LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

Interfacing "Soft" and "Hard" Matter with Exquisite Chemical Control

Y.-H. Woo, J. A. Camarero

January 30, 2006

Disclaimer

Interfacing "Soft" and "Hard" Matter with Exquisite Chemical Control

Youn-Hi Woo and Julio A. Camarero*

*Chemical Biology and Nuclear Sciences Division, L-231,

Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 94550.

Tel. 925-422 6807, fax 925-422-3160, e-mail: camarero1@llnl.gov.

Abstract. The present paper reviews the recent development of new chemical and biological technologies for the site-specific immobilization of proteins onto inorganic materials and their potential applications to the fields of micro and nanotechnology.

Keywords: Protein immobilization, Native Chemical Ligation, Protein Splicing, Protein Chips.

Introduction

Nano-biotechnology is an exciting emerging field that lies at the interface of recent advancements in nanoscale science/technology and biotechnology [1]. The biological, biomedical, and medical applications of nanotechnology are some of the most promising, most exciting, and potentially most rewarding. Sensing and therapeutics using tools from micro and nanotechnology [2], the use of nanoscale drug delivery devices for targeted therapy [3], development of devices using micro-nano fabrication and scaffolding techniques [4], and miniature drug screening and discovery [5] are only some of the rapidly emerging possibilities.

The key for the success of all these applications relies, however, on interfacing the typically "soft" biomolecules found in biological systems (i.e., mainly proteins and DNA) with the generally "hard" inorganic materials found in micro and nanotechnology, ensuring that the interfaced biomolecules retain their remarkable biological properties. Although there has been an enormous amount of progress made in interfacing/immobilizing DNA biomolecules onto different inorganic materials [6], the immobilization of proteins has been a challenging task. This is mainly due to the heterogeneous chemical nature of protein surfaces and the marginal stability of the native active structure over the denatured inactive random coil structure.

Most of the available methods for the immobilization of proteins onto inorganic materials have relied on the use of non-specific adsorption of proteins or on the reaction of naturally occurring chemical groups within proteins (mostly amines and carboxylic acids) with complementary reactive groups chemically introduced onto the corresponding inorganic material [4, 7, 8]. In both cases, the corresponding proteins are attached onto

the surface in random orientations, which may cause the loss of the protein's biological activity [9]. The use of recombinant affinity tags addresses the orientation issue. However, in most cases, the interactions of the tags are reversible and unstable over time [10-14].

Site-specific covalent immobilization, on the other hand, allows the proteins to be arranged in a definite, controlled fashion. The reaction between these two groups should be highly chemoselective, thus behaving like a molecular '*Velcro*' [15, 16]. Finally, the use of hydrophilic spacers and linkers may also help to minimize the potentially detrimental interaction between the protein and the inorganic support.

The scope of this work is to review some of the most commonly employed methods as well as the latest developments for the chemoselective immobilization of biologically active proteins onto inorganic supports.

Immobilization of proteins onto inorganic substrates

Most methods used for chemoselective immobilization of proteins are based on ligation methods originally developed for the synthesis, semi-synthesis, and selective derivatization of proteins by chemical means (see Table 1) [16]. All these methods involve the derivatization of a protein with a unique chemical group at a defined position, which will later react chemoselectively with a complementary group previously introduced into the inorganic surface.

Surface modification

Silicon, metals (mainly Au and Ag) and semiconductor (i.e., Ag_2S, CdS, CdSe, and TiO_2) based substrates are among the most common materials used to immobilize proteins in micro- and nano-biotechnology. Si-based substrates are usually modified using derivatized trialkoxysilanes such as (3–aminopropyl)-trialkoxysilane (APS) or (3-mercaptopropyl)-trialkoxysilane, which are able to introduce an amino or thiol groups, respectively. These functionalities can be further chemically modified to introduce appropriate linkers where the proteins can be covalently attached in a chemoselective fashion.

Sulfur- and selenium-containing compounds can also be used for the modification of substrates based on several transition metals (Au, Ag and Pt) [17, 18] or semiconductor materials (e.g.., Ag_2S, CdS and CdSe) [7]. The most studied system, however, is the use of alkanethiols on gold surfaces. Chemisorption of alkanethiols as well as alkyl disulfides on clean gold gives rise to similar levels of surface coverage, although thiols react faster than disulfides [17, 19]. Alkylaminothiols are the most common species used for the preparation of functionalized gold surfaces. Among them, cysteamine ($HS–CH_2–CH_2–NH_2$) is probably the most used [20-23], in part due to its availability. Mercaptoalkyl carboxylic acids ($HS–(CH_2)_n–CO_2H$)) can also be used to introduce a reactive group onto the gold substrate [7, 23].

In our group, we have developed a new efficient synthetic solid-phase scheme for the rapid generation of modified alkanethiols [24, 25]. We have used this approach for the chemical synthesis of different modified alkane thiols that were successfully used to

immobilize different biological functional proteins onto Si-based and Au surfaces [24-26].

Immobilization of thiol-containing proteins

Cysteine is the only naturally occurring amino acid containing a thiol group in its side-chain, and its relative abundance in the average protein is rather small ($< 2\%$). Thiols have a pK_a of around 8.5, and they are nucleophilic at pH 7. Under these conditions, they can be reacted with high selectivity with some chemical groups like α-haloacetyl and maleimide-containing compounds to form a stable thioether covalent. There are a great variety of commercially available reagents that can be used for introducing maleimide as well as iodoacetyl groups on amine-derivatized inorganic supports [7, 24-29]. This provides a unique window of reactivity for the chemoselective attachment of proteins to surfaces through the thiol group of the amino acid Cys. The only requirement to have control over the orientation during the protein immobilization process is the presence of a unique and reactive Cys residue on the protein. Reactive Cys residues should not be involved in any structural element and should be exposed in a solvent accessible region of the protein. In the absence of an endogenous Cys residue, it will be necessary to introduce a Cys residue through mutation of a native residue. The effect of this Cys mutation on the structure and function of the protein can be easily minimized by following a few simple rules:

(1) If the tertiary structure of the protein is known, choose a region remote from the active site of the protein, preferably in a flexible surface loop.

(2) Choose the mutation to be as conservative as possible, e.g., Ser → Cys or Ala → Cys.

(3) Avoid mutation of residues that are conserved across a gene family.

(4) Take advantage of any known mutational data on the system since the effect on structure/function of a specific residue mutation may already be known.

In our laboratory, we used a genetically modified Cow Pea Mosaic Virus (CPMV) to create assembled viral nanostructures on maleimide-containing surface templates [24]. In this work, the CPMV was genetically engineered to present Cys residues at geometrically equivalent positions in the solvent-exposed βE-βF loop of the viral capsomer [30]. The maleimide containing surface templates were created by microprinting techniques and scanning probe nanolithography on a gold-coated mica surface using a PEGylated amino thiol linker. The amino group was then reacted with *N*-(maleimido-propionyloxy)-succinimide ester (MPS) to yield the corresponding maleimide-containing surface template.

