy local structures for oligopeptides (i.e., the MOLS method), was applied to the protein structure prediction. At the first step, the protein sequence was divided into short overlapping fragments, and then their structural libraries were built using the MOLS method. At the second step, the genetic algorithm exploited the libraries of fragment structures and predicted the single best structure for the protein sequence. In the application of this combined method to peptides and small proteins, such as the avian pancreatic polypeptide (36 AAs), the villin headpiece (36 AAs), melittin (26 AAs), the transcriptional activator Myb (52 AAs) and the Trp zipper (16 AAs), it could predict their near-native structures [358].

The computer-aided rational design methods for fusion proteins are promising because these methods allow us to easily predict the desired conformation and placement of the functional units and linker structures of fusion proteins, and consequently select suitable candidate linker sequences. However, it is difficult to determine the unique conformation of flexible linkers due to many local minima in free energy. Furthermore, if changes in the conformation or arrangement of functional units are essential to display their activity, the linker conformation should also be changed to allow the movement of functional units, e.g., the N-terminal ATP-binding domain and unfolded substrate protein-binding domain connected with a hydrophobic peptide linker in heat shock protein 70 [359]. This complicated conformational transition issue makes it difficult to design optimum linkers for fusion proteins with multiple conformations. Therefore, the rational design of fusion proteins with desired properties and predictable behavior remains a daunting challenge.

4 Conclusion

This review highlighted some of the recent developments in studies related to nanobio/bionanotechnology, including the applications of engineered biological molecules combined with functional nanomaterials in therapy, diagnosis, biosensing, bioanalysis and biocatalysis. Furthermore, this review focused on recent advances in biomolecular engineering for nanobio/bionanotechnology, such as nucleic acid engineering, gene engineering, protein engineering, chemical and enzymatic conjugation technologies, and linker engineering.

Based on creative chemical and biological technologies, manipulation protocols for biomolecules, especially nucleic acids, peptides, enzymes and proteins, were described. We also summarized the main strategies adopted in nucleic acid engineering, gene engineering, protein engineering, chemical and enzymatic conjugation technologies and linker engineering. Nucleic acid engineering based on the base-pairing and self-assembly characteristics of nucleic acids was highlighted as a key technology for DNA/RNA nanotechnologies, such as DNA/RNA origami, aptamers, ribozymes. Gene engineering includes direct manipulation technologies for genes, such as gene mutagenesis, DNA sequence amplification, DNA shuffling and gene fusion, which are powerful tools for generating enzymes, proteins, entire metabolic pathways, or even entire genomes with desired or improved properties. Two general strategies for protein engineering, i.e., rational protein design and directed evolution (i.e., high-throughput library screening- or selection-based approaches) were discussed. Conjugation technologies to site-specifically modify proteins with diverse natural and unnatural functionalities have been developed in the last two decades. These technologies range from classical chemical bioconjugation technologies, bioorthogonal chemical conjugations, protein chemical ligations and enzymatic conjugations, which were overviewed. Linker engineering for controlling the distance, orientation and interaction between functional components crosslinked in conjugates is also an important technology. The design and optimization strategies of chemical and biological linkers, such as oligonucleotides and polypeptides, were overviewed.

A variety of strategies are now available for designing and fabricating novel nanobiomaterials with highly ordered dimension and complexity based on biomolecular self-assembly characteristics governed by molecular interactions among nucleotides, peptides, proteins, lipids and small ligands, each of which focuses on design simplicity, high structural and functional control, or high fabrication accuracy [16–20, 106, 127, 132, 360–365]. Fundamentally, these properties are not mutually

exclusive, and the relative weaknesses of each approach will be solved in the near future. Given the rapid recent progress in the biomolecular engineering and nanotechnology fields, the design of completely novel biomaterial-based molecular devices and systems with functions tailored for specific applications seems to be much easier and more feasible than before.

Competing interests

The author declares that he has no competing interests.

Funding

This research was supported partly by Grants-in-Aid for Scientific Research (A) from Japan Society for the Promotion of Science (JSPS) (15H02319), the Center for NanoBio Integration (CNBI) in The University of Tokyo, and Translational System Biology and Medicine Initiative from The Ministry of Education, Culture, Sports, Science and Technology (MEXT).

Publisher's Note

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

Received: 17 February 2017 Accepted: 29 March 2017

Published online: 24 April 2017

References

1. W.C.W. Chan, Bionanotechnology progress and advances. *Biol. Blood Marrow Transplant.* **12**, 87–91 (2006). doi:[10.1016/j.bbmt.2005.10.004](https://doi.org/10.1016/j.bbmt.2005.10.004)
2. M.K. Kovarik, S.C. Jacobson, Nanofluidics in lab-on-a-chip devices. *Anal. Chem.* **81**, 7133–7140 (2009). doi:[10.1021/ac900614k](https://doi.org/10.1021/ac900614k)
3. J. West, M. Becker, S. Tombrink, A. Manz, Micro total analysis systems: latest achievements. *Anal. Chem.* **80**, 4403–4419 (2008). doi:[10.1021/ac800680j](https://doi.org/10.1021/ac800680j)
4. P. Yager, T. Edward, E. Fu, K. Helton, K. Nelson, M.R. Tam, B.H. Weigl, Microfluidic diagnostic technologies for global public health. *Nature* **442**, 412–418 (2006). doi:[10.1038/nature05064](https://doi.org/10.1038/nature05064)
5. C.D. Chin, V. Linder, S.K. Sia, Commercialization of microfluidic point-of-care diagnostic devices. *Lab Chip* **12**, 2118–2134 (2012). doi:[10.1039/c2lc21204h](https://doi.org/10.1039/c2lc21204h)
6. A. Gencoglu, A.R. Minerick, Electrochemical detection techniques in micro- and nanofluidic devices. *Microfluid. Nanofluid.* **17**, 781–807 (2014). doi:[10.1007/s10404-014-1385-z](https://doi.org/10.1007/s10404-014-1385-z)
7. M.P. McRae, G. Simmons, J. Wong, J.T. McDevitt, Programmable bio-nanochip platform: a point-of-care biosensor system with the capacity to learn. *Acc. Chem. Res.* **49**, 1359–1368 (2016). doi:[10.1021/acs.accounts.6b00112](https://doi.org/10.1021/acs.accounts.6b00112)
8. S. Howorka, Z. Siwy, Nanopore analytics: sensing of single molecules. *Chem. Soc. Rev.* **38**, 2360–2384 (2009). doi:[10.1039/b813796j](https://doi.org/10.1039/b813796j)
9. O.V. Salata, Applications of nanoparticles in biology and medicine. *J. Nanobiotechnol.* **2**, 3 (2004). doi:[10.1186/1477-3155-2-3](https://doi.org/10.1186/1477-3155-2-3)
10. W.T. Liu, Nanoparticles and their biological and environmental applications. *J. Biosci. Bioeng.* **102**, 1–7 (2006). doi:[10.1263/jbb.102.1](https://doi.org/10.1263/jbb.102.1)
11. M. Rawat, D. Singh, S. Saraf, S. Saraf, Nanocarriers: promising vehicle for bioactive drugs. *Biol. Pharm. Bull.* **29**, 1790–1798 (2006). doi:[10.1248/bpb.29.1790](https://doi.org/10.1248/bpb.29.1790)
12. D. Knopp, D.P. Tang, R. Niessner, Bioanalytical applications of biomolecule-functionalized nanometer-sized doped silica particles. *Anal. Chim. Acta* **647**, 14–30 (2009). doi:[10.1016/j.aca.2009.05.037](https://doi.org/10.1016/j.aca.2009.05.037)
13. R.A. Petros, J.M. Desimone, Strategies in the design of nanoparticles for therapeutic applications. *Nat. Rev. Drug Discov.* **9**, 615–627 (2010). doi:[10.1038/nrd2591](https://doi.org/10.1038/nrd2591)
14. M. Colombo, S. Carregal-Romero, M.F. Casula, L. Gutierrez, M.P. Morales, I.B. Boehm, J.T. Heverhagen, D. Prosperi, W.J. Parak, Biological applications of magnetic nanoparticles. *Chem. Soc. Rev.* **41**, 4306–4334 (2012). doi:[10.1039/c2cs15337h](https://doi.org/10.1039/c2cs15337h)

