CHAPTER II

Artificial Polypeptide Scaffold for Protein Immobilization

Abstract

An artificial polypeptide scaffold composed of surface anchor and protein capture domains was designed and expressed *in vivo*. By using a mutant *E. coli* phenylalanyl-tRNA synthetase, the photoreactive amino acid *para*-azidophenylalanine was incorporated into the surface anchor domain. Octyltrichlorosilane-treated surfaces were functionalized with this polypeptide by spin coating and photocrosslinking. The resulting protein films were shown to immobilize recombinant proteins through formation of coiled coil heterodimers.

2.1 Introduction

Protein microarray technologies are beginning to advance the field of proteomics by providing miniaturized platforms to probe the interactions and functions of proteins.[1] Presenting proteins in dense arrays enables the rapid screening of thousands of molecular events in a single experiment. This capability should facilitate the elucidation of protein profiles in organisms, the discovery of novel protein functions, and the development of systems-level understanding of biological phenomena.[2] The utility of microarrays to probe protein-protein interactions has been demonstrated by Zhu et al., where new calmodulin- and phospholipid-binding proteins were identified by screening a full scale yeast proteome microarray.[3] More recently, Nielsen et al. used antibody microarrays to profile the activation of receptor tyrosine kinases and analyze signal transduction networks in mammalian cells.[4]

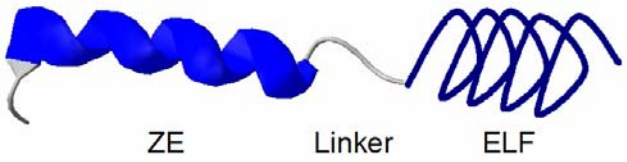
Despite the growing success of protein microarrays, it remains a central challenge to develop simple and general techniques to immobilize functional proteins onto solid supports. In certain cases, conventional immobilization methods (such as physical adsorption or covalent binding through lysine and cysteine residues) render active sites inaccessible or even denature proteins.[5] This difficulty may be overcome by engineering site specific attachment to a substrate through expression of recombinant fusion proteins bearing affinity tags. For example, the immobilization of his tag fusion proteins onto Ni-NTA functionalized slides has been shown to maintain higher protein activity than direct attachment to aldehyde slides.[3] Nevertheless, his tags do

not necessarily provide a general approach to array fabrication, because the binding interaction is sensitive to pH and to some common buffer components.[6] To solve this problem, a variety of alternative strategies are being developed. One such approach uses the strong interaction between avidin and biotin to immobilize proteins in combination with *in vivo* or *in vitro* biotinylation.[7] Other methods introduce polypeptide tags to effect selective and covalent attachment.[8] Optimal immobilization schemes should be characterized by simple cloning schemes, efficient protein expression, selective affinity and simple surface chemistry. This formidable challenge requires the design of new biomaterials that maintain protein architecture and allow specific chemistries to be utilized for immobilization.[9]

2.2 Results and discussion

We have approached this problem by creating an artificial polypeptide scaffold **1** that can be used to immobilize recombinant proteins on substrates (Figure 2.1). The polypeptide contains separate surface anchor and protein capture domains, and uses an artificial amino acid to covalently crosslink the polypeptide to surfaces. The protein capture domain functions through coiled coil association of a designed parallel heterodimeric leucine zipper pair, designated ZE and ZR. These structures are based on the sequences developed by Vinson et al.[10] with minor modifications (see materials and methods). Vinson et al. showed that this leucine zipper system has a heterodimerization affinity of 10^{-15} M, while homodimerization affinities are in the

micromolar range. The acidic component ZE is incorporated into **1** as the protein capture domain and the basic portion ZR is fused to target proteins as an affinity tag.