As shown in Fig. (**2**), the Cys-mutated CPMV was chemoselectively attached only on those areas containing the maleimide function. Mirkin and co-workers [31] have also used a similar scheme to template viral deposition on chemically modified surfaces.

The use of the Cys thiol group is one of the most common method used by the scientific community for achieving ordered immobilization of different proteins and viruses onto inorganic supports (reviewed in detail in [16]). It should be noted, however, that is not general method and can only be applied when there is only one reactive Cys residue, either artificially introduced or naturally occurring, on the protein to be attached.

When the protein contains multiple Cys residues, it is better to use alternative chemoselective methods.

Immobilization of proteins by using Expressed Protein Ligation

One the most efficient ways to site-specifically immobilize biologically active proteins onto solid supports is by using Expressed Protein Ligation (EPL) [32-35]. Key to this approach is the use of protein α-thioesters (see Table 1), which can be efficiently attached to surfaces containing *N*-terminal Cys residues through Native Chemical Ligation (NCL) as shown in Fig. (**3**) [36, 37]. In this reaction, two fully unprotected polypeptides, one containing a C-terminal *α*-thioester group and the other an *N*-terminal Cys residue, react chemoselectively under neutral aqueous conditions with the formation of a native peptide bond at the ligation site.

We recently immobilized several biologically active proteins onto modified glass surfaces through their *C*-termini using this approach [25]. In this work, two fluorescent proteins (EGFP and DsRed) and a SH3 domain protein C-terminal α-thioesters were readily expressed in *E. coli*, using an intein expression system [38]. The α-thioester proteins were then immobilized onto a *N*-terminal Cys-containing glass slide. The chemical modification of the glass slide was accomplished first by silanization with (3-acryloxypropyl)-trimethoxysilane and then reacting with a mixture of PEGylated thiol linkers **1** and **2**, shown in Fig. (**4**), in a molar ratio of 1:5, respectively. Linker **1** contained a protected *N*-terminal Cys residue for the selective attachment of the α-thioester proteins. Linker **2** was used as diluent to control the number of reactive sites on the surface. Linker **1** contains a longer PEG moiety than linker **2** to ensure that the reactive

Cys groups were readily available to react with the corresponding protein α-thioester in solution.

When the derivatization was complete, the corresponding protecting groups of the Cys residue from linker **1** were removed by a brief treatment with trifluoroacetic acid. The surface was rinsed, neutralized, and quickly used for micro-spotting, see Fig. (**5**). The ligation reaction was kept at 36 h in the dark at room temperature, and the protein-modified slide was then extensively washed.

As shown in Fig. (**5**), only specific attachment between the α-thioester proteins and the Cys-containing glass surface was observed. No fluorescence signal was observed where the control EGFP protein lacking an α-thioester was spotted. It is interesting to note that the immobilized DsRed protein, which only has red fluorescence as a tetramer, retained its red fluorescence thereby indicating that its quaternary architecture was unaffected by the attachment to the PEGylated glass surface.

Yao and co-workers have also used NCL and EPL, for the selective immobilization of N-terminally Cys-containing polypeptide [39] and proteins [40] onto α-thioester coated glass slides. In this case, the polypeptide/proteins are site-specifically immobilized through their N-termini, which may be convenient in cases where the C-terminal immobilization, described earlier, affects the activity of the protein.

Immobilization of azide-containing proteins

The azide function is not present in any naturally occurring protein. However, Bertozzi and co-workers,[41], have reported a novel method for incorporation of azide groups into recombinant proteins. They showed the unnatural amino acid azido-

homoalanine, a surrogate of methionine, can be loaded into the methionyl-tRNA synthetase of *E. coli* and incorporated in recombinant proteins expressed in methionine depleted bacterial cultures by using a methionine auxotroph *E. coli* strain.

Azido-containing proteins can be chemoselectively immobilized onto solid supports modified with a suitable phosphine via a modified version of the Staudinger ligation reaction [42-45]. This reaction allows the formation of an amide bond between an arylphosphine moiety and azide group (see Fig. (**6**)). This reaction is highly chemoselective and works with better yields when Z is –CH$_2$- and X is sulfur (i.e., a thioester function). The introduction of the arylphosphine derivative in a carboxylic-containing surface can be conveniently carried out (see Fig. (**6B**)) by using an activated carboxylic surface and the diphenylphosphinomethanethiol, the synthesis which was recently reported by Raines and co-workers [10] for the immobilization of small azide-containing synthetic polypeptides onto glass slides.

Azido-containing proteins can also be selectively immobilized onto alkine-coated supports through a highly specific azide-alkyne cycloadition [46, 47]. This reaction has been used in many bioconjugation reactions, including labeling of proteins in bacterial [48] and mammalian cells [47] as well as viral particles [49].

More recently, Tirrell and co-workers [50] have shown that several alkyne-containing amino acids can be incorporated into recombinant proteins in a residue-specific manner and therefore it should also be possible to immobilize alkyne-containg proteins onto azide-derivatized surfaces.

Chemoenzymatic methods for the site-specific immobilization of proteins

All the methods described so far rely on pure chemoselective reactions with little or no activation at all. That means that the efficiency of these reactions depends on the concentration of the reagents (i.e., on the concentration of the protein to be attached to the corresponding surface) to bring close enough both reactants in order to allow them to react in an efficient way.

A way to overcome this intrinsic barrier and make ligation reactions more efficient, even under high dilution, is through the introduction of complementary moieties in the protein and the surface, which can form a stable and specific intermolecular complex. Once formed, this complex brings both reactive groups in close proximity, thus increasing the local effective concentration of both reactants.

The main advantages of using this approach are:

(1) The requirement of considerable less protein since the ligation reaction works efficiently even under high dilution conditions.

(2) No need for purification since at high dilution the only protein that will react with the surface will be the one having the complementary affinity and reactive tag (see Fig. (**7**)).

Immobilization by using active site-directed capture ligands

The idea of using reactive ligands to capture proteins has been used by Meares and co-workers [51] for creating antibodies with infinite affinity. In this interesting work, the authors created an antibody against a metal-complex ligand, which contained a reactive electrophile close to the binding site. When the antibody and the ligand are apart, their complementary groups do not react (mainly due to the dilution effect). However, when the antibody specifically binds the ligand, the effective concentration of their complementary groups is greatly increased, thus leading to the irreversible formation of a covalent bond.

More recently, Mrksich and co-workers [52] have used this same principle for the selective attachment of protein onto surfaces with total control over the orientation. In their approach, they used the protein calmodulin fused with the enzyme cutinase as a capture protein. Cutinase is a 22 kDa serine esterase that is able to form a site-specific covalent adduct with chlorophosphonate ligands [53]. The chlorophosphonate group mimics the tetrahedral transition state of an ester hydrolysis. When it binds specifically to the active site of the enzyme, the hydroxyl group of the catalytic serine residue reacts covalently with the chlorophosphonate to yield a stable covalent adduct that is resistant to hydrolysis.