15. D.E. Lee, H. Koo, I.C. Sun, J.H. Ryu, K. Kim, I.C. Kwon, Multifunctional nanoparticles for multimodal imaging and theragnosis. *Chem. Soc. Rev.* **41**, 2656–2672 (2012). doi:[10.1039/c2cs15261d](https://doi.org/10.1039/c2cs15261d)
16. A.V. Pinheiro, D. Han, W.M. Shih, H. Yan, Challenges and opportunities for structural DNA nanotechnology. *Nat. Nanotechnol.* **6**, 763–772 (2011). doi:[10.1038/NNANO.2011.187](https://doi.org/10.1038/NNANO.2011.187)
17. P. Guo, The emerging field of RNA nanotechnology. *Nat. Nanotechnol.* **5**, 833–842 (2010). doi:[10.1038/NNANO.2010.231](https://doi.org/10.1038/NNANO.2010.231)
18. E.H.C. Bromley, K. Channon, E. Moutevelis, D. Woolfson, Peptide and protein building blocks for synthetic biology: from programming biomolecules to self-organized biomolecular systems. *ACS Chem. Biol.* **3**, 38–50 (2008). doi:[10.1021/cb700249v](https://doi.org/10.1021/cb700249v)
19. D. Papapostolou, S. Howorka, Engineering and exploiting protein assemblies in synthetic biology. *Mol. Biosyst.* **5**, 723–732 (2009). doi:[10.1039/b902440a](https://doi.org/10.1039/b902440a)
20. S. Mashaghi, T. Jadidi, G. Koenderink, A. Mashaghi, Lipid nanotechnology. *Int. J. Mol. Sci.* **14**, 4242–4282 (2013). doi:[10.3390/ijms14024242](https://doi.org/10.3390/ijms14024242)
21. Y. Miura, Y. Hoshino, H. Seto, Glycopolymer nanobiotechnology. *Chem. Rev.* **116**, 1673–1692 (2016). doi:[10.1021/acs.chemrev.5b00247](https://doi.org/10.1021/acs.chemrev.5b00247)
22. K.Y. Baik, S.Y. Park, S. Namgung, D. Kim, D.G. Cho, M. Lee, S. Hong, Synthetic nanowire/nanotube-based solid substrates for controlled cell growth. *Nano Converg.* **1**, 28 (2014). doi:[10.1186/s40580-014-0028-0](https://doi.org/10.1186/s40580-014-0028-0)
23. M.A. Kafi, H.Y. Cho, J.W. Choi, Engineered peptide-based nanobio-materials for electrochemical cell chip. *Nano Converg.* **3**, 17 (2016). doi:[10.1186/s40580-016-0077-7](https://doi.org/10.1186/s40580-016-0077-7)
24. S.T.D. Chueng, L. Yang, Y. Zhang, K.B. Lee, Multidimensional nanomaterials for the control of stem cell fate. *Nano Converg.* **3**, 23 (2016). doi:[10.1186/s40580-016-0083-9](https://doi.org/10.1186/s40580-016-0083-9)
25. J.H. Choi, J. Lee, W. Shin, J.W. Choi, H.J. Kim, Priming nanoparticle-guided diagnostics and therapeutics towards human organs-on-chips microphysiological system. *Nano Converg.* **3**, 24 (2016). doi:[10.1186/s40580-016-0084-8](https://doi.org/10.1186/s40580-016-0084-8)
26. W. Jo, U.K. Cheang, M.J. Kim, Development of flagella bio-templated nanomaterials for electronics. *Nano Converg.* **1**, 10 (2014). doi:[10.1186/s40580-014-0010-x](https://doi.org/10.1186/s40580-014-0010-x)
27. G. Faccio, K. Gajda-Schranz, J. Ihssen, F. Boudoire, Y. Hu, B.S. Mun, D.K. Bora, L. Thöny-Meyer, A. Braun, Charge transfer between photo-synthetic proteins and hematite in bio-hybrid photoelectrodes for solar water splitting cells. *Nano Converg.* **2**, 9 (2015). doi:[10.1186/s40580-014-0040-4](https://doi.org/10.1186/s40580-014-0040-4)
28. D.B.T.G. Raj, N.A. Khan, Designer nanoparticle: nanobiotechnology tool for cell biology. *Nano Converg.* **3**, 22 (2016). doi:[10.1186/s40580-016-0082-x](https://doi.org/10.1186/s40580-016-0082-x)
29. F. Gu, R. Karnik, A. Wang, F. Alexis, E.L. Nissenbaum, S. Hong, R. Langer, O.C. Farokhzad, Targeted nanoparticles for cancer therapy. *Nano Today* **2**, 14–21 (2007). doi:[10.1016/S1748-0132\(07\)70083-X](https://doi.org/10.1016/S1748-0132(07)70083-X)
30. M.K. Yu, J. Park, S. Jon, Targeting strategies for multifunctional nanoparticles in cancer imaging and therapy. *Theranostics* **2**, 3–44 (2012). doi:[10.7150/thno.3463](https://doi.org/10.7150/thno.3463)
31. T. Sun, Y.S. Zhang, B. Pang, D.C. Hyun, M. Yang, Y. Xia, Engineered nanoparticles for drug delivery in cancer therapy. *Angew. Chem. Int. Ed.* **53**, 12320–12364 (2014). doi:[10.1002/anie.201403036](https://doi.org/10.1002/anie.201403036)
32. P.K. Ba, B.H. Chung, Multiplexed detection of various breast cancer cells by perfluorocarbon/quantum dot nanoemulsions conjugated with antibodies. *Nano Converg.* **1**, 23 (2014). doi:[10.1186/s40580-014-0023-5](https://doi.org/10.1186/s40580-014-0023-5)
33. W.K. Rhim, M. Kim, K.L. Hartman, K.W. Kang, J.M. Nam, Radionuclide-labeled nanostructures for in vivo imaging of cancer. *Nano Converg.* **2**, 10 (2015). doi:[10.1186/s40580-014-0041-3](https://doi.org/10.1186/s40580-014-0041-3)
34. D.E. Owens III, N.A. Peppas, Opsonization, biodistribution, and pharmacokinetics of polymeric nanoparticles. *Int. J. Pharm.* **307**, 93–102 (2006). doi:[10.1016/j.ijpharm.2005.10.010](https://doi.org/10.1016/j.ijpharm.2005.10.010)
35. M.M. Kemp, R.J. Linhardt, Heparin-based nanoparticles. *WIREs Nanomed. Nanobiotechnol.* **2**, 77–87 (2010). doi:[10.1002/wnan.068](https://doi.org/10.1002/wnan.068)
36. K. Alhareth, C. Vauthier, F. Bourasset, C. Gueutin, G. Ponchel, F. Moussa, Conformation of surface-decorating dextran chains affects the pharmacokinetics and biodistribution of doxorubicin-loaded nanoparticles. *Eur. J. Pharm. Biopharm.* **81**, 453–457 (2012). doi:[10.1016/j.ejpb.2012.03.009](https://doi.org/10.1016/j.ejpb.2012.03.009)
37. P.S. Low, W.A. Henne, D.D. Doorneweerd, Discovery and development of folic-acid-based receptor targeting for imaging and therapy of cancer and inflammatory diseases. *Acc. Chem. Res.* **41**, 120–129 (2008). doi:[10.1021/ar7000815](https://doi.org/10.1021/ar7000815)
38. K. Chen, X. Chen, Integrin targeted delivery of chemotherapeutics. *Theranostics* **1**, 189–200 (2011). doi:[10.7150/thno.v01p0189](https://doi.org/10.7150/thno.v01p0189)
39. R. Stoltenburg, C. Reinemann, B. Strehlitz, SELEX—a (r)evolutionary method to generate high-affinity nucleic acid ligands. *Biomol. Eng.* **24**, 381–403 (2007). doi:[10.1016/j.bioeng.2007.06.001](https://doi.org/10.1016/j.bioeng.2007.06.001)
40. X. Wu, J. Chen, M. Wu, J.X. Zhao, Aptamers: active targeting ligands for cancer diagnosis and therapy. *Theranostics* **5**, 322–344 (2015). doi:[10.7150/thno.10257](https://doi.org/10.7150/thno.10257)
41. T.R. Daniels, T. Delgado, G. Helguera, M.L. Penichet, The transferrin receptor part II: targeted delivery of therapeutic agents into cancer cells. *Clin. Immunol.* **121**, 159–176 (2006). doi:[10.1016/j.clim.2006.06.006](https://doi.org/10.1016/j.clim.2006.06.006)
42. E.T. Boder, W. Jiang, Engineering antibodies for cancer therapy. *Annu. Rev. Chem. Biomol. Eng.* **2**, 53–75 (2011). doi:[10.1146/annurev-chembioeng-061010-114142](https://doi.org/10.1146/annurev-chembioeng-061010-114142)
43. C. Bechara, S. Sagan, Cell-penetrating peptides: 20 years later, where do we stand? *FEBS Lett.* **587**, 1693–1702 (2013). doi:[10.1016/j.febslet.2013.04.031](https://doi.org/10.1016/j.febslet.2013.04.031)
44. A.K. Varkouhi, M. Scholte, G. Storm, H.J. Haisma, Endosomal escape pathways for delivery of biological. *J. Control. Release* **151**, 220–228 (2011). doi:[10.1016/j.jconrel.2010.11.004](https://doi.org/10.1016/j.jconrel.2010.11.004)
45. A.E. Sayed, S. Futaki, H. Harashima, Delivery of macromolecules using arginine-rich cell-penetrating peptides: ways to overcome endosomal entrapment. *AAPS J.* **11**, 13–22 (2009). doi:[10.1208/s12248-008-9071-2](https://doi.org/10.1208/s12248-008-9071-2)
46. Y.H. Li, J. Wang, M.G. Wientjes, J.L.S. Au, Delivery of nanomedicines to extracellular and intracellular compartments of a solid tumor. *Adv. Drug Deliv. Rev.* **64**, 29–39 (2012). doi:[10.1016/j.addr.2011.04.006](https://doi.org/10.1016/j.addr.2011.04.006)
47. W. Viricel, A. Mbarek, J. Leblond, Switchable lipids: conformational change for fast pH-triggered cytoplasmic delivery. *Angew. Chem. Int. Ed.* **54**, 12743–12747 (2015). doi:[10.1002/anie.201504661](https://doi.org/10.1002/anie.201504661)
48. D.H. Kim, I. Hwang, Direct targeting of proteins from the cytosol to organelles: the ER versus endosymbiotic organelles. *Traffic* **14**, 613–621 (2013). doi:[10.1111/tra.12043](https://doi.org/10.1111/tra.12043)
49. L.D. Field, J.B. Delehanty, Y.C. Chen, I.L. Medintz, Peptides for specifically targeting nanoparticles to cellular organelles: *Quo vadis?* *Acc. Chem. Res.* **48**, 1380–1390 (2015). doi:[10.1021/ar500449v](https://doi.org/10.1021/ar500449v)
50. S. Negi, S. Pandey, S.M. Srinivasan, A. Mohammed, C. Guda, LocSigDB: a database of protein localization signals. *Database* **2015**, 1–7 (2015). doi:[10.1093/database/bav003](https://doi.org/10.1093/database/bav003)
51. P. Verderio, S. Avvakumova, G. Alessio, M. Bellini, M. Colombo, E. Galbiati, S. Mazzucchelli, J.P. Avila, B. Santini, D. Prosperi, Delivering colloidal nanoparticles to mammalian cells: a nano–bio interface perspective. *Adv. Healthcare Mater.* **3**, 957–976 (2014). doi:[10.1002/adhm.201300602](https://doi.org/10.1002/adhm.201300602)
52. D. Schmaljohann, Thermo- and pH-responsive polymers in drug delivery. *Adv. Drug Deliv. Rev.* **58**, 1655–1670 (2006). doi:[10.1016/j.addr.2006.09.020](https://doi.org/10.1016/j.addr.2006.09.020)
53. E. Fleige, M.A. Quadir, R. Haag, Stimuli-responsive polymeric nanocarriers for the controlled transport of active compounds: concepts and applications. *Adv. Drug Deliv. Rev.* **64**, 866–884 (2012). doi:[10.1016/j.addr.2012.01.020](https://doi.org/10.1016/j.addr.2012.01.020)
54. C. Brieke, F. Rohrbach, A. Gottschalk, G. Mayer, A. Heckel, Light-controlled tools. *Angew. Chem. Int. Ed.* **51**, 8446–8476 (2012). doi:[10.1002/anie.201202134](https://doi.org/10.1002/anie.201202134)
55. P. Klan, T. Solomek, C.G. Bochet, A. Blanc, R. Givens, M. Rubina, V. Popik, A. Kostikov, J. Wirz, Photoremovable protecting groups in chemistry and biology: reaction mechanisms and efficacy. *Chem. Rev.* **113**, 119–191 (2013). doi:[10.1021/cr300177k](https://doi.org/10.1021/cr300177k)
56. Y.H. Yang, H. Aloysius, D. Inoyama, Y. Chen, L.Q. Hu, Enzyme-mediated hydrolytic activation of prodrugs. *Acta Pharmaceutica Sinica B* **1**, 143–159 (2011). doi:[10.1016/j.japsb.2011.08.001](https://doi.org/10.1016/j.japsb.2011.08.001)
57. D.H. Nguyen, J.S. Lee, J.W. Bae, J.H. Choi, Y. Lee, J.Y. Son, K.D. Park, Targeted doxorubicin nanotherapy strongly suppressing growth of multidrug resistant tumor in mice. *Int. J. Pharm.* **495**, 329–335 (2015). doi:[10.1016/j.ijpharm.2015.08.083](https://doi.org/10.1016/j.ijpharm.2015.08.083)
58. X. Zhao, P. Liu, Biocompatible graphene oxide as a folate receptor-targeting drug delivery system for the controlled release of anti-cancer drugs. *RSC Adv.* **4**, 24232–24239 (2014). doi:[10.1039/c4ra02466d](https://doi.org/10.1039/c4ra02466d)
59. H. Hosoya, A.S. Dobroff, W.H.P. Driessen, V. Cristini, L.M. Brinker, F.I. Staquicini, M. Cardó-Vila, S. D'Angelo, F. Ferrara, B. Proneth, Y.S. Lin, D.R. Dunphy, P. Dogra, M.P. Melancon, R.J. Stafford, K. Miyazono, J.G. Gelovani, K. Kataoka, C.J. Brinker, R.L. Sidman, W. Arapc, R. Pasqualini,

- Integrated nanotechnology platform for tumor-targeted multimodal imaging and therapeutic cargo release. *Proc. Natl. Acad. Sci. USA* **113**, 1877–1882 (2016). doi:[10.1073/pnas.1525796113](https://doi.org/10.1073/pnas.1525796113)
60. Z. Zhang, Y. Liu, C. Jarreau, M.J. Welch, J.S.A. Taylor, Nucleic acid-directed self-assembly of multifunctional gold nanoparticle imaging agents. *Biomater. Sci.* **1**, 1055–1064 (2013). doi:[10.1039/c3bm60070j](https://doi.org/10.1039/c3bm60070j)
 61. K. Minamihata, Y. Maeda, S. Yamaguchi, W. Ishihara, A. Ishiwatari, S. Takamori, S. Yamahira, T. Nagamune, Photosensitizer and polycationic peptide-labeled streptavidin as a nano-carrier for light-controlled protein transduction. *J. Biosci. Bioeng.* **120**, 630–636 (2015). doi:[10.1016/j.jbiosc.2015.04.001](https://doi.org/10.1016/j.jbiosc.2015.04.001)
 62. A. Ishiwatari, S. Yamaguchi, S. Takamori, S. Yamahira, K. Minamihata, T. Nagamune, Photolytic protein aggregates: versatile materials for controlled release of active proteins. *Adv. Healthcare Mater.* **5**, 1002–1007 (2016). doi:[10.1002/adhm.201500957](https://doi.org/10.1002/adhm.201500957)
 63. M. Mishina, K. Minamihata, K. Moriyama, T. Nagamune, Peptide tag-induced horseradish peroxidase-mediated preparation of a streptavidin-immobilized redox-sensitive hydrogel. *Biomacromol* **17**, 1978–1984 (2016). doi:[10.1021/acs.biomac.6b00149](https://doi.org/10.1021/acs.biomac.6b00149)
 64. X. Zhang, Q. Guo, D. Cui, Recent advances in nanotechnology applied to biosensors. *Sensors* **9**, 1033–1053 (2009). doi:[10.3390/s9021033](https://doi.org/10.3390/s9021033)
 65. M. Holzinger, A.L. Goff, S. Cosnier, Nanomaterials for biosensing applications: a review. *Front. Chem.* **2**, 63 (2014). doi:[10.3389/fchem.2014.00063](https://doi.org/10.3389/fchem.2014.00063)
 66. S. Vaddiraju, I. Tomazos, D.J. Burgess, F.C. Jaing, F. Papadimitrakopoulos, Emerging synergy between nanotechnology and implantable biosensors: a review. *Biosens. Bioelectron.* **25**, 1553–1565 (2010). doi:[10.1016/j.bios.2009.12.001](https://doi.org/10.1016/j.bios.2009.12.001)
 67. J. Nam, N. Won, J. Bang, H. Jin, J. Park, S. Jung, S. Jung, Y. Park, S. Kim, Surface engineering of inorganic nanoparticles for imaging and therapy. *Adv. Drug Deliv. Rev.* **65**, 622–648 (2013). doi:[10.1016/j.addr.2012.08.015](https://doi.org/10.1016/j.addr.2012.08.015)
 68. S.S. Agasti, S. Rana, M.H. Park, C.K. Kim, C.C. You, V.M. Rotello, Nanoparticles for detection and diagnosis. *Adv. Drug Deliv. Rev.* **62**, 316–328 (2010). doi:[10.1016/j.addr.2009.11.004](https://doi.org/10.1016/j.addr.2009.11.004)
 69. H.S. Song, O.S. Kwon, S.H. Lee, S.J. Park, U.K. Kim, J. Jang, T.H. Park, Human taste receptor-functionalized field effect transistor as a human-like nanobioelectronic tongue. *Nano Lett.* **13**, 172–178 (2013). doi:[10.1021/nl3038147](https://doi.org/10.1021/nl3038147)
 70. O.S. Kwon, H.S. Song, S.J. Park, S.H. Lee, J.H. An, J.W. Park, H. Yang, H. Yoon, J. Bae, T.H. Park, J. Jang, An ultrasensitive, selective, multiplexed superbioelectronic nose that mimics the human sense of smell. *Nano Lett.* **15**, 6559–6567 (2015). doi:[10.1021/acs.nanolett.5b02286](https://doi.org/10.1021/acs.nanolett.5b02286)
 71. K. Balasubramanian, M. Burghard, Biosensors based on carbon nanotubes. *Anal. Bioanal. Chem.* **385**, 452–468 (2006). doi:[10.1007/s00216-006-0314-8](https://doi.org/10.1007/s00216-006-0314-8)
 72. B.A. Kairdolf, A.M. Smith, T.H. Stokes, M.D. Wang, A.N. Young, S. Nie, Semiconductor quantum dots for bioimaging and biodiagnostic applications. *Annu. Rev. Anal. Chem.* **6**, 143–162 (2013). doi:[10.1146/annurev-anchem-060908-155136](https://doi.org/10.1146/annurev-anchem-060908-155136)
 73. T.H. Shin, Y. Choi, S. Kim, J. Cheon, Recent advances in magnetic nanoparticle-based multi-modal imaging. *Chem. Soc. Rev.* **44**, 4501–4516 (2015). doi:[10.1039/c4cs00345d](https://doi.org/10.1039/c4cs00345d)
 74. E. Boisselier, D. Astruc, Gold nanoparticles in nanomedicine: preparations, imaging, diagnostics, therapies and toxicity. *Chem. Soc. Rev.* **38**, 1759–1782 (2009). doi:[10.1039/b806051g](https://doi.org/10.1039/b806051g)
 75. S.K. Vashist, A.G. Venkatesh, K. Mitsakakis, G. Czilwik, G. Roth, F. von Stetten, R. Zengerle, Nanotechnology-based biosensors and diagnostics: technology push versus industrial/healthcare requirements. *Bionanoscience* **2**, 115–126 (2012). doi:[10.1007/s12668-012-0047-4](https://doi.org/10.1007/s12668-012-0047-4)
 76. J.F. Rusling, C.V. Kumar, J.S. Gutkind, V. Patel, Measurement of biomarker proteins for point-of-care early detection and monitoring of cancer. *Analyst* **135**, 2496–2511 (2010). doi:[10.1039/c0an00204f](https://doi.org/10.1039/c0an00204f)
 77. S.T. Sanjay, G. Fu, M. Dou, F. Xu, R. Liu, H. Qie, X.J. Li, Biomarker detection for disease diagnosis using cost-effective microfluidic platforms. *Analyst* **140**, 7062–7081 (2015). doi:[10.1039/c5an00780a](https://doi.org/10.1039/c5an00780a)
 78. D.M. Bruls, T.H. Evers, J.A.H. Kahlman, P.J.W. van Lankvelt, M. Ovsyanko, E.G.M. Pelssers, J.J.H.B. Schleipen, F.K. de Theije, C.A. Verschuren, T. van der Wijk, J.B.A. van Zon, W.U. Dittmer, A.H.J. Immink, J.H. Nieuwenhuis, M.W.J. Prins, Rapid integrated biosensor for multiplexed immunoassays based on actuated magnetic nanoparticles. *Lab Chip* **9**, 3504–3510 (2009). doi:[10.1039/b913960e](https://doi.org/10.1039/b913960e)
 79. N.P. Sardesai, J.C. Barron, J.F. Rusling, Carbon nanotube microwell array for sensitive electrochemiluminescent detection of cancer biomarker proteins. *Anal. Chem.* **83**, 6698–6703 (2011). doi:[10.1021/ac201292q](https://doi.org/10.1021/ac201292q)
 80. N.P. Sardesai, K. Kadimisetty, R. Faria, J.F. Rusling, A microfluidic electrochemiluminescent device for detecting cancer biomarker proteins. *Anal. Bioanal. Chem.* **405**, 3831–3838 (2013). doi:[10.1007/s00216-012-6656-5](https://doi.org/10.1007/s00216-012-6656-5)
 81. Y.E. Choi, J.W. Kwak, J.W. Park, Nanotechnology for early cancer detection. *Sensors* **10**, 428–455 (2010). doi:[10.3390/s100100428](https://doi.org/10.3390/s100100428)
 82. R.V. Devi, M. Doble, R.S. Verma, Nanomaterials for early detection of cancer biomarker with special emphasis on gold nanoparticles in immunoassays/sensors. *Biosens. Bioelectron.* **68**, 688–698 (2015). doi:[10.1016/j.bios.2015.01.066](https://doi.org/10.1016/j.bios.2015.01.066)
 83. X. Yang, X.W. Shi, Y. Liu, W.E. Bentley, G.F. Payne, Orthogonal enzymatic reactions for the assembly of proteins at electrode addresses. *Langmuir* **25**, 338–344 (2009). doi:[10.1021/la802618q](https://doi.org/10.1021/la802618q)
 84. J.B. Kim, J.W. Grate, P. Wang, Nanobiocatalysis and its potential applications. *Trends Biotechnol.* **26**, 639–646 (2008). doi:[10.1016/j.tibtech.2008.07.009](https://doi.org/10.1016/j.tibtech.2008.07.009)
 85. P. Wang, Nanoscale biocatalyst systems. *Curr. Opin. Biotechnol.* **17**, 574–579 (2006). doi:[10.1016/j.copbio.2006.10.009](https://doi.org/10.1016/j.copbio.2006.10.009)
 86. M.L. Verma, C.J. Barrow, M. Puri, Nanobiotechnology as a novel paradigm for enzyme immobilisation and stabilisation with potential applications in biodiesel production. *Appl. Microbiol. Biotechnol.* **97**, 23–39 (2013). doi:[10.1007/s00253-012-4535-9](https://doi.org/10.1007/s00253-012-4535-9)
 87. D. Cardinale, N. Carette, T. Michon, Virus scaffolds as enzyme nano-carriers. *Trends Biotechnol.* **30**, 369–376 (2012). doi:[10.1016/j.tibtech.2012.04.001](https://doi.org/10.1016/j.tibtech.2012.04.001)
 88. M.L. Verma, M. Puri, C.J. Barrow, Recent trends in nanomaterials immobilised enzymes for biofuel production. *Crit. Rev. Biotechnol.* **36**, 108–119 (2016). doi:[10.3109/07388551.2014.928811](https://doi.org/10.3109/07388551.2014.928811)
 89. A.A. Homaei, R. Sariri, F. Vianello, R. Stevanato, Enzyme immobilization: an update. *J. Chem. Biol.* **6**, 185–205 (2013). doi:[10.1007/s12154-013-0102-9](https://doi.org/10.1007/s12154-013-0102-9)
 90. J. Kim, J.W. Grate, Single-enzyme nanoparticles armored by a nanometer-scale organic/inorganic network. *Nano Lett.* **3**, 1219–1222 (2003). doi:[10.1021/nl034404b](https://doi.org/10.1021/nl034404b)
 91. A. Butt, A. Farrukh, A. Ghaffar, H. Duran, Z. Oluz, H. ur Rehman, T. Hussain, R. Ahmad, A. Tahir, B. Yameen, Design of enzyme-immobilized polymer brush-grafted magnetic nanoparticles for efficient nematocidal activity. *RSC Adv.* **5**, 77682–77688 (2015). doi:[10.1039/c5ra10063a](https://doi.org/10.1039/c5ra10063a)
 92. G. Bayramoglu, M.Y. Arica, Reversible immobilization of catalase on fibrous polymer grafted and metal chelated chitosan membrane. *J. Mol. Catal. B Enzym.* **62**, 297–304 (2010). doi:[10.1016/j.molcatb.2009.11.013](https://doi.org/10.1016/j.molcatb.2009.11.013)
 93. J. Fu, J. Reinhold, N.W. Woodbury, Peptide-modified surfaces for enzyme immobilization. *PLoS ONE* **6**, e18692 (2011). doi:[10.1371/journal.pone.0018692](https://doi.org/10.1371/journal.pone.0018692)
 94. A. Care, P.L. Bergquist, A. Sunna, Solid-binding peptides: smart tools for nanobiotechnology. *Trends Biotechnol.* **33**, 259–268 (2015). doi:[10.1016/j.tibtech.2015.02.005](https://doi.org/10.1016/j.tibtech.2015.02.005)
 95. Y. Takatsujia, R. Yamasakia, A. Iwanaga, M. Lienemann, M.B. Linder, T. Haruyama, Solid-support immobilization of a "swing" fusion protein for enhanced glucose oxidase catalytic activity. *Colloids Surf. B Biointerfaces* **112**, 186–191 (2013). doi:[10.1016/j.colsurfb.2013.07.051](https://doi.org/10.1016/j.colsurfb.2013.07.051)
 96. U.B. Sleytr, B. Schuster, E.M. Egelseer, D. Pum, S-layers: principles and applications. *FEMS Microbiol. Rev.* **38**, 823–864 (2014). doi:[10.1111/1574-6976.12063](https://doi.org/10.1111/1574-6976.12063)
 97. M.H. Yang, B.G. Choi, T.J. Park, N.S. Heo, W.H. Hong, S.Y. Lee, Site-specific immobilization of gold binding polypeptide on gold nanoparticle-coated graphene sheet for biosensor application. *Nanoscale* **3**, 2950–2956 (2011). doi:[10.1039/c1nr10197h](https://doi.org/10.1039/c1nr10197h)
 98. H. Tschiggerl, A. Breitwieser, G. de Roo, T. Verwoerd, C. Schaffer, U.B. Sleytr, Exploitation of the S-layer self-assembly system for site directed immobilization of enzymes demonstrated for an extremophilic laminarinase from *Pyrococcus furiosus*. *J. Biotechnol.* **133**, 403–411 (2008). doi:[10.1016/j.jbiotec.2007.09.018](https://doi.org/10.1016/j.jbiotec.2007.09.018)
 99. S. Schoffelen, J.C.M. van Hest, Chemical approaches for the construction of multi-enzyme reaction systems. *Curr. Opin. Struct. Biol.* **23**, 613–621 (2013). doi:[10.1016/j.sbi.2013.06.010](https://doi.org/10.1016/j.sbi.2013.06.010)
 100. A. Kuchler, M. Yoshimoto, S. Luginbühl, F. Mavelli, P. Walde, Enzymatic reactions in confined environments. *Nat. Nanotechnol.* **11**, 409–420 (2016). doi:[10.1038/NNANO.2016.54](https://doi.org/10.1038/NNANO.2016.54)