def gabcdef gabcdef gabcdef gabcdef gabcdef gabcd

ZE LEI EAAALEQ ENTALET EVAELEQ EVQRLEN IVSQYRT RYGPL

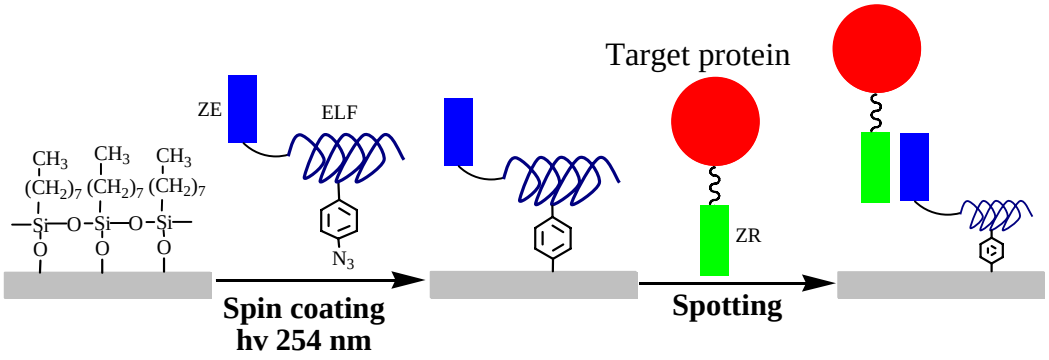
ZR LEI RAAALRR RNTALRT RVAELRQ RVQRLRN EVSQYET RYGPL

Linker $G(GS)_6G$ **ELF** $[(VPGVG)_2VPGFG(VPGVG)_2]_5VPGC$

Figure 2.1 Design of the artificial polypeptide scaffold **1** and related amino acid sequence.

An important part of the scaffold design is introduction of an elastin mimetic domain ELF for surface anchorage. ELF consists of five repeats of 25 amino acids with the sequence $(VPGVG)_2VPG\mathit{F}G(VPGVG)_2$. Because of its hydrophobic character, ELF provides strong adhesion to hydrophobic surfaces.[11] Moreover, because **1** is expressed in a bacterial host harboring a mutant *E. coli* phenylalanyl-tRNA synthetase (A294G), the phenylalanine residues in the ELF domain are partially replaced by a photoreactive nonnatural amino acid, *para*-azidophenylalanine.[12] This moiety can be used to generate covalent linkages to substrates upon UV irradiation. In order to reduce possible steric hindrance, the ZE and ELF domains are linked by a flexible spacer of 14 amino acids. The designed protein sequence was reverse-translated based

on the codons most often used in *E. coli* and expressed in the phenylalanine auxotroph strain AF-IQ.[13] Typical yields were 50 mg/L and the rate of incorporation of *para*-azidophenylalanine was approximately 45% as determined by amino acid analysis.



Scheme 2.1 Surface functionalization and coiled-coil mediated immobilization of proteins.

The successful design, *in vivo* expression and purification of **1** allowed us to prepare a functionalized surface for protein immobilization (Scheme 2.1). In this procedure, a solution of **1** (0.8 mg/ml in 50% trifluoroethanol) was spin-coated on glass slides that were pretreated with octyltrichlorosilane (OTS) to make them hydrophobic. Once dry, the protein films were irradiated with UV light.[14] Irradiation of the films covalently crosslinked the protein to the substrate through photodecomposition of the arylazide groups[15] of *para*-azidophenylalanine. Any noncovalently bound protein was removed by sonicating in 80% DMSO for 20 minutes. Measurements of the water static contact angle indicated marked changes in

wettability upon formation of the protein film; the contact angle was 60^o after photocrosslinking and sonication as compared to 107^o for the initial OTS substrates. In a final step, films were blocked with 1% casein solution to reduce nonspecific protein adsorption.

Green fluorescent protein (GFP) and glutathione-S-transferase (GST) were chosen as model systems to test the efficacy of the functionalized surface for protein immobilization. These proteins were expressed *in vivo* with the ZR tag fused to their C-termini. Proteins lacking the fusion tag were expressed as controls. Purified proteins at a concentration of 5 μM were spotted onto the surface to generate protein microarrays (Fig 2.2a). The arrays were incubated in a humid chamber for 1 hour and then thoroughly washed twice with PBS-Tween buffer (PBS plus 0.5% Tween-20) to remove nonspecifically bound protein. Each array was probed with a mixture of cy3-anti-GST and alexa647-anti-GFP (4 μg/ml each), washed, and scanned with a Genepix microarray scanner. As expected, spots containing fusion proteins showed much stronger signals than control proteins. The average signal to noise ratio SNR[16] of GST-ZR was 196 ± 20 and that of GFP-ZR was 43 ± 4. Without the ZR fusion, GST spots yielded weaker detectable signals (SNR 15 ± 3[17]), and GFP spots could not be distinguished from background. The sensitivity of the method is high; SNR ratios of 4 or 2 (GST-ZR or GFP-ZR) were obtained when proteins were spotted at concentrations as low as 50 nM. These qualities encouraged us to examine immobilization of ZR tagged proteins directly from crude cell lysates. Cell lysates containing overexpressed fusion or control proteins were spotted onto functionalized surface and detected by the procedure described above. As shown in Fig 2.2b, significant protein attachment occurred only when the complementary zipper fusion