In this case, the authors used a gold surface to immobilize the cutinase inhibitor. The attachment is extremely selective and can be carried with the whole crude *E. coli* periplasmic lysate containing the cutinase fusion protein, thus demonstrating that there is no need to purify the protein before the immobilization step [52]. This approach has also

been used for the preparation of antibody arrays on self-assembled monolayers presenting a phosphonate capture ligand [54].

Walsh and co-workers [13] have also recently reported a very elegant scheme for the chemoenzymatic site-specific modification of proteins. In their approach, the target proteins are expressed as fusions to a peptide carrier protein (PCP) excised from a nonribosomal peptide synthetase (NRPS). NRPS PCPs are 8-10 kDa autonomously folded, compact, and stable domains. These domains contain one specific Ser residue that can be catalytically phosphorylated by the phosphopantetheinyl (Ppant) transferase SFP using CoA (Coenzyme A) as substrate. Using the Ppant transferase SFP from *B. subtilis*, the authors were able to specifically label proteins with Ppant-biotin using biotin-CoA as substrate. These biotin-labeled proteins were used to produce protein microarrays onto an avidin-coated glass slide. Johnsson and co-workers [55] have used a similar approach involving the transfer of phosphopantetheine derivatives to a peptide-acyl carrier protein fused to the protein of interest. In a similar way, this approach could be used for the site-specific immobilization of PCP-fusion proteins onto surfaces derivatized by CoA.

Johnsson's group [56] has also developed a novel approach for the site-specific labeling of recombinant proteins using a mutant O^6-alkylguanine-DNA alkyltransferase (AGT). This modified enzyme can efficiently transfer a benzyl group to itself when presented with O^6-benzylguanine derivatives (see Fig. (**8**)). The mutant enzyme is promiscuous with respect to the substituents appended to the benzyl group, enabling a range of probes (including inorganic supports) to be used for site-specific labeling.

Immobilization by using protein trans-splicing

The main limitation of the previous methods is that the linker between the protein of interest and the surface is always another protein or protein domain. In some cases, the presence of such a big linker could potentially give rise to some problems, especially in those applications where the attached protein will be involved in studying protein/protein interactions with complex protein mixtures [57, 58]. This is extremely important in biosensing applications.

In order to address this problem, our group has developed a new traceless capture ligand approach for the selective immobilization of proteins to surfaces based on the protein trans-splicing process (see Fig. (**9**)) [26]. This process is similar to protein splicing [59, 60] with the only difference being the intein self-processing domain is split in two fragments (called *N*-intein and *C*-intein, respectively) [61, 62].

In our approach, the C-intein fragment is covalently immobilized onto a glass surface through a PEGylated-peptide linker while the N-intein fragment is fused to the C-terminus of the protein to be attached to surface. When both intein fragments interact, they form an active intein domain, which ligates the protein of interest to the surface at the same time the split intein is spliced out into solution (see Fig. (**9**)).

Key to our approach is the use of the naturally split DnaE intein from *Synechocystis sp*. PCC6803 [63]. The C- and N-intein fragments of the DnaE intein are able to self-assemble spontaneously (K_d = 0.1-0.2 μM) not requiring any refolding step [26, 64]. The DnaE intein-mediated trans-splicing reaction is also very efficient under physiological-like conditions ($\tau_{1/2} \approx$ 4 h and trans-splicing yields ranging from 85% to almost quantitative) [26].

Using this approach, we have successfully immobilized Maltose Binding Protein (MBP) and Enhanced Green Fluorescent Protein (EGFP) to chemically modified SiO_2-based substrates. Both proteins were modified at the DNA level to append the DnaE N-intein fragment (I_N, residues 1-123 of the DnaE intein) at their C-termini. The two fusion proteins (MBP-I_N and EGFP-I_N) were readily expressed in *E. coli* and purified by affinity chromatography.

In order to enable the site-specific attachment of the I_N fusion proteins onto a glass support, an amine-functionalized glass slide was first treated with maleimidopropionic acid N-hydroxysuccinimide ester (MPS) and then reacted with a mixture of PEGylated thiol linkers **3** and **4**, in a molar ratio of 3:97, respectively (Fig. (**9B**)). Linker **3** contains the DnaE C-intein fragment (I_C, residues 124-159 of the DnaE intein) followed by the corresponding C-extein sequence CFNK. The Cys residue in the C-extein sequence was protected with a S-^{t}Bu protecting group. This was required to ensure that the I_C-containing linker was selectively immobilized through its PEGylated terminal thiol onto the maleimido-coated glass surface.

Linker **4** was used as diluent to control the density of reactive sites on the modified glass surface. As shown in Fig. (**9B**), different concentrations of pure EGFP-I_N and MBP-I_N fusion proteins were spotted onto I_C-coated glass slides (Fig. (**10**)). As a control, a solution of EGFP with no I_N fragment was spotted in the same slide (Fig. (**10A**)). In addition, a solution of MBP-I_N was spotted onto a glass slide derivatized with only the non-functional linker **4** (Fig. (**10B**)). The trans-splicing reaction was incubated for 16 h in a humidified chamber at 37°C. As shown in Fig. (**10A**), only specific immobilization of the proteins containing the I_N polypeptide to the I_C-containing glass

surfaces was observed. No fluorescent signal was detected from the control protein EGFP, lacking the I_N polypeptide, after washing. In addition, the MBP immobilization was minimal when the trans-splicing active MBP-I_N fusion protein was spotted onto a control glass slide coated with the non-functional linker **3** (Fig. (**10B**)). The trans-splicing mediated attachment of EGFP and MBP in both cases was efficient in a range of concentrations. It is interesting to note that the attached EGFP retained its characteristic green fluorescence, indicating that its tertiary structure was unaffected by the attachment to the PEGylated glass surface.

We have also explored the ability to selectively immobilize I_N-containing fusion proteins from complex mixtures through protein trans-splicing. We have shown that cellular fractions of *E. coli* overexpressing MBP-I_N can be selectively immobilized onto I_C-coated glass slides with minimal background, as described above (Fig. (**11A**). The use of MBP-I_N expressed using an *E. coli*-based *in vitro* transcription/translation (IVT) systems also gave similar results (Fig. (**11B**)).

Protein trans-splicing provides a unique method for the site-specific immobilization of proteins onto solid supports. The immobilization of the protein is highly specific and efficient, thus eliminating the need for the purification and/or reconcentration of the proteins prior to the immobilization step. More importantly, once the protein is immobilized to the surface, both intein fragments are spliced out into solution, providing a completely traceless method of attachment. Finally, this methodology can be easily interfaced with cell-free protein expression systems, allowing for rapid access to the high-throughput production of protein chips and other types of biosensors.