101. J.A. Mullaney, B.H.A. Rehm, Design of a single-chain multi-enzyme fusion protein establishing the polyhydroxybutyrate biosynthesis pathway. *J. Biotechnol.* **147**, 31–36 (2010). doi:[10.1016/j.jbiotec.2010.02.021](https://doi.org/10.1016/j.jbiotec.2010.02.021)
102. M.C.M. van Oers, F.P.J.T. Rutjes, J.C.M. van Hest, Cascade reactions in nanoreactors. *Curr. Opin. Biotechnol.* **28**, 10–16 (2014). doi:[10.1016/j.copbio.2013.10.011](https://doi.org/10.1016/j.copbio.2013.10.011)
103. R. Chen, Q. Chen, H. Kim, K.H. Siu, Q. Sun, S.L. Tsai, W. Chen, Biomolecular scaffolds for enhanced signaling and catalytic efficiency. *Curr. Opin. Biotechnol.* **28**, 59–68 (2014). doi:[10.1016/j.copbio.2013.11.007](https://doi.org/10.1016/j.copbio.2013.11.007)
104. T.W. Muir, Semisynthesis of proteins by expressed protein ligation. *Annu. Rev. Biochem.* **72**, 249–289 (2003). doi:[10.1146/annurev.biochem.72.121801.161900](https://doi.org/10.1146/annurev.biochem.72.121801.161900)
105. T. Heck, G. Faccio, M. Richter, L. Thöny-Meyer, Enzyme-catalyzed protein crosslinking. *Appl. Microbiol. Biotechnol.* **97**, 461–475 (2013). doi:[10.1007/s00253-012-4569-z](https://doi.org/10.1007/s00253-012-4569-z)
106. H. Hirakawa, T. Haga, T. Nagamune, Artificial protein complexes for biocatalysis. *Top. Catal.* **55**, 1124–1137 (2012). doi:[10.1007/s11244-012-9900-5](https://doi.org/10.1007/s11244-012-9900-5)
107. O. Idan, H. Hess, Engineering enzymatic cascades on nanoscale scaffolds. *Curr. Opin. Biotechnol.* **24**, 606–611 (2013). doi:[10.1016/j.copbio.2013.01.003](https://doi.org/10.1016/j.copbio.2013.01.003)
108. H. Hirakawa, N. Kamiya, T. Tanaka, T. Nagamune, Intramolecular electron transfer in a cytochrome P450cam system with a site-specific branched structure. *Protein Eng. Des. Sel.* **20**, 453–459 (2007). doi:[10.1093/protein/gzm045](https://doi.org/10.1093/protein/gzm045)
109. H. Hirakawa, N. Kamiya, Y. Kawarabayasi, T. Nagamune, Artificial self-sufficient P450 in reversed micelles. *Molecules* **15**, 2935–2948 (2010). doi:[10.3390/molecules15052935](https://doi.org/10.3390/molecules15052935)
110. M.C. Good, J.G. Zalatan, W.A. Lim, Scaffold proteins: hubs for controlling the flow of cellular information. *Science* **332**, 680–686 (2011). doi:[10.1126/science.1198701](https://doi.org/10.1126/science.1198701)
111. H. Hirakawa, T. Nagamune, Molecular assembly of P450 with ferredoxin and ferredoxin reductase by fusion to PCNA. *ChemBioChem* **11**, 1517–1520 (2010). doi:[10.1002/cbic.201000226](https://doi.org/10.1002/cbic.201000226)
112. W.R. Whitaker, J.E. Dueber, Metabolic pathway flux enhancement by synthetic protein scaffolding. *Methods Enzymol.* **497**, 447–468 (2011). doi:[10.1016/B978-0-12-385075-1.00019-6](https://doi.org/10.1016/B978-0-12-385075-1.00019-6)
113. C. You, S. Myung, Y.H.P. Zhang, Facilitated substrate channeling in a self-assembled trifunctional enzyme complex. *Angew. Chem. Int. Ed.* **51**, 8787–8790 (2012). doi:[10.1002/ange.201202441](https://doi.org/10.1002/ange.201202441)
114. J.E. Dueber, G.C. Wu, G.R. Malmirchegini, T.S. Moon, C.J. Petzold, A.V. Ullal, K.L.J. Prather, J.D. Keasling, Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* **27**, 753–761 (2009). doi:[10.1038/nbt.1557](https://doi.org/10.1038/nbt.1557)
115. H. Watanabe, H. Hirakawa, T. Nagamune, Phosphite-driven self-sufficient cytochrome P450. *ChemCatChem* **5**, 3835–3840 (2013). doi:[10.1002/cctc.201300445](https://doi.org/10.1002/cctc.201300445)
116. S. Morais, Y. Barak, J. Caspi, Y. Hadar, R. Lamed, Y. Shoham, D.B. Wilson, E.A. Bayer, Cellulase-xylanase synergy in designer cellulosomes for enhanced degradation of a complex cellulosic substrate. *mBio* **1**, e00285–10 (2010). doi:[10.1128/mBio.00285-10](https://doi.org/10.1128/mBio.00285-10)
117. C.M. Niemeyer, Semisynthetic DNA–protein conjugates for biosensing and nanofabrication. *Angew. Chem. Int. Ed.* **49**, 1200–1216 (2010). doi:[10.1002/anie.200904930](https://doi.org/10.1002/anie.200904930)
118. C. Teller, I. Willner, Organizing protein–DNA hybrids as nanostructures with programmed functionalities. *Trends Biotechnol.* **28**, 619–628 (2010). doi:[10.1016/j.tibtech.2010.09.005](https://doi.org/10.1016/j.tibtech.2010.09.005)
119. J. Fu, M. Liu, Y. Liu, H. Yan, Spatially-interactive biomolecular networks organized by nucleic acid nanostructures. *Acc. Chem. Res.* **45**, 1215–1226 (2012). doi:[10.1021/ar200295q](https://doi.org/10.1021/ar200295q)
120. Y.R. Yang, Y. Liu, H. Yan, DNA nanostructures as programmable biomolecular scaffolds. *Bioconjugate Chem.* **26**, 1381–1395 (2015). doi:[10.1021/acs.bioconjchem.5b00194](https://doi.org/10.1021/acs.bioconjchem.5b00194)
121. M. Erkelenz, C.H. Kuo, C.M. Niemeyer, DNA-mediated assembly of cytochrome P450 BM3 subdomains. *J. Am. Chem. Soc.* **133**, 16111–16118 (2011). doi:[10.1021/ja204993s](https://doi.org/10.1021/ja204993s)
122. O.I. Willner, Y. Weizmann, R. Gill, O. Lioubashevski, R. Freeman, I. Willner, Enzyme cascades activated on topologically programmed DNA scaffolds. *Nat. Nanotechnol.* **4**, 249–254 (2009). doi:[10.1038/NNANO.2009.50](https://doi.org/10.1038/NNANO.2009.50)
123. J. Fu, M. Liu, Y. Liu, N.W. Woodbury, H. Yan, Interenzyme substrate diffusion for an enzyme cascade organized on spatially addressable DNA nanostructures. *J. Am. Chem. Soc.* **134**, 5516–5519 (2012). doi:[10.1021/ja300897h](https://doi.org/10.1021/ja300897h)
124. Y. Fu, D. Zeng, J. Chao, Y. Jin, Z. Zhang, H. Liu, D. Li, H. Ma, Q. Huang, K.V. Gothelf, C. Fan, Single-step rapid assembly of DNA origami nanostructures for addressable nanoscale bioreactors. *J. Am. Chem. Soc.* **135**, 696–702 (2013). doi:[10.1021/ja3076692](https://doi.org/10.1021/ja3076692)
125. V. Linko, M. Eerikainen, M.A. Kostianen, A modular DNA origami-based enzyme cascade nanoreactor. *Chem. Commun.* **51**, 5351–5354 (2015). doi:[10.1039/c4cc08472a](https://doi.org/10.1039/c4cc08472a)
126. C.J. Delebecque, A.B. Lindner, P.A. Silver, F.A. Aldaye, Organization of intracellular reactions with rationally designed RNA assemblies. *Science* **333**, 470–474 (2011). doi:[10.1126/science.1206938](https://doi.org/10.1126/science.1206938)
127. N.C. Seeman, Nanomaterials based on DNA. *Annu. Rev. Biochem.* **79**, 65–87 (2010). doi:[10.1146/annurev-biochem-060308-102244](https://doi.org/10.1146/annurev-biochem-060308-102244)
128. N.C. Seeman, P.S. Lukema, Nucleic acid nanostructures: bottom-up control of geometry on the nanoscale. *Rep. Prog. Phys.* **68**, 237–270 (2005). doi:[10.1088/0034-4885/68/1/R05](https://doi.org/10.1088/0034-4885/68/1/R05)
129. H. Kim, Y. Park, J. Kim, J. Jeong, S. Han, J.S. Lee, J.B. Lee, Nucleic acid engineering: RNA following the trail of DNA. *ACS Comb. Sci.* **18**, 87–99 (2016). doi:[10.1021/acscombsci.5b00108](https://doi.org/10.1021/acscombsci.5b00108)
130. P.W.K. Rothmund, Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006). doi:[10.1038/nature04586](https://doi.org/10.1038/nature04586)
131. M. Endo, Y. Yang, H. Sugiyama, DNA origami technology for biomaterials applications. *Biomater. Sci.* **1**, 347–360 (2013). doi:[10.1039/c2bm00154c](https://doi.org/10.1039/c2bm00154c)
132. E.J. Cho, J.W. Lee, A.D. Ellington, Applications of aptamers as sensors. *Annu. Rev. Anal. Chem.* **2**, 241–264 (2009). doi:[10.1146/annurev.anchem.1.031207.112851](https://doi.org/10.1146/annurev.anchem.1.031207.112851)
133. M. Darmostuk, S. Rimpelova, H. Gbelcova, T. Ruml, Current approaches in SELEX: an update to aptamer selection technology. *Biotechnol. Adv.* **33**, 1141–1161 (2015). doi:[10.1016/j.biotechadv.2015.02.008](https://doi.org/10.1016/j.biotechadv.2015.02.008)
134. H. Sun, Y. Zu, A highlight of recent advances in aptamer technology and its application. *Molecules* **20**, 11959–11980 (2015). doi:[10.3390/molecules200711959](https://doi.org/10.3390/molecules200711959)
135. M. Blind, M. Blank, Aptamer selection technology and recent advances. *Mol. Ther. Nucleic Acids* **4**, e223 (2015). doi:[10.1038/mtna.2014.74](https://doi.org/10.1038/mtna.2014.74)
136. A.D. Keefe, S. Pai, A. Ellington, Aptamers as therapeutics. *Nat. Rev. Drug Discov.* **9**, 537–550 (2010). doi:[10.1038/nrd3141](https://doi.org/10.1038/nrd3141)
137. G.M. Stovall, R.S. Bedenbaugh, S. Singh, A.J. Meyer, P.J. Hatala, A.D. Ellington, B. Hall, In vitro selection using modified or unnatural nucleotides. *Curr. Protoc. Nucleic Acid Chem.* **56**, 9.6.1–9.6.33 (2014). doi:[10.1002/0471142700.nc0906s56](https://doi.org/10.1002/0471142700.nc0906s56)
138. S. Jäger, G. Rasched, H. Kornreich-Leshem, M. Engeser, O. Thum, M. Famulok, A versatile toolbox for variable DNA functionalization at high density. *J. Am. Chem. Soc.* **127**, 15071–15082 (2005). doi:[10.1021/ja051725b](https://doi.org/10.1021/ja051725b)
139. S. Song, L. Wang, J. Li, J. Zhao, C. Fan, Aptamer-based biosensors. *Trends Anal. Chem.* **27**, 108–117 (2008). doi:[10.1016/j.trac.2007.12.004](https://doi.org/10.1016/j.trac.2007.12.004)
140. Y. Liu, C. Lin, H. Li, H. Yan, Aptamer-directed self-assembly of protein arrays on a DNA nanostructure. *Angew. Chem. Int. Ed.* **44**, 4333–4338 (2005). doi:[10.1002/anie.200501089](https://doi.org/10.1002/anie.200501089)
141. J. Mulhbachter, P. St-Pierre, D.A. Lafontaine, Therapeutic applications of ribozymes and riboswitches. *Curr. Opin. Pharmacol.* **10**, 551–556 (2010). doi:[10.1016/j.coph.2010.07.002](https://doi.org/10.1016/j.coph.2010.07.002)
142. K.Y. Jen, A.M. Gewirtz, Suppression of gene expression by targeted disruption of messenger RNA: available options and current strategies. *Stem Cells* **18**, 307–319 (2000). doi:[10.1634/stemcells.18-5-307](https://doi.org/10.1634/stemcells.18-5-307)
143. J. Rossi, Ribozymes, genomeics and therapeutics. *Chem. Biol.* **6**, R33–R37 (1999). doi:[10.1016/S1074-5521\(99\)80001-5](https://doi.org/10.1016/S1074-5521(99)80001-5)
144. D.S. Wilson, J.W. Szostak, In vitro selection of functional nucleic acids. *Annu. Rev. Biochem.* **68**, 611–647 (1999). doi:[10.1146/annurev.biochem.68.1.611](https://doi.org/10.1146/annurev.biochem.68.1.611)
145. G.F. Joyce, Directed evolution of nucleic acid enzymes. *Annu. Rev. Biochem.* **73**, 791–836 (2004). doi:[10.1146/annurev.biochem.73.011303.073717](https://doi.org/10.1146/annurev.biochem.73.011303.073717)
146. A. Jäschke, Artificial ribozymes and deoxyribozymes. *Curr. Opin. Struct. Biol.* **11**, 321–326 (2001). doi:[10.1016/S0959-440X\(00\)00208-6](https://doi.org/10.1016/S0959-440X(00)00208-6)
147. J.H. Jeong, Y.W. Cho, B. Jung, K. Park, J.-D. Kim, Self-assembled nanoparticles of ribozymes with poly(ethylene glycol)-b-poly(L-lysine) block copolymers. *Jpn. J. Appl. Phys.* **45**, 591–595 (2006). doi:[10.1143/JJAP.45.591](https://doi.org/10.1143/JJAP.45.591)