tag was present.

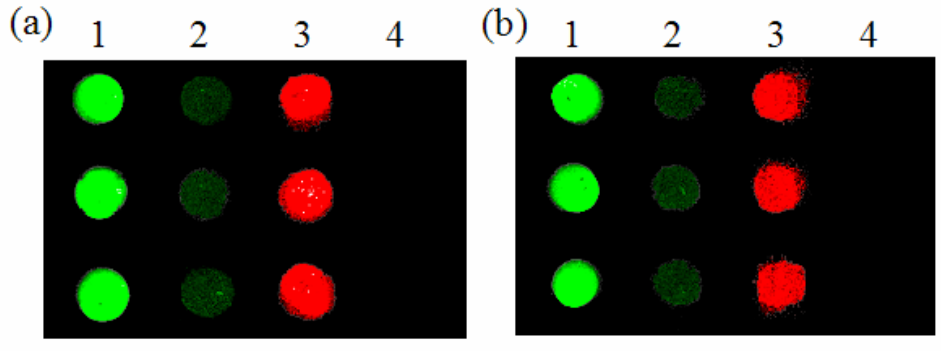


Figure 2.2 Immobilization and detection of proteins on polypeptide functionalized surfaces. Printed spots were detected with a mixture of cy3-anti-GST and alexa647-anti-GFP. (a) 5 μM purified proteins (1) GST-ZR, (2) GST, (3) GFP-ZR, (4) GFP. & (b) Cell lysates (1) GST-ZR, (2) GST, (3) GFP-ZR, (4) GFP. The spots are 200 μm in diameter.

The protein immobilization method presented here has several advantages over traditional methods. First, spin coating plus photoimmobilization provides a simple and convenient route to uniform protein films. This procedure requires a fabrication time of minutes and yields dense surface coverage. Second, the heterodimeric association of this leucine zipper system is highly specific and stable. In fact, we have found that the heterodimer forms even in 8 M urea solution over a wide range of pH values (tested from pH 4.0 to 8.0). This stability expands the range of working conditions to stringent situations where other methods are not applicable. Third, considering the relatively small size of the zipper tag (43 amino acids), it is unlikely

that the function of the fusion proteins will be compromised. Finally, direct immobilization of fusion proteins from crude cell lysates makes it feasible to fabricate protein arrays in high throughput fashion by eliminating time consuming and costly purification steps.

2.3 Materials and methods

(This section appeared as supporting information for Zhang, K. C.; Diehl, M. R.; Tirrell, D. A. J. Am. Chem. Soc. ***2005****, 127, 10136-10137.)*

2.3.1 Cloning and expression of the polypeptide scaffold

(a) Sequence optimization

def gabcdef gabcdef gabcdef gabcdef gabcdef gabcd

ZE LEI EAAALEQ ENTALET EVAELEQ EVQRLEN IVSQYRT RYGPL

ZR LEI RAAALRR RNTALRT RVAELRQ RVQRLRN EVSQYET RYGPL

Figure 2.3 Sequence modifications of the heterodimeric leucine zipper.

Compared to the leucine zipper sequences reported by Vinson et al.,[10] some modifications (shown in red) were made to enhance the application described here. Phenylalanine residues in the first heptads of both coils were replaced by alanine. In order to provide an additional salt bridge, glutamic acid in the fifth heptad of ZE was changed to arginine and isoleucine in ZR was changed to glutamic acid.