Summary

In this review, we have summarized several of the most commonly used methods as well as the most recent developments for interfacing "soft" and "hard" matter. The ability to interface these two worlds plays an extremely important role in the development of new technologies, such as the creation of protein chips [9, 25, 26, 65-67]. Like DNA chips, protein chips will allow the analysis and screening of thousands of proteins simultaneously, which will promote the discovery of new drug targets. Protein chips can also be used as a diagnosis tool for profiling protein expression in order to find potentially relevant biomarkers [67]. Another potential application for ordered protein films on surfaces will be the production of optimized biosensors [68, 69]. An ordered protein film has a higher activity density than a random protein film, where a significant percentage of the protein molecules bound to the surface are potentially inactive conformations. This will allow biosensor miniaturization without losing sensitivity.

The combination of recent nano-lithography techniques combined with the ability to bind proteins in an extremely ordered fashion will allow the creation of protein nanopatterns, which could be used as templates for the crystallization of biomolecules [24, 31, 70]. This emerging technology is expected to have a significant impact in the structural biology field where the main bottleneck is obtaining suitable crystals for physical studies.

In summary, proteins represent an extremely fertile territory in micro- and particularly nano-biotechnology because they possess a variety of properties ideal for

engineering. They have sophisticated architectures at nanoscale dimensions, complex chemistry composition, and a wide variety of enzymatic activities.

Proteins are capable of carrying out complex tasks in living cells. The flagellar motors of bacteria, the linear motors of muscle and cytoskeleton, voltage-gated ion channels, DNA replication complexes, or the photosynthetic reaction centers are just a few examples. The development of new chemical/biochemical technologies for the engineering of proteins with absolute control will shape the future of nanotechnology, allowing us to harness the full power of proteins to create new components for materials and devices.

Acknowledgments

References

[1]   Whitesides, G. M. *Nat. Biotechnol.*, **2003**, *21*, 1161-1165.

[2]   Haruyama, T. *Adv. Drug. Deliv. Rev.*, **2003**, *55*, 393-401.

[3]   Kayser, O., Lemke, A. and Hernandez-Trejo, N. *Curr. Pharm. Biotechnol.*, **2005**, *6*, 3-5.

[4]   Lee, K. B., Park, S. J., Mirkin, C. A., Smith, J. C. and Mrksich, M. *Science*, **2002**, *295*, 1702-1705.

[5]   Jain, K. K. *Drug Discov. Today*, **2005**, *10*, 1435-1442.

[6]   Seliger, H., Hinz, M. and Happ, E. *Curr. Pharm. Biotechnol.*, **2003**, *4*, 379-395.

[7]   Chan, W. C. and Nie, S. *Science*, **1998**, *281*, 2016-2018.

[8]   Liu, G. Y. and Amro, N. A. *Proc. Natl. Acad. Sci. U S A*, **2002**, *99*, 5165-5170.

[9]   Phizicky, E., Bastiaens, P. I., Zhu, H., Snyder, M. and Fields, S. *Nature*, **2003**, *422*, 208-215.

[10]   Soellner, M. B., Dickson, K. A., Nilsson, B. L. and Raines, R. T. *J. Am. Chem. Soc.*, **2003**, *125*, 11790-11791.

[11]   Kohn, M., Wacker, R., Peters, C., Schroder, H., Soulere, L., Breinbauer, R., Niemeyer, C. M. and Waldmann, H. *Angew. Chem. Int. Ed. Engl.*, **2003**, *42*, 5830-5834.

[12]   Lue, R. Y., Chen, G. Y., Hu, Y., Zhu, Q. and Yao, S. Q. *J. Am. Chem. Soc.*, **2004**, *126*, 1055-1062.

[13]   Yin, J., Liu, F., Li, X. and Walsh, C. T. *J. Am. Chem. Soc.*, **2004**, *126*, 7754-7755.

[14]   Ramachandran, N., Hainsworth, E., Bhullar, B., Eisenstein, S., Rosen, B., Lau, A. Y., Walter, J. C. and LaBaer, J. *Science*, **2004**, *305*, 86-90.

[15]   Schnolzer, M. and Kent, S. B. *Science*, **1992**, *256*, 221-225.

[16]   Camarero, J. A. *Biophys. Rev. Lett.*, **2006**, *1*, 1-28.

[17]   Sullivan, T. P. and Huck, W. T. S. *Eur. J. Org. Chem.*, **2003**, *2003*, 17-29.

[18]   Ulman, A. *Chem. Rev.*, **1996**, *96*, 1533-1554.

[19]   Biebuyck, H. A., Bian, C. D. and Whitesides, G. M. *Langmuir*, **1994**, *10*, 1825-1831.

[20]   Wagner, P., Hegner, M., Guntherodt, H.-J. and Semenza, G. *Langmuir*, **1995**, *11*, 3867-3875.

[21] Yam, C.-M., Pradier, C. M., Salmain, M., Marcus, P. and Jaouen, G. *J. Colloid Interface Sci.*, **2001**, *235*, 183-189.

[22] Uvdal, K. and Vikinge, T. P. *Langmuir*, **2001**, *17*, 2008-2012.

[23] Liu, Y. C., Wang, C. M. and Hsiung, K. P. *Anal. Biochem.*, **2001**, *299*, 130-135.

[24] Cheung, C. L., Camarero, J. A., Woods, B. W., Lin, T. W., Johnson, J. E. and De Yoreo, J. J. *J. Am. Chem. Soc.*, **2003**, *125*, 6848-6849.

[25] Camarero, J. A., Kwon, Y. and Coleman, M. A. *J. Am. Chem. Soc.*, **2004**, *126*, 14730-14731.

[26] Kwon, Y., Coleman, M. A. and Camarero, J. A. *Angew. Chem. Int. Ed.*, **2006**, in press.

[27] MacBeath, G., Koehler, A. N. and Schreiber, S. L. *J. Am. Chem. Soc.*, **1999**, *121*, 7967-7968.

[28] Xiao, S. J., Textor, M., Spencer, N. D. and Sigrist, H. *Langmuir*, **1998**, *14*, 5507-5516.

[29] Cheng, Y. Y., Chang, H. C., Hoops, G. and Su, M. C. *J. Am. Chem. Soc.*, **2004**, *126*, 10828-10829.

[30] Wang, Q., Tianwei, L., Liang, T., Johnson, J. E. and Finn, M. G. *Angew. Chem. Int. Ed.*, **2002**, *41*, 459-462.

[31] Smith, J. C., Lee, K. B., Wang, Q., Finn, M. G., Johnson, J. E., Mrksich, M. and Mirkin, C. A. M. *Nano Lett.*, **2003**, *3*, 883-886.

[32] Muir, T. W., Sondhi, D. and Cole, P. A. *Proc. Natl. Acad. Sci. U S A*, **1998**, *95*, 6705-6710.

[33] Camarero, J. A. and Muir, T. W. *Current Protocols in Protein Science*, **1999**, 1-21.