148. J.L. Rouge, T.L. Sita, L. Hao, F.M. Kouri, W.E. Briley, A.H. Stegh, C.A. Mirkin, Ribozyme-spherical nucleic acids. *J. Am. Chem. Soc.* **137**, 10528–10531 (2015). doi:[10.1021/jacs.5b07104](https://doi.org/10.1021/jacs.5b07104)
149. R.C. Cadwell, G.F. Joyce, Randomization of genes by PCR mutagenesis. *Genome Res.* **2**, 28–33 (1992). doi:[10.1101/gr.2.1.28](https://doi.org/10.1101/gr.2.1.28)
150. H. Ueda, M. Nakanishi, E. Suzuki, T. Nagamune, Enhancement of mutation frequency with nucleotide triphosphate analogs in PCR random mutagenesis. *J. Ferment. Technol.* **79**, 303–305 (1995). doi:[10.1016/0922-338X\(95\)90624-9](https://doi.org/10.1016/0922-338X(95)90624-9)
151. M. Zaccolo, D.M. Williams, D.M. Brown, E. Gherardi, An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues. *J. Mol. Biol.* **255**, 589–603 (1996). doi:[10.1006/jmbi.1996.0049](https://doi.org/10.1006/jmbi.1996.0049)
152. R. Fujii, M. Kitaoka, K. Hayashi, Error-prone rolling circle amplification: the simplest random mutagenesis protocol. *Nat. Protoc.* **1**, 2493–2497 (2006). doi:[10.1038/nprot.2006.403](https://doi.org/10.1038/nprot.2006.403)
153. Y.P. Lai, J. Huang, L.F. Wang, J. Li, Z.R. Wu, A new approach to random mutagenesis in vitro. *Biotechnol. Bioeng.* **86**, 622–627 (2004). doi:[10.1002/bit.20066](https://doi.org/10.1002/bit.20066)
154. B. Sullivan, A.Z. Walton, J.D. Stewart, Library construction and evaluation for site saturation mutagenesis. *Enzyme Microb. Technol.* **53**, 70–77 (2013). doi:[10.1016/j.enzmictec.2013.02.012](https://doi.org/10.1016/j.enzmictec.2013.02.012)
155. W.P.C. Stemmer, DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA* **91**, 10747–10751 (1994). doi:[10.1073/pnas.91.22.10747](https://doi.org/10.1073/pnas.91.22.10747)
156. M. Kikuchi, K. Ohnishi, S. Harayama, Novel family shuffling methods for the in vitro evolution of enzymes. *Gene* **236**, 159–167 (1999). doi:[10.1016/S0378-1119\(99\)00240-1](https://doi.org/10.1016/S0378-1119(99)00240-1)
157. M. Kikuchi, K. Ohnishi, S. Harayama, An effective family shuffling method using single-stranded DNA. *Gene* **243**, 133–137 (2000). doi:[10.1016/S0378-1119\(99\)00547-8](https://doi.org/10.1016/S0378-1119(99)00547-8)
158. V. Sieber, C.A. Martinez, F.H. Arnold, Libraries of hybrid proteins from distantly related sequences. *Nat. Biotechnol.* **19**, 456–460 (2001). doi:[10.1038/88129](https://doi.org/10.1038/88129)
159. S. Lutz, M. Ostermeier, S.J. Benkovic, Rapid generation of incremental truncation libraries for protein engineering using α -phosphothioate nucleotides. *Nucleic Acids Res.* **29**, 16e (2001). doi:[10.1093/nar/29.4.e16](https://doi.org/10.1093/nar/29.4.e16)
160. S. Lutz, M. Ostermeier, G.L. Moore, C.D. Maranas, S.J. Benkovic, Creating multiple-crossover DNA libraries independent of sequence identity. *Proc. Natl. Acad. Sci. USA* **98**, 11248–11253 (2001). doi:[10.1073/pnas.201413698](https://doi.org/10.1073/pnas.201413698)
161. J.A. Brannigan, A.J. Wilkinson, Protein engineering 20 years on. *Nat. Rev. Mol. Cell Biol.* **3**, 964–970 (2002). doi:[10.1038/nrm975](https://doi.org/10.1038/nrm975)
162. J.A. Kolkman, W.P.C. Stemmer, Directed evolution of proteins by exon shuffling. *Nat. Biotechnol.* **19**, 423–428 (2001). doi:[10.1038/88084](https://doi.org/10.1038/88084)
163. N. Doi, H. Yanagawa, Insertional gene fusion technology. *FEBS Lett.* **457**, 1–4 (1999). doi:[10.1016/S0014-5793\(99\)00991-6](https://doi.org/10.1016/S0014-5793(99)00991-6)
164. S. Elleuche, Bringing functions together with fusion enzymes—from nature's inventions to biotechnological applications. *Appl. Microbiol. Biotechnol.* **99**, 1545–1556 (2015). doi:[10.1007/s00253-014-6315-1](https://doi.org/10.1007/s00253-014-6315-1)
165. S.A. Navai, N. Ahmed, Targeting the tumour profile using broad spectrum chimaeric antigen receptor T-cells. *Biochem. Soc. Trans.* **44**, 391–396 (2016). doi:[10.1042/BST20150266](https://doi.org/10.1042/BST20150266)
166. A. Miyawaki, Y. Niino, Molecular spies for bioimaging—fluorescent protein-based probes. *Mol. Cell* **58**, 632–643 (2015). doi:[10.1016/j.molcel.2015.03.002](https://doi.org/10.1016/j.molcel.2015.03.002)
167. H. Mattoussi, J.M. Mauro, E.R. Goldman, G.P. Anderson, V.C. Sundar, F.V. Mikulec, M.G. Bawendi, Self-assembly of CdSe–ZnS quantum dot bioconjugates using an engineered recombinant protein. *J. Am. Chem. Soc.* **122**, 12142–12150 (2000). doi:[10.1021/ja002535y](https://doi.org/10.1021/ja002535y)
168. R.A. Studer, B.H. Dessailly, C.A. Orengo, Residue mutations and their impact on protein structure and function: detecting beneficial and pathogenic changes. *Biochem. J.* **449**, 581–594 (2013). doi:[10.1042/BJ20121221](https://doi.org/10.1042/BJ20121221)
169. M. Suárez, A. Jaramillo, Challenges in the computational design of proteins. *J. R. Soc. Interface* **6**, S477–S491 (2009). doi:[10.1098/rsif.2008.0508.focus](https://doi.org/10.1098/rsif.2008.0508.focus)
170. J.G. Saven, Computational protein design: engineering molecular diversity, nonnatural enzymes, nonbiological cofactor complexes, and membrane proteins. *Curr. Opin. Chem. Biol.* **15**, 452–457 (2011). doi:[10.1016/j.cbpa.2011.03.014](https://doi.org/10.1016/j.cbpa.2011.03.014)
171. R.J. Pantazes, M.J. Grisewood, C.D. Maranas, Recent advances in computational protein design. *Curr. Opin. Struct. Biol.* **21**, 467–472 (2011). doi:[10.1016/j.sbi.2011.04.005](https://doi.org/10.1016/j.sbi.2011.04.005)
172. M.S. Packer, D.R. Liu, Methods for the directed evolution of proteins. *Nat. Rev. Genet.* **16**, 379–394 (2015). doi:[10.1038/nrg3927](https://doi.org/10.1038/nrg3927)
173. H. Leemhuis, V. Stein, A.D. Griffiths, F. Hollfelder, New genotype–phenotype linkages for directed evolution of functional proteins. *Curr. Opin. Struct. Biol.* **15**, 472–478 (2005). doi:[10.1016/j.sbi.2005.07.006](https://doi.org/10.1016/j.sbi.2005.07.006)
174. P.A.G. Tizei, E. Csibra, L. Torres, V.B. Pinheiro, Selection platforms for directed evolution in synthetic biology. *Biochem. Soc. Trans.* **44**, 1165–1175 (2016). doi:[10.1042/BST20160076](https://doi.org/10.1042/BST20160076)
175. H.M.E. Azzazya, W.E. Highsmith Jr., Phage display technology: clinical applications and recent innovations. *Clin. Biochem.* **35**, 425–445 (2002). doi:[10.1016/S0009-9120\(02\)00343-0](https://doi.org/10.1016/S0009-9120(02)00343-0)
176. A.R. Mäkelä, C. Oker-Blom, Baculovirus display: a multifunctional technology for gene delivery and eukaryotic library development. *Adv. Virus Res.* **68**, 91–112 (2006). doi:[10.1016/S0065-3527\(06\)68003-2](https://doi.org/10.1016/S0065-3527(06)68003-2)
177. S.Y. Lee, J.H. Choi, Z. Xu, Microbial cell-surface display. *Trends Biotechnol.* **21**, 45–52 (2003). doi:[10.1016/S0167-7799\(02\)00006-9](https://doi.org/10.1016/S0167-7799(02)00006-9)
178. M. Desvaux, E. Dumas, I. Chafsey, M. Hébraud, Protein cell surface display in Gram-positive bacteria. *FEMS Microbiol. Lett.* **256**, 1–15 (2006). doi:[10.1111/j.1574-6968.2006.00122.x](https://doi.org/10.1111/j.1574-6968.2006.00122.x)
179. T. Tanaka, R. Yamada, C. Ogino, A. Kondo, Recent developments in yeast cell surface display toward extended applications in biotechnology. *Appl. Microbiol. Biotechnol.* **95**, 577–591 (2012). doi:[10.1007/s00253-012-4175-0](https://doi.org/10.1007/s00253-012-4175-0)
180. S.A. Gai, K.D. Wittrup, Yeast surface display for protein engineering and characterization. *Curr. Opin. Struct. Biol.* **17**, 467–473 (2007). doi:[10.1016/j.sbi.2007.08.012](https://doi.org/10.1016/j.sbi.2007.08.012)
181. K.D. Wittrup, Protein engineering by cell-surface display. *Curr. Opin. Biotechnol.* **12**, 395–399 (2001). doi:[10.1016/S0958-1669\(00\)00233-0](https://doi.org/10.1016/S0958-1669(00)00233-0)
182. M. He, Cell-free protein synthesis: applications in proteomics and biotechnology. *New Biotechnol.* **25**, 126–131 (2008). doi:[10.1016/j.nbt.2008.08.004](https://doi.org/10.1016/j.nbt.2008.08.004)
183. E.D. Carlson, R. Gan, C.E. Hodgman, M.C. Jewett, Cell-free protein synthesis: applications come of age. *Biotechnol. Adv.* **30**, 1185–1194 (2012). doi:[10.1016/j.biotechadv.2011.09.016](https://doi.org/10.1016/j.biotechadv.2011.09.016)
184. W.D. Fessner, Systems biocatalysis: development and engineering of cell-free “artificial metabolisms” for preparative multi-enzymatic synthesis. *New Biotechnol.* **32**, 658–664 (2015). doi:[10.1016/j.nbt.2014.11.007](https://doi.org/10.1016/j.nbt.2014.11.007)
185. S.M. Schinn, A. Broadbent, W.T. Bradley, B.C. Bundy, Protein synthesis directly from PCR: progress and applications of cell-free protein synthesis with linear DNA. *New Biotechnol.* **33**, 480–487 (2016). doi:[10.1016/j.nbt.2016.04.002](https://doi.org/10.1016/j.nbt.2016.04.002)
186. D. Lipovsek, A. Plückthun, In-vitro protein evolution by ribosome display and mRNA display. *J. Immunol. Methods* **290**, 51–67 (2004). doi:[10.1016/j.jim.2004.04.008](https://doi.org/10.1016/j.jim.2004.04.008)
187. R. Odegrip, D. Coomber, B. Eldridge, R. Hederer, P.A. Kuhlman, C. Ullman, K. FitzGerald, D. McGregor, CIS display: In vitro selection of peptides from libraries of protein–DNA complexes. *Proc. Natl. Acad. Sci. USA* **101**, 2806–2810 (2004). doi:[10.1073/pnas.0400219101](https://doi.org/10.1073/pnas.0400219101)
188. J. Bertschinger, D. Neri, Covalent DNA display as a novel tool for directed evolution of proteins in vitro. *Protein Eng. Des. Sel.* **17**, 699–707 (2004). doi:[10.1093/protein/gzh082](https://doi.org/10.1093/protein/gzh082)
189. N. Doi, H. Yanagawa, STABLE: protein–DNA fusion system for screening of combinatorial protein libraries in vitro. *FEBS Lett.* **457**, 227–230 (1999). doi:[10.1016/S0014-5793\(99\)01041-8](https://doi.org/10.1016/S0014-5793(99)01041-8)
190. O. Nord, M. Uhlén, P.A. Nygren, Microbead display of proteins by cell-free expression of anchored DNA. *J. Biotechnol.* **106**, 1–13 (2003). doi:[10.1016/j.jbiotec.2003.09.002](https://doi.org/10.1016/j.jbiotec.2003.09.002)
191. D.S. Tawfik, A.D. Griffiths, Man-made cell-like compartments for molecular evolution. *Nat. Biotechnol.* **16**, 652–656 (1998). doi:[10.1038/nbt0798-652](https://doi.org/10.1038/nbt0798-652)
192. H. Lin, V.W. Cornish, Screening and selection methods for large-scale analysis of protein function. *Angew. Chem. Int. Ed.* **41**, 4402–4425 (2002). doi:[10.1002/1521-3773\(20021202\)41:23<4402:AID-ANIE4402>3.0.CO;2-H](https://doi.org/10.1002/1521-3773(20021202)41:23<4402:AID-ANIE4402>3.0.CO;2-H)
193. S. Becker, H.U. Schmoldt, T.M. Adams, S. Wilhelm, H. Kolmar, Ultra-high-throughput screening based on cell-surface display and fluorescence-activated cell sorting for the identification of novel biocatalysts. *Curr. Opin. Biotechnol.* **15**, 323–329 (2004). doi:[10.1016/j.copbio.2004.06.001](https://doi.org/10.1016/j.copbio.2004.06.001)