(b) Vector construction

The genes for ZE and ZR were constructed by assembling synthetic oligonucleotides through PCR. A *BamH*I fragment encoding ZE with a C-terminal flexible linker $(Gly\text{-}Ser)_6$ was ligated in frame with a synthetic gene encoding ELF in a pQE60 (Qiagen) plasmid developed in this laboratory. To avoid *in vivo* degradation at 37 ^{o}C and enhance the expression yield, a dicistronic construct (Figure 2.4) was used to express ZE$(gs)_6$ELF. First, a *Nhe*I fragment encoding mutant *E. coli* PheRS (A294G) was inserted into expression plasmid pQE60 to yield pQE-FS*. Then the ZR gene was ligated into the *Bam*HI site of pQE-FS* to yield ZR-pQE-FS*. Finally, a DNA fragment encoding ZE$(gs)_6$ELF and containing a ribosome binding site was amplified and ligated into the *Bpu*1102I site of ZR-pQE-FS*. The resulting plasmid (ZRhis-ZE$(gs)_6$ELF) was transformed into phenylalanine auxotroph *E. coli* strain AF-IQ.

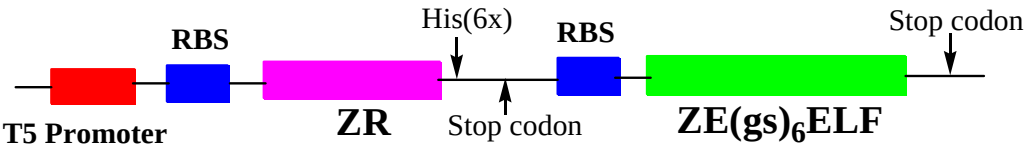


Figure 2.4 Dicistronic expression construct for the polypeptide scaffold.

(c) Protein expression and purification

Cultures were grown at 37 ^{o}C in M9 minimal medium supplemented with glucose (0.4 wt %), thiamine (5 mg L^{-1}), $MgSO_4$ (1 mM), $CaCl_2$ (0.1 mM), 20 amino

acids (15 mg L^{-1} Phe, 40 mg L^{-1} other amino acids), and antibiotics (200 μg/ml ampicillin and 35 μg/ml chloramphenicol). At an optical density of 1.0 at 600 nm (OD_{600}), the culture was supplemented with *para*-azidophenylalanine (0.3 g L^{-1}). After incubation for 15 minutes, IPTG (2 mM) was added to induce protein expression at 37°C for 5 hours. The cell pellet was lysed in 8 M urea solution, followed by freeze-thawing and sonication. The cell lysate was passed through a Ni-NTA column and washed with 8 M urea solution (pH=6.3). $ZE(gs)_6ELF$ was eluted with 6 M GuHCl solution (pH=7.0) and collected. The collected protein sample was subjected to dialysis against ddH_2O and freeze drying.

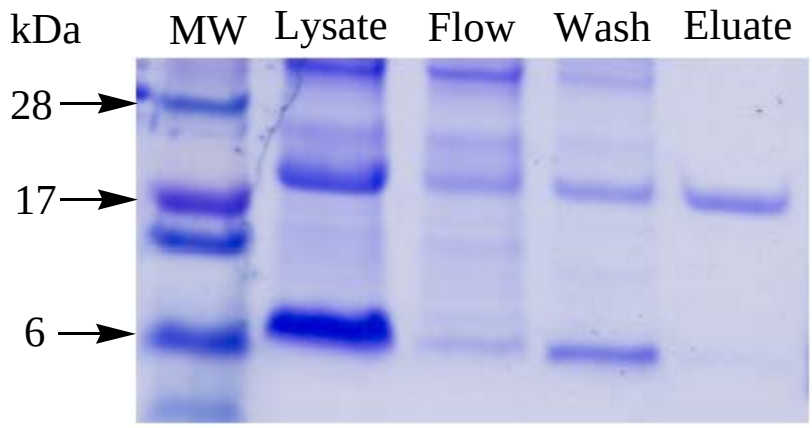


Figure 2.5 Purification of $ZE(gs)_6ELF$ by Ni-NTA affinity chromatography.