[34] Evans, T. C., Benner, J. and Xu, M.-Q. *J. Biol. Chem.*, **1999**, *274*, 3923-3926.

[35] Muir, T. W. *Annu. Rev. Biochem.*, **2003**, *72*, 249-289.

[36] Dawson, P. E., Muir, T. W., Clark-Lewis, I. and Kent, S. B. H. *Science*, **1994**, *266*, 776-779.

[37] Tam, J. P., Lu, Y. A., Liu, C. F. and Shao, J. *Proc. Natl. Acad. Sci. USA*, **1995**, *92*, 12485-12489.

[38] Noren, C. J., Wang, J. M. and Perler, F. B. *Angew. Chem. Int. Ed.*, **2000**, *39*, 451-456.

[39] Lesaicherre, M. L., Uttamchandani, M., Chen, G. Y. and Yao, S. Q. *Bioorg. Med. Chem. Lett.*, **2002**, *12*, 2079-2083.

[40] Girish, A., Sun, H., Yeo, D. S., Chen, G. Y., Chua, T. K. and Yao, S. Q. *Bioorg. Med. Chem. Lett.*, **2005**, *15*, 2447-2451.

[41] Kiick, K. L., Saxon, E., Tirrell, D. A. and Bertozzi, C. R. *Proc. Natl. Acad. Sci. USA*, **2002**, *99*, 19-24.

[42] Saxon, E., Armstrong, J. I. and Bertozzi, C. R. *Org. Lett.*, **2000**, *2*, 2141-2143.

[43] Saxon, E. and Bertozzi, C. R. *Science*, **2000**, *287*, 2007-2010.

[44] Nilsson, B. L., Kiessling, L. L. and Raines, R. T. *Org. Lett.*, **2001**, *3*, 9-12.

[45] Kohn, M. and Breinbauer, R. *Angew. Chem. Int. Ed. Engl.*, **2004**, *43*, 3106-3116.

[46] Lewis, W. G., Green, L. G., Grynszpan, F., Radic, Z., Carlier, P. R., Taylor, P., Finn, M. G. and Sharpless, K. B. *Angew. Chem. Int. Ed. Engl.*, **2002**, *41*, 1053-1057.

[47] Agard, N. J., Prescher, J. A. and Bertozzi, C. R. *J. Am. Chem. Soc.*, **2004**, *126*, 15046-15047.

[48] Link, A. J., Vink, M. K. and Tirrell, D. A. *J. Am. Chem. Soc.*, **2004**, *126*, 10598-
      10602.

[49] Sen Gupta, S., Raja, K. S., Kaltgrad, E., Strable, E. and Finn, M. G. *Chem.
      Commun. (Cam.b)*, **2005**, 4315-4317.

[50] Beatty, K. E., Xie, F., Wang, Q. and Tirrell, D. A. *J. Am. Chem. Soc.*, **2005**, *127*,
      14150-14151.

[51] Chmura, A. J., Orton, M. S. and Meares, C. F. *Proc Natl Acad Sci U S A*, **2001**, *98*,
      8480-8484.

[52] Hodneland, C. D., Lee, Y. S., Min, D. H. and Mrksich, M. *Proc. Natl. Acad. Sci.
      U.S.A.*, **2002**, *99*, 5048-5052.

[53] Mannesse, M. L., Boots, J. W., Dijkman, R., Slotboom, A. J., van der Hijden, H. T.,
      Egmond, M. R., Verheij, H. M. and de Haas, G. H. *Biochim. Biophys. Acta*, **1995**,
      1259, 56-64.

[54] Kwon, Y., Han, Z., Karatan, E., Mrksich, M. and Kay, B. K. *Anal. Chem.*, **2004**, *76*,
      5713-5720.

[55] George, N., Pick, H., Vogel, H., Johnsson, N. and Johnsson, K. *J. Am. Chem. Soc.*,
      2004, *126*, 8896-8897.

[56] Keppler, A., Pick, H., Arrivoli, C., Vogel, H. and Johnsson, K. *Proc. Natl. Acad.
      Sci. USA*, **2004**, *101*, 9955-9959.

[57] Zhu, H., Klemic, J. F., Chang, S., Bertone, P., Casamayor, A., Klemic, K. G.,
      Smith, D., Gerstein, M., Reed, M. A. and Snyder, M. *Nat. Genet.*, **2000**, *26*, 283-
      289.

[58] Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S., Houfek, T., Mitchell, T., Miller, P., Dean, R. A., Gerstein, M. and Snyder, M. *Science*, **2001**, *293*, 2101-2105.

[59] Chong, S., Shao, Y., Paulus, H., Benner, J., Perler, F. B. and Xu, M. Q. *J. Biol. Chem.*, **1996**, *271*, 22159-22168.

[60] Perler, F. B. and Adam, E. *Curr. Opin. Biotechnol.*, **2000**, 377-383.

[61] Lew, B. M., Mills, K. V. and Paulus, H. *J. Biol. Chem.*, **1998**, *273*, 15887-15890.

[62] Perler, F. B. *Trends Biochem Sci*, **1999**, *24*, 209-211.

[63] Wu, H., Hu, Z. and Liu, X. Q. *Proc. Natl. Acad. Sci. USA*, **1998**, *95*, 9226-9231.

[64] Shi, J. and Muir, T. W. *J. Am. Chem. Soc.*, **2005**, *127*, 6198-6206.

[65] Schweitzer, B., Predki, P. and Snyder, M. *Proteomics*, **2003**, *3*, 2190-2199.

[66] Predki, P. F. *Curr. Opin. Chem. Biol.*, **2004**, *8*, 8-13.

[67] Barry, R. and Soloviev, M. *Proteomics*, **2004**, *4*, 3717-3726.

[68] Wu, G., Datar, R. H., Hansen, K. M., Thundat, T., Cote, R. J. and Majumdar, A. *Nat. Biotechnol.*, **2001**, *19*, 856-860.

[69] Ligler, F. S., Taitt, C. R., Shriver-Lake, L. C., Sapsford, K. E., Shubin, Y. and Golden, J. P. *Anal. Bioanal. Chem.*, **2003**, *377*, 469-477.

[70] Vega, R. A., Maspoch, D., Salaita, K. and Mirkin, C. A. *Angew. Chem. Int. Ed. Engl.*, **2005**, *44*, 6013-6015.