194. G. Yang, S.G. Withers, Ultrahigh-throughput FACS-based screening for directed enzyme evolution. *ChemBioChem* **10**, 2704–2715 (2009). doi:[10.1002/cbic.200900384](https://doi.org/10.1002/cbic.200900384)
195. A. Currin, N. Swainston, P.J. Dayace, D.B. Kell, Synthetic biology for the directed evolution of protein biocatalysts: navigating sequence space intelligently. *Chem. Soc. Rev.* **44**, 1172–1239 (2015). doi:[10.1039/c4cs000351a](https://doi.org/10.1039/c4cs000351a)
196. S. Kumar, Engineering cytochrome P450 biocatalysts for biotechnology, medicine, and bioremediation. *Expert. Opin. Drug. Metab. Toxicol.* **6**, 115–131 (2010). doi:[10.1517/17425250903431040](https://doi.org/10.1517/17425250903431040)
197. H. Zhao, K. Chockalingam, Z. Chen, Directed evolution of enzymes and pathways for industrial biocatalysis. *Curr. Opin. Biotechnol.* **13**, 104–110 (2002). doi:[10.1016/S0958-1669\(02\)00291-4](https://doi.org/10.1016/S0958-1669(02)00291-4)
198. A. Fernandez-Gacio, M. Uguen, J. Fastrez, Phage display as a tool for the directed evolution of enzymes. *Trends Biotechnol.* **21**, 408–414 (2003). doi:[10.1016/S0167-7799\(03\)00194-X](https://doi.org/10.1016/S0167-7799(03)00194-X)
199. J. McCafferty, D. Schofield, Identification of optimal protein binders through the use of large genetically encoded display libraries. *Curr. Opin. Chem. Biol.* **26**, 16–24 (2015). doi:[10.1016/j.cbpa.2015.01.003](https://doi.org/10.1016/j.cbpa.2015.01.003)
200. I. Chen, B.M. Dorr, D.R. Liu, A general strategy for the evolution of bond-forming enzymes using yeast display. *Proc. Natl. Acad. Sci. USA* **108**, 11399–11404 (2011). doi:[10.1073/pnas.1101046108](https://doi.org/10.1073/pnas.1101046108)
201. M. Ho, I. Pastan, Mammalian cell display for antibody engineering. *Methods Mol. Biol.* **525**, 337–352 (2009). doi:[10.1007/978-1-59745-554-1_18](https://doi.org/10.1007/978-1-59745-554-1_18)
202. P.M. Bowers, R.A. Horlick, T.Y. Neben, R.M. Toobian, G.L. Tomlinson, J.L. Dalton, H.A. Jones, A. Chen, L. Altobelli III, X. Zhang, J.L. Macomber, I.P. Krampf, B.F. Wu, A. McConnell, B. Chau, T. Holland, A.D. Berkebile, S.S. Neben, W.J. Boyle, D.J. King, Coupling mammalian cell surface display with somatic hypermutation for the discovery and maturation of human antibodies. *Proc. Natl. Acad. Sci. USA* **108**, 20455–20460 (2011). doi:[10.1073/pnas.1114010108](https://doi.org/10.1073/pnas.1114010108)
203. R. Yoshida, M. Kawahara, T. Nagamune, A novel platform for antibody library selection in mammalian cells based on a growth signalobody. *Biotechnol. Bioeng.* **111**, 1170–1179 (2014). doi:[10.1002/bit.25173](https://doi.org/10.1002/bit.25173)
204. R. Yoshida, M. Kawahara, T. Nagamune, Domain structure of growth signalobodies critically affects the outcome of antibody library selection. *J. Biochem.* **157**, 497–506 (2015). doi:[10.1093/jb/mvv008](https://doi.org/10.1093/jb/mvv008)
205. K. Soga, H. Abo, S.Y. Qin, T. Kyoutou, K. Hiemori, H. Tateno, N. Matsumoto, J. Hirabayashi, K. Yamamoto, Mammalian cell surface display as a novel method for developing engineered lectins with novel characteristics. *Biomolecules* **5**, 1540–1562 (2015). doi:[10.3390/biom5031540](https://doi.org/10.3390/biom5031540)
206. A.D. Griffiths, D.S. Tawfik, Directed evolution of an extremely fast phosphotriesterase by in vitro compartmentalization. *EMBO J.* **22**, 24–35 (2003). doi:[10.1093/emboj/cdg014](https://doi.org/10.1093/emboj/cdg014)
207. M. Naimuddin, T. Kubo, A high performance platform based on cDNA display for efficient synthesis of protein fusions and accelerated directed evolution. *ACS Comb. Sci.* **18**, 117–129 (2016). doi:[10.1021/acscmbosci.5b00139](https://doi.org/10.1021/acscmbosci.5b00139)
208. S.A. Richman, D.M. Kranz, Display, engineering, and applications of antigen-specific T cell receptors. *Biomol. Eng.* **24**, 361–373 (2007). doi:[10.1016/j.bioeng.2007.02.009](https://doi.org/10.1016/j.bioeng.2007.02.009)
209. C. Jäckel, D. Hilvert, Biocatalysts by evolution. *Curr. Opin. Biotechnol.* **21**, 753–759 (2010). doi:[10.1016/j.copbio.2010.08.008](https://doi.org/10.1016/j.copbio.2010.08.008)
210. A.S. Bommarius, J.K. Blum, M.J. Abrahamson, Status of protein engineering for biocatalysts: how to design an industrially useful biocatalyst. *Curr. Opin. Chem. Biol.* **15**, 194–200 (2011). doi:[10.1016/j.cbpa.2010.11.011](https://doi.org/10.1016/j.cbpa.2010.11.011)
211. E.M. Brustad, F.H. Arnold, Optimizing non-natural protein function with directed evolution. *Curr. Opin. Chem. Biol.* **15**, 201–210 (2011). doi:[10.1016/j.cbpa.2010.11.020](https://doi.org/10.1016/j.cbpa.2010.11.020)
212. Y. Li, P.C. Cirino, Recent advances in engineering proteins for biocatalysis. *Biotechnol. Bioeng.* **111**, 1273–1287 (2014). doi:[10.1002/bit.25240](https://doi.org/10.1002/bit.25240)
213. A.C. Obermeyer, B.D. Olsen, Synthesis and application of protein-containing block copolymers. *ACS Macro Lett.* **4**, 101–110 (2015). doi:[10.1021/mz500732e](https://doi.org/10.1021/mz500732e)
214. M.T. Smith, A.K. Hawes, B.C. Bundy, Reengineering viruses and virus-like particles through chemical functionalization strategies. *Curr. Opin. Biotechnol.* **24**, 620–626 (2013). doi:[10.1016/j.copbio.2013.01.011](https://doi.org/10.1016/j.copbio.2013.01.011)
215. J.M. Gilmore, R.A. Scheck, A.P. Esser-Kahn, N.S. Joshi, M.B. Francis, N-terminal protein modification through a biomimetic transamination reaction. *Angew. Chem. Int. Ed.* **45**, 5307–5311 (2006). doi:[10.1002/anie.200600368](https://doi.org/10.1002/anie.200600368)
216. R.P. Lyon, J.R. Setter, T.D. Bovee, S.O. Doronina, J.H. Hunter, M.E. Anderson, C.L. Balasubramanian, S.M. Duniho, C.I. Leiske, F. Li, P.D. Senter, Self-hydrolyzing maleimides improve the stability and pharmacological properties of antibody-drug conjugates. *Nat. Biotechnol.* **32**, 1059–1062 (2014). doi:[10.1038/nbt.2968](https://doi.org/10.1038/nbt.2968)
217. N.S. Joshi, J.R. Whitaker, M.B. Francis, A three-component Mannich-type reaction for selective tyrosine bioconjugation. *J. Am. Chem. Soc.* **126**, 15942–15943 (2004). doi:[10.1021/ja0439017](https://doi.org/10.1021/ja0439017)
218. Y. Seki, T. Ishiyama, D. Sasaki, J. Abe, Y. Sohna, K. Oisaki, M. Kanai, Transition metal-free tryptophan-selective bioconjugation of proteins. *J. Am. Chem. Soc.* **138**, 10798–10801 (2016). doi:[10.1021/jacs.6b06692](https://doi.org/10.1021/jacs.6b06692)
219. C.C. Liu, P.G. Schultz, Adding new chemistries to the genetic code. *Annu. Rev. Biochem.* **79**, 413–444 (2010). doi:[10.1146/annurev.biochem.052308.105824](https://doi.org/10.1146/annurev.biochem.052308.105824)
220. J.A. Johnson, Y.Y. Lu, J.A.V. Deventer, D.A. Tirrell, Residue-specific incorporation of non-canonical amino acids into proteins. *Curr. Opin. Chem. Biol.* **14**, 774–780 (2010). doi:[10.1016/j.cbpa.2010.09.013](https://doi.org/10.1016/j.cbpa.2010.09.013)
221. A.R. Goerke, J.R. Swartz, High-level cell-free synthesis yields of proteins containing site-specific non-natural amino acids. *Biotechnol. Bioeng.* **102**, 400–416 (2009). doi:[10.1002/bit.22070](https://doi.org/10.1002/bit.22070)
222. T. Hoshaka, M. Sisido, Incorporation of non-natural amino acids into proteins. *Curr. Opin. Chem. Biol.* **6**, 809–815 (2002). doi:[10.1016/S1367-5931\(02\)00376-9](https://doi.org/10.1016/S1367-5931(02)00376-9)
223. T. Passioura, H. Suga, Reprogramming the genetic code in vitro. *Trends Biochem. Sci.* **39**, 400–4008 (2014). doi:[10.1016/j.tibs.2014.07.005](https://doi.org/10.1016/j.tibs.2014.07.005)
224. K. Lang, J.W. Chin, Bioorthogonal reactions for labeling proteins. *ACS Chem. Biol.* **9**, 16–20 (2014). doi:[10.1021/cb4009292](https://doi.org/10.1021/cb4009292)
225. Y. Gong, L. Pan, Recent advances in bioorthogonal reactions for site-specific protein labeling and engineering. *Tetrahedron Lett.* **56**, 2123–2132 (2015). doi:[10.1016/j.tetlet.2015.03.065](https://doi.org/10.1016/j.tetlet.2015.03.065)
226. P.E. Dawson, S.B.H. Kent, Synthesis of native proteins by chemical ligation. *Annu. Rev. Biochem.* **69**, 923–960 (2000). doi:[10.1146/annurev.biochem.69.1.923](https://doi.org/10.1146/annurev.biochem.69.1.923)
227. L.R. Malins, R.J. Payne, Recent extensions to native chemical ligation for the chemical synthesis of peptides and proteins. *Curr. Opin. Chem. Biol.* **22**, 70–78 (2014). doi:[10.1016/j.cbpa.2014.09.021](https://doi.org/10.1016/j.cbpa.2014.09.021)
228. V. Muralidharan, T.W. Muir, Protein ligation: an enabling technology for the biophysical analysis of proteins. *Nat. Methods* **3**, 429–438 (2006). doi:[10.1038/nmeth886](https://doi.org/10.1038/nmeth886)
229. H. Tian, A. Fürstenberg, T. Huber, Labeling and single-molecule methods to monitor G protein-coupled receptor dynamics. *Chem. Rev.* **117**, 186–245 (2017). doi:[10.1021/acs.chemrev.6b00084](https://doi.org/10.1021/acs.chemrev.6b00084)
230. M. Rashidian, J.K. Dozier, M.D. Distefano, Enzymatic labeling of proteins: techniques and approaches. *Bioconjugate Chem.* **24**, 1277–1294 (2013). doi:[10.1021/bc400102w](https://doi.org/10.1021/bc400102w)
231. J.M. McFarland, D. Rabuka, Recent advances in chemoenzymatic bioconjugation methods. *Org. Chem. Insights* **5**, 7–14 (2015). doi:[10.4137/OCS17957](https://doi.org/10.4137/OCS17957)
232. S.A. Walper, K.B. Turner, I.L. Medintz, Enzymatic bioconjugation of nanoparticles: developing specificity and control. *Curr. Opin. Biotechnol.* **34**, 232–241 (2015). doi:[10.1016/j.copbio.2015.04.003](https://doi.org/10.1016/j.copbio.2015.04.003)
233. P. Wu, W. Shui, B.L. Carlson, N. Hu, D. Rabuka, J. Lee, C.R. Bertozzi, 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). doi:[10.1073/pnas.0807820106](https://doi.org/10.1073/pnas.0807820106)
234. J.K. Dozier, S.L. Khatwani, J.W. Wollack, Y.C. Wang, C. Schmidt-Dannert, M.D. Distefano, Engineering protein farnesyltransferase for enzymatic protein labeling applications. *Bioconjugate Chem.* **25**, 1203–1212 (2014). doi:[10.1021/bc500240p](https://doi.org/10.1021/bc500240p)
235. W.P. Heal, S.R. Wickramasinghe, P.W. Bowyer, A.A. Holder, D.F. Smith, R.J. Leatherbarrow, E.W. Tate, Site-specific N-terminal labelling of proteins in vitro and in vivo using *N*-myristoyl transferase and bioorthogonal ligation chemistry. *Chem. Commun.* 480–482 (2008). doi:[10.1039/B716115H](https://doi.org/10.1039/B716115H)
236. I. Chen, M. Howarth, W.Y. Lin, A.Y. Ting, Site-specific labeling of cell surface proteins with biophysical probes using biotin ligase. *Nat. Methods* **2**, 99–104 (2005). doi:[10.1038/nmeth735](https://doi.org/10.1038/nmeth735)
237. I. Chen, Y.A. Choi, A.Y. Ting, Phage display evolution of a peptide substrate for yeast biotin ligase and application to two-color quantum