2.3.2 Cloning and expression of the fusion proteins

A GFP gene fragment was PCR amplified from pGFPuv (Clontech) and cloned into the *Pst*I site of pQE9 (Qiagen) to yield pQE9-GFP. The ZR gene assembled from synthetic oligonucleotides was digested by *Hind*III and ligated into pQE9-GFP to

yield the expression vector pQE9-GFPZR. To construct the GST-ZR fusion protein, the ZR fragment was inserted into the *Bam*HI site of expression vector pGEX-2TK (Amersham Pharmacia Biotech). The resulting expression plasmids were transformed into *E. coli* strain BL-21 and protein expression was performed in 2XYT rich medium induced by 1 mM IPTG. GFP and GFP-ZR, both of which carry N-terminal hexahistidine tags, were purified on Ni-NTA spin columns, while GST and GST-ZR were purified on glutathione columns.

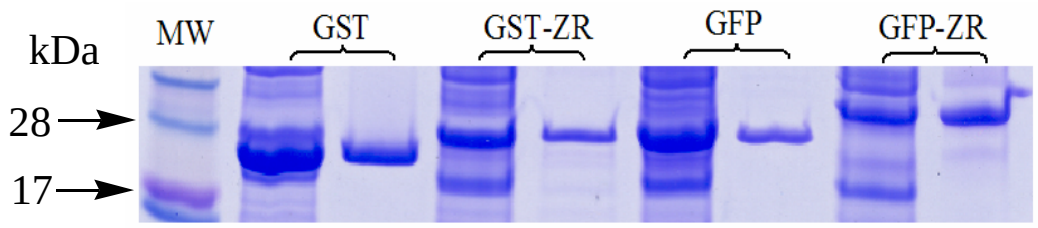


Figure 2.6 Expression and purification of model proteins.

2.3.3 Protein array procedure

(a) Substrate preparation

Standard glass slides (Corning) were immersed into concentrated H_2SO_4 for 1 hour. After washing thoroughly with water, they were immersed into a boiling solution of 1/1/5 (v/v) H_2O_2 (30%) /NH_4OH (30%)/H_2O for 30 minutes. Then the slides were gently shaken in 1% octyltrichlorosilane in toluene for 30 minutes. Finally they were cleaned twice in methanol and twice in DI water. The functionalized slides were cured at 110^{o}C for 30 minutes.

(b) Polypeptide film preparation

A solution of $ZE(gs)_6ELF$ (80 µl, 0.8 mg/ml) in 50% trifluoroethanol was applied to a glass slide and spun at 1,500 RPM for 45 seconds. The resulting film was irradiated in a photochemical reactor equipped with 254 nm wavelength UV lamps for 5 minutes. After thorough washing with 80% DMSO and ddH_2O, the slides were dried for microarray experiments. They were blocked with 1% casein for 1 hour and then stored at 4^oC for later use.

(c) Protein spotting and detection

GFP and GST fusion or control proteins diluted in printing buffer (0.5% casein, 0.5% polyvinylpyrrolidone, 15% sorbitol, 0.05% sarkosyl, 1X PBS) to a final concentration of 5 µM were spotted on slides with a custom-built microarrayer equipped with a MicroQuill® 2000 Array Pin (Majer Precision). After incubation for 1 hour in a humid chamber (saturated K_2SO_4), the slides were washed twice, each time for 15 minutes with PBS-Tween buffer (PBS plus 0.5% Tween-20). Finally, the immobilized proteins were probed with a mixture of cy3-antiGST and alexa647-antiGFP (4 µg/ml each) for 2 hours. After thorough washing with PBS-Tween buffer, fluorescence scans were acquired on a Genepix 4200A microarray scanner .

For cell lysate experiments, cell pellets from 100 ml 2XYT expression cultures were lysed with 5 ml 8 M urea solution and centrifuged to obtain clear supernatants.

Printing solutions were prepared by diluting the supernatants 20-fold into printing buffer. The protocol for microarray preparation was the same as that described above.

Acknowledgment. This work was supported by the NSF-sponsored Center for the Science and Engineering of Materials at the California Institute of Technology.