Figures and Figure Captions

Table (1). Summary of different chemoselective methods commonly used for the site-specific immobilization of proteins onto inorganic substrates.

| Protein				Surface		Type of covalent bond
Reactive group	Modification	Location	Reactivity	Reactive group	Reactivity	
R–SH	Natural Cys residue or introduced by mutation	C-terminal N-terminal Accessible loops	Nucleophile	(iodoacetamide / maleimide)	Electrophile	Thioether
R–CHO	Oxidation of glycoproteins and N-terminal Ser, Cys or Thr residues	N-terminal and glycosylation site	Electrophile	H₂N-HN / H₂N-O	Nucleophile	Hydrazone or oxime
R–CO–SR''	Express protein as C-fusion with intein then treatment with soluble R''-SH	C-terminal	Electrophile	HS / H₂N / H₂N-HN / H₂N-O	Nucleophile	Amide, hydrazide or oxamate
R–CO–NHNH₂	Express protein as C-fusion with intein then treatment with hydrazine	C-terminal	Nucleophile	OHC	Electrophile	Hyfrazone
R–CO–SH	Express protein as C-fusion with intein then treatment with bisulfide	C-terminal	Nucleophile	X (X: I, Br) / maleimide	Electrophile	Thioester
HX–Z / H₂N–R (X: O,S Z: -H,-CH₃)	Natural N-terminal Ser, Thr or Cys residues, otherwise introduced by mutation	N-terminal	Nucleophile	OHC	Electrophile	thiazolidine or oxazolidine
R–N₃	Replacement of Met residues with azidohomoalanine	Anywhere, preferentially accessible loops	Electrophile	(phosphine)	Nucleophile	Amide
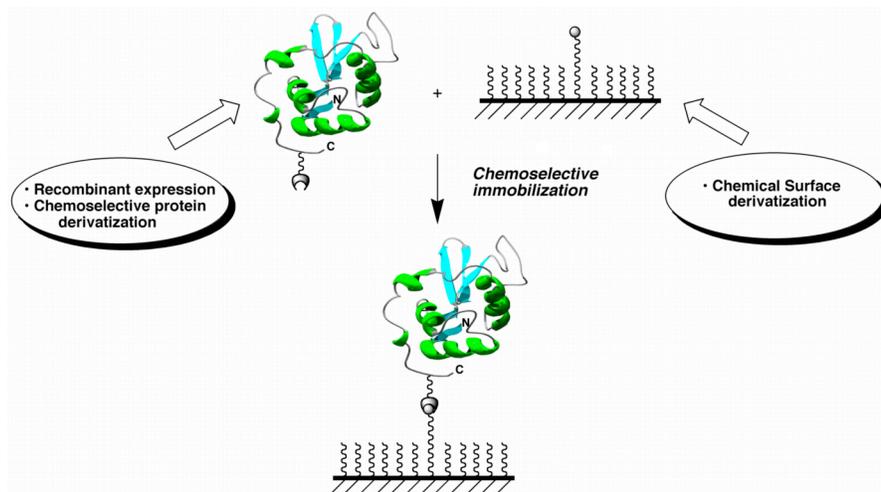
Fig. (1). Principle of chemoselectivity for the site-specific covalent immobilization of a protein onto an inorganic support.
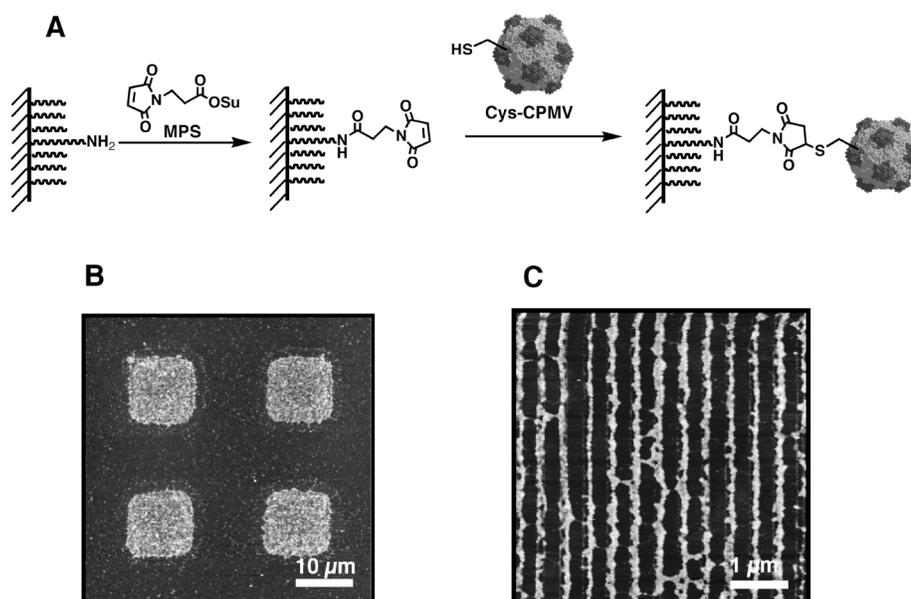