- dot labeling of cell surface proteins. *J. Am. Chem. Soc.* **129**, 6619–6625 (2007). doi:[10.1021/ja071013g](https://doi.org/10.1021/ja071013g)
238. S.A. Slavoff, I. Chen, Y.A. Choi, A.Y. Ting, Expanding the substrate tolerance of biotin ligase through exploration of enzymes from diverse species. *J. Am. Chem. Soc.* **130**, 1160–1162 (2008). doi:[10.1021/ja076655i](https://doi.org/10.1021/ja076655i)
 239. J.D. Cohen, P. Zou, A.Y. Ting, Site-specific protein modification using lipoic acid ligase and bis-aryl hydrazone formation. *ChemBioChem* **13**, 888–894 (2012). doi:[10.1002/cgic.201100764](https://doi.org/10.1002/cgic.201100764)
 240. A. Fontana, B. Spolaore, A. Mero, F.M. Veronese, Site-specific modification and PEGylation of pharmaceutical proteins mediated by transglutaminase. *Adv. Drug Deliv. Rev.* **60**, 13–28 (2008). doi:[10.1016/j.addr.2007.06.015](https://doi.org/10.1016/j.addr.2007.06.015)
 241. N. Kamiya, T. Tanaka, T. Suzyuki, T. Takazawa, S. Takeda, K. Watanabe, T. Nagamune, S-peptide as a potent peptidyl linker for protein crosslinking by microbial transglutaminase from *Streptomyces mobaraensis*. *Bioconjug. Chem.* **14**, 351–357 (2003). doi:[10.1021/bc025610y](https://doi.org/10.1021/bc025610y)
 242. T. Tanaka, N. Kamiya, T. Nagamune, Peptidyl linkers for protein heterodimerization catalyzed by microbial transglutaminase. *Bioconjug. Chem.* **15**, 491–497 (2004). doi:[10.1021/bc034209o](https://doi.org/10.1021/bc034209o)
 243. T. Takazawa, N. Kamiya, H. Ueda, T. Nagamune, Enzymatic labeling of a single chain variable fragment of an antibody with alkaline phosphatase by microbial transglutaminase. *Biotechnol. Bioeng.* **86**, 399–404 (2004). doi:[10.1002/bit.20019](https://doi.org/10.1002/bit.20019)
 244. T. Tanaka, N. Kamiya, T. Nagamune, N-terminal glycine-specific protein conjugation catalyzed by microbial transglutaminase. *FEBS Lett.* **579**, 2092–2096 (2005). doi:[10.1016/j.febslet.2005.02.064](https://doi.org/10.1016/j.febslet.2005.02.064)
 245. S. Dasgupta, S. Samantaray, D. Sahal, R.P. Roy, Isopeptide ligation catalyzed by quintessential sortase A: mechanistic cues from cyclic and branched oligomers of indolicidin. *J. Biol. Chem.* **286**, 23996–24006 (2011). doi:[10.1074/jbc.M111.247650](https://doi.org/10.1074/jbc.M111.247650)
 246. S. Pritz, Y. Wolf, O. Kraetke, J. Klose, M. Bienert, M. Beyermann, Synthesis of biologically active peptide nucleic acid-peptide conjugates by sortase-mediated ligation. *J. Org. Chem.* **72**, 3909–3912 (2007). doi:[10.1021/jo062331i](https://doi.org/10.1021/jo062331i)
 247. X. Huang, A. Aulabaugh, W. Ding, B. Kapoor, L. Alksne, K. Tabei, G. Ellestad, Kinetic mechanism of *Staphylococcus aureus* sortase SrtA. *Biochemistry* **42**, 11307–11315 (2003). doi:[10.1021/bi034391g](https://doi.org/10.1021/bi034391g)
 248. T. Tanaka, T. Yamamoto, S. Tsukiji, T. Nagamune, Site-specific protein modification on living cells catalyzed by sortase. *ChemBioChem* **9**, 802–807 (2008). doi:[10.1002/cbic.200700614](https://doi.org/10.1002/cbic.200700614)
 249. T. Yamamoto, T. Nagamune, Expansion of the sortase-mediated labeling method for site-specific N-terminal labeling of cell surface proteins on living cells. *Chem. Commun.* 1022–1024 (2009). doi:[10.1039/b818792d](https://doi.org/10.1039/b818792d)
 250. S. Tsukiji, T. Nagamune, Sortase-mediated ligation: a gift from gram-positive bacteria to protein engineering. *ChemBioChem* **10**, 787–798 (2009). doi:[10.1002/cbic.200800724](https://doi.org/10.1002/cbic.200800724)
 251. N. Hirota, D. Tasuda, T. Hashidate, T. Yamamoto, S. Yamaguchi, T. Nagamune, T. Nagase, T. Shimizu, M. Nakamura, Amino acid residues critical for endoplasmic reticulum export and trafficking of platelet-activating factor receptor. *J. Biol. Chem.* **285**, 5931–5940 (2010). doi:[10.1074/jbc.M109.066282](https://doi.org/10.1074/jbc.M109.066282)
 252. W. Lan, S. Yamaguchi, T. Yamamoto, S. Yamahira, M. Tan, N. Murakami, J. Zhang, M. Nakamura, T. Nagamune, Visualization of the pH-dependent dynamic distribution of G2A in living cells. *FASEB J.* **28**, 3965–3974 (2014). doi:[10.1096/fj.14-252999](https://doi.org/10.1096/fj.14-252999)
 253. T. Lee, J. Min, H. Hirakawa, T. Nagamune, J.W. Choi, Fusion protein bilayer fabrication composed of recombinant azurin/cytochrome P450 by the sortase-mediated ligation method. *Colloids Surf. B Biointerfaces* **120**, 215–221 (2014). doi:[10.1016/j.colsurfb.2014.03.034](https://doi.org/10.1016/j.colsurfb.2014.03.034)
 254. Y. Yamamura, H. Hirakawa, S. Yamaguchi, T. Nagamune, Enhancement of sortase A-mediated protein ligation by inducing a β -hairpin structure around the ligation site. *Chem. Commun.* **47**, 4742–4744 (2011). doi:[10.1039/c0cc05334a](https://doi.org/10.1039/c0cc05334a)
 255. H. Hirakawa, S. Ishikawa, T. Nagamune, Ca^{2+} -independent sortase-A exhibits high selective protein ligation activity in the cytoplasm of *Escherichia coli*. *Biotechnol. J.* **10**, 1487–1492 (2015). doi:[10.1002/biot.201500012](https://doi.org/10.1002/biot.201500012)
 256. C. Zhang, A.M. Spokoyny, Y. Zou, M.D. Simon, B.L. Pentelute, Enzymatic “click” ligation: selective cysteine modification in polypeptides enabled by promiscuous glutathione S-transferase. *Angew. Chem. Int. Ed.* **52**, 14001–14005 (2013). doi:[10.1002/anie.201306430](https://doi.org/10.1002/anie.201306430)
 257. J.O. Fierer, G. Veggiani, M. Howarth, SpyLigase peptide–peptide ligation polymerizes antibodies to enhance magnetic cancer cell capture. *Proc. Natl. Acad. Sci. USA* **111**, E1176–E1181 (2014). doi:[10.1073/pnas.1315776111](https://doi.org/10.1073/pnas.1315776111)
 258. A. Juillerat, T. Gronemeyer, A. Keppler, S. Gendreizig, H. Pick, H. Vogel, K. Johnsson, Directed evolution of O^6 -alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules in vivo. *Chem. Biol.* **10**, 313–317 (2003). doi:[10.1016/S1074-5521\(03\)00068-1](https://doi.org/10.1016/S1074-5521(03)00068-1)
 259. A. Gautier, A. Juillerat, C. Heinis, I.R. Corrêa Jr., M. Kindermann, F. Beauflis, K. Johnsson, An engineered protein tag for multiprotein labeling in living cells. *Chem. Biol.* **15**, 128–136 (2008). doi:[10.1016/j.chembiol.2008.01.007](https://doi.org/10.1016/j.chembiol.2008.01.007)
 260. G.V. Los, L.P. Encell, M.C. McDougall, D.D. Hartzell, N. Karassina, C. Zimprich, M.G. Wood, R. Learish, R.F. Ohana, M. Urh, D. Simpson, J. Mendez, K. Zimmerman, P. Otto, G. Vidugiris, J. Zhu, A. Darzins, D.H. Klauert, R.F. Bulleit, K.V. Wood, HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem. Biol.* **3**, 373–382 (2008). doi:[10.1021/cb800025k](https://doi.org/10.1021/cb800025k)
 261. Z. Chen, C. Jing, S.S. Gallagher, M.P. Sheetz, V.W. Cornish, Second-generation covalent TMP-tag for live cell imaging. *J. Am. Chem. Soc.* **134**, 13692–13699 (2012). doi:[10.1021/ja303374p](https://doi.org/10.1021/ja303374p)
 262. K. Katayama, S. Tsukiji, T. Furuta, T. Nagamune, A bromocoumarin-based linker for synthesis of photocleavable peptidocjugates with high photosensitivity. *Chem. Commun.* 5399–5401 (2008). doi:[10.1039/b812058g](https://doi.org/10.1039/b812058g)
 263. S. Yamaguchi, Y. Chen, S. Nakajima, T. Furuta, T. Nagamune, Light-activated gene expression from site-specific caged DNA with a biotinylated photolabile protection group. *Chem. Commun.* **46**, 2244–2246 (2010). doi:[10.1039/b922502a](https://doi.org/10.1039/b922502a)
 264. Y. Ohno, H. Ueda, N. Shibata, T. Nagamune, Enhanced fluorescence resonance energy transfer immunoassay with improved sensitivity based on the Fab'-based immunoconjugates. *Anal. Biochem.* **360**, 266–272 (2007). doi:[10.1016/j.ab.2006.10.025](https://doi.org/10.1016/j.ab.2006.10.025)
 265. S. Takamori, S. Yamaguchi, N. Ohashia, T. Nagamune, Sterically bulky caging for light-inducible protein activation. *Chem. Commun.* **49**, 3013–3015 (2013). doi:[10.1039/c3cc38026b](https://doi.org/10.1039/c3cc38026b)
 266. Y. Oshiba, T. Tamaki, H. Ohashi, H. Hirakawa, S. Yamaguchi, T. Nagamune, T. Yamaguchi, Effect of length of molecular recognition moiety on enzymatic activity switching. *J. Biosci. Bioeng.* **116**, 433–437 (2013). doi:[10.1016/j.jbiosc.2013.04.003](https://doi.org/10.1016/j.jbiosc.2013.04.003)
 267. T. Ando, S. Tsukiji, T. Tanaka, T. Nagamune, Construction of a small-molecule-integrated semisynthetic split intein for in vivo protein ligation. *Chem. Commun.* 4995–4997 (2007). doi:[10.1039/b712843f](https://doi.org/10.1039/b712843f)
 268. H.A. Chung, K. Tajima, K. Kato, N. Matsumoto, K. Yamamoto, T. Nagamune, Modulating the actions of NK cell-mediated cytotoxicity using lipid-PEG(n) and inhibitory receptor-specific antagonistic peptide conjugates. *Biotechnol. Prog.* **21**, 1226–1230 (2005). doi:[10.1021/bp049646b](https://doi.org/10.1021/bp049646b)
 269. U. Tomita, S. Yamaguchi, Y. Sugimoto, S. Takamori, T. Nagamune, Poly(ethylene glycol)-lipid-conjugated antibodies enhance dendritic cell phagocytosis of apoptotic cancer cells. *Pharmaceuticals* **5**, 405–416 (2012). doi:[10.3390/ph5050405](https://doi.org/10.3390/ph5050405)
 270. U. Tomita, S. Yamaguchi, Y. Maeda, K. Chujo, K. Minamihata, T. Nagamune, Protein cell-surface display through in situ enzymatic modification of proteins with a poly(ethylene glycol)-lipid. *Biotechnol. Bioeng.* **110**, 2785–2789 (2013). doi:[10.1002/bit.24933](https://doi.org/10.1002/bit.24933)
 271. P. Farkaš, B. Systrický, Chemical conjugation of biomacromolecules: a mini-review. *Chem. Pap.* **64**, 683–695 (2010). doi:[10.2478/s11696-010-0057-z](https://doi.org/10.2478/s11696-010-0057-z)
 272. J. Kalia, R.T. Raines, Advances in bioconjugation. *Curr. Org. Chem.* **14**, 138–147 (2010)
 273. N. Stephanopoulos, M.B. Francis, Choosing an effective protein bioconjugation strategy. *Nat. Chem. Biol.* **7**, 876–884 (2011). doi:[10.1038/nchembio.720](https://doi.org/10.1038/nchembio.720)
 274. S.S. van Berkel, M.B. van Eldijk, J.C.M. van Hest, Staudinger ligation as a method for bioconjugation. *Angew. Chem. Int. Ed.* **50**, 8806–8827 (2011). doi:[10.1002/anie.201008102](https://doi.org/10.1002/anie.201008102)
 275. E. Hemaprabha, Chemical crosslinking of proteins: a review. *J. Pharm. Sci. Innov.* **1**, 22–26 (2012)
 276. C.S. McKay, M.G. Finn, Click chemistry in complex mixtures: bioorthogonal bioconjugation. *Chem. Biol.* **21**, 1075–1101 (2014). doi:[10.1016/j.chembiol.2014.09.002](https://doi.org/10.1016/j.chembiol.2014.09.002)