2.4 References

(1) (a) MacBeath, G.; Schreiber, S. L. *Science* **2000**, *289*, 1760. (b) Zhu, H.; Klemic, J. F.; Chang, S.; Bertone, P.; Casamayor, A.; Klemic, K. G.; Smith, D.; Gerstein, M.; Reed, M. A.; Snyder, M. *Nat. Genet.* **2000**, *26*, 283. (c) Haab, B. B.; Dunham, M. J.; Brown, P. O. *Genome Biol.* **2001**, *2*, 1.

(2) (a) MacBeath, G. *Nat. Genet.* **2002**, *32(suppl.)*, 526. (b) Phizicky, E.; Bastiaens, P. I.; Zhu, H.; Snyder, M.; Fields, S. *Nature* **2003**, *422*, 208.

(3) Zhu, H.; Bigin, 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.; Snyder, M. *Science* **2001**, *293*, 2101.

(4) Nielsen, U. B.; Cardone, M. H.; Sinskey, A. J.; MacBeath, G.; Sorger, P. K. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9330.

(5) Zhu, H.; Snyder, M. *Curr. Opin. Chem. Biol.* **2003**, *7*, 55.

(6) Paborsky, L. R.; Dunn, K. E.; Gibbs, C. S.; Dougherty, J. P. *Anal. Biochem.***1996**, *234*, 60.

(7) (a) Boer, E.; Rodriguez, P.; Bonte, E.; Krijgsveld, J.; Katsantoni, E.; Heck, A.; Grosveld, F.; Strouboulis, J. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 7480. (b) Lesaicherre, M. L.; Lue, R. Y.; Chen, G. Y.; Zhu, Q.; Yao, S. Q. *J. Am. Chem. Soc.* **2002**, *124*, 8768. (c) Lue, R. Y.; Chen, G. Y.; Hu, Y.; Zhu, Q.; Yao, S. Q. *J. Am. Chem. Soc.* **2004**, *126*, 1055. (d) Yin, J.; Liu, F.; Li, X.; Walsh, C. T. *J. Am. Chem. Soc.* **2004**, 126, 7754.

(8) (a) Hodneland, C. D.; Lee, Y. S.; Min, D. H.; Mrksich, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5048. (b) Kindermann, M.; George, N.; Johnsson, N.; Johnsson, K. *J. Am. Chem. Soc.* **2003**, *125*, 7810. (c) Camarero, J. A.; Kwon, Y.; Coleman, M. A. *J. Am. Chem. Soc.* **2004**, *126*, 14730.

(9) Langer, R.; Tirrell, D. A. *Nature* **2004**, *428*, 487.

(10) Moll, J. R.; Ruvinov, S. B.; Pastan, I.; Vinson, C. *Protein Sci.* **2001**, *10*, 649.

(11) (a) Shimazu, M.; Mulchandani, A.; Chen, W.; *Biotechnol. Bioeng.* **2003**, *81*, 74. (b) Nath, N.; Chilkoti, A. *Anal. Chem.* **2003**, *75*, 709. (c) Wada, A.; Mie, M.; Aizawa, M.; Lahoud, P.; Cass, A. E.; Kobatake, E. *J. Am. Chem. Soc*. **2003**, *125*, 16228. (d) Hyun, J.; Lee, W. K.; Nath, N.; Chilkoti, A.; Zauscher, S. *J. Am. Chem. Soc*. **2004**, 126, 7330.

(12) Kirshenbaum, K.; Carrico, I. S.; Tirrell, D. A. *ChemBioChem.* **2002**, *3*, 235.

(13) Sharma, N.; Furter, R.; Kast, P.; Tirrell, D. A. *FEBS. Lett.* **2000**, *467*, 37.

(14) Photocrosslinking was done in a Rayonet photoreactor equipped with 254-nm lamps for 5 minutes at room temperature.

(15) Borden, W. T.; Gritsan, N. P.; Hadad, C. M.; Karney, W. L.; Kemnitz, C. R; Platz, M. S. *Acc. Chem. Res.* **2000**, 33, 765.

(16) SNR= (signal-background)/ (standard deviation of background).

(17) The nonspecific adsorption signal was caused by rapid drying of microspots during printing. We didn't observe signals from GST spots if we used a manual pipette to spot 1 μL protein solutions.