Fig. (2). Chemoselective attachment of genetically modified CPMV virus with unique cysteine residues (Cys-CPMV). (B) Atomic force microscope (AFM) height image of Cys-CPMV virus assembled on a micron-sized template containing the maleimide function. (C) AFM height image of a monolayer-thick of virions assembled on a parallel line pattern created by nanografting with a maleimide-containing linker.
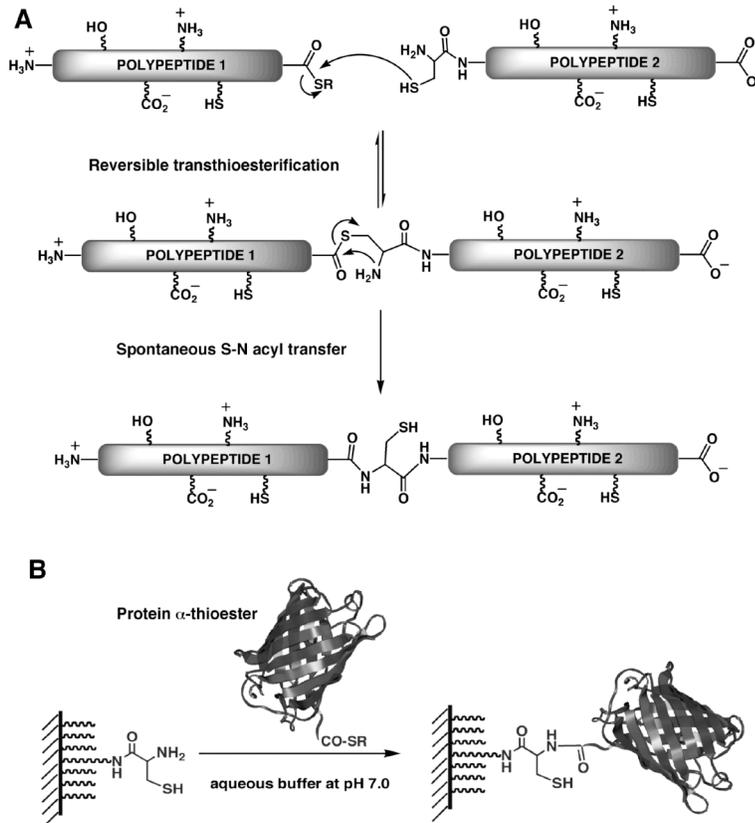
Fig. (3). (A) Principle of Native Chemical Ligation (NCL). (B) Site-specific attachment of a protein α-thioester through its C-terminus by Expressed Protein Ligation (EPL).
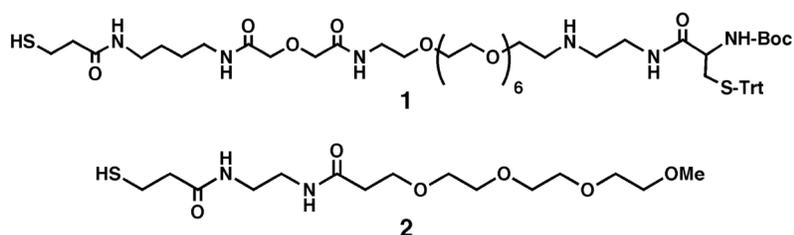
Fig. (4). PEGylated thiol linkers used to attach *C*-terminal α-thioester proteins to acryloxy-modified glass surfaces (see Fig. (**5**)).
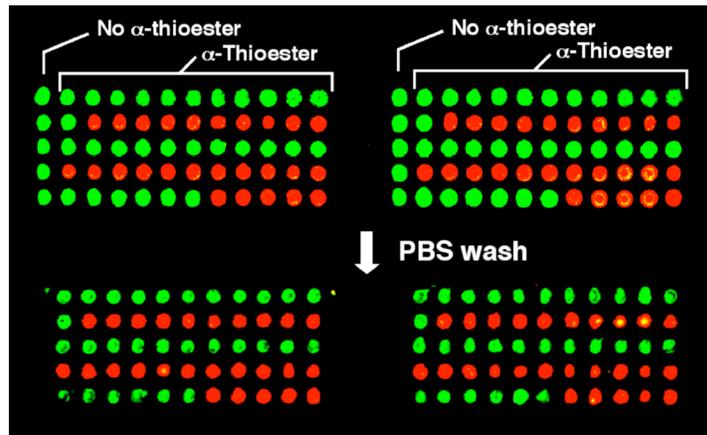
Fig. (5). Selective attachment of EGFP (green) and DsRed (red) α-thioesters onto a Cys-containing glass slide. Epifluorescence image of the glass slide after the protein micro-spotting (top) and after PBS (phosphate buffer solution) washes (bottom). EGFP with no α-thioester was used as control. Spotting was carried out using 100 µM protein solutions. Dot size is ≈ 100 µm.
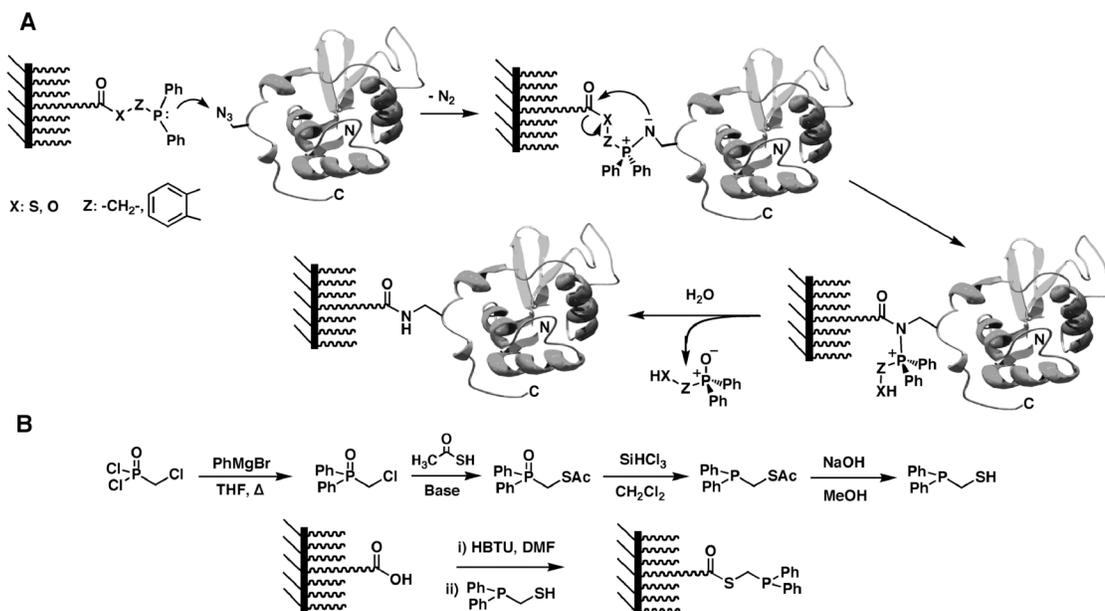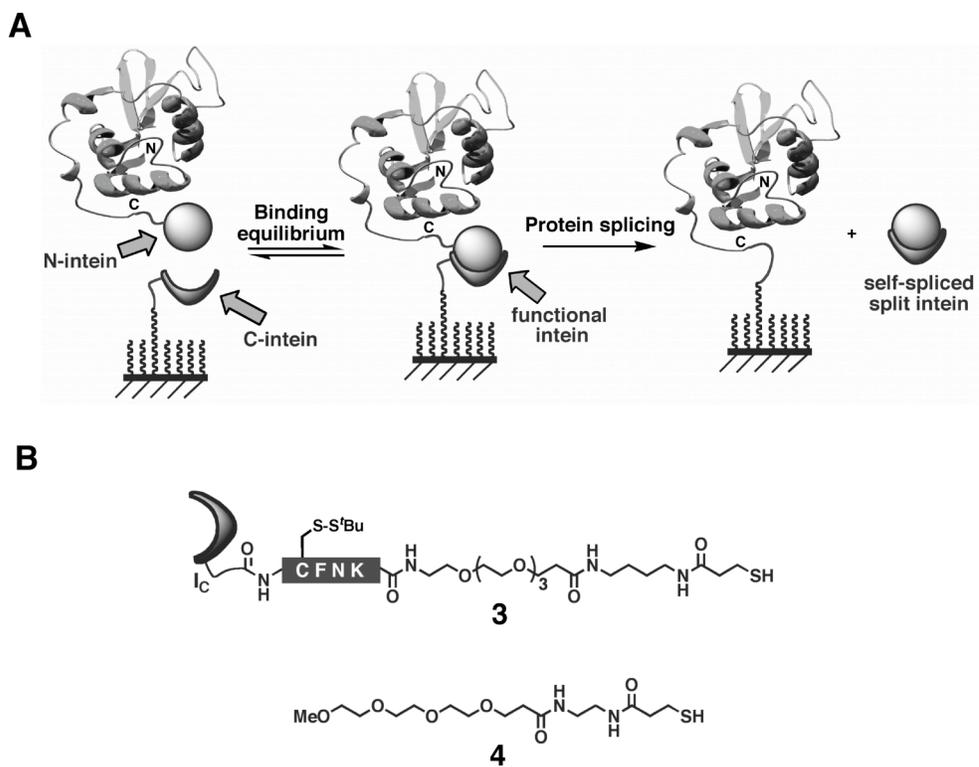
Fig. (6). (A) Chemoselective attachment of proteins to surfaces by using a modification of the Staudinger reaction. (B) Synthetic scheme for modifying a carboxylic surface with a suitable phosphinothiol for the Staudinger ligation reaction.
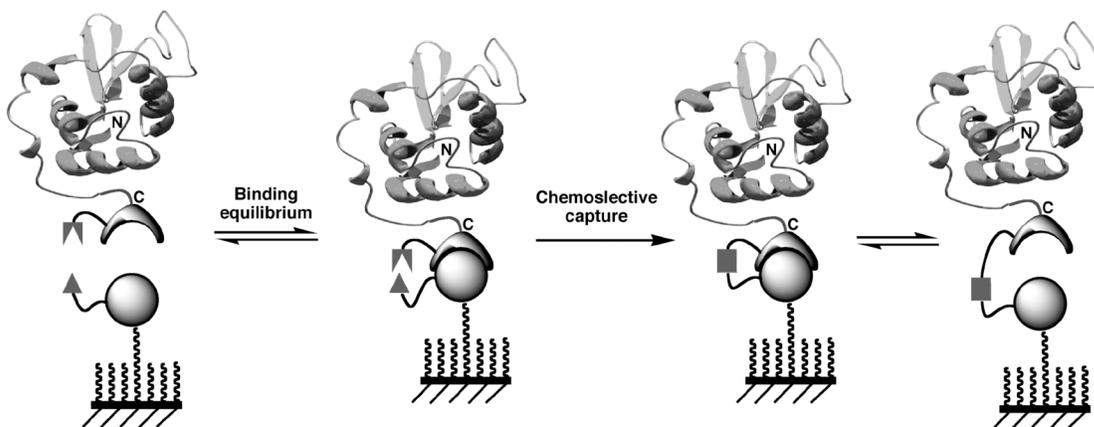
Fig. (7). Attachment of a protein to a surface by using an affinity assisted chemoselective reaction.
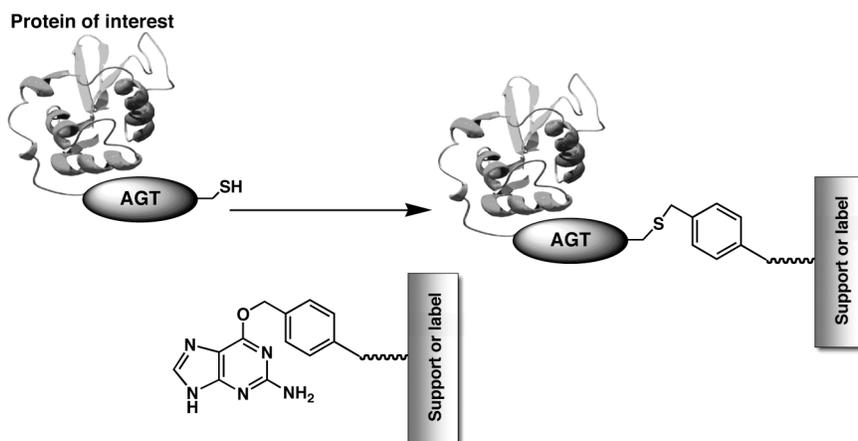