277. D.C. Goldstein, P. Thordarson, J.R. Peterson, The bioconjugation of redox proteins to novel electrode materials. *Aust. J. Chem.* **62**, 1320–1327 (2009). doi:[10.1071/CH09240](https://doi.org/10.1071/CH09240)
278. E. Jabbari, Bioconjugation of hydrogels for tissue engineering. *Curr. Opin. Biotechnol.* **22**, 655–660 (2011). doi:[10.1016/j.copbio.2011.01.003](https://doi.org/10.1016/j.copbio.2011.01.003)
279. J.B. Blanco-Canosa, M. Wu, K. Susumu, E. Petryayeva, T.L. Jennings, P.E. Dawson, W.R. Algar, I.L. Medintz, Recent progress in the bioconjugation of quantum dot. *Coord. Chem. Rev.* **263–264**, 101–137 (2014). doi:[10.1016/j.ccr.2013.08.030](https://doi.org/10.1016/j.ccr.2013.08.030)
280. C.M. Niemeyer, Self-assembled nanostructures based on DNA: towards the development of nanobiotechnology. *Curr. Opin. Chem. Biol.* **4**, 609–618 (2000). doi:[10.1016/S1367-5931\(00\)00140-X](https://doi.org/10.1016/S1367-5931(00)00140-X)
281. T.A. Taton, R.C. Mucic, C.A. Mirkin, R.L. Letsinger, The DNA-mediated formation of supramolecular mono- and multilayered nanoparticle structures. *J. Am. Chem. Soc.* **122**, 6305–6306 (2000). doi:[10.1021/ja0007962](https://doi.org/10.1021/ja0007962)
282. P.E. Nielsen, Peptide nucleic acid: a versatile tool in genetic diagnostics and molecular biology. *Curr. Opin. Biotechnol.* **12**, 16–20 (2001). doi:[10.1016/S0958-1669\(00\)00170-1](https://doi.org/10.1016/S0958-1669(00)00170-1)
283. J. Kurreck, Antisense technologies: improvement through novel chemical modifications. *Eur. J. Biochem.* **270**, 1628–1644 (2003). doi:[10.1046/j.1432-1033.2003.03555.x](https://doi.org/10.1046/j.1432-1033.2003.03555.x)
284. R.P. Singh, B.K. Oh, J.W. Choi, Application of peptide nucleic acid towards development of nanobiosensor arrays. *Bioelectrochemistry* **79**, 153–161 (2010). doi:[10.1016/j.bioelechem.2010.02.004](https://doi.org/10.1016/j.bioelechem.2010.02.004)
285. N. Shi, R.J. Boado, W.M. Pardridge, Antisense imaging of gene expression in the brain in vivo. *Proc. Natl. Acad. Sci. USA* **97**, 14709–14714 (2000). doi:[10.1073/pnas.250332397](https://doi.org/10.1073/pnas.250332397)
286. P. Hongmanee, H. Stender, O.F. Rasmussen, Evaluation of a fluorescence in situ hybridization assay for differentiation between tuberculous and nontuberculous *Mycobacterium* species in smears of Lowenstein–Jensen and mycobacteria growth indicator tube cultures using peptide nucleic acid probes. *J. Clin. Microbiol.* **39**, 1032–1035 (2001). doi:[10.1128/JCM.39.3.1032-1035.2001](https://doi.org/10.1128/JCM.39.3.1032-1035.2001)
287. H. Aoki, Y. Umezawa, Trace analysis of an oligonucleotide with a specific sequence using PNA-based ion-channel sensors. *Analyst* **128**, 681–685 (2003). doi:[10.1039/B300465A](https://doi.org/10.1039/B300465A)
288. D.B. Wetlaufer, Nucleation, rapid folding, and globular intrachain regions in proteins. *Proc. Natl. Acad. Sci. USA* **70**, 697–701 (1973)
289. M. Go, Correlation of DNA exonic regions with protein structural units in haemoglobin. *Nature* **291**, 90–92 (1981). doi:[10.1038/291090a0](https://doi.org/10.1038/291090a0)
290. K. Yu, C. Liu, B.G. Kim, D.Y. Lee, Synthetic fusion protein design and applications. *Biotechnol. Adv.* **33**, 155–164 (2015). doi:[10.1016/j.biotechadv.2014.11.005](https://doi.org/10.1016/j.biotechadv.2014.11.005)
291. K. Terpe, Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* **60**, 523–533 (2003). doi:[10.1007/s00253-002-1158-6](https://doi.org/10.1007/s00253-002-1158-6)
292. S. Costa, A. Almedia, A. Castro, L. Domingues, Fusion tags for protein solubility, purification and immunogenicity in *Escherichia coli*: the novel Fh8 system. *Front. Microbiol.* **5**, 63 (2014). doi:[10.3389/fmicb.2014.00063](https://doi.org/10.3389/fmicb.2014.00063)
293. A. Müller-Taubenberger, K.I. Anderson, Recent advances using green and red fluorescent protein variants. *Appl. Microbiol. Biotechnol.* **77**, 1–12 (2007). doi:[10.1007/s00253-007-1131-5](https://doi.org/10.1007/s00253-007-1131-5)
294. H. Yang, L. Liu, F. Xu, The promises and challenges of fusion constructs in protein biochemistry and enzymology. *Appl. Microbiol. Biotechnol.* **100**, 8273–8281 (2016). doi:[10.1007/s00253-016-7795-y](https://doi.org/10.1007/s00253-016-7795-y)
295. G. Gilardi, Y.T. Mehareenna, G.E. Tsotsou, S.J. Sadeghi, M. Fairhead, S. Gianini, Molecular Lego: design of molecular assemblies of P450 enzymes for nanobiotechnology. *Biosens. Bioelectron.* **17**, 133–145 (2002). doi:[10.1016/S0956-5663\(01\)00286-X](https://doi.org/10.1016/S0956-5663(01)00286-X)
296. S.J. Sadeghi, G. Gilardi, Chimeric P450 enzymes: activity of artificial redox fusions driven by different reductases for biotechnological applications. *Biotechnol. Appl. Biochem.* **60**, 102–110 (2013). doi:[10.1002/bab.1086](https://doi.org/10.1002/bab.1086)
297. K.Y. Choi, E.O. Jung, D.H. Jung, B.P. Pandey, N. Lee, H. Yun, H. Park, B.G. Kim, Novel iron–sulfur containing NADPH-reductase from *Nocardia farcinica* IFM10152 and fusion construction with CYP51 lanosterol demethylase. *Biotechnol. Bioeng.* **109**, 630–636 (2012). doi:[10.1002/bit.24359](https://doi.org/10.1002/bit.24359)
298. B. Lee, S. Takeda, K. Nakajima, J. Noh, J.W. Choi, M. Hara, T. Nagamune, Rectified photocurrent in a protein based molecular photo-diode consisting of a cytochrome b_{562} -green fluorescent protein chimera. *Biosens. Bioelectron.* **19**, 1169–1174 (2004). doi:[10.1016/j.bios.2003.11.016](https://doi.org/10.1016/j.bios.2003.11.016)
299. J.W. Choi, Y.S. Nam, B.H. Lee, D.J. Ahn, T. Nagamune, Charge trap in self-assembled monolayer of cytochrome b_{562} -green fluorescent protein chimera. *Curr. Appl. Phys.* **6**, 760–765 (2006). doi:[10.1016/j.cap.2005.04.035](https://doi.org/10.1016/j.cap.2005.04.035)
300. Y. Lu, M.B. Lerner, Z.J. Qi, J.J. Mitala Jr., J.H. Lim, B.M. Discher, A.T.C. Johnson, Graphene-protein bioelectronic devices with wavelength-dependent photoresponse. *Appl. Phys. Lett.* **100**, 033110 (2012). doi:[10.1063/1.3678024](https://doi.org/10.1063/1.3678024)
301. S.R. Schmidt, Fusion-proteins as biopharmaceuticals: applications and challenges. *Curr. Opin. Drug Discov. Dev.* **12**, 284–295 (2009). doi:[10.13140/2.1.4445.5366](https://doi.org/10.13140/2.1.4445.5366)
302. N. Ilk, E.M. Egelseer, U.B. Sleytr, S-layer fusion proteins—construction principles and applications. *Curr. Opin. Biotechnol.* **22**, 824–831 (2011). doi:[10.1016/j.copbio.2011.05.510](https://doi.org/10.1016/j.copbio.2011.05.510)
303. C. Suzuki, H. Ueda, E. Suzuki, T. Nagamune, Construction and characterization of anti-hapten scFv-alkaline phosphatase chimeric protein. *J. Biochem.* **122**, 322–329 (1997). (PMID: [9378709](https://pubmed.ncbi.nlm.nih.gov/9378709/))
304. C. Suzuki, H. Ueda, K. Tsumoto, W.C. Mahoney, I. Kumagai, T. Nagamune, Open sandwich ELISA with VH-/VL-alkaline phosphatase fusion proteins. *J. Immunol. Methods* **224**, 171–184 (1999). doi:[10.1016/S0022-1759\(99\)00020-4](https://doi.org/10.1016/S0022-1759(99)00020-4)
305. T. Yokozeki, H. Ueda, R. Arai, W.C. Mahoney, T. Nagamune, A homogeneous noncompetitive immunoassay for the detection of small haptens. *Anal. Chem.* **74**, 2500–2504 (2002). doi:[10.1021/ac015743x](https://doi.org/10.1021/ac015743x)
306. Y. Ohiro, R. Arai, H. Ueda, T. Nagamune, A homogeneous and noncompetitive immunoassay based on the enhanced fluorescence resonance energy transfer by leucine zipper interaction. *Anal. Chem.* **74**, 5786–5792 (2002). doi:[10.1021/ac0203387](https://doi.org/10.1021/ac0203387)
307. N. Komiya, H. Ueda, Y. Ohiro, T. Nagamune, Homogeneous sandwich immunoassay based on the enzymatic complementation induced by single-chain Fv fragments. *Anal. Biochem.* **327**, 241–246 (2004). doi:[10.1016/j.ab.2004.01.035](https://doi.org/10.1016/j.ab.2004.01.035)
308. R. Abe, H. Ohashi, I. Iijima, M. Ihara, H. Takagi, T. Hoshida, H. Ueda, “Quenchbodies”: quench-based antibody probes that show antigen-dependent fluorescence. *J. Am. Chem. Soc.* **133**, 17386–17394 (2011). doi:[10.1021/ja205925j](https://doi.org/10.1021/ja205925j)
309. K.N. Islam, M. Ihara, J. Dong, N. Kasagi, T. Mori, H. Ueda, Direct construction of an open-sandwich enzyme immunoassay for one-step noncompetitive detection of thyroid hormone T4. *Anal. Chem.* **83**, 1008–1014 (2011). doi:[10.1021/ac102801r](https://doi.org/10.1021/ac102801r)
310. H.J. Jeong, Y. Ohmuro-Matsuyama, H. Ohashi, F. Ohsawa, Y. Tatsu, M. Inagaki, H. Ueda, Detection of vimentin serine phosphorylation by multicolor Quenchbodies. *Biosens. Bioelectron.* **40**, 17–23 (2013). doi:[10.1016/j.bios.2012.06.030](https://doi.org/10.1016/j.bios.2012.06.030)
311. T. Sogo, M. Kawahara, H. Ueda, M. Otsu, M. Onodera, H. Nakauchi, T. Nagamune, T cell growth control using hapten-specific antibody/interleukin-2 receptor chimera. *Cytokine* **46**, 127–136 (2009). doi:[10.1016/j.cyto.2008.12.020](https://doi.org/10.1016/j.cyto.2008.12.020)
312. M. Kawahara, H. Ueda, T. Nagamune, Engineering cytokine receptors to control cellular functions. *Biochem. Eng. J.* **48**, 283–294 (2010). doi:[10.1016/j.bej.2009.09.010](https://doi.org/10.1016/j.bej.2009.09.010)
313. M. Kawahara, J. Chen, T. Sogo, J. Teng, M. Otsu, M. Onodera, H. Nakauchi, H. Ueda, T. Nagamune, Growth promotion of genetically modified hematopoietic progenitors using an antibody/c-Mpl chimera. *Cytokine* **55**, 402–408 (2011). doi:[10.1016/j.cyto.2011.05.024](https://doi.org/10.1016/j.cyto.2011.05.024)
314. M. Kawahara, T. Nagamune, Engineering of mammalian cell membrane proteins. *Curr. Opin. Chem. Eng.* **1**, 411–417 (2012). doi:[10.1016/j.coche.2012.05.002](https://doi.org/10.1016/j.coche.2012.05.002)
315. K. Saka, M. Kawahara, T. Nagamune, Reconstitution of a cytokine receptor scaffold utilizing multiple different tyrosine motifs. *Biotechnol. Bioeng.* **110**, 3197–3204 (2013). doi:[10.1002/bit.24973](https://doi.org/10.1002/bit.24973)
316. Y. Tone, M. Kawahara, D. Kawaguchi, H. Ueda, T. Nagamune, Death signalbody: inducing conditional cell death in response to a specific antigen. *Hum. Gene Ther. Methods* **24**, 141–150 (2013). doi:[10.1089/hgtb.2012.147](https://doi.org/10.1089/hgtb.2012.147)
317. Y. Tone, M. Kawahara, J. Hayashi, T. Nagamune, Cell fate conversion by conditionally switching the signal-transducing domain of signalbodies. *Biotechnol. Bioeng.* **110**, 3219–3226 (2013). doi:[10.1002/bit.24985](https://doi.org/10.1002/bit.24985)
318. M. Kawahara, A. Hitomi, T. Nagamune, S-Fms signalbody enhances myeloid cell growth and migration. *Biotechnol. J.* **9**, 954–961 (2014). doi:[10.1002/biot.201300346](https://doi.org/10.1002/biot.201300346)

319. H. Nakabayashi, S. Aoyama, M. Kawahara, T. Nagamune, Differentiation signalobody: demonstration of antigen-dependent osteoclast differentiation from a progenitor cell line. *J. Biosci. Bioeng.* **122**, 357–363 (2016). doi:[10.1016/j.jbiosc.2016.02.010](https://doi.org/10.1016/j.jbiosc.2016.02.010)
320. T. Miura, T. Nagamune, M. Kawahara, Ligand-induced dimeric antibody for selecting antibodies against a membrane protein based on mammalian cell proliferation. *Biotechnol. Bioeng.* **113**, 1113–1123 (2016). doi:[10.1002/bit.25858](https://doi.org/10.1002/bit.25858)
321. S. Lee, Y. Kaku, S. Inoue, T. Nagamune, M. Kawahara, Growth signalobody selects functional intrabodies in the mammalian cytoplasm. *Biotechnol. J.* **11**, 565–573 (2016). doi:[10.1002/biot.201500364](https://doi.org/10.1002/biot.201500364)
322. J. Winkler, Nanomedicines based on recombinant fusion proteins for targeting therapeutic siRNA oligonucleotides. *Ther. Deliv.* **2**, 891–905 (2011)
323. W.M. Pardridge, Blood–brain barrier drug delivery of IgG fusion proteins with a transferrin receptor monoclonal antibody. *Expert Opin. Drug Deliv.* **12**, 207–222 (2015). doi:[10.1517/17425247.2014.952627](https://doi.org/10.1517/17425247.2014.952627)
324. X. Chen, J.L. Zaro, W.C. Shen, Fusion protein linkers: property, design and functionality. *Adv. Drug Deliv. Rev.* **65**, 1357–1369 (2013). doi:[10.1016/j.addr.2012.09.039](https://doi.org/10.1016/j.addr.2012.09.039)
325. P. Argos, An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. *J. Mol. Biol.* **211**, 943–958 (1990). doi:[10.1016/0022-2836\(90\)90085-Z](https://doi.org/10.1016/0022-2836(90)90085-Z)
326. S. Ohnishi, H. Kamikubo, M. Onitsuka, M. Kataoka, D. Shortle, Conformational preference of polyglycine in solution to elongated structure. *J. Am. Chem. Soc.* **128**, 16338–16344 (2006). doi:[10.1021/ja066008b](https://doi.org/10.1021/ja066008b)
327. R. Bird, K. Hardman, J. Jacobson, S. Johnson, B. Kaufman, S. Lee, T. Lee, S. Pope, G. Riordan, M. Whitlow, Single-chain antigen-binding proteins. *Science* **242**, 423–426 (1988). doi:[10.1126/science.3140379](https://doi.org/10.1126/science.3140379)
328. S. Marqusee, R.L. Baldwin, Helix stabilization by Glu[−]—Lys⁺ salt bridges in short peptides of de novo design. *Proc. Natl. Acad. Sci. USA* **84**, 8898–8902 (1987). doi:[10.1073/pnas.84.24.8898](https://doi.org/10.1073/pnas.84.24.8898)
329. R.A. George, J. Heringa, An analysis of protein domain linkers: their classification and role in protein folding. *Protein Eng.* **15**, 871–879 (2002). doi:[10.1093/protein/15.11.871](https://doi.org/10.1093/protein/15.11.871)
330. C.L. Young, Z.T. Britton, A.S. Robinson, Recombinant protein expression and purification: a comprehensive review of affinity tags and microbial applications. *Biotechnol. J.* **7**, 620–626 (2012). doi:[10.1002/biot.201100155](https://doi.org/10.1002/biot.201100155)
331. D.L. Crimmins, S.M. Mische, N.D. Denslow, Chemical cleavage of proteins in solution. *Curr. Protoc. Protein Sci.* Chapter 11, 11.4.1–11.4.11 (2005). doi:[10.1002/0471140864.ps1104s40](https://doi.org/10.1002/0471140864.ps1104s40)
332. Y. Li, Self-cleaving fusion tags for recombinant protein production. *Biotechnol. Lett.* **33**, 869–881 (2011). doi:[10.1007/s10529-011-0533-8](https://doi.org/10.1007/s10529-011-0533-8)
333. W.J. Bradshaw, A.H. Davies, C.J. Chambers, A.K. Roberts, C.C. Shone, K.R. Acharya, Molecular features of the sortase enzyme family. *FEBS J.* **282**, 2097–2114 (2015). doi:[10.1111/febs.13288](https://doi.org/10.1111/febs.13288)
334. X. Chen, Y. Bai, J. Zaro, W.C. Shen, Design of an in vivo cleavable disulfide linker in recombinant fusion proteins. *Biotechniques* **49**, 513–518 (2010). doi:[10.2144/000113450](https://doi.org/10.2144/000113450)
335. H.L. Zhao, C. Xue, J.L. Du, M. Ren, S. Xia, Z.M. Liu, Balancing the pharmacokinetics and pharmacodynamics of interferon-alpha2b and human serum albumin fusion protein by proteolytic or reductive cleavage increases its in vivo therapeutic efficacy. *Mol. Pharm.* **9**, 664–670 (2012). doi:[10.1021/mp200347q](https://doi.org/10.1021/mp200347q)
336. M. Gustavsson, J. Lehtiö, S. Denman, T.T. Teeri, K. Hult, M. Martinelle, Stable linker peptides for a cellulose-binding domain-lipase fusion protein expressed in *Pichia pastoris*. *Protein Eng.* **14**, 711–715 (2001). doi:[10.1093/protein/14.9.711](https://doi.org/10.1093/protein/14.9.711)
337. N. Adlakha, S. Sawant, A. Anil, A. Lali, S.S. Yazdani, Specific fusion of β -1,4-endoglucanase and β -1,4-glucosidase enhances cellulolytic activity and helps in channeling of intermediates. *Appl. Environ. Microbiol.* **78**, 7447–7454 (2012). doi:[10.1128/AEM.01386-12](https://doi.org/10.1128/AEM.01386-12)
338. G.G. Yang, X.Y. Xu, Y. Ding, Q.Q. Cui, Z. Wang, Q.Y. Zhang, S.H. Shi, Z.Y. Lv, X.Y. Wang, J.H. Zhang, R.G. Zhang, C.S. Xu, Linker length affects expression and bioactivity of the onconase fusion protein in *Pichia pastoris*. *Genet. Mol. Res.* **14**, 19360–19370 (2015). doi:[10.4238/2015.December.29.46](https://doi.org/10.4238/2015.December.29.46)
339. Y. Maeda, H. Ueda, J. Kazami, G. Kawano, E. Suzuki, T. Nagamune, Engineering of functional chimeric protein G-*Vargula* luciferase. *Anal. Biochem.* **249**, 147–152 (1997). doi:[10.1006/abio.1997.2181](https://doi.org/10.1006/abio.1997.2181)
340. H. Ueda, T. Yokozeki, R. Arai, K. Tsumoto, I. Kumagai, T. Nagamune, An optimized homogeneous noncompetitive immunoassay based on the antigen-driven enzymatic complementation. *J. Immunol. Methods* **279**, 209–218 (2003). doi:[10.1016/S0022-1759\(03\)00256-4](https://doi.org/10.1016/S0022-1759(03)00256-4)
341. K. Saka, M. Kawahara, T. Nagamune, Quantitative control of intracellular signaling activity through chimeric receptors incorporating multiple identical tyrosine motifs. *Biotechnol. Bioeng.* **111**, 948–955 (2014). doi:[10.1002/bit.25151](https://doi.org/10.1002/bit.25151)
342. W. Liu, M. Kawahara, H. Ueda, T. Nagamune, Construction of a fluorescein-responsive chimeric receptor with strict ligand dependency. *Biotechnol. Bioeng.* **101**, 975–984 (2008). doi:[10.1002/bit.21961](https://doi.org/10.1002/bit.21961)
343. T. Haga, H. Hirakawa, T. Nagamune, Fine tuning of spatial arrangement of enzymes in a PCNA-mediated multienzyme complex using a rigid poly-L-proline linker. *PLoS ONE* **8**, e75114 (2013). doi:[10.1371/journal.pone.0075114](https://doi.org/10.1371/journal.pone.0075114)
344. F. Xue, Z. Gu, J.A. Feng, LINKER: a web server to generate peptide sequences with extended conformation. *Nucleic Acids Res.* **32**, W562–W565 (2004). doi:[10.1093/nar/gkh422](https://doi.org/10.1093/nar/gkh422)
345. P. Lu, M.G. Feng, Bifunctional enhancement of a β -glucanase-xylanase fusion enzyme by optimization of peptide linkers. *Appl. Microbiol. Biotechnol.* **79**, 579–587 (2008). doi:[10.1007/s00253-008-1468-4](https://doi.org/10.1007/s00253-008-1468-4)
346. R. Arai, H. Ueda, A. Kitayama, N. Kamiya, T. Nagamune, Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Eng.* **14**, 529–532 (2001). doi:[10.1093/protein/14.8](https://doi.org/10.1093/protein/14.8)
347. R. Arai, W. Wriggers, Y. Nishikawa, T. Nagamune, T. Fujisawa, Conformations of variably linked chimeric proteins evaluated by synchrotron X-ray small-angle scattering. *Proteins* **57**, 829–838 (2004). doi:[10.1002/prot.20244](https://doi.org/10.1002/prot.20244)
348. N. Kobayashi, K. Yanase, T. Sato, S. Unzai, M.H. Hecht, R. Arai, Self-assembling nano-architectures created from a protein nano-building block using an intermolecularly folded dimeric de novo protein. *J. Am. Chem. Soc.* **137**, 11285–11293 (2015). doi:[10.1021/jacs.5b03593](https://doi.org/10.1021/jacs.5b03593)
349. C.J. Crasto, J.A. Feng, LINKER: a program to generate linker sequences for fusion proteins. *Protein Eng.* **13**, 309–312 (2000). doi:[10.1093/protein/13.5.309](https://doi.org/10.1093/protein/13.5.309)
350. C. Liu, J.X. Chin, D.Y. Lee, SynLinker: an integrated system for designing linkers and synthetic fusion proteins. *Bioinformatics* **31**, 3700–3702 (2015). doi:[10.1093/bioinformatics/btv447](https://doi.org/10.1093/bioinformatics/btv447)
351. J. Chiang, I. Li, E. Pham, K. Truong, FPMOD: a modeling tool for sampling the conformational space of fusion proteins. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **1**, 4111–4114 (2006). doi:[10.1109/EMBS.2006.259224](https://doi.org/10.1109/EMBS.2006.259224)
352. E. Pham, J. Chiang, I. Li, W. Shum, K. Truong, A computational tool for designing FRET protein biosensors by rigid-body sampling of their conformational space. *Structure* **15**, 515–523 (2007). doi:[10.1016/j.str.2007.03.009](https://doi.org/10.1016/j.str.2007.03.009)
353. A.M. Wollacott, A. Zanghellini, P. Murphy, D. Baker, Prediction of structures of multidomain proteins from structures of the individual domains. *Protein Sci.* **16**, 165–175 (2007). doi:[10.1110/ps.062270707](https://doi.org/10.1110/ps.062270707)
354. T.M. Cheng, T.L. Blundell, J. Fernandez-Recio, Structural assembly of two-domain proteins by rigid-body docking. *BMC Bioinform.* **9**, 441 (2008). doi:[10.1186/1471-2105-9-441](https://doi.org/10.1186/1471-2105-9-441)
355. Y. Zhang, I-TASSER server for protein 3D structure prediction. *BMC Bioinform.* **9**, 40 (2008). doi:[10.1186/1471-2105-9-40](https://doi.org/10.1186/1471-2105-9-40)
356. S. Lovell, I. Davis, B. Arendall III, P. Bakker, M. Word, M. Prisant, J. Richardson, D. Richardson, Structure validation by $C\alpha$ -geometry: ϕ , Ψ and $C\beta$ deviation. *Proteins Struct. Funct. Bioinform.* **50**, 437–450 (2003). doi:[10.1002/prot.10286](https://doi.org/10.1002/prot.10286)
357. S. Shamriz, H. Ofoghi, N. Moazami, Effect of linker length and residues on the structure and stability of a fusion protein with malaria vaccine application. *Comput. Biol. Med.* **76**, 24–29 (2016). doi:[10.1016/j.combiomed.2016.06.015](https://doi.org/10.1016/j.combiomed.2016.06.015)
358. J. Arunachalam, V. Kanagasabai, N. Gautham, Protein structure prediction using mutually orthogonal Latin squares and a genetic algorithm. *Biochem. Biophys. Res. Commun.* **342**, 424–433 (2006). doi:[10.1016/j.bbrc.2006.01.162](https://doi.org/10.1016/j.bbrc.2006.01.162)
359. A. Nicolai, P. Delarue, P. Senet, Decipher the mechanisms of protein conformational changes induced by nucleotide binding through free-energy landscape analysis: ATP binding to Hsp70. *PLoS Comput. Biol.* **9**, e1003379 (2013). doi:[10.1371/journal.pcbi.1003379](https://doi.org/10.1371/journal.pcbi.1003379)
360. N.P. King, Y.T. Lai, Practical approaches to designing novel protein assemblies. *Curr. Opin. Struct. Biol.* **23**, 632–638 (2013). doi:[10.1016/j.sbi.2013.06.002](https://doi.org/10.1016/j.sbi.2013.06.002)

361. R. Wang, Y. Zhang, D. Lu, J. Ge, Z. Liu, R.N. Zare, Functional protein-organic/inorganic hybrid nanomaterials. *WIREs Nanomed. Nanobiotechnol.* **5**, 320–328 (2013). doi:[10.1002/wnan.1210](https://doi.org/10.1002/wnan.1210)
362. L.H.L. Lua, N.K. Connors, F. Sainsbury, Y.P. Chuan, N. Wibowo, A.P.J. Middeberg, Bioengineering virus-like particles as vaccines. *Biotechnol. Bioeng.* **111**, 425–440 (2014). doi:[10.1002/bit.25159](https://doi.org/10.1002/bit.25159)
363. D. Romera, P. Couleaud, S.H. Mejias, A. Aires, A.L. Cortajarena, Biomolecular templating of functional hybrid nanostructures using repeat protein scaffolds. *Biochem. Soc. Trans.* **43**, 825–831 (2015). doi:[10.1042/BST20150077](https://doi.org/10.1042/BST20150077)
364. Q. Luo, C. Hou, Y. Bai, R. Wang, J. Liu, Protein assembly: versatile approaches to construct highly ordered nanostructures. *Chem. Rev.* **116**, 13571–13632 (2016). doi:[10.1021/acs.chemrev.6b00228](https://doi.org/10.1021/acs.chemrev.6b00228)
365. S. Hiyama, Y. Moritani, R. Gojo, S. Takeuchi, K. Sutoh, Biomolecular-motor-based autonomous delivery of lipid vesicles as nano- or micro-scale reactors on a chip. *Lab Chip* **10**, 2741–2748 (2010). doi:[10.1039/c004615a](https://doi.org/10.1039/c004615a)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