Fig. (8). Site-specific immobilization of C-terminal mutant O^6-alkylguanine-DNA alkyltransferase (AGT)-fusion proteins. In this enzymatic alkylation approach, the target protein is fused to an AGT of $\approx$ 200 residues. Selective immobilization can be accomplished by using O^6-benzylguanine-coated solid supports.

Fig. (9). A. Site-specific immobilization of proteins onto a solid support through split-intein mediated protein trans-splicing. B. Structures of linkers used for the derivatization of a glass substrate.
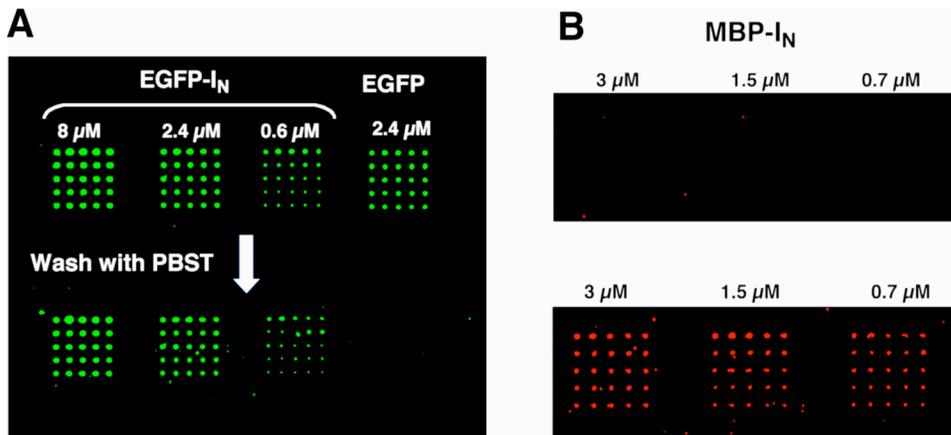
Fig. (10). Selective attachment of EGFP-I$_N$ (green) and MBP-I$_N$ (red) onto derivatized glass slides. A. Epifluorescence image of an I$_C$-coated glass slide after the spotting of different concentrations of EGFP-I$_N$ and EGFP (top) and after buffer washes (bottom). B. Different concentrations of MBP-I$_N$ were spotted onto a PEG- (top) and I$_C$-coated (bottom) glass slides. After washing, the immobilized MBP was detected by immunofluorescence. Dot size is ≈ 100 μm.
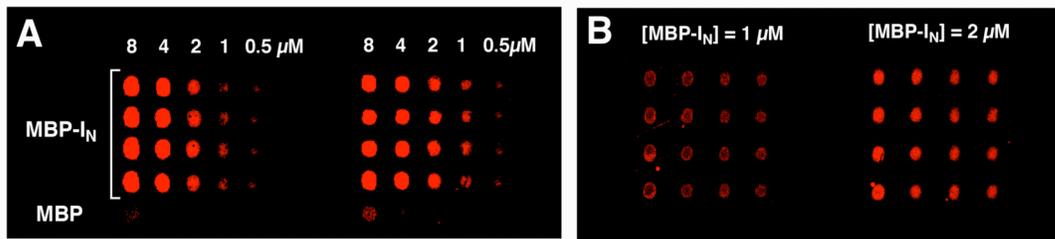
Fig. (11). Selective immobilization of MBP-I$_N$ from complex mixtures. A. Soluble cellular fraction of *E. coli* cells overexpressing MBP-I$_N$. B. MBP-I$_N$ expressed *in vitro* using a cell-free system. Protein concentrations in the cell lysate and IVT crude reaction were estimated by Western Blotting. In both cases, MBP was detected by immunofluorescence after washing. The level of non-specific background can be seen at the bottom of the slide when MBP with no DnaE N-intein was spotted.